3

Mechanisms of Carcinogenesis

3.1 Molecular Hallmarks of Cancer

Summary

> Cancer is a multi-step process in which cells undergo metabolic and behavioural changes, leading them to proliferate in an excessive and untimely way

> These changes arise through modifications in mechanisms that control cell proliferation and lifespan, relationships with neighbouring cells, and capacity to escape the immune system

> Modifications that lead to cancer include genetic changes that modify the DNA sequence. Another way to change the programme of cells is to modify the conformation of chromatin, the structure that wraps up DNA and regulates its access by DNA reading, copying and repair machineries. Such changes are called "epigenetic"

> Among the 23 000 or so genes that constitute the human genome, a few hundred are commonly targeted by genetic or epigenetic changes. These genes are parts of networks of genes that regulate cell division, differentiation and life span

> The emergence of technologies for genome-wide analysis of genetic and epigenetic changes is advancing our capacity to map patterns of alterations specific for each particular cancer, paving the way to personalised medicine based on molecular diagnosis

> Advances in cancer molecular biology also identify new ways to inhibit the growth of cancer, leading to new, more selective and less toxic forms of cancer chemotherapy

Cancer is a complex disease that is very variable in its presentation, development and outcome from one patient to the other. The same heterogeneity and variability exist at the cellular and molecular level. Cancer is a multi-step process during which cells undergo profound metabolic and behavioural changes, leading them to proliferate in an excessive and untimely way, to escape surveillance by the immune system, and ultimately to invade distant tissues to form metastases [1]. These changes arise through the accumulation of modifications in the genetic programmes that control cell proliferation and lifespan, relationships with neighbouring cells, and capacity to escape the immune system. This process results in the formation of a mass of deregulated cells, which can be qualified as "outlaw" because they do not obey the rules that control normal cell growth and behaviour. Such a mass may be asymptomatic for a long time. However, it will ultimately grow to perturb physiological functions, giving rise to multiple symptoms, depending upon the location and size of the mass, and of the spread of cancer cells within the organism.

Cancer genes

The genetic programmes targeted by cancer are present in scores of genes that are dispersed throughout the human genome. It is believed that human DNA contains about 23 000 genes. Several thousand of these genes (3000–5000) encode proteins involved in genetic programmes that are deregulated in cancer. A dysfunctional gene can result in the production of abnormal levels of a critical protein (either too much or too little), the production of an aberrant protein (with either gain or loss of function), or the complete absence of the protein. For example, mutation in a gene called KRAS turns a small protein located just inside the cell membrane into an amplifier of cell growth signalling. This protein normally works as a signalling intermediate between receptors for growth factors at the cell surface and the molecular wiring systems that send growth signals to the cell nucleus to enact cell division. When the KRAS gene is mutant,

the corresponding protein is behaving like a switch locked in the "on" position generating a permanent cell division signal. Mutations of KRAS are common in many cancers, such as colorectal cancers (in 30–40% of the cases) or adenocarcinomas of the lung (in 20–30% of the cases). Such an activated gene is called an "oncogene" because it promotes cell proliferation. In contrast, some genes contribute to cancer development when they are inactivated. This is, for example, the case of the TP53 gene. This gene encodes a protein that naturally acts like an "emergency brake" to avoid inappropriate cell division. Mutation in this gene disrupts the protein, which becomes unable to stop the proliferation of cells when needed. Mutations in TP53 are found in almost every kind of cancer. Such a gene that contributes to cancer development through loss of its function is called a tumour suppressor, because in normal conditions its active products work as a brake to suppress cancer growth [2].

Cellular origin and progression of cancer

Many cancers arise from just one cell (or from a small number of cells) [3]. To become cancerous, this cell must acquire several changes in oncogenes and tumour suppressor genes that will make the cell capable of proliferating well beyond its normal limit. This process will result in the formation of a clone of "outlaw" cells. If such a clone is tolerated by the organism and allowed to remain unperturbed, it may continue to proliferate and, during this process, the cells it contains will accumulate more and more modifications. In such a disrupted context, only the fittest and the most aggressive cells will thrive, taking the place of other, less disorganised cells. This is how tumours develop to become malignant. This is also why cancer is difficult to treat: when patients are given a drug that kills cancer cells with great efficacy, the few cells. This approach is that survive are those that have undergone changes that make them resistant to the drug. This very small group of residual cells may be enough for the cancer to relapse in a form that is worse than its initial form.

For the oncologist and the cancer pathologist, cancer is best described as a progressive disease. It starts as a small, inconspicuous lesion that generally remains confined to its tissue of origin and is considered as clinically benign since, when detected at an early stage, it may be completely resected and may not cause the patient's death. Sometimes, these small lesions appear within a tissue area that is affected by a chronic inflammatory disease, such as cirrhosis in the liver, gastritis in the stomach or intestinal metaplasia (Barrett's oesophagus) in the lower oesophagus. These chronic diseases that represent a favourable terrain for cancer occurrence are called "precursor diseases". When undetected at an early, benign stage, cancer has a chance to develop and progress not only in size but also in its capacity to interfere and perturb neighbouring cells. Such cancers will not be confined any longer: they will spread within the organ affected and then will disseminate to neighbouring organs. They also enter the lymphatic vessels to spread into lymph nodes. Through the lymph or bloodstream, they can travel to distant organs and form colonies, the metastases, in general located in bone, lung, liver or brain. Tumour dissemination is often facilitated by the fact that cancer cells promote angiogenesis, that is, the synthesis of new small blood vessels dedicated to tumour vascularisation, thus improving tumour supply in oxygen and nutrients.

Disseminated cancers are much more difficult to treat. In addition to localised therapy targeting tumour foci (surgery, radiotherapy) they require systemic therapies using cytotoxic drugs (chemotherapy). Chemotherapy is based on the use of toxic substances that interfere with DNA and cell division to preferentially kill cancer cells. This approach is based on the hypothesis that the latter may be more sensitive than normal cells since they replicate their DNA more often in order to divide, thus providing a larger window of opportunity for cell killing through DNA damage. Recently, new methods using molecules that target one particular type of cancer-related molecular changes have been introduced clinically [4]. They include,

for example, antibodies directed against cell-surface molecules expressed by cancer cells (such as trastuzumab, an antibody that inhibits a cell-surface receptor called HER-2 which is often over-expressed in breast cancer) or drugs that block the activity of enzymes activated in cancer cells (such as imatinib, which blocks an enzyme activated in many tumours of the gastrointestinal stroma, or erlotinib, which inhibits the enzyme activity associated with the receptor of the epidermal growth factor. Thus, the notion of time and progression is critical in carcinogenesis. This concept has been confirmed in studies using laboratory animals. These have shown that the development of cancer in mice or rats exposed to carcinogens occurred through different steps, with "initiation" (during which the carcinogen creates mutations in the DNA of normal cells), being followed by "promotion" (in which the initiated cells develop a growth advantage over their neighbours allowing them to produce a distinct lesion, and then by progression (in which tumours become more and more aggressive through the accumulation of supplementary genetic and epigenetic modifications). As a result, a number of oncogenes and tumour suppressors are frequently altered in many cancers, irrespective of the organ site or the cause of the disease. The products of these genes are all part of a network of factors that work together to control cell proliferation, differentiation and survival

How do genes become disrupted?

Cancer may start when small changes called mutations occur in the DNA sequence[5]. They can be limited to a single base change, thus changing one of the 3 bases that define a codon and leading to the selection of a different amino-acid to be integrated into a protein. In some cases, this is enough to dramatically change the activity of that protein. Other DNA alterations can affect a large number of bases, sometimes removing from the genome a large stretch of DNA that contains several genes, or translocating it elsewhere into the genome to form new genes made of the fusion of non-contiguous DNA segments, thus leading

to the synthesis of new, abnormal proteins. Such changes, whatever their size, are called "genetic alterations" or "mutations". Significantly, these changes can be detected by sequencing the DNA of cancer cells.

The way DNA is read and copied is critically dependent upon the way in which DNA is compacted, packaged and organised. There are "closed" DNA areas (which are locked for editing and copying) and "open" ones (which the cell can copy, read, and use to produce RNA and proteins). Thus, another way to change the programme of cells, aside from DNA mutation, is to modify the overall packaging in order to shut down genes that lie in open areas or to switch on those that lie in closed areas. Such changes cannot be detected just by sequencing DNA. They require the analysis of chemical modifications that regulate the accessibility and readability of DNA. These changes are called "epigenetic".

The role of genetic changes in cancer has been recognised for over 50 years and scientists have now built a long catalogue of genes that are mutated in cancer. In contrast, the role of epigenetic changes has been recognised only relatively recently [6]. In describing the mechanisms of cancer, genetic and epigenetic changes have to be considered together, as two sides of the same coin. They respond to each other and influence each other, generating pathways of sequential changes that determine how a given cell will progressively acquire the characteristics of a cancer cell. Both genetic and epigenetic changes are universally present in human cancer: they induce changes in gene expression that dividing cells can transmit to their daughters over many cell generations. Disrupted epigenetic states may result in functional consequences equivalent to those induced by genetic alterations.

The causes of genetic and epigenetic changes are numerous. The genome contains 3 billion base pairs and, as any molecule in or body, each of them can be modified by reaction with a variety of chemical, physical or toxicological

agents. For example, a number of chemicals classified as carcinogens can attack DNA bases, bind to them and induce modifications in the coding sequence. UV light can form typical changes by bridging together adjacent cytosines to form a dipyrimidine dimer. This results in double mutations, where two Cs are replaced by two Ts. These mutations are "signature" of mutagenesis by UV light in skin cancers. Ionizing radiations induce single or double DNA strand breaks. The main cause of DNA damage, however, is "hiccups" in the physiological processes of DNA replication and repair. Each time a cell divides, it has to produce a perfectly accurate copy of the 3 billion base pairs of its DNA. This process is tightly controlled by very elaborate DNA proofreading and repair systems. However, errors may occur and remain unrepaired. If such error occurs in a gene involved in cancer, this may result in its disruption, conferring on cells a new property that may make them better adapted to life within the deregulated system of a growing cancer mass. The cell will thus thrive in these conditions, and progressively take preeminence to become part of a malignant tumour.

What does it take for a cell to become a cancer cell?

In many respects, a cancer cell is a rogue cell that escapes the laws and rules governing cell community life, thus attaining an independent survival advantage. In doing so, cancer cells strive to adapt and fight off the defence systems of the organism and progressively adopt an aggressive, invasive behaviour. Cancer cells become able to travel within the body and to home preferentially in hospitable organ environments as metastases. Metastatic cancer cells have become so good at adapting themselves to new conditions that they resist many attempts at killing them, including cytotoxic drugs or radiation treatments. This is the reason why most cancers are best treated at an early stage, at a time when cancer cells still have limited adaptive capacity and are thus unable to bypass the effects of treatment.

Recent experimental studies have identified the minimum number of steps needed to develop a fully cancerous cell [7]. Three fundamental rules must be violated. The first is that cells should proceed to divide only when they receive appropriate signals. To break this rule, the cell has to permanently activate cell division by switching on the circuits that become normally activated when the cell is stimulated by a hormone or a growth factor. Rule number two specifies that when confronted by stressful or improper conditions for DNA replication, cells activate self-destruction programmes rather than allow DNA replication to proceed in conditions where genes may become damaged. To bypass these auto-destruction programmes, the cell has to get rid of its safety brakes, which normally prevent aberrant or excessive cell division. These brakes are controlled by two master genes, RB1 (also called the Retinoblastoma gene) and TP53 (which produces the p53 protein, a stress sensor that normally prevents cells from dividing when their environment is disturbed). When these two brakes are removed by mutation, cells can not only divide but also avoid entering programmed cell death, thus allowing the formation of a tumour mass. Rule number 3 determines that normal cells divide only a limited, fixed number of times. In other words, cells have a "division counter" that prevents them from replicating their DNA beyond a certain, predefined number of rounds. Normal cells are only capable of replicating their DNA and dividing for a finite number of times, due to a particular structure at the end of each chromosome called the telomere. The telomere is made of small repeats of DNA sequence which become eroded each time the cell divides. When all repeats are gone, the cell cannot divide any longer and becomes a senescent cell. In the cancer cell, the activation of an enzyme called telomerase allows the addition of new repeats at the end of chromosomes, thus allowing the cell to divide well past the finite number of divisions it has been programmed to make. This process is equivalent to the acquisition of a form of "permanent cell youth".

Achieving these three functional changes is enough for the cell to become cancerous. But

at the molecular level, this is not a simple operation. Taken separately, each of these changes may upset normal cell function and lead it to destruction in a process called apoptosis, a type of "cell suicide" that eliminates abnormal cells. Thus, the central problem of a would-be cancer cell is to operate all these changes in a coordinated manner. This is where both genetic susceptibility and environmental changes play a major role. Genetic susceptibility can confer on the normal cells of some people a greater capacity to make rapid changes, thus increasing the chances that they can occur simultaneously in a single cell. Environmental changes may act as natural selection to allow the survival of abnormal cells that appear to be fitter than normal cells in perturbed conditions. This is why cancer is a disease in which both genetic and environmental changes play such important roles. From the molecular point of view, these roles cannot be separated.

The Cancer Box

Despite their intrinsic diversity, cells operate along common schemes in the conduct of the basic processes that control cell proliferation and death. As a result, a number of oncogenes and tumour suppressors are frequently altered in many cancers, irrespective of the organ site or the cause of the disease (Table 3.1.1). The products of these genes are all part of a network of factors that work together to control cell proliferation, differentiation and survival. Figure 3.1.1 represents the outline of what can be defined as "the cancer box", that is, the core network of genes and processes that have to be altered in any cancer cells. This cancer box involves three main signalling processes. Two of them are growth-promoting processes and one is a growth-suppressive mechanism. One of the growth-promoting processes uses as its main effector a protein called beta-catenin. This protein has several intracellular roles. It can locate at the intracellular face of the cell membrane where it plays a role as a component of cell-to-cell junctions and of the intracellular fibre skeleton of the cell. It can also locate in the cytoplasm, where it forms part of

a complex that receives growth and proliferation signals captured at the cell surface. When activated by such signals, it can relocate into the nucleus where it stimulates the expression of genes involved in cell proliferation. The gene encoding beta-Catenin, CTNNB1, is disrupted by mutation in 10 to 20% of various epithelial cancers (e.g. lung, breast, liver or colon cancers). Other genes involved in this process are APC (often mutated in colon cancer), MYC and CCND1 (encoding cyclin D1). The other major growth signalling process involved in the cancer box includes cell-surface receptors such as EGFR (the epidermal growth factor receptor), a protein that extends on both sides of the cell membrane. On the external side, it captures growth factors present in the bloodstream or in intercellular spaces. On the internal side, it possesses an enzymatic activity, a tyrosine kinase, which becomes activated when the receptor binds a growth factor. The tyrosine kinase then initiates a cascade of intracellular signals akin to a chain reaction that propagates through amplifying molecules such as the product of the KRAS gene. The ultimate effect of these signals is to activate cell proliferation by stimulation of the progression of the cell cycle. To counteract these signals, the main anti-proliferative process is controlled by the TP53 gene. Its product, p53, can be best described as a sensor of stress, in particular DNA damaging stress. When the cell DNA is damaged and not repaired, p53 senses this abnormality, accumulates in the nucleus and activates a large number of anti-proliferative mechanims, often simultaneously. These anti proliferative mechanisms can block cell cycle progression (thus counteracting the proliferative effects of the two processes described earlier), push the cells to differentiate (thus driving them towards a status were they do not proliferate) or induce a cell suicide programme called apoptosis (leading the cell to self-destruct its DNA and other components to leave only small bodies that are eliminated by the macrophages, the specialised, 'garbage'-collecting cells in the tissues).

The key to the cancer box lies in the way these three processes are interconnected. The main

connection is ensured by a very special chromosomal locus, located at the far end of the short arm of chromosome 9. It contains a gene called CDKN2a. This locus is quite unique in the fact that it is made of two overlapping genes that use the same DNA segments as templates for RNA and protein synthesis. In other words, this locus can direct the synthesis of two different proteins that do not have a single common amino-acid. One is called p16 and is a negative regulator of cell cycle (thus exerting anti-proliferative effects). P16 belongs to a family of regulators called CDK

inhibitors, that is, factors that inhibit enzymes (cyclin-dependent kinases) that drive cells to a higher rate of proliferation. By blocking these enzymes, CDK inhibitors prevent cell division and induce a mechanism called cell cycle arrest. The other is called p14ARF (for Alternative Reading Frame) and controls the activation of p53. Thus, through its two products, this single gene controls the connection between the various components of the cancer box. Therefore, it is not unexpected that the CDKN2a gene is altered by several mechanisms in almost every cancer.
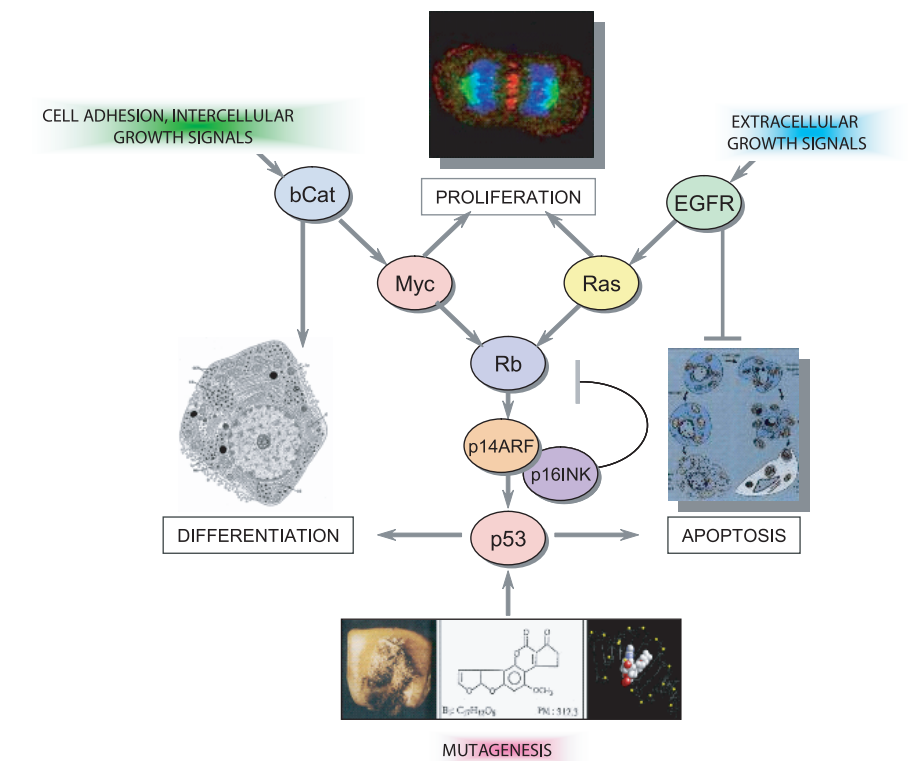

CELL ADHESION, INTERCELLULAR GROWTH SIGNALS

EXTRACELLULAR GROWTH SIGNALS

PROLIFERATION

bCat EGFR

Myc Ras

Rb

p14ARF

p16INK

DIFFERENTIATION p53 APOPTOSIS

MUTAGENESIS

Fig. 3.1.1 *The Cancer Box*
This figure illustrates how several genes may cooperate in cancer development. The three pictures illustrate the 3 phases of a cell's life: division (top), differentiation (left) and programmed cell death (right). Important genes and their cooperation are represented by arrows. Cell adhesion signals are transmitted through betaCatenin towards the components of the cell division machinery, which converge on the RB1 gene (a regulator of cell cycle). Growth signals also converge towards the same control points through genes such as cell surface receptors with tyrosine kinase activity (RTK) and their intracellular transmitters (RAS). At the bottom of the figure is represented the universal "brake" of cell division control: the p53 protein. TP53, the gene encoding p53, is often the target of environmental mutagens, as for example aflatoxin, a contaminant of the diet that generates base changes into DNA. This mutation eliminates the brake effect and allows uncontrolled proliferation.
Source: Pierre Hainaut, unpublished

Genetic changes

Genetic changes are the cornerstone of cancer. The sequencing of the entire human genome has made it possible to identify genetic alterations in cancers in unprecedented detail. About 300 different genes have been shown to be mutated at some frequency in human cancers. Within this catalogue, a shortlist of 20 or 30 genes appear to be frequently mutated in almost any type of cancer (including those of the "cancer box") (Table 3.1.1). These genes may be seen as "master genes" that control very basic functions essential for cell division control.

Detecting mutations in cancer cells has many potential implications for research and therapy. First, mutations can be informative of the evolution of cancer and provide clinically interesting prognostic or predictive information. Second, the first mutations that contribute to cancer occur, by definition, prior to the development of a lesion. Detecting these mutations may thus help in early cancer diagnosis. Finally, in several cases mutations may be good indicators of therapeutic responses, and may help to select therapies that have greater chances of success. This is the case, for example, for mutations in the EGF receptor (EGFR) that are found in 20–40% of lung cancers in never smokers. These mutations constitutively activate the receptor, generating a constant cell proliferation signal. The signal can be blocked by small inhibitory drugs such as gefinitib (Iressa) or erlotinib (Tarceva). These drugs have interesting therapeutic effects in patients with EGFR mutations, but are poorly effective in most other patients.

Cells of many cancers accumulate mutations at a rate significantly higher than in normal cells, a property referred to as "Mutator Phenotype". This property of transformed cells is believed to be critical for the development of cancer as well as for the development of resistance to cancer treatments [8]. The Mutator Phenotype is the consequence of mutations in genes that normally control DNA repair and integrity. Cells with such mutations become unable to correctly repair mutation-inducing DNA damage, and

thus accumulate mutations at a much higher rate than normal cells. Molecular mechanisms underlying the Mutator Phenotype may include defects in DNA repair, gene transcription, cell cycle control and cell death.

TP53 tumour suppressor: an example of common genetic change

The most studied of all cancer genes is TP53, which encodes the p53 protein, a tumour suppressor that is mutated in about half of all human cancer cases. A database of all these mutations is maintained at the International Agency for Research on Cancer [9]. It compiles about 24 000 TP53 mutations detected in almost every type of human cancer. We now have a very good understanding of the molecular effects of these mutations. Most of them fall within a part of the protein that binds to DNA, allowing p53 to regulate several dozen of other genes. The mutations are often single base substitution, leading to the replacement of one amino-acid in the protein by another one.

This small change is enough to perturb protein folding and to prevent it to bind to DNA, thus inducing a loss of function.

Close examination of the distribution of these mutations show that they occur in a non-random fashion and there are significant differences in mutation patterns among cancers that are strongly associated with exposure to environmental mutagens. These differences are due to the fact that different mutagens can damage DNA in particular ways, thus leading to different types of mutations. Thus, mutations in TP53 can be seen as "molecular signatures" of mutagenic events that contribute to cancer. This makes the TP53 mutation profile a potent biomarker in molecular epidemiology, as a potential reporter of specific mutagenic exposures. This is supported by the evidence showing that mutation patterns in common cancers differ significantly depending on geographic variations in incidence, indicative of differences in exposure to specific environmental carcinogens [10]. For example, in liver cancers, the type and

frequency of mutations is very different between patients in Europe and the USA and those in many countries of Africa or Southeast Asia. The difference is due to the impact of a particular mutagen, aflatoxin, which is produced by a fungus that contaminates many food components in tropical areas. This mutagen is virtually absent from the western diet, but induces a characteristic TP53 mutation that contributes to liver cancer in regions of sub-Saharan Africa and South East Asia.

In many cancers, presence of a mutation in TP53 is correlated with a rather poor prognosis and bad response to treatment. So far, this fact has had only limited impact in the clinics because the same information could be deduced from other markers routinely scored by the pathologist: the size of the tumour, its grade, the extension of the disease into lymph nodes, etc. However, it has recently been recognised that TP53 mutations may help to distinguish between tumours that, to the pathologist, look the same. For example, in breast cancers, presence of a TP53 mutation allows the identification of tumours that are at a high risk of progression among tumours classified by the pathologist as of "good prognosis". It is therefore possible to single out those patients and offer them more aggressive treatment. It would not be justified to give such a treatment to all patients because the risk of secondary effects would outweigh the benefits, since most of them actually do not need such treatment.

TP53 mutations can also be detected outside tumour tissues, in particular in body fluids such as blood. The presence of mutant TP53 in blood is due to the fact that cancer masses release small amounts of dead cells that originate from the tumour and therefore contain the same mutation. Detection of such mutations in the plasma may be exploited for early cancer detection. Indeed, TP53 mutations in plasma DNA have been reported in patients with cancers of the colon, pancreas, lung and liver. For example, the aflatoxin-induced TP53 mutation mentioned above is detectable in the plasma of non-cancer subjects from China who are chronic carriers of Hepatitis B virus,
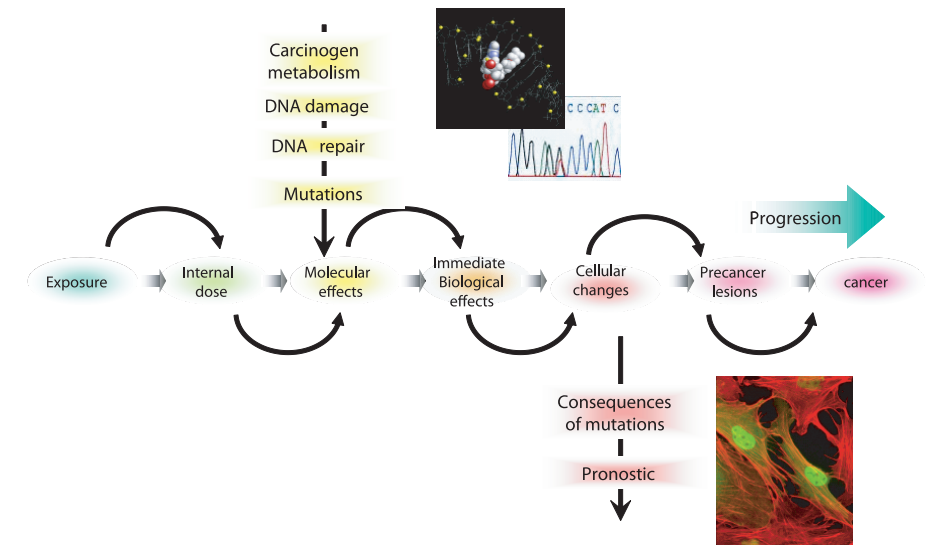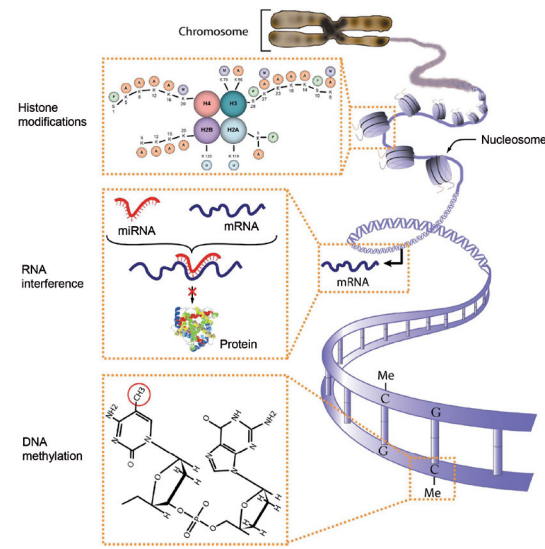

Fig. 3.1.2 Epigenetic regulation of gene expression, transcription and repair
The figure shows, from top to bottom, how DNA is packaged into a supra-molecular structure, the chromatin, which controls how cellular or environmental signals can act on DNA. Changes in these structures have a critical impact on how gene works. These changes are called "epigenetic".
Source: Zdenko Herceg and Thomas Vaissiere, unpublished

Fig. 3.1.3 How cancer progresses, from environmental risk factors to overt disease
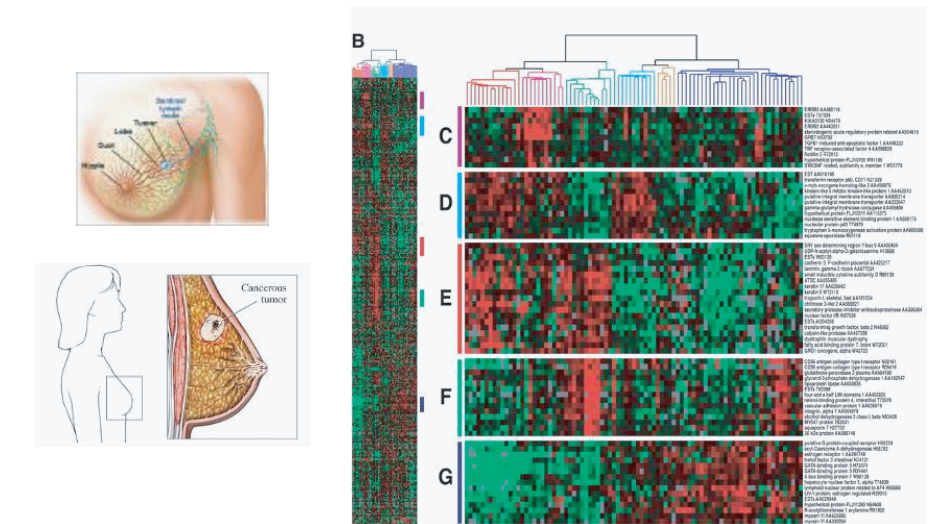Source: Pierre Hainaut, unpublished

Fig. 3.1.4 Gene expression patterns in breast cancer. Each individual dot represents the expression of one gene. The colour indicates the level of expression (red: high; green: low)
From Sorlie [16] assembled from several published figures

up to 5 years ahead of the development of liver cancer [11]. Studies on TP53 also illustrate another characteristic of genetic mutations that cause cancer: some of them can occur in the germline and be transmitted as inherited traits. Thus, the persons who inherit these mutations are born with a cancer-causing gene that is present in all the cells of the body. These persons are at very high risk of developing cancer. Inheritance of a TP53 mutation causes a serious familial cancer syndrome, the Li-Fraumeni Syndrome (LFS). Members of families who inherit the mutation often have tumours in their childhood or teenage years, and require specific medical surveillance and care to detect and cure these cancers at the earliest possible stage.

Epigenetic changes

The field of epigenetics is one of the most rapidly expanding fields of modern biology, with enormous implication on our thinking and understanding of biological phenomena and diseases, notably cancer. Historically, the term epigenetics was used to describe all biological phenomena that do not follow normal genetic principle. The term was coined by Conrad Waddington in 1942 to describe the discipline in biology which studies "the interactions of genes with their environment that bring the phenotype into being". Since then, a number of biological events that are not coded in DNA sequence itself have been considered epigenetic phenomena. These include imprinting (the conditioning of parental genomes during gametogenesis ensuring that a specific locus is exclusively expressed from either maternal or paternal genome in the offspring), paramutations (heritable changes in one allele induced by another allele) in plants, and X-chromosome inactivation in females. One of the most remarkable recent discoveries is that different epigenetic events may share common underlying molecular mechanisms. These advances turned academic and public attention to the potential application of epigenetic mechanisms to biomedical research and important public health issues.

In a broader sense, epigenetics can be considered as an interface between genotype and phenotype. In other words, epigenetics encompasses mechanisms that modify the final outcome of the genetic code without altering the underlying DNA sequence. The importance of epigenetic principle is highlighted by the fact that all cells in any given organism share an identical genome with other cell types, yet they can exhibit strikingly different morphological and functional properties. Therefore, it is obvious that epigenetic events define the identity and proliferation potential of different cells in the body, the features that are typically deregulated in cancer. Nowadays, epigenetics may be defined as the study of all changes that are stably transmitted over many rounds of cell divisions, but that do not alter the nucleotide sequence (genetic code). Epigenetic inheritance includes DNA methylation, histone modifications and RNA-mediated silencing, all of which are essential mechanisms that allow the stable propagation of gene activity states from one generation of cells to the next. Consistent with the importance of epigenetic mechanisms, deregulation of epigenetic states is intimately linked to human diseases, most notably cancer [6,12]
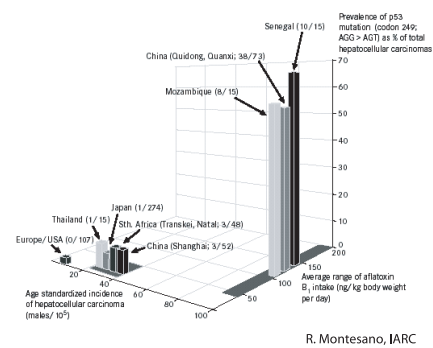
DNA methylation

The best-studied epigenetic mechanism is DNA methylation. The methylation of DNA refers to the covalent addition of a methyl group to the 5-carbon (C^5) position of cytosine bases that are located 5′ to a guanosine base. This is a very small chemical modification of the DNA molecule that while it does not alter the DNA code, may have major regulatory consequences. Aberrant DNA methylation is tightly connected to a wide variety of human cancer. Two forms of aberrant DNA methylation are found in human cancer: the overall loss
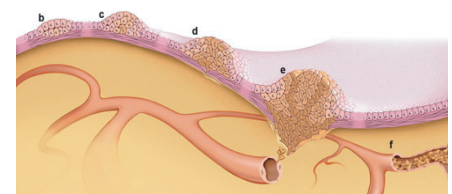

Fig. 3.1.6 Steps in cancer formation in an epithelial tissue, from first cancer cell to invasion

Fig. 3.1.5 Mutation in TP53 at codon 249 in liver cancers induced by exposure to aflatoxin, a contaminant of the diet in many low-resource countries
The graph compiles data from various countries on incidence of liver cancer, on levels of aflatoxin contamination in food and on the prevalence of the mutation at codon 249 in liver cancers.
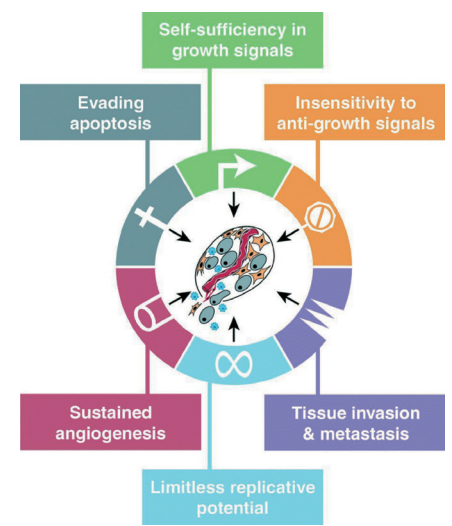

Fig. 3.1.7 Hallmarks of cancer, after Hanahan [7]
To become cancerous, a cell has to acquire complementary changes in 6 basic mechanisms

of 5-methyl-cytosine (global hypomethylation) and gene promoter-associated (CpG island-specific) hypermethylation [13,14]. While the precise consequences of genome-wide hypomethylation are still debated (activation of cellular proto-oncogenes, induction of chromosome instability), hypermethylation of gene promoters is in turn associated with gene inactivation. When hypermethylated, gene promoters become unable to bind the factors that are responsible for gene expression. The gene thus becomes inactivated. A large number of studies indicated that the silencing of tumour suppressor genes and other cancer-related genes may occur through hypermethylation of their promoters.

Unscheduled hypermethylation of gene promoters represents an attractive target for early diagnosis, risk assessment and cancer prevention. For example, the genes that are the target of DNA hypermethylation early in tumour development, in a high percentage of cases, and specific to cancer type, are of particular interest. A number of studies showed that the $p16^{INK4a}$ (CDKN2A) tumour suppressor gene is among the most frequently silenced cancer-associated genes in human cancer, and that this silencing is associated with promoter hypermethylation. Unscheduled addition of methyl markers (de novo methylation) at the $p16^{INK4a}$ promoter is one of the most frequent epigenetic alterations detected in a wide range of human cancers. In addition, silencing of $p16^{INK4a}$ by promoter
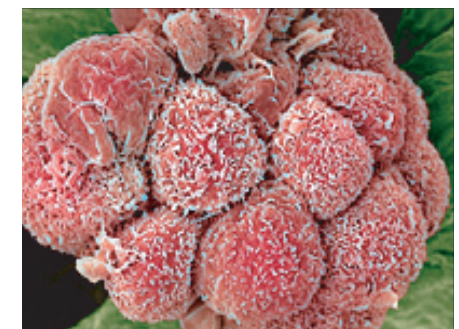

Fig. 3.1.8 Breast cancer cells

hypermethylation is highly tumour-specific and appears to be the earliest event in some cancer types, making this gene an attractive target for preventive strategies. While cancer epigenetics have focused primarily on DNA methylation changes as biomarker, cancer-specific modifications of chromatin proteins (histones) and the expression profiles of microRNAs (a family of small non-coding transcripts important for stable repression of specific genes) as potential biomarkers remain largely unexplored.

The CIMP phenotype

The studies on DNA methylation involving multiple genes revealed that some cancer types exhibit concurrent methylation of groups of cancer-associated genes, a phenomenon known as the CpG island methylator (CIMP) phenotype [15]. Although the CIMP phenotype has been studied primarily in colorectal cancer, other studies provided evidence that the CIMP phenotype may also be present in different cancer types including hepatocellular carcinoma, gastric cancer, pancreatic cancer, glioblastomas, oral cancer, leukaemias and solid tumours [15]. However, it should be noted that the CIMP-positive tumours represent only a subset of cancers with distinct epigenotype. Analogous to the contribution of the Mutator phenotype to genetic changes, the presence of the CIMP phenotype may explain the simultaneous occurrence of methylation of many genes in some (but not all) cancer types. This may be exploited for prognostic purposes but also in the design of "epigenetic therapy".

Despite a wealth of studies providing evidence for an association between abnormal DNA methylation patterns in a variety of human cancers, the causes and underlying mechanism of this phenomenon remain unclear. Specific agents (epimutagens) or combinations thereof in the environment, diet or lifestyle may promote, and/or relieve resistance against, unscheduled methylation and/or histone modifications, leading to altered gene expression and oncogenic process. Large population-based cohorts and case−control studies may offer

excellent opportunities to test the contribution of repeated and chronic exposure to epimutagens in the environment and nutrition to abnormal levels and patterns of DNA methylation in specific cancers.

Perspectives for combating cancer

Until recently, genetic and epigenetic studies on cancer have so far been exploited primarily for improving the knowledge of the mechanisms of cancer development. However, the recent emergence of powerful technologies for genome-wide analysis of genetic and epigenetic changes is dramatically advancing our capacity to identify multiple changes in gene expression as well as genetic or epigenetic
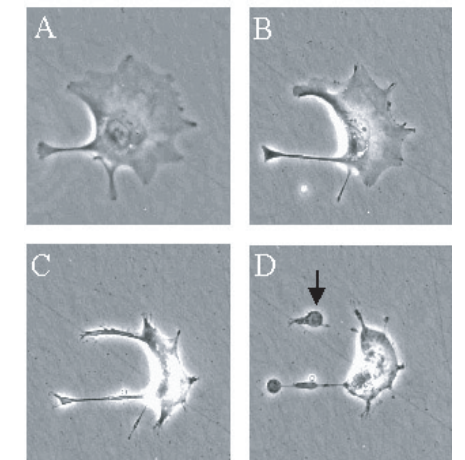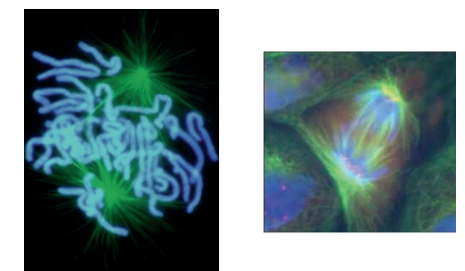

Fig. 3.1.9 Steps in apoptosis

Fig. 3.1.10 Two microphotographs of dividing cells

signatures in specific cancers. This affords an opportunity to map the pattern of genetic and epigenetic alterations that is specific for each particular cancer. This information is the first step towards a form of personalised medicine where patients will be treated and followed up according to protocols that will take into account the molecular properties of a given cancer. The road to such personalised medicine is however still very long. It will require that tests for genetic and epigenetic alterations become part of routine and affordable practice in the hospital. It will also necessitate the performance of clinical trials to show which treatment protocol gives the best results according to the pattern of genetic and epigenetic changes. The near future will likely bring insights into which alterations or combinations thereof can be interpreted as reliable biomarkers of exposure to cancer risk factors and tumorigenesis. This in turn will enhance priority setting in selecting new drugs to be developed to combat cancer.

Suggested further reading and links

Weinberg, R.A. How cancer arises, Scientific American, Sept 1996, pp 62-70; link to pdf: http://www.bme.utexas.edu/research/orly/teaching/BME303/Weinberg.pdf

Inside Cancer: a multimedia guide to cancer biology. http://www.insidecancer.org/

ONCOGENES
PDGF Codes for platelet-derived growth factor. Involved in glioma (a brain cancer)
EGFR Codes for the receptor for epidermal growth factor. Involved in glioblastoma (a brain cancer) and breast cancer
HER-2 or ERBB2. Codes for a growth factor receptor. Involved in breast, salivary gland and ovarian cancers
RET Codes for a growth factor receptor. Involved in thyroid cancer
KRAS Involved in lung, ovarian, colon and pancreatic cancers
NRAS Involved in leukaemias
MYC1 Involved in leukaemias and breast, stomach and lung cancers
NMYC Involved in neuroblastoma (a nerve cell cancer) and glioblastoma
LMYC Involved in lung cancer
BCL2 Codes for a protein that normally blocks cell suicide. Involved in follicular B cell lymphoma
CCND1 or **PRAD1** Codes for cyclin D1, a stimulatory component of the cell cycle clock. Involved in breast, head and neck cancers
CTNB1 Codes for beta-catenin, involved in liver cancers
MDM2 Codes for an antagonist of the p53 tumor suppressor protein. Involved in sarcomas (connective tissue cancers) and other cancers

TUMOUR SUPPRESSOR GENES
APC Involved in colon and stomach cancers
DPC4 Codes for a relay molecule in a signalling pathway that inhibits cell division. Involved in pancreatic cancer
NF-1 Codes for a protein that inhibits a stimulatory (Ras) protein. Involved in neurofibroma and pheochromocytoma (cancers of the peripheral nervous system) and myeloid leukemia
NF-2 Involved in meningioma and ependymoma (brain cancers) and schwannoma (affecting the wrapping around peripheral nerves)
CDKN2A or MTS1 Codes for the p16 protein, a braking component of the cell cycle clock. Involved in a wide range of cancers
RB1 Codes for the pRB protein, a master brake of the cell cycle. Involved in retinoblastoma and bone, bladder, small cell lung and breast cancer
TP53 Codes for the p53 protein, which can halt cell division and induce abnormal cells to kill themselves. Involved in a wide range of cancers
WT1 Involved in Wilms' tumour of the kidney
BRCA1 Involved in breast and ovarian cancers
BRCA2 Involved in breast cancer
VHL Involved in renal cell cancer

Table 3.1.1 *Common oncogenes and tumour suppressor genes involved in human cancer From Weinberg, Scientific American 1996, with modifications*

REFERENCES

1. Merlo LM, Pepper JW, Reid BJ, et al. (2006). Cancer as an evolutionary and ecological process. *Nat Rev Cancer* 6: 924-935.

2. Olivier M, Petitjean A, Marcel V, et al. Recent advances in p53 research: an interdisciplinary perspective. *Cancer Gene Ther* Advance online publication 19 September 2008. DOI: 10.1038/cgt.2008.69.

3. Nowell PC (1976). The clonal evolution of tumor cell populations. *Science* 194: 23-28.

4. Baselga J (2006). Targeting tyrosine kinases in cancer: the second wave. *Science* 312: 1175-1178.

5. Herceg Z and Hainaut P (2007). Genetic and epigenetic alterations as biomarkers for cancer detection, diagnosis and prognosis. *Molecular Oncology* 1: 26-41.

6. Cortez CC and Jones PA (2008). Chromatin, cancer and drug therapies. *Mutat Res.*

7. Hanahan D and Weinberg RA (2000). The hallmarks of cancer. *Cell* 100: 57-70.

8. Loeb LA, Bielas JH, Beckman RA (2008). Cancers exhibit a mutator phenotype: clinical implications. *Cancer Res* 68: 3551-3557.

9. IARC TP53 Mutation Database, R13, November 2008. Available from: http://www-p53.iarc.fr/ Date Accessed: November 14, 2008.

10. Hainaut P and Hollstein M (2000). p53 and human cancer: the first ten thousand mutations. *Adv Cancer Res* 77: 81-137.

11. Jackson PE, Kuang SY, Wang JB, et al. (2003). Prospective detection of codon 249 mutations in plasma of hepatocellular carcinoma patients. *Carcinogenesis* 24: 1657-1663.

12. Herranz M and Esteller M (2007). DNA methylation and histone modifications in patients with cancer: potential prognostic and therapeutic targets. *Methods Mol Biol* 361: 25-62.

13. Feinberg AP and Tycko B (2004). The history of cancer epigenetics. *Nat Rev Cancer* 4: 143-153.

14. Jones PA and Baylin SB (2002). The fundamental role of epigenetic events in cancer. *Nat Rev Genet* 3: 415-428.

15. Issa JP (2004). CpG island methylator phenotype in cancer. *Nat Rev Cancer* 4: 988-993.

16. Sorlie T, Perou CM, Tibshirani R, et al. (2001). Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc Natl Acad Sci U S A* 98: 10869-10874.

DNA Damage Response and DNA Repair

Summary

> Even under normal cellular conditions, genomic DNA is under constant threat from DNA damage and DNA breaks that are constantly produced by endogenous and exogenous (environmental) genotoxic agents

> DNA repair machineries represent an arsenal of tools devised by cells to repair DNA damage and hence defend themselves against constant challenge to genomic integrity

> One of the major achievements of the last two decades has been the isolation and characterisation of the genes and their protein products involved in different pathways of DNA repair

> Recent studies revealed that all types of DNA repair (including nucleotide excision repair, base excision repair, and double strand break repair) are complex and dynamic processes that require careful orchestration of many enzymes, adapter proteins and chromatin-modifying activities

> Defects in key players and pathways involved in DNA damage response and DNA repair can lead to cancer and other human diseases

Sources of DNA damage

Exogenous sources. There are a number of agents or phenomena capable of inducing DNA lesions (Figure 3.2.1). They result in DNA base modifications, formation of covalent bridges between complementary strands, single strand breaks (SSBs) and double strand breaks (DSBs) [1]; note that SSBs may also be transformed into DSBs. These insults can have several sources, such as food, water, chemical products, radiation and others. For instance, the aflatoxins are group of carcinogenic fungal metabolites that are found in foodstuffs contaminated with *Aspergillus* strains. The detection of both DNA and protein adducts in humans exposed to 1,3-Butadiene provided key pieces of evidence to support the evaluation by an IARC group that concluded that 1,3-Butadeine was probably a human carcinogen (group 2A) [2].

Some electrophyl molecules carrying alkyl functions have a high affinity for nitrogen molecules of purine and pyrimidine bases. The resulting alkylation can lead to the creation of punctual mutations, but also destabilises the binding between two adjacent nucleotides, thus generating both SSBs and DSBs. A number of alkylating agents are present in cigarette smoke, explaining in part the cancerigenic properties of the cigarette. One of these agents is methylmethane sulfonate (MMS), which is frequently used in laboratories in order to study DSBs. Cigarette smoke contains several other genotoxic agents, including benzo[a]pyrene, that have been found to increase the frequency of mutations [3,4].

Different types of radiation can also produce DNA damage. The ultraviolet (UV) radiation from the sun, for example, can produce a covalent linkage between two adjacent pyrimidine bases in DNA to form, among others, thymine dimers. Ionizing radiation can also create DNA damage either directly or indirectly; directly by producing the formation of radicals on sugars that are present in the DNA chain, and indirectly by provoking the radiolysis of water molecules, generating radical ions that act on phosphodiester bonds that link the nucleotides. If the distance between the two sites of damage created by two radicals is short, the DNA chain may break, which is why the DSBs are the most frequent types of breaks caused by gamma irradiation.

Endogenous sources. Even under normal cellular conditions, genomic DNA is subject to spontaneous endogenous changes (Figure 3.2.1). The chemical events that lead to DNA damage include hydrolysis, oxidation and electrophilic attack. These reactions are triggered by exposure of cells to exogenous chemicals (e.g. environmental agents, food constituents), but they can also result from endogenous metabolic processes.

The most frequent spontaneous chemical reactions that create DNA damage in cells are depurination and deamination. Depurination, caused by thermal fluctuations, represents the linkage of the N-glycosyl of purine bases to deoxyribose hydrolyze. About 5000 purine bases (adenine and guanine) are lost every day from the DNA of each human cell because of depurination. Deamination transforms cytosine to uracile or thymine at a rate of 100 bases per cell per day. In addition, it has long been hypothesised that the presence of methylated CpG sequences *per se* are the major cause of mutability in mammalian genomes. Considerable attention has focused on the cause of CpG sites because this can be a common site of mutations or DNA methylation, detected in a range of genetic diseases as well as in many cancers [5,6]. All hypotheses and experimental studies seem to agree on the importance of methylation of cytosine residues. Methylation increases the rate of hydrolytic deamination and also increases the reactivity of neighbouring guanines to electrophiles [5,7].

Other examples of endogenous chemical events leading to DNA damage are apurinic/apyrimidinic (AP) sites that can be produced by spontaneous hydrolysis, alkylation-induced hydrolysis or glycosylase-catalysed base-excision repair. Oxygen radical attack on DNA leads to a plethora of oxidised bases, as well as strand scission. Chromosomal alterations are initiated *inter alia* by double-strand breaks resulting from oxidative cleavage of the DNA backbone or enzymatic cleavage during chromatin remodelling (e.g. by topoisomerase II). DNA replication itself contributes about ten double-strand breaks per cell cycle in the form of stalled or blocked replication forks [8].

DNA repair

Genomic DNA within each human cell is constantly exposed to an array of damaging agents of both environmental origin, exemplified by sunlight and tobacco smoke, and of endogenous origin, including water and oxygen [9]. This scenario necessitates constant surveillance so that damaged nucleotides may be removed and replaced before their presence in a DNA strand at the time of replication leads to the generation of mutations [10]. Restoration of normal DNA structure is achieved in human cells by one of several DNA repair enzymes that cut out the damaged or inappropriate bases and replace them with the normal nucleotide sequence. This type of cellular response is referred to as "excision repair", and there are two major repair pathways which function in this manner: "base excision repair", which works mainly on modifications caused by endogenous agents, and "nucleotide excision repair", which removes lesions caused by environmental mutagens. UV light is probably the most common exogenous mutagen to which human cells are exposed, and the importance of the nucleotide excision repair pathway in protecting against UV-induced carcinogenesis is clearly demonstrated in the inherited disorder xeroderma pigmentosum. Individuals who have this disease lack one of the enzymes involved in nucleotide excision repair and have a thousandfold greater risk of developing skin cancer following exposure to sunlight than do other individuals. Seven genes, ranging from XPA to XPG, are defective in XP syndrome [11].

One of the great achievements of the last two decades has been the isolation and characterisation of the genes, and their protein products, involved in base excision repair and nucleotide excision repair. It has become apparent that certain proteins so identified are not exclusively involved in DNA repair but play an integral part in other cellular processes such as DNA replication and recombination.

Excision repair

The first step in both base excision repair and nucleotide excision repair is the recognition of a modification in DNA by enzymes that detect either specific forms of damage or a distortion in the DNA helix. Recognition of damage is followed by an excision step in which DNA containing the modified nucleotide is removed. Gap-filling DNA synthesis and ligation of the free ends complete the repair process.

Nucleotide excision repair may occur in the non-transcribed (non-protein-coding) regions of DNA (Figure 3.2.2). A distortion in DNA is recognised, probably by the XPC-hHR23B protein (I). An open bubble structure is then formed around the lesion in a reaction that uses the ATP-dependent helicase activities of XPB and XPD (two of the subunits of TFIIH) and also involves XPA and RPA (II-III). The XPG and ERCC1-XPF nucleases excise and release a 24- to 32-residue oligonucleotide (IV), and the gap is filled in by PCNA-dependent polymerases (POL) epsilon and delta and sealed by a DNA ligase, presumed to be LIG1 (V). Nucleotide excision repair in regions that are transcribed (and hence code for proteins) requires the action of TFIIH [12].

DNA base excision repair (Figure 3.2.3) involves the removal of a single base by cleavage of the sugar-base bond by a damage-specific DNA glycosylase (e.g. hNth1 or uracil DNA glycosylase) and incision by an apurinic/apyrimidinic nuclease (human AP1) [13]. Gap-filling may proceed by replacement of a single base or by resynthesis of several bases in the damaged strand (depending on the pathway employed).

More complex and unusual forms of damage to DNA, such as double strand breaks, clustered sites of base damage and non-coding lesions
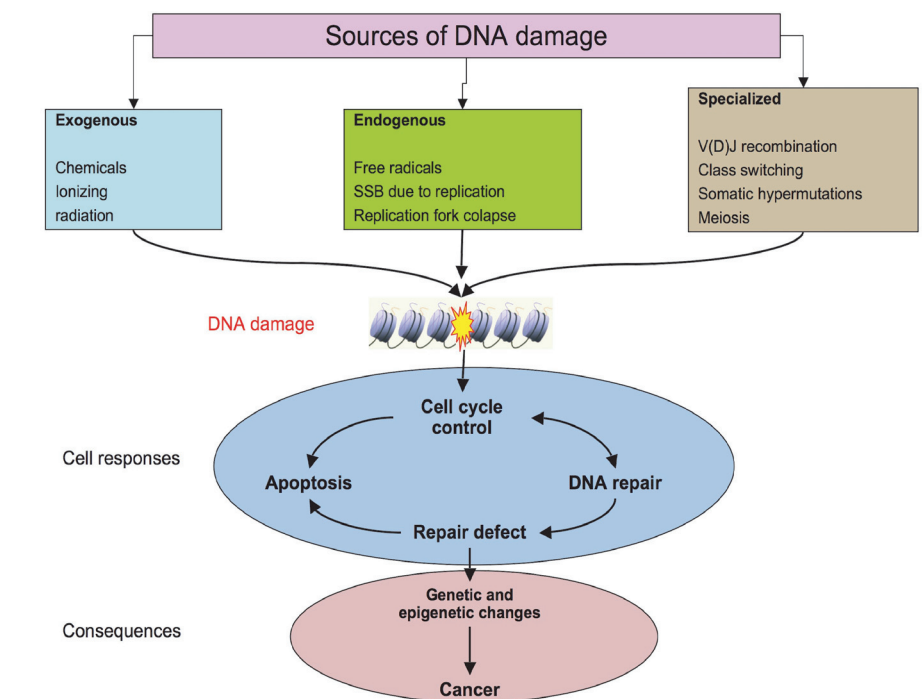

Fig. 3.2.1 Causes, cellular responses and consequences of DNA damage. Source: Zdenko Herceg, unpublished

that block the normal replication machinery are dealt with by alternative mechanisms. Inherited human diseases in which the patient shows extreme sensitivity to ionizing radiation and altered processing of strand breaks, such as ataxia telangiectasia and Nijmegen breakage syndrome, constitute useful models to study the repair enzymes involved in these processes. Indeed, if elucidation of base excision repair and nucleotide excision repair was the great achievement of the late 1990s in this field, then understanding strand break repair will probably be the great achievement of the next decade. This will have important consequences. Certain cancers are often treated with radiotherapy, and a small percentage of patients show considerable sensitivity to their treatment, with the result that treatment schedules are reduced to try to avoid adverse reactions. A better understanding of the possible causes of this radiosensitivity, including characterisation of the enzymes involved in the repair of DNA damage produced by ionizing radiation, may lead to better tailoring of radiotherapy doses to individual patients.

DNA double-strand breaks

DSBs are arguably the most dangerous type of DNA lesions that are constantly generated during the life of a cell. DSBs result from both exogenous and endogenous factors and represent an important threat to the integrity of the genome as they can lead to mutation induction, oncogenic transformation or cell death (Table 3.2.1). Mammalian cells may be subject to at least 10 000 different lesions every day, which may well represent a low estimate. In order to efficiently deal with DSBs the cell has evolved multiple cellular processes that are initiated in response to DNA damage including checkpoint activation, DNA repair, and changes in gene transcription. These processes will counteract the omnipresent DNA-damaging effects of endogenous and environmental genotoxic insults.

Detection and signalling of DSBs. The DNA damage checkpoint can be defined as a network of interacting pathways operating in concert to recognise damage in the DNA and elicit the response. It shares characteristics of a signal transduction pathway, and the participating proteins can be formally divided into sensors, transducers and effectors. Sensor proteins recognize DNA damage, directly or indirectly, and function to signal the presence of these abnormalities and initiate the biochemical cascade. Transducers are typically protein kinases that relay and amplify the damage signal from the sensors by phosphorylating other kinases or downstream target proteins; this can also be identified as signalling. Effector proteins include the ultimate downstream targets of the transducer protein kinases, and these can actually be the repair proteins.
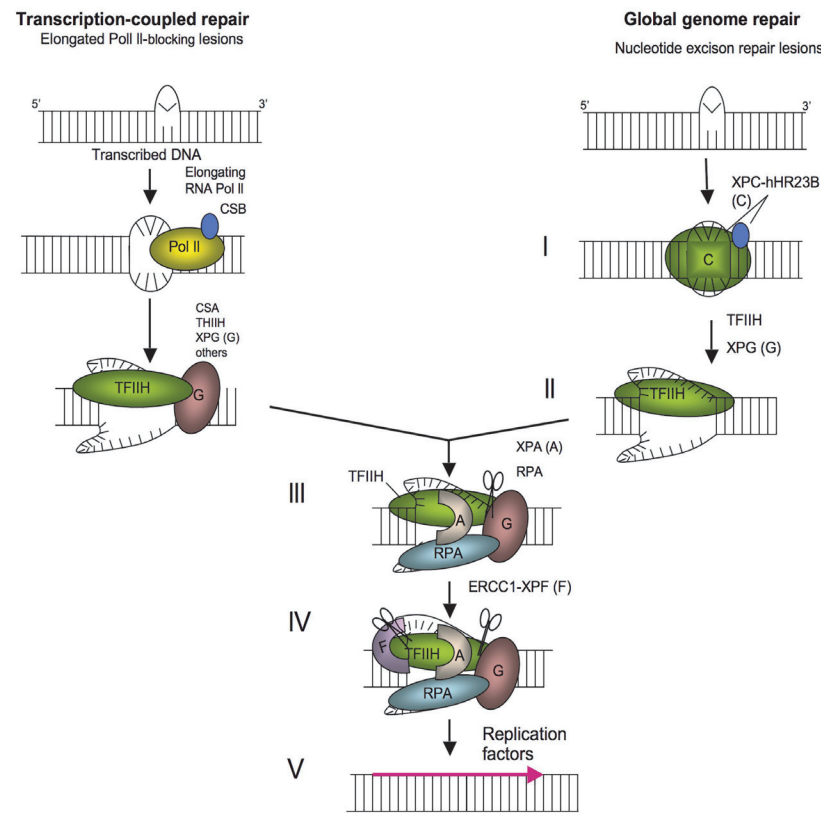

Fig. 3.2.2 *Stages of nucleotide excision repair (NER). There are two NER pathways that are involved in the removal of UV-induced and carcinogen-induced DNA lesions. The NER pathway involves a large number of proteins including damage recognition proteins (XPA, XPC, RPA), helicases (XPB, XPD) and nucleases (XPF, XPG), which interact to form a complex protein machinery. In transcription-coupled NER pathway, DNA lesions are recognized when they stall RNA polymerase II, whereas in global genome NER, the lesions are recognized by the proteins XPC and hHR23B. After the lesions are detected, both pathways are similar. The XPB and XPD helicases of the multi-subunit transcription factor TFIIH unwind DNA around the lesion (II). Single-stranded binding protein RPA stabilizes the intermediate structure (III). XPG and ERCC1-XPF cleave the borders of the damaged strand, generating a 24-32 base oligonucleotide containing the lesion (IV). The DNA replication machinery then fills in the gap (V). Mutations in the genes encoding proteins involved in NER lead to the condition known as xeroderma pigmentosum in humans, which is characterised by extreme UV sensitivity and a wide spectrum of other abnormalities*

In order for DSBs to be repaired efficiently, DNA damage first must be detected and the information transmitted to the effectors and DNA repair proteins through a signalling pathway. Of all these steps, however, the detection of DNA breaks is one of the least known. The major genes that can act as detectors and transducers in DNA damage response are the ATM (*Ataxia Telangiectasia* Mutated) superfamily of kinases and p53, activation of which seems to be important for DNA damage detection and cell cycle arrest. Defects in these important players and pathways can lead to cancer and other human diseases. One of the first DNA damage signalling events and the most easily detectable in DNA damage response is the phosphorylation of at the H2A variant, H2AX at serine 139 by the phosphatidylinositol-3 kinase-like family of kinases (PI3K) at DSB sites. The presence of γH2AX is important for both types of DSBs: repair HR (Homologous Recombination) and NHEJ (Non Homologous End Joining) and is required for the retention/accumulation of repair proteins at the break site. ATM, ATR (ATM and Rad3 related protein), and probably DNA-PKcs, all members of the PIKK family, are responsible for the phosphorylation of H2AX, and thus can represent the detectors of DNA damage. The activation of these detectors could be explained in two ways. First, DNA breaks produce a modification in the chromatin structure or spatial organisation, and this modification appears to be sufficient for the autophosphorylation of ATM and its activation.

In the case of NHEJ repair, the heterodimer Ku70/Ku80 seems to be the first detector, because it instantly binds to the damaged ends and recruits DNA-PK. TRRAP is a member of the PIKK superfamily, indicating that it, like other members of this family, it may have a role in DNA damage response. However, TRRAP lacks the kinase catalytic activity, preventing it from phophorylating downstream targets, but it still has an effect in later stages of DNA damage response through the P53 pathway. P53 plays a critical role in the control of cellular proliferation through the checkpoint
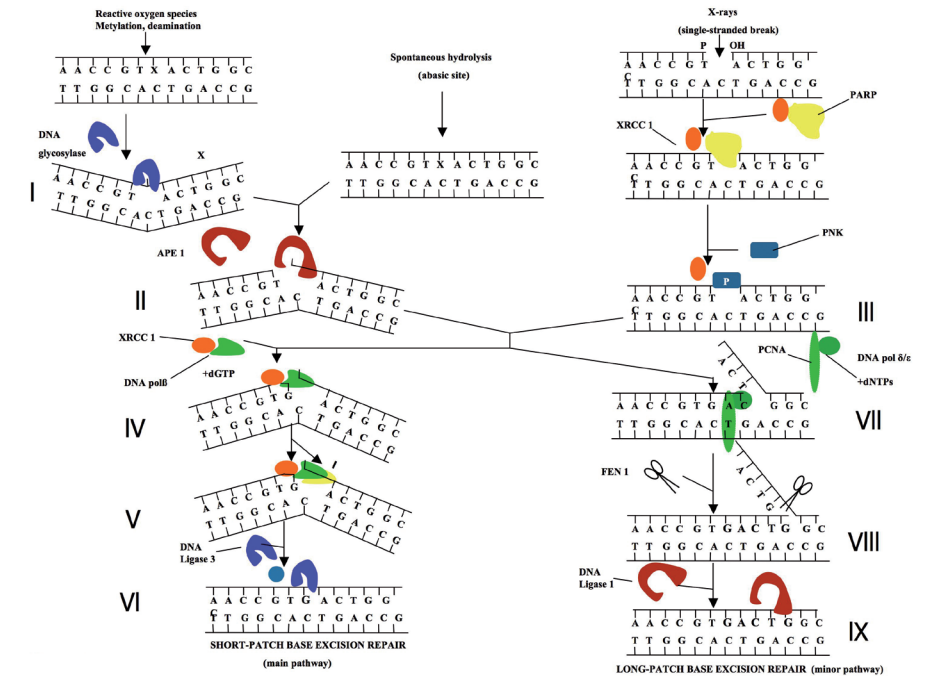

Fig. 3.2.3 *Stages of base excision repair. Many glycosylases, each of which deals with a relatively narrow spectrum of lesions, are involved. The glycosylase compresses the DNA backbone to flip the suspect base out of the DNA helix. Inside the glycosylase, the damaged base is cleaved, producing an "abasic" site (I). APE1 endonuclease cleaves the DNA strand at the abasic site (II). In the repair of single-stranded breaks, poly(ADP-ribose)polymerase (PARP) and polynucleotide kinase (PNK) may be involved. In the "short-patch" pathway, DNA polymerase b fills the single nucleotide gap and the remaining nick is sealed by DNA ligase 3. The "long-patch" pathway requires the proliferating cell nuclear antigen (PCNA) and polymerases b, e and d fill the gap of 2-10 nucleotides. Flap endonuclease (FEN-1) is required to remove the flap of DNA containing the damage and the strand is sealed by DNA ligase 3*

activation. Following DNA damage, several kinases including Chk2, CAK, ATM, ATR and DNA-PK can phosphorylate p53 leading to its stimulation and to an enhance binding to DNA. Mutations in the genes involved in DNA damage detection and signalling are frequently found in human cancer.

- *Repair of DSBs.* Two major types of DSB repair, homologous HR and NHEJ, have evolved to deal with the DNA damage constantly generated [14]:
- HR reconstitutes the missing DNA using a homologous copy, usually the sister chromatid. Thus this type of DSB repair occurs more often in G2 phase after DNA replication where the two chromatids are present.

NHEJ is less complicated then HR but is error prone. It processes DNA ends and religates them without any modifications, thus often creates errors.

There also exists a third type of DSB repair that is used less often: Single strand annealing (SSA), which shares components with both NHEJ and HRR and utilises a limited cohesion zone of several base pairs in order to religate DNA ends, in the same way as NHEJ.

The mammals predominantly use the HR whereas lower eukaryotes use NHEJ more often. It is believed that that NHEJ plays a more important role than HR in mitotically replicating cells. HR may play a more prominent role

Gene	Cellular phenotype	Mouse-knockout phenotype
Non-Homologous end-joining (NHEJ)		
Ku70	Radiosensitivity, impaired V(D)J recombination	Radiosensitivity, SCID phenotype, T-cell tumours, growth retardation
Ku80	Radiosensitivity, impaired V(D)J recombination	Radiosensitivity, SCID phenotype, T-cell tumours, growth retardation
DNA-PKcs	Radiosensitivity, impaired V(D)J recombination	Radiosensitivity, SCID phenotype, T-cell tumours
XRCC4	Radiosensitivity, impaired V(D)J recombination	Embryonic lethality, apoptosis of post-mitotic neurons
LIG4 (Ligase IV gene)	Radiosensitivity, impaired V(D)J recombination	Embryonic lethality, apoptosis of post-mitotic neurons
Homology-directed repair (HDR)		
ATM	Radiosensitivity, chromosomal instability, radioresistant DNA synthesis	Radiosensitivity, T-cell tumours, neurological dysfunction, growth retardation, infertility
BRCA1	Not viable	Embryonic lethality
BRCA2	Not viable	Embryonic lethality
RAD51	Not viable, chromosomal aberrations	Embryonic lethality
RAD52	Impaired homologous recombination	Moderate impairment of homologous recombination
RAD54	Radiosensitivity, MMC sensitivity, impaired homologous recombination	Radiosensitivity
XRCC2	Radiosensitivity	Embryonic lethality, chromosomal instability
NHEJ and HDR combined		
MRE11	Not viable, chromosomal instability	-
RAD50	Not viable	Embryonic lethality

Table 3.2.1 *Cellular and mouse-knockout phenotypes associated with mutations in DNA double strand break repair genes*

during meiosis and when sister chromatids are available during late S and G2 stages of the cell cycle, whereas NHEJ is more important during G1 and early S stages. To simplify, it is generally accepted that the predominance of one mechanism over the other is dependent on the cell cycle stage and the type of DSB [14].

DNA repair by NHEJ is more error prone but less demanding then homologous recombination. This type of repair does not need to match the damaged sequence to its intact copy on the homologous chromosome (which is typically at a distant site in the nucleus) or bring the two into close proximity. NHEJ repair is divided into several stages, starting with the processing of the damaged ends, then the establishment of molecular bridge between the strands that facilitates the ligation and finishing by the filling

of missing bases and ligation (Figure 3.2.4). The KU heterodimer, consisting of a tight complex of KU70 (70 kDa) and KU80 (86 kDa), would first recognise the DNA break, and tends to cling to the DNA end. There is considerable evidence that cellular end-joining systems have activities capable of processing all these aberrations. These activities are carried out by specialised endonucleases like Tdp1 and APE1. This cleaning step is very important in eliminating the chemical modifications, thus preventing religation.

The Ku70/Ku80 heterodimer then binds to DNA-PKcs forming the so-called DNA-PK holoenzyme. Formation of the DNA-PK holoenzyme complex on a DNA end results in activation of its kinase activity. This activation involves localised denaturation of the extreme end of

DNA and threading of a few bases of the single strands into defined channels bringing the two DNA strands closer to each other.

The next step is the recruitment of ligase IV through the DNA-PK. The effects of Ku and DNA-PK on the activity of XRCC4/ ligase IV are complex, but it is known that all of these proteins, either alone or in a concert, promote end-to-end association of linear DNAs WRN, Artemis and MRE11, all nucleases with putative roles in end-joining. Their catalytic activity could be important to process or "clean" the damaged bases on the break site before the ligation takes place.

The final step of NHEJ consists of first replacing the damaged bases that have been eliminated and then religating the DNA ends. Additional

factors involved in NHEJ repair include PNKP and BRCA1. BRCA1 binds to the MRN complex, and this interaction seems to be important for end joining in vitro. However the exact *in vitro* function of BRCA1 during NHEJ is still not clear.

Homologous recombination (HR). Generally, repair of DSBs by HR is considered most active in the late stages of the cell cycle (late S and G2), because homologous sequences in the form of sister chromatids, homologous chromosomes or DNA repeats are required. The homologous sequence of the damaged sequence is used as template and no single base is lost or changed, making this type of repair error-free. However, even though it rarely generates errors, HR may result in crossovers and lost of heterozygosity (LOH), and HR events can be classified according to whether or not they result in crossing-over between the homologous sequences. HRR is performed by the RAD52 epistasis group of proteins, which includes the products of *RAD50−55, RAD57,* and *RAD59,* the RAD51 paralogs RAD51b, c, d *and MRE11.* In addition, HR also involves BRCA proteins (BRCA1, BRCA2), XRCC proteins (XRCC2 and XRCC3), and the MRN complex.

The first event believed to occur during HRR is resection of the DNA to yield single strand overhangs. In yeast, the resection is thought to involve the MRN complex (Figure 3.2.4). However, this complex has an endonuclease activity, but no 5'-3' exonuclease activity essential for the resection. NBS1, a member of the MRN complex, is phosphorylated by ATM on serine 343. The NBS1 subunit appears to be important for transmitting signals from DNA damage sensors to MRN. It is thus possible that the phoshorylation of NBS1 activates the exonuclease property of the MRN complex. Other cofactors may also be needed to help the MRN complex resecting the DNA.

The second step in HR is the unwinding of the two complementary strands (Figure 3.2.4). The RAD50 subunit of MRN has ATPase activity that is believed to facilitate DNA unwinding. Then

the RPA protein (human single strand binding (SSB) protein) is recruited to single-stranded DNA, and is thought to protect it from the nuclease activity and to help the activity of RAD51, the equivalent of the bacterial RECA protein. RAD51 forms nucleoprotein complexes on SS DNA tails coated by RPA to initiate the intersection and exchange between the damaged DNA strand and the intact one situated on the sister chromatid. The RAD51 paralogs RAD55 and RAD57 act as accessory proteins and are believed to facilitate action of RAD51. Similarly, Rad52 helps RAD51 to form DNA exchange intermediates, and it is believed that Rad54

may help in unwinding the DNA at the DSB to facilitate access of other repair factors. BRCA1 and BRCA2 are believed to be important at early points of HR and perhaps coordinate repair with other cellular processes.

The activities of all these different proteins lead to the formation of a so-called "Holliday junction" (Figure 3.2.4). In this context, each single strand of the damaged strand is coupled with a homologous region of the model DNA strands. High fidelity is provided by this crossing over of different strands. HR can then go in either of two directions. Non-crossing-over, resulting from dis-
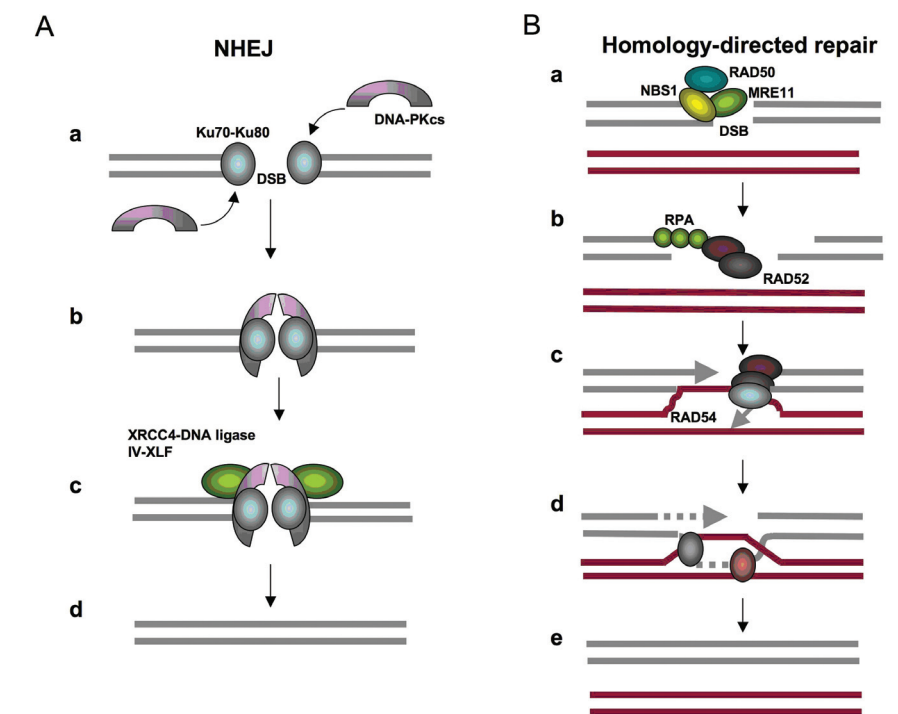

Fig. 3.2.4 *Pathways of DNA double strand break (DSB) repair. In Non Homologous End Joining (NHEJ), a single Ku heterodimer binds to each free DNA end of a DSB and recruits DNA-PKcs, resulting in the formation of the DNA-PK complex. Subsequently, Xrcc4/Ligase IV complex binds to each of the DNA ends and interacts to form a tetramer that may serve to bridge the DNA ends. Other repair proteins are likely involved in this pathway, including the MRE11/RAD50/NBS1 (Xrs2) complex. In Homologous Recombination (HR) repair, DSB is recognised by MRN (a complex of MRE11, RAD50 and NBS1), followed by the resection of DNA ends allowing binding of RPA, RAD51 and RAD52 to the single stranded DNA. A homologous region in an intact chromosome is then invaded (mediated by RAD52) followed by pairing (mediated by RAD54). DNA synthesis occurs from the invading end of the damaged DNA, extending the repair region and forming a Holliday junction (a cross-stranded structure that occurs between four strands of DNA during recombination). This junction translocates along the DNA in a process mediated by a branch migration complex and is cleaved by a resolvase. Source: Zdenko Herceg, unpublished*
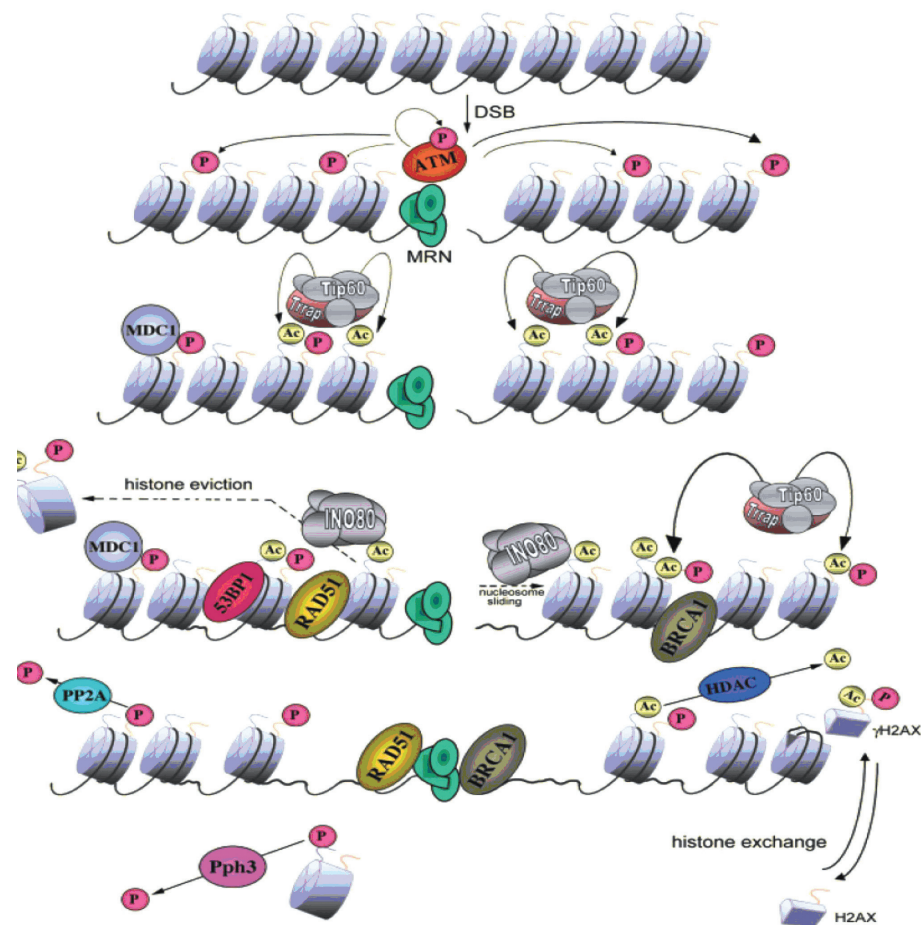
Fig. 3.2.5 *The interplay between chromatin modifying/remodelling activities and DNA repair machinery during the repair of DSBs. In response to a DSB, MRE11-RAD50-NBS1 (MRN) complex and ATM are recruited to the site of DNA break. Activated ATM kinase phosphorylates histone H2AX that allows binding of the early response proteins such as MDC1. This is followed by the recruitment of the TRRAP/TIP60 HAT complex that acetylates histone H4. Histone acetylation unwinds chromatin and/or serves as a binding platform, both of which facilitate the recruitment of remodelling complexes, such as the INO80 and SWR1, and late DNA repair proteins, such as RAD51 and BRCA1. The presence of INO80 may facilitate the eviction or sliding of the nucleosomes in the immediate vicinity of the break site to allow 5'-3' resecting and generation of a 3' single-strand DNA (ssDNA) overhang. This allows RAD51 and BRCA1 to stimulate DSB repair through homologous recombination. After the DSB is repaired, dephosphorylation of incorporated or evicted γH2AX may be mediated by Pph3 and PP2A. The third mechanism of attenuation of γH2AX signal involves the exchange of γH2AX (after its prior acetylation by TRRAP/TIP60 complex) with the unphosphorylated form. Finally, deacetylation of histones occurs to allow chromatin reassembly after DNA break is repaired [16]*

engagement of the Holliday junction followed by DNA pairing and gap filling in the damaged homologue, appears to be strongly favoured during HR in mammalian cells. The other possibility is to go through the classical path, in which the Holliday junctions are resolved by endonucleolytic cleavage, with an equal probability of yielding either a crossover or a noncrossover event. The last step would be filling of the gaps by DNA polymerase and the sealing of the breaks by ligase. However, the DNA polymerase and ligase necessary for this step have not yet been identified.

Repair of DSBs in the context of chromatin. Repair of DSBs is the processes that utilises DNA as substrate and subsequently needs to access the naked DNA. However, in a cell, naked DNA is vulnerable to nuclease digestion and other insults, and without some type of organisation, it would occupy more volume than necessary. Consequently, eukaryotic cells compact DNA into chromatin. The structure of chromatin fulfils essential functions not only by condensing and protecting DNA, but also in preserving genetic information and controlling gene expression. However, given its compacted structure, chromatin hinders several important cellular processes including the detection and repair of DNA breaks [15]. Repair of DSBs, either through HR or NHEJ, is a complex and dynamic process that requires careful orchestration of many enzymes and adapter proteins. In addition, a major hurdle is compacted chromatin, which must first be relaxed to allow access of the DNA repair machinery to damaged DNA. To achieve this, cellular mechanisms that alter the structure of chromatin must first function so that the broken DNA is made accessible to repair factors. Recent studies provided evidence on how the repair machinery gains access to broken DNA in highly condensed chromatin and how the repair process is coordinated with other chromatin-based processes, such as transcription [16]. These studies showed that chromatin modifying/remodelling activities have been associated with DNA repair. Biochemical and molecular studies have revealed different histone modifications associated with DNA repair and identified molecular players responsible for these modifications (Figure 3.2.5). Chromatin modifying/remodelling activities may thus be a part of an arsenal of tools devised by cells to facilitate repair of DNA breaks and hence defend themselves against constant challenge to genomic integrity. These activities include post-translational modifications of histones and ATP-dependent nucleosome mobilisation (chromatin remodelling). An additional mechanism that may facilitate DNA repair by altering chromatin structure involves exchange of histone variants into nucleosomes around break sites. Reversal of chromatin modifications also requires specific enzymatic activities to restore the structure of chromatin once DNA repair has been completed.

Consistent with the roles of chromatin modifying factors in critical cellular processes, growing evidence suggests that aberrant chromatin modification/remodelling is associated with cancer. Recent studies underscore the fact that DNA repair and other DNA-based processes, such as gene transcription and DNA replication, require elaborate coordination of chromatin modifying/remodelling activities. These studies reveal histone proteins as key carriers of epigenetic information, constituting a fundamental and critical regulatory system that extends beyond the genetic information. Therefore, these findings are the foundation for further investigation into the role of chromatin-based mechanisms in critical cellular processes and human cancer.

Other repair pathways

Human cells, in common with other eukaryotic and prokaryotic cells, can also perform one very specific form of damage reversal, the conversion of the methylated adduct, O6-methylguanine, in DNA back to the normal base. O6-Methylguanine is a miscoding lesion: both RNA and DNA polymerases "read" it incorrectly when they transcribe or replicate a DNA template containing it. As this modified base can pair with both the base cytosine (its correct partner) and the base thymine (an incorrect partner), its presence in DNA can give rise to transition mutations by mispairing of relevant bases. A specific protein, O6-alkylguanine-DNA-alkyltransferase, catalyses transfer of the methyl group from the guanine base to a cysteine amino acid residue located at the active site of the protein [17]. This error-free process restores the DNA to its original state but results in the inactivation of the repair protein. Consequently, repair can be saturated when cells are exposed to high doses of alkylating agents, and synthesis of the transferase protein is required before repair can continue.

Mismatched bases in DNA arising from errors in DNA replication, for instance guanine paired with thymine rather than cytosine, are repaired by several pathways involving either specific glycosylases, which remove the mismatched bases, or long-patch mismatch repair involving homologues of the bacterial genes MUTS and MUTL. Insertion or deletion loops at microsatellite sequences can be recognised by hMutSa (a heterodimer of hMSH2 and hMSH6) or hMutSb (a heterodimer of hMSH2 and hMSH3). Subsequent recruitment of hMutLa (a heterodimer of hMLH1 and hPMS2) to the altered DNA targets the area for repair, which requires excision, resynthesis and ligation. Single nucleotide mispairing events require hMutSa function for recognition. One important requirement of such repair processes is that they are able to distinguish the correct base from the incorrect one in the mispair. Since both bases are normal constituents of DNA, this cannot be achieved by an enzyme that scans the DNA for a lesion or structure that is not a normal constituent of the DNA. Defects in at least four of the genes whose products are involved in mismatch repair, namely hMSH2, hMLH1, hPMS1 and hPMS2, have been associated with hereditary nonpolyposis colorectal cancer. This is one of the most common genetic diseases, affecting as many as 1 in 200 individuals, and may account for 4–13% of all colorectal cancers. Affected individuals also develop tumours of the endometrium, ovary and other organs. The DNA of hereditary nonpolyposis colorectal cancer tumours is characterised by instabilities in simple mono-, di- and trinucleotide repeats which are common in the human genome. This instability is also seen in certain sporadic colorectal tumour cells and arises directly from alterations in the proteins involved in mismatch repair [18]. Generally speaking, genomic instability is considered as an indicator of, and fundamental to the nature of, malignant cell growth.

Sun Yat-Sen University Cancer Center was founded in 1964. It is the largest specialised Cancer centre integrated with cancer treatment, training, research and cancer prevention in southern China. Since 1980, the Cancer Center has been designated as the WHO Collaborating Center for Research on Cancer, and also houses the South China State Key Laboratory for Cancer Research.

Sun Yat-Sen University Cancer Center was one of the four initial cancer centres in China and now is a top national cancer institute. The editorial department of the *Chinese Journal of Cancer* is based here; the *Journal* is one of the national kernel academic journals and is published monthly.

Currently there are 1051clinic eds and more than 1500 staff members in the Center, including 150 senior professionals. As a renowned tertiary care centre, it accepts 24 000 inpatients and 300 000 outpatients each year from all over China and Southeast Asia.

Professor Yi-xin Zeng, the present director of the Cancer Center, is a member of the Chinese Academy of Sciences.

website: www.sysucc.com/en/index_en.htm

REFERENCES

1. van Gent DC, Hoeijmakers JH, Kanaar R (2001). Chromosomal stability and the DNA double-stranded break connection. *Nat Rev Genet* 2: 196-206.

2. IARC (1999). Hexachlorobutadiene. In: *IARC monographs on the evaluation of carcinogenic risks to humans, Volume 73, Some chemicals that cause tumours of the kidney or urinary bladder in rodents and some other substances.* Lyon, France: International Agency for Research on Cancer 277-294.

3. IARC (1986). Preamble. In: *IARC monographs on the evaluation of the carcinogenic risk of chemicals to humans, Volume 39, Some chemicals used in plastics and elastomers.* Lyon, France: International Agency for Research on Cancer 13-32.

4. Hoffmann D, Hoffmann I, El-Bayoumy K (2001). The less harmful cigarette: a controversial issue. a tribute to Ernst L. Wynder. *Chem Res Toxicol* 14: 767-790.

5. Fryxell KJ and Moon WJ (2005). CpG mutation rates in the human genome are highly dependent on local GC content. *Mol Biol Evol* 22: 650-658.

6. Hainaut P and Hollstein M (2000). p53 and human cancer: the first ten thousand mutations. *Adv Cancer Res 77*: 81-137.

7. Pfeifer GP, Tang M, Denissenko MF (2000). Mutation hotspots and DNA methylation. *Curr Top Microbiol Immunol* 249: 1-19.

8. Haber JE (1999). DNA recombination: the replication connection. *Trends Biochem Sci* 24: 271-275.

9. Lindahl T and Wood RD (1999). Quality control by DNA repair. *Science* 286: 1897-1905.

10. Aguilera A and Gomez-Gonzalez B (2008). Genome instability: a mechanistic view of its causes and consequences. *Nat Rev Genet* 9: 204-217.

11. Dworaczek H and Xiao W (2007). Xeroderma pigmentosum: a glimpse into nucleotide excision repair, genetic instability, and cancer. *Crit Rev Oncog* 13: 159-177.

12. Dip R, Camenisch U, Naegeli H (2004). Mechanisms of DNA damage recognition and strand discrimination in human nucleotide excision repair. *DNA Repair (Amst)* 3: 1409-1423.

13. David SS, O'Shea VL, Kundu S (2007). Base-excision repair of oxidative DNA damage. *Nature* 447: 941-950.

14. Valerie K and Povirk LF (2003). Regulation and mechanisms of mammalian double-strand break repair. *Oncogene* 22: 5792-5812.

15. Peterson CL and Cote J (2004). Cellular machineries for chromosomal DNA repair. *Genes Dev* 18: 602-616.

16. Loizou JI, Murr R, Finkbeiner MG, et al. (2006). Epigenetic information in chromatin: the code of entry for DNA repair. *Cell Cycle* 5: 696-701.

17. Margison GP, Butt A, Pearson SJ, et al. (2007). Alkyltransferase-like proteins. *DNA Repair (Amst)* 6: 1222-1228.

18. Imai K and Yamamoto H (2008). Carcinogenesis and microsatellite instability: the interrelationship between genetics and epigenetics. *Carcinogenesis* 29: 673-680.

The Cell Cycle

Summary

> The control of cell division is critical to normal tissue structure and function. It is regulated by a complex interplay of many genes that control the cell cycle, with DNA replication (S phase) and mitosis as major checkpoints

> The cell cycle is tightly regulated to minimise transmission of genetic damage to subsequent cell generations

> Progression through the cell cycle is primarily controlled by cyclins, associated kinases and their inhibitors. Retinoblastoma (RB) and p53 are major suppressor genes involved in the G1/S checkpoint control

> Cancer may be perceived as the consequence of loss of cell cycle control and progressive genetic instability

Cell proliferation occurs through a series of stages that are collectively termed the cell cycle. The "cell cycle" refers to the set of ordered molecular and cellular processes during which genetic material is replicated and segregates between two newly generated daughter cells via the process of mitosis. The cell cycle can be divided into two phases of major morphological and biochemical change: M phase ("mitosis"), during which division is evident morphologically, and S phase ("synthesis"), during which DNA is replicated. These two phases are separated by so-called G ("gap") phases. G1 precedes S phase and G2 precedes M phase.

During progression through this division cycle, the cell has to resolve a number of critical challenges. These include ensuring that sufficient ribonucleotides are available to complete DNA synthesis, proof-reading, editing and correcting the newly-synthesised DNA; that genetic mate-

rial is not replicated more than once; that the spatial organisation of the mitotic spindle apparatus is operational; that the packing and the condensation of chromosomes is optimal; and that there is equal distribution of cellular materials between the daughter cells. Moreover, immediately before or after the cell cycle, various factors interact to determine whether the cell divides again or whether the cell becomes committed to a programme of differentiation or of cell death. Therefore, the term "cell cycle" is often used in a broad sense to refer to, as well as the basic, self-replicating cellular process, a number of connected processes which determine pre- and post-mitotic commitments. These may include the commitment to stop dividing in order to enter a quiescent state, to undergo senescence or differentiation, or to leave the quiescent state to re-enter mitosis.

Molecular architecture of the cell cycle

The molecular ordering of the cell cycle is a complex biological process dependent upon the sequential activation and inactivation of molecular effectors at specific points of the cycle. Most current knowledge of these processes stems from experiments carried out in the oocyte of the frog, *Xenopus laevis*, or in yeast, either *Saccharomyces cerevisiae* (budding yeast) or *Schizosaccharomyces pombe* (fission yeast). The Xenopus oocyte is, by many criteria, one of the easiest cells to manipulate in the laboratory. Its large size (over a millimetre in diameter) means that cell cycle progression can be monitored visually in single cells. Microinjections can be performed for the purpose of interfering with specific functions of the biochemical machinery of the cell cycle. The Xenopus oocyte has proven to be an invaluable tool in the study of the biochemistry of the cell cycle, allowing, among other findings, the elucidation of the composition and regulation of maturation promoting factor (MPF), a complex enzyme comprising a kinase (p34cdc2) and a regulatory subunit (cyclin B), which drives progression from G2 to M phase[1]. In contrast, the exceptional genetic plasticity of yeast has allowed the

identification of scores of mutants with defects in cell cycle progression; in mammalian cells, these mutations would have been lethal and it would therefore have been impossible to characterise them. These mutants were called "cdc", for cell division cycle mutants, and many of them have been accorded wider recognition through the application of their names to the mammalian homologues corresponding to the yeast genes.
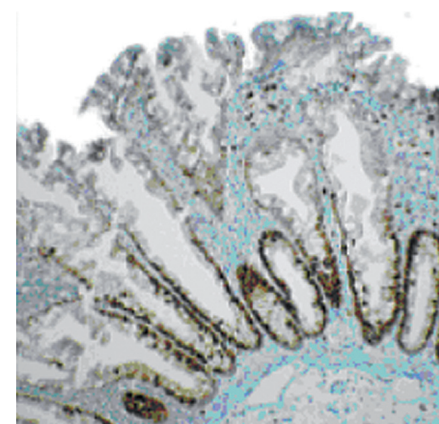

Fig. 3.3.1 Proliferating cells in the basal parts of the colonic crypts, visualised by immunohistochemistry (stained brown)
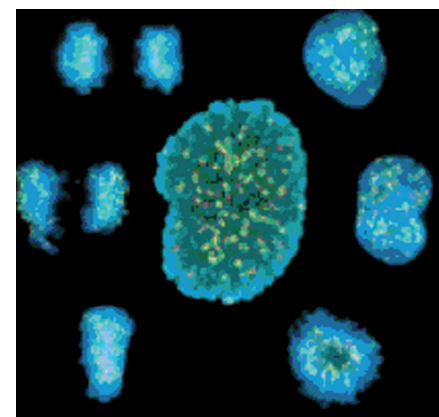

Fig. 3.3.2 A human osteosarcoma cell nucleus during mitosis. Cell division proceeds clockwise from upper right through interphase, prophase (centre), prometaphase, metaphase, anaphase and telophase. During the cycle, the chromosomes are replicated, segregated and distributed equally between the two daughter cells

One of the earliest genes to be identified in this way was cdc2. Isolated in *S. pombe*, cdc2 was determined to be able to correct a G2 cell cycle arrest defect. The product of this gene, a serine-threonine kinase of molecular weight 32–34 000 daltons, was subsequently shown to be the yeast homologue of the kinase contained in the Xenopus MPF. This enzyme became the paradigm of a class of enzymes now called cyclin-dependent kinases (CDKs). In their active form, CDKs form heterodimers with cyclins, a class of molecules synthesised in a time-dependent manner during the cell cycle. The progression of the cell cycle depends upon the sequential activation and inactivation of cyclin/CDK complexes [1], a process which requires the synthesis of cyclins, the formation of a complex between a specific cyclin and a CDK and post-translational modification of the CDK to convert the enzyme to an active form (Figure 3.3.3).

Progression through the cell cycle as mediated by cyclins is, in turn, determined by factors categorised as having either regulatory (upstream) or effector (downstream) roles. Upstream of cyclin/CDKs are regulatory factors called cyclin-dependent kinase inhibitors (CDKIs), which regulate the assembly and the activity of cyclin/CDK complexes. Downstream of cyclin/CDKs are effector molecules, essentially transcription factors, which

control the synthesis of proteins that mediate the molecular and cellular changes occurring during each phase.

CDKIs are small proteins that form complexes with both CDKs and cyclins [2]. Their role is primarily to inhibit the activities of cyclin/CDK complexes and to negatively regulate cell cycle progression. They constitute the receiving end of many of the molecular cascades signalling growth promotion or suppression of growth. Thus CDKIs may be considered as the interface between the cell cycle machinery and the network of molecular pathways which signal proliferation, death or stress responses. However, by virtue of their complexing properties, some CDKIs also play a positive role in cell cycle progression by facilitating the assembly of cyclin/CDK complexes. For example, p21, the product of the CDKN1A gene (also known as WAF1/CIP1), promotes the assembly of cyclin D/cdk2 complexes in G1 at a stoichiometric 1:1 ratio, but inhibits the activities of these complexes when expressed at higher levels. There are three main families of CDKIs, each with distinct structural and functional properties: the WAF1/CIP1 family (p21), the KIP family (p27, p57) and the INK4 family (p16, p15, p18) (Figure 3.3.3).

Downstream effectors of cyclin/CDKs include proteins mediating three main functional categories: (1) those involved in the control of the enzymes responsible for DNA replication, proof-reading and repair, (2) those involved in chromosome and chromatin remodelling and in the control of genomic integrity, and (3) those involved in the mechanics of cell division (including the formation of the centrosome and the mitotic spindle, and in the resorption of the nuclear membrane). These processes require the coordinated synthesis of hundreds of cellular proteins. Transcription factors of the E2F family play a critical role in the control of gene transcription during cell cycle progression (Figure 3.3.4). In G1, factors of the E2F family are bound to their DNA targets but are maintained in a transcriptionally inactive state by the binding of proteins of the retinoblastoma (pRb) protein family. At the G1/S transition, the sequential phosphorylation of pRb by several cyclin/CDKs dissociates pRb from the complexes, allowing E2Fs to interact with transcription co-activators and to initiate mRNA synthesis [3].

Through this mechanism, E2Fs exert a dual function both as transcriptional repressors in G1, when bound to pRb, and as transcriptional activators in G1/S and in S phase, after dissociation of pRb from the complex. Recent observations suggest that transcriptional repression by E2Fs is essential to prevent the premature activation of cell cycle effectors,

Gene (chromosome)	Product	Type of alteration	Role in cell cycle	Involvement in cancer
p53 (17p13)	p53	Mutations, deletions	Control of p21, 14-3-3σ, etc.	Altered in over 50% of all cancers
CDKN2A (9p22)	p16 and p19arf	Mutations, deletions, hypermethylation	Inhibition of CDK4 and 6	Altered in 30-60% of all cancers
RB1 (13q14)	pRb	Deletions	Inhibition of E2Fs	Lost in retinoblastomas, altered in 5-10% of other cancers.
CCND1	Cyclin D1	Amplification	Progression into G1	10-40% of many carcinomas
CDC25A, CDC25B	cdc25	Overexpression	Progression in G1, G2	10-50% of many carcinomas
KIP1	p27	Down-regulation	Progression in G1/S	Breast, colon and prostate cancers

Table 3.3.1 Cell cycle regulatory genes commonly altered in human cancers

which would scramble the temporal sequence of molecular events and preclude cell cycle progression.

Cell cycle checkpoints

The notion of "cell cycle checkpoints" is also derived from early studies in Xenopus oocytes and in yeast mutants. In *S. cerevisiae*, commitment to the mitotic cycle requires the crossing of a "restriction point" called the start transition. Failure to cross this transition results in cells being blocked in the G1 phase of the cycle. Another control point has been clearly identified after S phase, at the transition between G2 and M phases. Cells unable to cross this checkpoint may remain blocked in a pre-mitotic, tetraploid state. Physiologically, this checkpoint is active in germ cells during the second division of meiosis: cells that have undergone the first, asymmetric division of the meiotic cycle arrest in G2 until completing the second division, which is triggered by fertilisation. This concept of "cell cycle checkpoints" was later extended to all mammalian cells[4,5]. It is now common to envisage the mammalian cell cycle as a succession of checkpoints that must be negotiated in order for division to be achieved. There is no clear agreement on how many such checkpoints exist in the mammalian cell cycle, or on their exact position.

Control of cdk1 at G2/M transition

The regulation of the complex between cdk1 (also called p34cdc2) and cyclin B exemplifies how different factors co-operate to control the activation of cyclin/CDK complexes at a cell cycle checkpoint. This activation process requires co-operation between three levels of regulation: association between the two partners of the complex, post-translational modifications of the kinase and of the cyclin, and escape from the negative regulation exerted by the CDKIs.

In early G2, cdk1 is in an inactive form. Its activation requires first association with cyclin B, followed by post-translational modification of
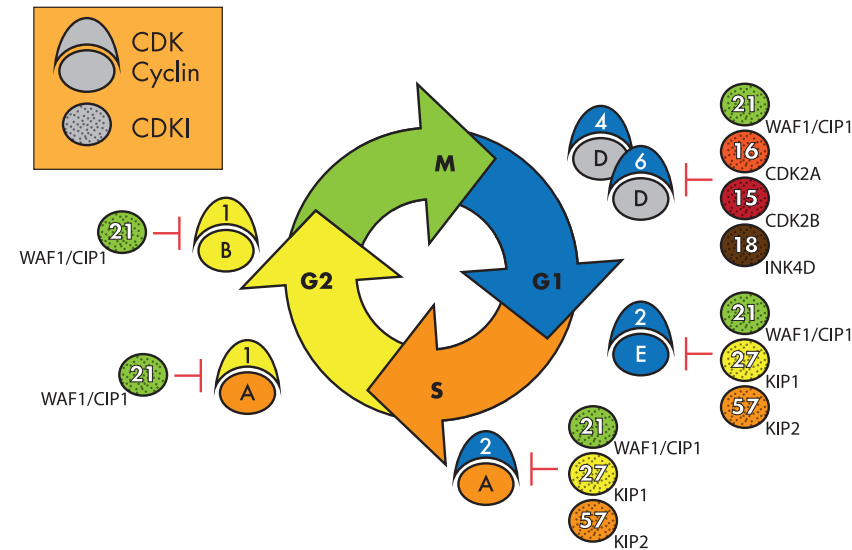

Fig. 3.3.3 *The progression of the cell cycle depends upon the sequential activation and inactivation of cyclin/CDK complexes. This process requires the synthesis of cyclins, the formation of a complex between a specific cyclin and a CDK, and modification of the CDK to convert this enzyme to an active form. The enzyme's activity may be disrupted by a specific inhibitor, a CDKI*
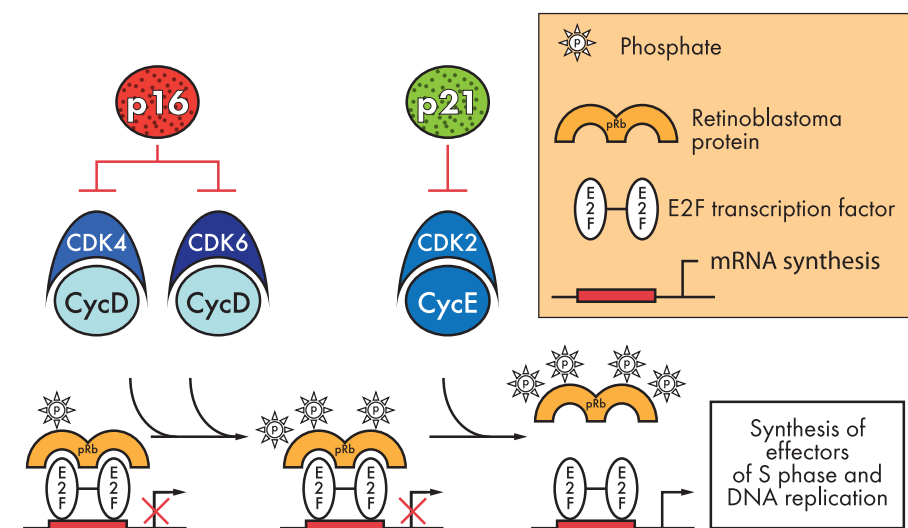

Fig. 3.3.4 *Progression from G1 to S phase is regulated by phosphorylation of the retinoblastoma protein (pRb), in the absence of which DNA replication cannot proceed*

the kinase itself. This modification includes phosphorylation of a conserved threonine residue (Thr161) by a kinase complex called CAK (CDK-activating kinase), as well as dephosphorylation of two residues localised within the active site of the enzyme, a threonine (Thr14) and a tyrosine (Tyr15). The removal of these phosphate groups is carried out by the dual-specificity phosphatases of the cdc25 group, comprising three isoforms in humans (A, B and C). Activation of these phosphatases is therefore crucial for the activation of cyclin B/cdk1 complexes. The phosphatase is directly controlled by a number of regulators, including plk1 (polo-like kinase), an activating kinase, pp2A, (protein phosphatase 2A), an inhibitory phosphatase and 14-3-3s, a signal transduction molecule which complexes with cdc25, sequesters it in the cytoplasm and thus prevents it from dephosphorylating its nuclear targets. Of course, the action of cdc25 phosphatases is counteracted by kinases that restore the phosphorylation of Thr14 and Tyr15, named wee1 and mik1 [2].

Following the activation process outlined above, the cyclin B/cdk1 complex is potentially able to catalyse transfer of phosphates to substrate proteins. However, in order to achieve this, it has to escape the control exerted by CDKIs, such as p21. The function of this CDKI is itself controlled by several activators, including BRCA1, the product of a breast cancer susceptibility gene. The p21 protein is removed from the complex by a still poorly understood phosphorylation process, which also drives rapid degradation of the protein by the proteasome. This leaves the cyclin B/cdk1 complex ready to function, after a final step of autophosphorylation, in which cdk1 phosphorylates cyclin B. The complex is now fully active and ready to phosphorylate many different substrates, such as nuclear lamins, during entry into mitosis.

Regulation of the cell cycle and control of genetic stability

During the cell cycle, a number of potential problems may result in damage to the genome. These problems may arise at three distinct stages: (1)

during DNA replication, especially if the cell is under conditions of stress that favour the formation of DNA damage (irradiation, exposure to carcinogens, etc.); (2) following the termination of DNA replication, when the cell effectively "switches off" its DNA synthesis machinery; and (3) during M phase, when the cell must negotiate the delicate task of segregating chromatids equally. A tight coupling between these processes and cell cycle regulation is therefore crucial to allow the cell to pause during the cell cycle in order to afford the time necessary for the successful completion of all the operations of DNA and chromosome maintenance. Failure to do this may result in both genetic and genomic instabilities, which are hallmarks of cancer. Genetic instability is characterised by an increased rate of gene mutation, deletion or recombination (essentially due to defects in DNA repair). Genomic instability results in chromosome translocations, loss or duplication of large chromosome fragments and aberrant chromosome numbers (aneuploidy).

Tens of molecules have been identified as components of the signalling cascades which couple detection of DNA damage and regulation of the cell cycle. One of these is the product of the tumour suppressor gene p53. p53 is specifically activated after various forms of direct DNA damage (such as single or double strand breaks in DNA) and regulates the transcription of several inhibitors of cell cycle progression, particularly at the G1/S and G2/M transitions [6,7]. Other important molecules in this coupling process include the checkpoint kinases chk1 and chk2. Chk1 is activated after replication blockage during S-phase. In turn, chk1 activates wee1 and mik1, two kinases that counteract the action of cdc25 and keep cdk1 in an inactive form. Thus, through activation of chk1, the cell triggers an emergency mechanism that ensures that cells with incompletely replicated DNA cannot enter mitosis.

The cell cycle and cancer

Genes involved in cell cycle control are important among those subject to the genetic altera-

tions that give rise to cancer [8,9]. However, the proliferation of cancer cells requires that the cells retain functional cell cycle processes. The cell cycle alterations seen in cancer are mainly confined to two major sets of regulators: those involved in the negative control of cell cycle progression (inactivation of which leads to accelerated and unchecked cell proliferation) and those involved in coupling the maintenance of genome integrity to the cell cycle (inactivation of which results in cells having gene alterations that progressively accumulate during carcinogenesis) (Table 3.3.1) [8]. Most of the genes corresponding to these two categories fall within the group of tumour suppressors, and many of them are also direct participants in DNA repair processes.

The gene which encodes p16 (CDKN2A/INK4A) has been established as a tumour suppressor gene [10], and mutations and deletions at this site are commonly found in primary human tumours, especially melanoma (although the contribution of another protein encoded by the same locus on chromosome 9p, p14ARF, to suppressor activity remains to be determined). Unlike the CDKN2A/INK4A gene, the CDKN1A gene (encoding p21) is rarely disrupted in cancer. As p21 plays many roles in the negative regulation of almost all phases of the cell cycle, loss of this function might be expected to result in uncontrolled cell division. This is apparently not the case, as

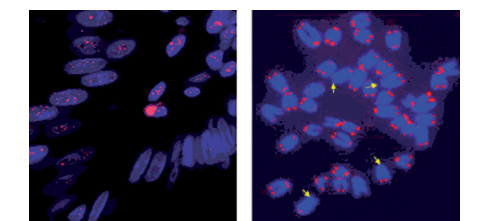Fig. 3.3.5 *Telomeres contain repetitive DNA sequences that cap the ends of chromosomes. Fluorescence in situ hybridization analysis of human interphase (left) and mouse metaphase chromosome spread (right) is shown, using oligonucleotide probes specific for telomere (red) DNA sequences, and the DNA dye DAPI (blue). Arrows indicate the loss of telomere signal in cells obtained from knockout mice containing mutated telomerase gene*

mice lacking the CDKN1A gene do not show an increased frequency of cancer. This observation illustrates one of the most important characteristics of cell cycle regulatory mechanisms: there is a large degree of redundancy and overlap in the function of any particular effector. Therefore, cancer-causing deregulation of the cell cycle requires a combination of many alterations in genes encoding proteins that, either alone or in concert, are critical for the control of cell division.

Apart from inactivation of negative regulators, a few cell cycle genes may be activated as oncogenes, in that their alteration results in enhanced activity leading to accelerated cell proliferation. The best example of such a cell cycle oncogene is CCND1, the gene encoding cyclin D1, a G1-specific cyclin [11]. This gene is located on chromosome 11p13, within a large region that is amplified in up to 20% of several carcinomas (e.g. breast, head and neck, oesophageal and lung cancers).

There is also limited evidence for transcriptional activation of cyclin A (an S-phase cyclin) and for activating mutations of CDK4 (one of the partners of cyclin D1) in some cancers. Indeed, the high complexity of cell cycle effectors provides an extremely diverse range of possibilities for cancer-associated alterations. In this respect, cancer can be seen as, fundamentally, a disease of the cell cycle.

Telomeres and telomerase

Telomeres are specialised structures at the ends of eukaryotic chromosomes (Figure 3.3.5). These structures contain many copies of G-rich repeats that are highly conserved in most eukaryotic species. Telomeres have arisen as an evolutionary response to the problem posed by the development of linear chromosomes. Chromosome ends may be recognised as DNA breaks; thus cells needed a mechanism to protect natural chromosome ends by "hiding" them from DNA damage recognition machinery. A wealth of studies demonstrated that telomeres function to protect

chromosome ends from fusion events and are therefore critical for chromosome stability and genomic integrity.

Studies in the early 1960s suggested that specific human cells can only divide about 50 times, a phenomenon known as the Hayflick limit [12]. This was based on the assumption that the number of divisions is determined by both the initial length of the telomeres and the rate of telomere shortening. Experimental studies demonstrated that the telomeres of normal human somatic cells shorten by 50 to 150 base pairs every time cell division occurs, supporting the theory of the Hayflick limit. Subsequent studies provided evidence that critically short telomeres cease to function as protective structure and cause the cell to trigger cell suicide (apoptosis) or undergo senescence (permanent arrest of cell proliferation) (Figure 3.3.6). Therefore telomerase shortening during cell division appears to act as a cell division counting mechanism. Normal cells thus have a limited proliferative capacity, and this acts as a major barrier against carcinogenesis.

While it is now recognised that telomeres have many more functions than simply protecting chromosome ends, the initial concept of a replication barrier is still valid. Consistent with this notion, recent studies showed that most human cancer activate telomerase at some point during the process of tumour development and progression, a phenomenon typically absent in normal cells. Cells that have accumulated some carcinogenic changes are unable to form clinically significant cancers unless this proliferation barrier imposed by telomere clock is breached. This is supported by the evidence that more than 85% of all cancers achieve this by expressing an enzyme, telomerase, that synthesises new telomeric DNA to replace the sequences lost during cell division [13].

The catalytic subunit of human telomerase, hTERT, was cloned in 1997 [14]. It has subsequently been shown that genetic manipulations of hTERT which result in inhibition of telomerase activity in tumour cells limit their proliferation

and often result in cell death. This raises the possibility that telomerase inhibitors may be a very useful form of therapy for many or most types
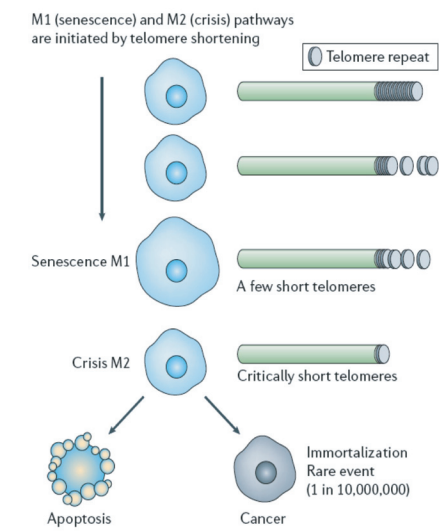

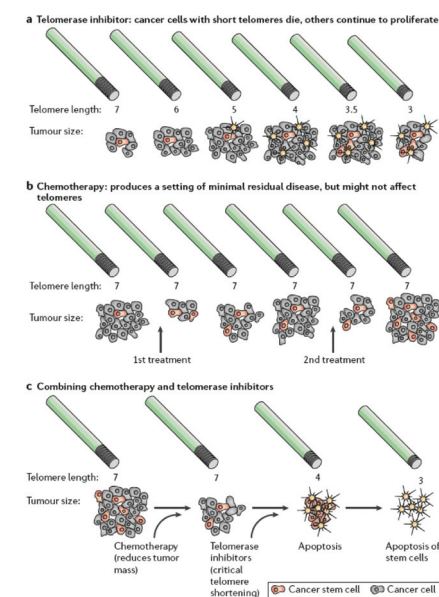**Fig. 3.3.6** *Telomere-mediated senescence and tumorigenesis [13]*

Fig. 3.3.7 *Cancer therapy targeting telomerase [13]*

of cancer. However, in tumours with long telomeres, it may take many cell divisions before telomerase inhibitors exert an anti-tumour effect. When such drugs are developed they will therefore need to be carefully integrated with other anticancer treatments.

It is interesting to note that not all tumours need to activate telomerase. Studies showed that approximately 10% of human tumours rely on a telomerase-independent mechanism to maintain their telomeres. This phenomenon, known

as the alternative lengthening of telomeres (ALT) mechanism, relies on recombination between telomeres [15].

Telomerase assays have not yet entered routine clinical practice, but there is considerable interest in their possible use for cancer diagnosis and prognosis. For example, telomerase assays of urine sediments may be useful for diagnosis of urinary tract cancer, and telomerase activity levels may be a predictor of outcome in neuroblastoma [16]. In summary, there have been many

important discoveries in the field of telomere research, suggesting that telomeres may be an attractive target for the development of therapeutic intervention in different types of human cancer (Figure 3.3.7). However, more clinical studies are needed to test telomerase-targeted approaches that could lead to effective cancer interventions with minimal side effects. Studies aiming to close many gaps in our understanding of telomere-maintenance mechanism will be instrumental in these endeavours.

REFERENCES

1. Bartek J and Lukas J (2007). DNA damage checkpoints: from initiation to recovery or adaptation. *Curr Opin Cell Biol* 19: 238-245.

2. Boutros R, Lobjois V, Ducommun B (2007). CDC25 phosphatases in cancer cells: key players? Good targets? *Nat Rev Cancer* 7: 495-507.

3. Du W and Pogoriler J (2006). Retinoblastoma family genes. *Oncogene* 25: 5190-5200.

4. Hartwell LH and Weinert TA (1989). Checkpoints: controls that ensure the order of cell cycle events. *Science* 246: 629-634.

5. Orford KW and Scadden DT (2008). Deconstructing stem cell self-renewal: genetic insights into cell cycle regulation. *Nat Rev Genet* 9: 115-128.

6. Hainaut P and Hollstein M (2000). p53 and human cancer: the first ten thousand mutations. *Adv Cancer Res* 77: 81-137.

7. Stark GR and Taylor WR (2006). Control of the G2/M transition. *Mol Biotechnol* 32: 227-248.

8. Lobrich M and Jeggo PA (2007). The impact of a negligent G2/M checkpoint on genomic instability and cancer induction. *Nat Rev Cancer* 7: 861-869.

9. Malumbres M and Barbacid M (2001). To cycle or not to cycle: a critical decision in cancer. *Nat Rev Cancer* 1: 222-231.

10. Strohmaier H, Spruck CH, Kaiser P, et al. (2001). Human F-box protein hCdc4 targets cyclin E for proteolysis and is mutated in a breast cancer cell line. *Nature* 413: 316-322.

11. Santamaria D and Ortega S (2006). Cyclins and CDKS in development and cancer: lessons from genetically modified mice. *Front Biosci* 11: 1164-1188.

12. Hayflick L (1980). Recent advances in the cell biology of aging. *Mech Ageing Dev* 14: 59-79.

13. Shay JW and Wright WE (2006). Telomerase therapeutics for cancer: challenges and new directions. *Nat Rev Drug Discov* 5: 577-584.

14. Lingner J and Cech TR (1998). Telomerase and chromosome end maintenance. *Curr Opin Genet Dev* 8: 226-232.

15. Muntoni A and Reddel RR (2005). The first molecular details of ALT in human tumor cells. *Hum Mol Genet* 14 Spec No. 2: R191-R196.

16. Hiyama E and Hiyama K (2002). Clinical utility of telomerase in cancer. *Oncogene* 21: 643-649.

Cell Death

3.4

Summary

> The term apoptosis refers to a type of cell death that occurs both physiologically and in response to external stimuli, including X-rays and anticancer drugs

> Apoptotic cell death is characterised by distinctive morphological changes different from those occurring during necrosis, which follows ischaemic injury or toxic damage

> Apoptosis is regulated by several distinct signalling pathways. Dysregulation of apoptosis may result in disordered cell growth and thereby contribute to carcinogenesis

> Selective induction of apoptosis in tumour cells is among current strategies for the development of novel cancer therapies

In the adult organism, the number of cells is kept relatively constant through cell death and division, and deregulation of this balance (homeostasis) may trigger pathological conditions such as neurodegenerative diseases and cancer. Apoptosis and necrosis are two forms of cell death, with distinct morphological and biochemical features. While apoptosis accounts for most physiological cell deaths, necrosis is usually induced in pathological situations by accidental, acute damage to cells. Apoptosis, or programmed cell death, is a tightly regulated process under normal conditions that facilitates fundamental processes such as development (for example, by removal of unwanted tissue during embryogenesis) and the immune response (for example, by elimination of self-reactive T cells). This type of cell death is distinguished from necrosis both morphologically (Figures 3.4.1 and 3.4.2) and functionally. Specifically, apoptosis involves single cells rather than areas of tissue and does not provoke inflammation (Figures 3.4.3 and

3.4.4). Tissue homeostasis is dependent on controlled elimination of unwanted cells, often in the context of a continuum in which specialisation and maturation is ultimately succeeded by cell death in what may be regarded as the final phase of differentiation. Apart from elimination in a physiological context, cells that have been lethally exposed to cytotoxic drugs or radiation may be subject to apoptosis.

The process of apoptosis can be described by reference to distinct phases, termed "regulation", "effector" and "engulfing" [1]. The regulatory phase includes all the signalling pathways that culminate in commitment to cell death. Some of these pathways regulate only cell death, but many of them have overlapping roles in the control of cell proliferation, differentiation, responses to stress and homeostasis. Critical to apoptosis signalling are the "initiator" caspases (including caspase-8, caspase-9 and caspase-10) whose role is to activate the more abundant "effector" caspases (including caspase-3 and caspase-7) which, in turn, bring about the morphological change indicative of apoptosis. Finally, the engulfing process involves the recognition of cellular "remains" and their elimination by the engulfing activity of surrounding cells.

Identification of genes mediating apoptosis in human cells has been critically dependent on definition of the ced genes in the nematode Caenorhabditis elegans, members of this gene family being homologous to human BCL2 (which suppresses apoptosis), APAF-1 (which mediates caspase activation) and the caspases themselves (proteases which mediate cell death). The centrality of apoptosis to cancer biology is indicated by excess tumorigenesis in BCL2-transgenic and p53-deficient mice. An appreciation of apoptosis provides a basis for the further development of novel and conventional cancer therapy.

The role of cell death in tumour growth

Apoptosis, or lack of it, may be critical to tumorigenesis [2]. BCL2, a gene mediating resistance to apoptotic stimuli, was discovered at the t(14:18) chromosomal translocation in low-grade B cell non-Hodgkin lymphoma. It thus became apparent that neoplastic cell expansion could be attributable to decreased cell death rather than rapid proliferation. Defects in apoptosis allow neoplastic cells to survive beyond senescence, thereby providing protection from hypoxia and oxidative stress as the tumour mass expands. Growth of tumours, specifically in response to chemical carcinogens, has been correlated with altered rates of apoptosis in affected tissues as cell populations with altered proliferative activity emerge. Paradoxically, growth of some cancers, specifically including breast, has been positively correlated with increasing apoptosis [3].

Interrelationships between mitogenic and apoptotic pathways

A dynamic relationship between regulation of growth/mitosis and apoptosis may be demonstrated using a variety of relevant signalling pathways. Many differing promoters of cell proliferation have been found to possess proapoptotic activity. Thus, ectopic expression of the C-MYC oncogene (normally associated with proliferative activity) causes apoptosis in cultured cells subjected to serum deprivation (which otherwise prevents proliferation). Oncogenes that stimulate mitogenesis can also activate apoptosis. These include oncogenic ras, myc and E2F. Mutations in E2F that prevent its interaction with the retinoblastoma protein (pRb) accelerate S phase entry and apoptosis. A function of pRb is to suppress apoptosis: pRb-deficient cells seem to be more susceptible to p53-induced apoptosis.

Agents such as radiation or cytotoxic drugs cause cell cycle arrest and/or cell death [4]. The DNA damage caused by radiation or drugs is detected by various means (Figure 3.4.2). DNA-dependent protein kinase and the ataxia telangiectasia mutated gene (ATM) (as well as the related ATR protein) bind to damaged DNA and initiate phosphorylation cascades to transmit damage signals. DNA-dependent protein kinase is believed to play

a key role in the response to double-stranded DNA breaks. ATM plays an important part in the response to DNA damage caused by ionizing radiation, controlling the initial phosphorylation of proteins such as p53, Mdm2, BRCA1, Chk2 and Nbs1. Other sensors of DNA damage include mammalian homologues of the PCNA-like yeast proteins Rad1, Rad9 and Hus1, as well as the yeast homologue of replication factor C, Rad17. Specific molecules detect nucleotide mismatch or inappropriate methylation. Following exposure of mammalian cells to DNA-damaging agents, p53 is activated and among many "targets" consequently upregulated are the cyclin-dependent kinase inhibitor p21 (which causes G1 arrest) and Bax (which induces apoptosis). Thus, the tumour suppressor gene p53 mediates two responses to DNA damage by radiation or cytotoxic drugs: cell cycle arrest at the G1 phase of the cell cycle and apoptosis. The serine/threonine kinase Chk2 is also able to positively interact with p53 and BRCA1. Chk2 and the functionally related Chk1 kinase appear to have a role in the inhibition of entry into mitosis via inhibition of the phosphatase Cdc25.

The regulatory phase

Two major apoptotic signalling pathways have been identified in mammalian cells (Figure 3.4.5). The "extrinsic" pathway depends upon the conformational change in certain cell surface receptors following the binding of respective ligands. The "intrinsic" pathway involves mitochondrial function and is initiated by growth factor deprivation, corticosteroids or DNA damage induced by radiation or cytotoxic drugs.

Cell surface receptors

Apoptosis may be induced by signalling molecules, usually polypeptides such as growth factors or related molecules, which bind to "death" receptors on the cell surface [2]. Such cell death was initially investigated in relation to the immune response, but has much wider

ramifications. The best-characterised receptors belong to the tumour necrosis factor (TNF) receptor gene superfamily [5] (Figure 3.4.5). In addition to a ligand-binding domain, death receptors contain homologous cytoplasmic sequence termed the "death domain". Members of the family include Fas/APO-1/CD95 and TNF-1 receptor (which binds TNFa). Activation of the Fas (or CD95) receptor by its specific ligand (FasL or CD95L) results in a conformational change such that the "death domain" interacts with the adaptor molecule FADD which then binds procaspase-8. In some cell types, drug-induced apoptosis is associated with Fas activation. Ultraviolet irradiation directly activates the Fas receptor in the absence of ligand. TRAIL (TNF-related apoptosis-inducing ligand, Apo-2L) has 28% amino acid identity to FasL. TRAIL induces cell death only in tumorigenic or transformed cells and not in normal cells [5].

The regulation of apoptosis by BCL2 family genes

While the members of the "death receptor" family and their ligands have structural elements in common, agents and stimuli initiating the mitochondrial pathway to apoptosis are diverse. Common to these stimuli, however, is a change in mitochondrial function, often mediated by members of the BCL2 family [6]. In humans, at least 16 homologues of BCL2 have been identified. Several family members (including Bcl-2, Bcl-xL, Bcl-W) suppress apoptosis, while others induce apoptosis and may be subdivided on the basis of their ability to dimerize with Bcl-2 protein (Bad, Bik, Bid) or not (Bax, Bak). Phosphorylation of Bad protein by a specific (Akt/PKB) and other kinases prevents dimerization with Bcl-2 and promotes cell survival. At least two distinct mechanisms of action are rec-
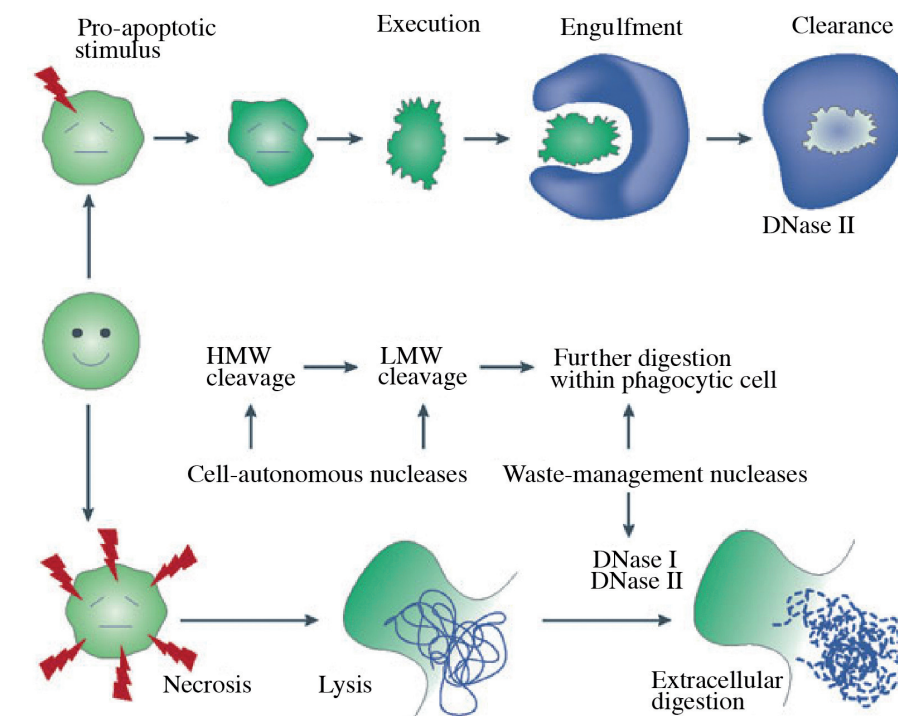

Fig. 3.4.1 *Apoptosis and necrosis are distinguished by characteristic morphological changes and molecular machineries involved in these processes*

ognised: the binding of Bcl-2 (or other members of the family) with either pro- or anti-apoptotic members of the Bcl-2 family or the formation of pores in mitochondrial membranes. Bcl-xL is a potent death suppressor that is upregulated in some tumour types. Bax is a death promoter that is inactivated in certain types of colon cancer, stomach cancer and in haematopoietic malignancies. By dint of relevant binding sites, Bax is under the direct transcriptional control of p53.

Involvement of mitochondria

Apoptosis induced by cytotoxic drugs is accompanied by critical changes in mitochondria [2,7]. Such apoptotic stimuli induce translocation of Bax from cytosol to mitochondria, which induces release of cytochrome c (Figure 3.4.5). Loss of transmembrane potential follows cytochrome c release and is dependent on caspase activation (see below), whereas cytochrome c release is not. Bcl-2 and Bcl-xL reside chiefly in the outer mitochondrial membrane. Bcl-2, Bcl-xL and Bax can form ion channels when they are added to synthetic membranes, and this may be related to their impact on mitochondrial biology [2,8].

In the cytosol after release from mitochondria, cytochrome c activates the caspases through formation of a complex (the "apoptosome") with Apaf-1 (apoptotic-protease activating factor-1), procaspase-9 and ATP. It appears that Bcl-2/Bcl-xL may suppress apoptosis by either preventing release of cytochrome c or interfering with caspase activation by cytochrome c and Apaf-1. Sustained production of nitric oxide (NO) may cause the release of mitochondrial cytochrome c into the cytoplasm and thus contribute to the activation of caspases. However, nitric oxide is involved in several aspects of apoptosis and may act both as a promoter and inhibitor depending on conditions [9].

The effector and engulfing phases

In mammals at least 13 proteases that mediate the breakdown of cell structure during apoptosis have been identified and are designated caspases-1 through -13 [10]. All possess an active

site cysteine and cleave substrates after aspartic acid residues. They exist as inactive zymogens, but are activated by different processes which most often involve cleavage of their pro-forms (designated procaspase-8, etc.) at particular sites, thereby generating subunits which form active proteases consisting of two large and two small subunits. Proteolytic cascades may occur with some caspases operating as upstream initiators (which have large N-terminal
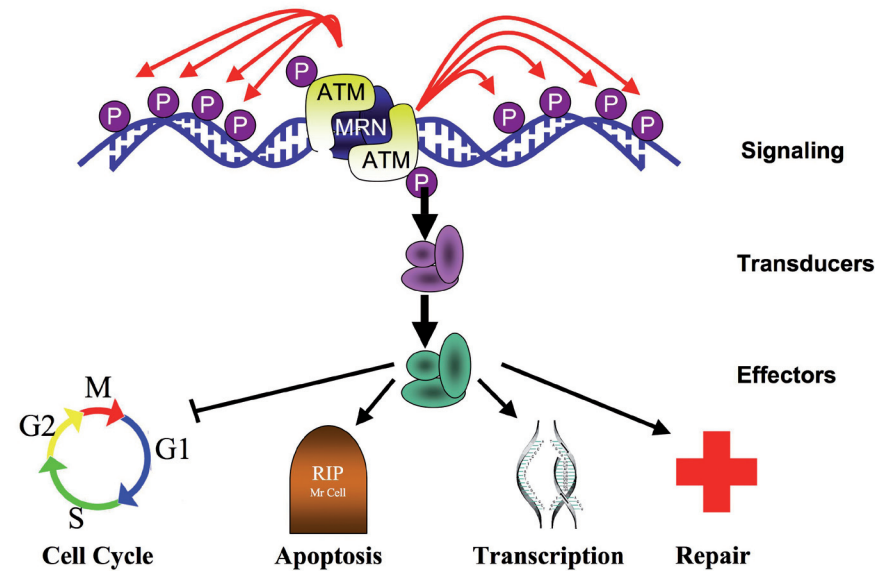

Fig. 3.4.2 *Programmed cell death (apoptosis) is a critical cellular process that may be triggered in response to DNA damage . Source: Zdenko Herceg and Rabih Murr, unpublished*
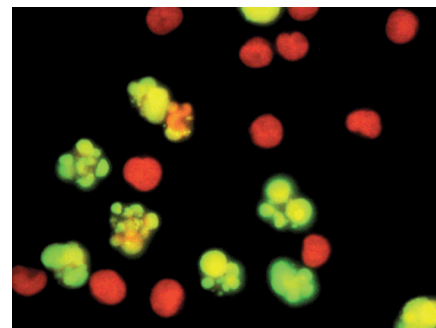

Fig. 3.4.3 *Cancer cells treated with chemotherapy agents often die via apoptosis. The image above shows human lymphoma cells treated with the chemotherapy agent camptothecin. The cells that are undergoing apoptosis appear yellow and show the characteristic membrane blebbing seen in cells dying via apoptosis*
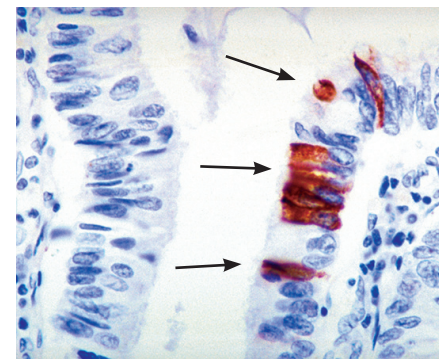

Fig. 3.4.4 *Apoptotic cells in an adenoma, visualised by immunohistochemistry (red). Apoptosis is restricted to single cells, unlike necrosis, which typically involves groups of cells. Apoptosis does not produce an inflammatory response.*

prodomains and are activated by protein-protein interaction) and others being downstream effectors (activated by protease cleavage). As noted earlier, at least two pathways of caspase activation can be discerned: one involving FADD or similar protein-protein complexes and the other mediated by release of cytochrome c. In the former, affinity labelling suggests that caspase-8 activates caspases-3 and -7 and that caspase-3 in turn may activate caspase-6. On the other hand, release of cytochrome c into the cytoplasm results in the activation of caspase-9 which in turn activates caspase-3.

Though the intrinsic pathway to caspase-3 activation may be distinguished from the extrinsic pathway (i.e. that activated by Fas, etc.), some interaction is demonstrable. Thus, caspase-9 is able to activate caspase-8. Nonetheless, the pathways are separate to the extent that caspase-8 null animals are resistant to Fas- or TNF-induced apoptosis while still susceptible to chemotherapeutic drugs; cells deficient in caspase-9 are sensitive to killing by Fas/TNF but show resistance to drugs and dexamethasone. Finally, death of some cells may occur independently of caspase-3. Caspases-3, -7 and -9 are inactivated by proteins of the inhibitor of apoptosis family (IAPs) which are suppressors conserved throughout evolution. The IAP protein "survivin" is overexpressed in a large proportion of human cancers. Little is known about the involvement of caspase mutations in cancer.

Caspase substrates and late stages of apoptosis

Apoptosis was initially defined by reference to specific morphological change. In fact, both mitosis and apoptosis are characterised by a loss of substrate attachment, condensation of chromatin and phosphorylation and disassembly of nuclear lamins. These changes are now attributable to caspase activation and its consequences.

Most of the more than 60 known caspase substrates are specifically cleaved by caspase-3 and caspase-3 can process procaspases-2,

-6, -7 and -9 [11]. Despite the multiplicity of substrates, protease activity mediated by caspases is specific and seems likely to account for much of the morphological change associated with apoptosis. Caspases cleave key components of the cytoskeleton, including actin as well as nuclear lamins and other structural proteins. Classes of enzymes cleaved by caspases cover proteins involved in DNA metabolism and repair exemplified by poly(ADP-ribose) polymerase and DNA-dependent protein

kinase [12,13]. Other classes of substrates include various kinases, proteins in signal transduction pathways and proteins involved in cell cycle control, exemplified by pRb. Cleavage of some substrates is cell-type specific. Caspase activity accounts for internucleosomal cleavage of DNA, one of the first characterised biochemical indicators of apoptosis. ICAD/DFF-45 is a binding partner and inhibitor of the CAD (caspase-activated DNAase) endonuclease, and cleavage of ICAD by caspase-3 relieves
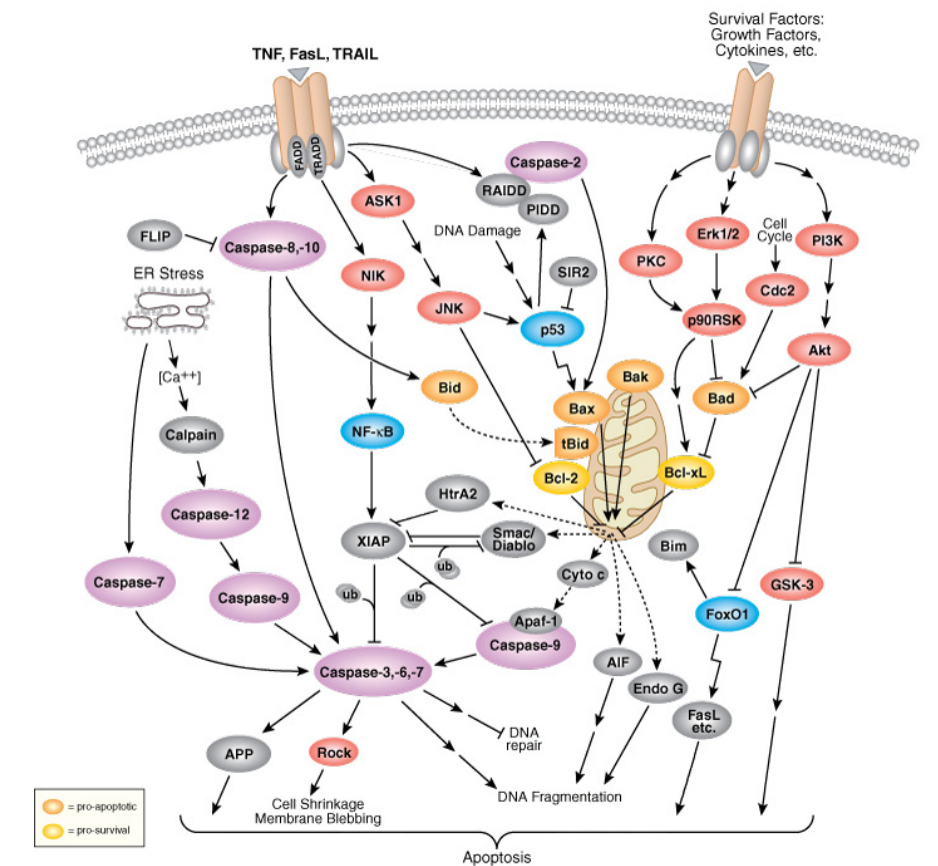
Fig. 3.4.5 *Apoptosis occurs when specific proteases (caspases) digest critical proteins in the cell. The caspases are normally present as inactive procaspases. Two pathways lead to their activation. The death receptor pathway (at the top and left side of the figure) is triggered when ligands bind to death receptors such as CD95/Fas. The mitochondrial pathway is triggered by internal insults such as DNA damage as well as by extracellular signals. In both pathways, procaspases are brought together. They then cleave each other to release active caspase. The binding of ligand (FasL or CD95L) to CD95 brings procaspase 8 molecules together; release of mitochondrial components bring procaspases 9 together. The active caspase 8 and 9 then activate other procaspases such as procaspase 3.*

Fig. 3.4.6 Neuroblastoma cells treated with ionizing radiation undergo apoptosis. The TUNEL assay was used to visualize apoptotic cells (green), before (A) and 24 hours after (B) treatment with X-rays (5 Gray). Close-up shows that the nuclei of the apoptotic cells are fragmented (C).

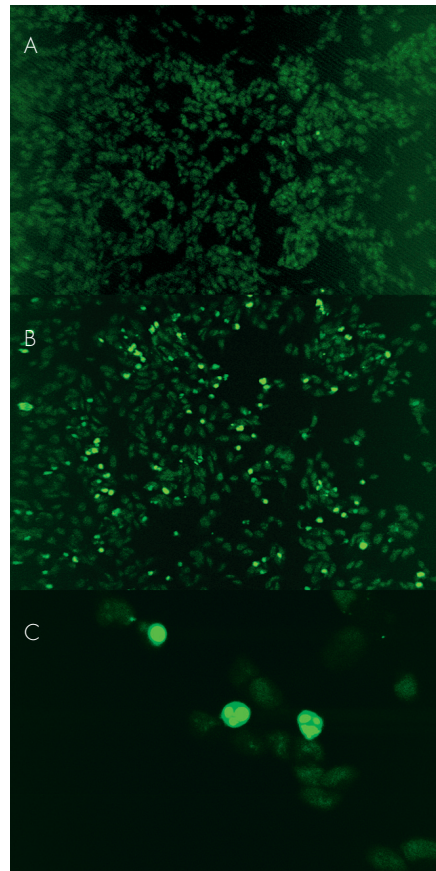*Fig. 3.4.7* Signaling pathways targeted by anticancer agents. Activation of cell signaling by a selected repertoire of protein tyrosine and serine/threonine kinases is the hallmark of many cancers. Certain tyrosine kinases and serine/threonine kinases have become targets for signal-transduction inhibitors. These are marked by colours according to their current status of development: approved drugs (red); drugs in the clinical trials (green); and drugs in preclinical trials (yellow). Green arrows denote direct transcriptional targets. Red lines show direct inhibitory pathways. Black arrows show direct activation events, and dashed arrows show events that are either indirect or questionable.

the inhibition and promotes the endonuclease activity of CAD (Figures 3.4.6 and 3.4.7).

Therapeutic implications

In theory, knowledge of critical signalling or effector pathways which bring about apoptosis provides a basis for therapeutic intervention, including the development of novel drugs to activate particular pathways.

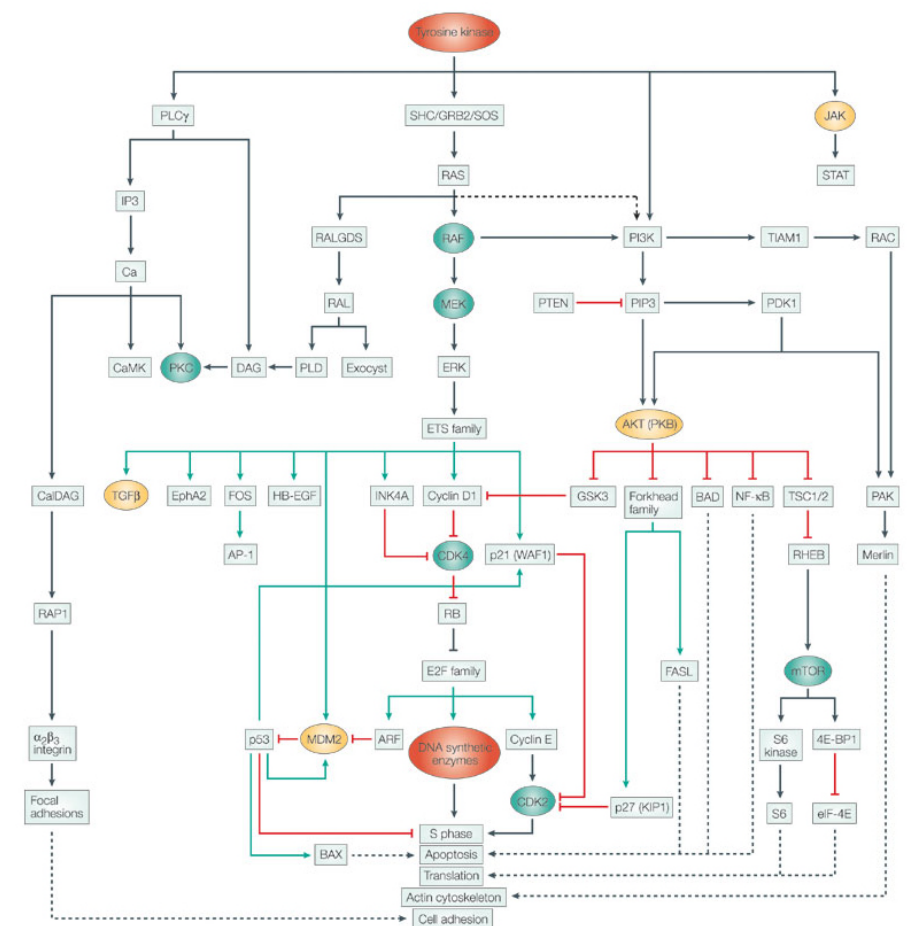Several options are under investigation [14]. More immediately, attempts are being made to exploit knowledge of apoptotic processes to increase the efficacy or specificity of currently available therapy. Simple answers have not emerged. Thus, for example, relatively increased expression of Bcl-2 (which, under many experimental conditions, inhibits apoptosis) is not necessarily indicative of poor prognosis, and the reverse appears true for some tumour types. In experimental systems, cells acquiring apoptosis defects (e.g. p53 mutations) can more readily survive hypoxic stress and the effects of cytotoxic drugs [2,9]. However, clinical studies have not consistently established that mutation of p53 is associated with poor response to chemotherapy [15].

The function of Bcl-2 family members may be subject to interference by small molecules

[16]. In preclinical animal models, suppression of Bcl-2 by an antisense oligonucleotide has been shown to retard tumour growth and the approach is currently subject to clinical trial. Likewise, antisense oligonucleotides directed at survivin are being evaluated. The possibility of using recombinant TRAIL to induce apoptosis in malignant cells is under investigation. TRAIL is implicated as the basis of all-trans-retinoic treatment of promyelocytic leukaemia [17]. Also noteworthy is the development of caspase inhibitors for the treatment of certain degenerative (non-cancerous) diseases characterized by excess apoptosis.

Drugs shown to induce apoptosis specifically include chemopreventive agents, exemplified by 4-hydroxyphenylretinamide. Butyrate, a short-chain fatty acid produced by bacterial fermentation of dietary fibre, inhibits cell growth in vitro and promotes differentiation; it also induces apoptosis. Both roles may contribute to its prevention of colorectal cancer. Moreover, cyclo-oxygenase enzyme (COX-2) expression may modulate intestinal apoptosis via changes in Bcl-2 expression. Aspirin and similar drugs which inhibit COX-2 may promote apoptosis and prevent tumour formation.

Drugs targeting signal transduction pathways

In complex multicellular organisms, cell proliferation, differentiation and survival are regulated by a number of extracellular hormones, growth factors and cytokines. These molecules are ligands for cellular receptors and communicate with the nucleus of the cell through a network of intracellular signalling pathways. In cancer cells, key components of these signal transduction pathways may be subverted by proto-oncogenes through over-expression or mutation, leading to unregulated cell signalling and cellular proliferation. Because a number of these components may be preferentially over-expressed or mutated in human cancers, the cell signalling cascade provides a variety of targets for anticancer therapy.

Different approaches have been used to attack these targets and include classical cytotoxic agents as well as small molecule drug inhibitors. In addition, antisense oligonucleotides, vaccines, antibodies, ribozymes and gene therapy approaches have been utilized.

The diagram in Figure 3.4.7 illustrates cell signalling pathways that are targeted by anticancer agents currently undergoing clinical testing. The drug imatinib is already in clinical use. It is hoped that in future, a combination of agents targeting parallel pathways, as well as combinations with classical cytotoxic agents will improve the outcome of cancer patients.

Classes of agents and their potential targets include:

- Inhibitors of ligands, such as recombinant human antibody to VEGF (rHu mAbVEGF)
- Receptors, anti-receptor antibodies and tyrosine kinase receptor inhibitors
- RAS farnesyltransferase inhibitors
- RAF inhibitors
- MEK inhibitors
- Rapamycin analogues
- Protein kinase C (PKC) inhibitors
- Inhibitors of protein degradation
- Inhibitors of protein trafficking

REFERENCES

1. Strasser A, O'Connor L, Dixit VM (2000). Apoptosis signaling. *Annu Rev Biochem* 69: 217-245.

2. Letai AG (2008). Diagnosing and exploiting cancer's addiction to blocks in apoptosis. *Nat Rev Cancer* 8: 121-132.

3. Parton M, Dowsett M, Smith I (2001). Studies of apoptosis in breast cancer. *BMJ* 322: 1528-1532.

4. Rich T, Allen RL, Wyllie AH (2000). Defying death after DNA damage. *Nature* 407: 777-783.

5. Mocellin S and Nitti D (2008). TNF and cancer: the two sides of the coin. *Front Biosci* 13: 2774-2783.

6. Cory S and Adams JM (2002). The Bcl2 family: regulators of the cellular life-or-death switch. *Nat Rev Cancer* 2: 647-656.

7. Youle RJ and Strasser A (2008). The BCL-2 protein family: opposing activities that mediate cell death. *Nat Rev Mol Cell Biol* 9: 47-59.

8. Matsuyama S, Llopis J, Deveraux QL, et al. (2000). Changes in intramitochondrial and cytosolic pH: early events that modulate caspase activation during apoptosis. *Nat Cell Biol* 2: 318-325.

9. Brown GC (2007). Nitric oxide and mitochondria. *Front Biosci* 12: 1024-1033.

10. Krieser RJ and White K (2002). Engulfment mechanism of apoptotic cells. *Curr Opin Cell Biol* 14: 734-738.

11. Riedl SJ and Salvesen GS (2007). The apoptosome: signalling platform of cell death. *Nat Rev Mol Cell Biol* 8: 405-413.

12. Herceg Z and Wang ZQ (2001). Functions of poly(ADP-ribose) polymerase (PARP) in DNA repair, genomic integrity and cell death. *Mutat Res* 477: 97-110.

13. Logue SE and Martin SJ (2008). Caspase activation cascades in apoptosis. *Biochem Soc Trans* 36: 1-9.

14. Nicholson DW (2000). From bench to clinic with apoptosis-based therapeutic agents. *Nature* 407: 810-816.

15. Selivanova G and Wiman KG (2007). Reactivation of mutant p53: molecular mechanisms and therapeutic potential. *Oncogene* 26: 2243-2254.

16. Zheng TS (2001). Death by design: the big debut of small molecules. *Nat Cell Biol* 3: E43-E46.

17. Altucci L, Rossin A, Raffelsberger W, et al. (2001). Retinoic acid-induced apoptosis in leukemia cells is mediated by paracrine action of tumor-selective death ligand TRAIL. *Nat Med* 7: 680-686.

3.5 Invasion and Metastasis

Summary

> Metastasis results from the spread of tumour cells from their original location to other organ sites; metastatic disease is the main cause of death from cancer

> The organ distribution of metastases depends upon the type and location of the primary tumour and the route of dissemination of metastatic cells.

> The formation of metastasis involves a series of steps during which cancer cells leave the original tumour, enter lymph or blood circulation, survive and migrate, and colonize distant organs; this complex process is driven by genetic and epigenetic changes.

> Metastasis may develop from a small number of "cancer stem cells" which can change shape and properties to disseminate into the organism and adapt to the conditions of different organs

> Treatment of metastasis often combines local therapy aimed at removing or neutralising the metastases, and systemic therapy aimed at destroying micrometastases as well as preventing the formation of additional ones

The ability of tumour cells to spread from their original location to invade and colonise distant organ sites is the main feature that distinguishes benign from malignant cancers. Metastatic disease is also the major cause of death from cancer. As long as the tumour remains confined to one specific location, it remains curable provided it can be removed surgically and that the tumour does not irreversibly destroy the function of a vital organ. Once tumour cells start to spread into the organism, however, they become more difficult to control. First, they may reach distant organ sites and form secondary tumours, called metastases. Second, the tumour cells that are capable of spreading have acquired special properties that make them more resistant to treatments and to destruction by the immune system. Therefore, detection of distant spread and metastases is often an indicator of poor prognosis for the patient. This is reflected in the TNM classification system, which provides a universal, simple system to describe the anatomic extent of a cancer (Table 3.5.1).

The term "metastasis" comes from the combination of two Greek words, "meta", meaning "next" or "beyond", and "stasis", meaning "location" or "position". A metastasis is therefore a misplaced lesion, a lesion that has changed position. The term "invasion" refers to the process by which a tumour can form metastases: it consists of a series of steps by which growing tumours disturb the architecture of the tissue where they arise, take the space and place of normal cells, infiltrate into healthy areas and cross vessel barriers to enter the lymphatic or blood circulation. Loco-regional invasion is in itself a factor of poor prognosis, but not as poor as distant metastases. A locally invasive (N1) tumour would normally not be counted as metastatic disease. It may remain treatable using the same protocols as non-metastatic, localised lesions, without the need for extended whole-body treatments.

For a long time, metastatic disease has been considered as the ultimate step in cancer progression. It was thought that the most transformed cancer cells acquire the capacity to become independent from their organ of origin, to invade other organs, to travel in the body and to form colonies. This view is challenged by recent discoveries on cancer stem cells, which are capable of self-renewal and also of generating daughter cells that evolve into different cell shapes and phenotypes depending upon interactions with their environment [1]. Thus, in a given cancer, several lines of cancer cell development may exist. Most cells may develop in a certain direction, which preserve traits of the general architecture of the tissue where they arise. These cells contribute to the growing local tumour mass. Other cells may assume different shapes and roles and undergo morphological transitions that allow them to cross barriers and invade other organs. These are the metastatic cells [2]. It follows that metastatic cells can occur even in very small, apparently "early" cancer lesions. This may explain why over 10% of patients presenting to oncology clinics may have metastases without an identified primary tumour. These patients are said to have Cancer of Unknown Primary origin (CUP) or Unknown Primary Tumours (UPT) [3]. In these patients, the primary tumour may be so small that it is not detectable even using sophisticated methods. Yet these occult primary tumours can be the site of formation of cells with a high capacity to spread to other organs and form aggressive colonies.

Organ preference of metastases

The organ distribution of metastases depends upon the type and location of the primary tumour (Table 3.5.2) [4]. In many instances, it is determined by the route of dissemination of metastatic cells. For example, sarcomas tend to metastasise to lungs because of the venous drainage of muscles; colon carcinoma cells enter the portal circulation thereby gaining access to the liver. Invasion usually starts in the first array of capillaries encountered within the tumour or in the immediate tumour neighbourhood. Invasion also develops in the first lymph nodes encountered as cancer cells leave their tissue of origin. Metastatic cells can hop from node to node through lymphatic channels and accumulate into draining nodes, from where they can flow into efferent lymph nodes towards many organs. Lymphatic channels may present less of a challenge to tumour cell entry than capillaries because of their scanty basement membrane. The propensity of a tumour cell to invade lymphatic vessels or through capillaries depends upon its ability to adhere to specific structures, such as reticular fibres in the subcapsular sinus of draining nodes or endothelial cells that line blood vessels. Interactions with these structures are dependent upon the types of adhesion molecules expressed by tumour cells, in particular the integrins.

The distribution of metastases is not only a matter of route of dissemination. The most common places for the metastases to develop are the liver, the brain, the bones, the lung and the adrenal glands. There is a propensity for certain tumours to seed in particular organs. This was first recognised by Stephen Paget in 1889, based on his observation from autopsies of 700 women who died from metastatic breast cancer. He formulated the "seed and soil" hypothesis, proposing that specific cancer cells (the seed) have an affinity for certain organs (the soil) [5]. For example, breast cancer cells that have a physiological need for calcium selectively metastasise to bone because they can use it as an abundant source of calcium. In general, cancer cells tend to metastasise to organs where blood and energy supplies are abundant (such as liver or lung) or that are separated from the immune system by a physical barrier (such as the brain).

Detecting these metastases is a major challenge, as it is virtually impossible to explore all possible organ locations in sufficient detail. Advances in medical imaging techniques are making it possible to locate lesions of very small size, thus lowering the threshold for detection of metastases [6]. Current research is also focussing on detecting single, disseminating cancer cells in lymphatic or blood vessels, and on identifying patterns of gene expression in primary tumours that may predict their propensity to form metastases.

Molecular biology of metastasis

The metastasic process consists of a series of steps during which cancer cells leave the original tumour site, enter lymph or blood circulation (a process called intravasasion), survive and migrate, and extravasate to colonise distant organs. This complex process implies that candidate metastatic cells acquire many properties through genetic or epigenetic changes [7, 8]. On the basis of their level of participation in the metastatic process, Nguyen and Massagué have distinguished three general classes of metastasis genes: metastasis initiation, metastasis progression, and metastasis virulence (Figure 3.5.1) [9]. Metastasis initiation genes are those that provide an advantage in primary tumours, paving the way for tumour cells to enter the circulation. Metastasis progression genes are those that fulfil certain rate-limiting functions in primary tumour growth, and other specific functions in metastatic colonisation. Metastasis virulence genes are those that provide a selective advantage in secondary sites but not in the primary tumour, thus participating in meta-

T = primary tumour	
TX	Primary tumour cannot be assessed
T0	No evidence of primary tumour
Tis	Carcinoma in situ
T1	Tumour invades submucosa
T2	Tumour invades muscularis propria
T3	Tumour invades through muscularis propria into subserosa or into non-peritonealized pericolic or perirectal tissues
T4	Tumour directly invades other organs or structures and/or perforates visceral peritoneum
N = regional lymph nodes	
NX	Regional lymph nodes cannot be assessed
N0	No regional lymph node metastasis
N1	Metastasis in 1 to 3 regional lymph nodes
N2	Metastasis in 4 or more regional lymph nodes
M = distant metastasis	
MX	Distant metastasis cannot be assessed
M0	No distant metastasis
M1	Distant metastasis

Table 3.5.1 *TNM classification of cancer of the colon and rectum*

Primary tumour	Site of metastasis
Bronchial cancer	Adrenal (often bilateral)
Breast ductal carcinoma	Liver
Breast lobular carcinoma	Diffuse peritoneal seeding
Breast	Bone, ovary
Lung	Brain
Ocular melanoma	Liver
Prostate	Bone
Melanoma	Brain

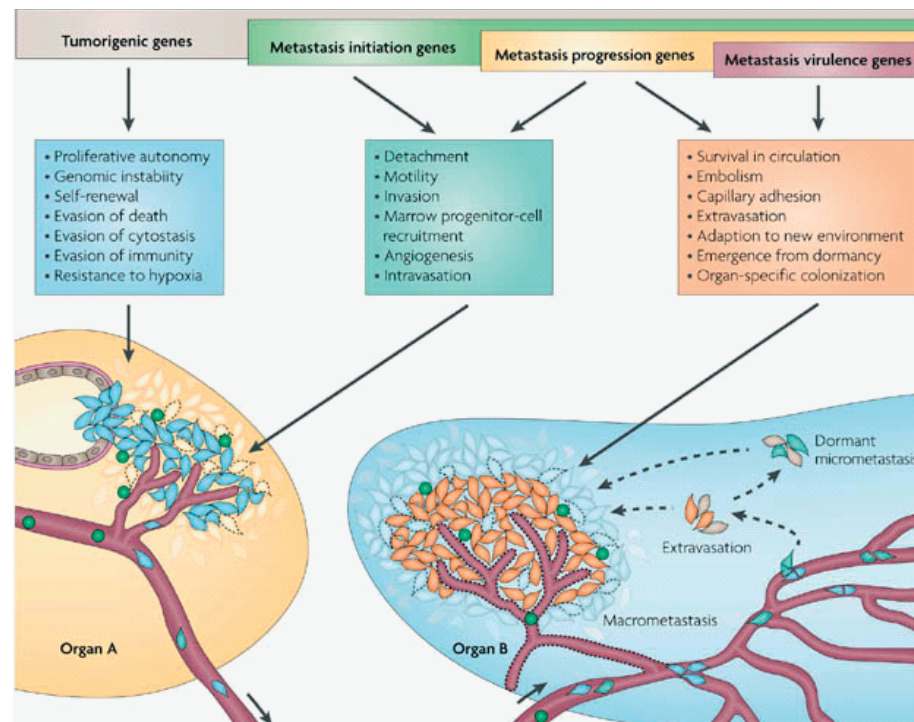**Table 3.5.2** *Site of metastasis of common cancers*

Fig. 3.5.1 *A general view of the sequence of biological events involved in the formation of metastases [9]*

static colonization but not in primary tumour development. However, before becoming candidates to metastasis, malignant cells must fulfil a number of tumorigenic conditions (see *Molecular hallmarks of cancer*, Chapter 3.1): they must be capable of unlimited proliferation, of evasion from the environmental constraints of their tissue of origin, and of attracting a blood supply through the formation of new capillaries and blood vessels, a process called angiogenesis [10]. As tumours grow, they must adapt and respond to environmental pressures such as those exerted by the immune response, the lowering oxygen tension and the increased acidic environment [11]. Such capacities are acquired during tumour initiation and local development, but must remain active throughout the development of metastatic disease, since they are critical for cancer cells to survive during their spread into the organism and during the development of distant colonies.

Metastasis initiation. Acquisition of metastastatic potential requires that candidate cells break away from the primary tumour and attach to and degrade the protein structures that make up the surrounding extracellular matrix (ECM). Most solid tumours arise from epithelial cells that are normally bounded by basement membranes which separate them from the underlying stroma and mesenchymal compartments. Breaching the basement membrane is the first step in the transition from in situ carcinoma to invasive, potentially metastatic cancer [12]. The basement membrane is composed of a complex of structural proteins including Collagen IV (the major component), laminin, entactin, and heparan sulfate proteoglycans. Interactions of tumour cells with basement membranes and ECM components comprise two critical phase phases: adhesion and matrix dissolution.

Adhesion. Epithelial cells are normally polarised and attached to each other via different types of cell-to-cell junctions, such as tight junctions, adherens junctions and desmosomes, as well as through intercellular adhesion molecules such as E-Cadherin. Initiation of metastasis requires releasing cells from cell-to-cell contacts that keep them into their proper place in the epithelium. Thus, cancer cells usually demonstrate multiple changes in the expression of cell adhesion components [13]. E-Cadherin, in particular, is a frequent target for genetic or epigenetic alterations that down-regulate its function, which may be considered as a tumour suppressor gene. First, its constitutive mutation predisposes to some forms of cancer (such as gastric cancer) and its re-introduction in metastatic cancer cells quench their invasive potential. Secondly, it interacts with beta-Catenin, an important oncogene, and provides a signalling connection between structural cell adhesion and cell proliferation. Loss of E-Cadherin frees beta-Catenin from its anchor at the cell membrane and makes it available for translocation into the nucleus, where it can activate transcription factors involved in stimulating cell proliferation.

Epithelial cells entertain contacts with the basal membranes and with the ECM through many other classes of molecules. Among them, integrins deserve special mention as changes in their expression patterns may have a profound influence on enabling cancer cells to adapt to changes in their micro-environment, a pre-requisite for successful migration. Integrins are cell surface receptors that mediate a dual, signalling and adhesion function [14]. Among the ligands of integrins are fibronectin, vitronectin, collagen, and laminin. Integrins are heterodimeric proteins containing two distinct chains, α (alpha) and β (beta). In mammals, 19 γ and 8 β subunits have been characterised. Through different combinations of alpha and beta subunits, about 24 unique integrins can be generated. The molecular mass of the integrin subunits can vary from 90 to 160 kDa, with the intracellular domain representing only a minor part (40 to 70 amino-acids). Integrins couple the ECM outside the cell to the intracellular cytoskeleton. This bond ensures that the cell can tightly adhere to ECM components without

being sheared and ripped away by movements of the ECM. Many differences in integrin expression between benign and malignant cells have been documented. They allow cells to develop different binding and adhesion properties, enabling them to detach from their original support and to seek novel adhesion points on ECM components as well as on cells lining blood or lymphatic vessels.

Other cell adhesion molecules implicated in cancer progression and metastasis include members of the immunoglobulin supergene family such as ICAM-1, ICAM-22, VECAM and PECAM. The latter are upregulated on activated endothelial cells, and can interact with integrins on leucocytes or circulating tumour cells to facilitate their extravasation. CD44 is an adhesion molecule normally present both at the surface of epithelial cells and of lymphocytes. On normal cells, however, CD44 is expressed as different splicing variants in both cell types. A change in splicing patterns from "epithelial-type" to "haematopoietic-type" may assist in carcinoma cell dissemination by providing recognition signals that lymphocytes normally use during their homing to specific tissues [15,16]. Thrombospondin mediates adhesion between circulating tumour cells, platelet and endothelial cells, promoting embolisation (vessel obstruction). This induces endothelial cells to retract, exposing the vessel's basement membrane and providing tumour cells an access for adhering to exposed proteins [17].

Matrix dissolution. Invasive cancer cells show increased expression of many enzymes, as well as decreased expression of their regulators, involved in the degradation of components of the ECM, thus physically opening up breaches that facilitate cancer cell dissemination [8]. One important group of such enzymes is the matrix metalloproteinases (MMP). The MMP family contains a diverse group of enzymes with different substrate preference (collagenase, gelatinase, stromelysin, proenzyme). All family members comprise a leader domain, a propeptide domain and a highly conserved catalytic domain containing a zinc atom involved in substrate binding. They play important roles during normal development

and morphogenesis, and their activities are tightly regulated. Activation of the leader domain is regulated by endogenous MMP inhibitors, which include a-2 macroglobulin and tissue inhibitors of metalloproteinases (TIMPs). An imbalance between MMPs and naturally occurring MMP inhibitors may cause an excess of extracellular matrix destruction, allowing cancer cells to invade surrounding tissues and metastasise. Two of the most studies MMPs are MMP-2 and -9 (gelatinase A and B, respectvely). There is clear evidence for increased levels of active forms of MMP-2 and/or 9 in bladder, breast, colon, protaste, lung, oesophageal and gastric cancer tissues. This increased expression can take place in cancer cells and/or in surrounding normal stromal cells, indicating that cancer cells can somehow induce stromal cells to secrete factors that facilitate migration, invasion and, ultimately, metastasis. Urokinase plasminogen activator (uPA) is also frequently upregulated in cancer. It controls the synthesis of plasmin, which degrades laminin and also activates gelatinases. Thus, upregulation of these enzymes in cancer can lead to proteolytic cascades that degrade the basement membranes and components of the stroma [18].

Besides their direct role in degrading ECM components, MMPs are also indirectly involved in promoting metastasis through their roles in angiogenesis. The formation of capillary sprouts is a physiological process that requires localised proteolysis of the stroma (mediated in part by MMP-2 and MMP-9 in addition to uPA). MMP-9 plays a role in the "angiogenic switch" that occurs during cancer progression by releasing VEGF by sequestration in the ECM. Furthermore, these proteases also contribute to sustained tumour growth by the ectodomain cleavage of membrane-bound pro-forms of growth factors, and the release of peptides which are mitogenic and chemotactic for cancer cells.

Metastatic dissemination. Dissemination starts when aggressive tumour cells enter the bloodstream through the newly formed vasculature that they have attracted. This process is facilitated by the particular, incomplete and leaky structure of

the blood vessels that is typical of many cancers. Intravasation is also enhanced by an epithelial-to-mesenchymal transition that confers to carcinoma cells plasticity and added motility similar to embryonic cells [19,20]. Although the rate of malignant cell shedding in the bloodstream generally increases with tumour size and grade, dissemination can occur from the early stages of the primary tumour. Only a minute proportion of cells that enter blood vessels or lymphatics will ultimately generate metastases. Indeed, disseminating cells are faced with multiple challenges. Among those, the most significant ones are the capacity to escape cell death due to the detachment of their support (a process called "anoikis", from the Greek word "oikos", meaning "home", preceded by the negative Greek prefix alpha), to escape recognition and destruction by the immune system and to recruit partners that facilitate their circulation and extravasation, including in particular the formation of partner aggregates [21]. Adhesion molecules that mediate attachment to capillary walls play a critical role in the dissemination process.

Numerous innate and adaptive immune effector cells and molecules participate in the recognition and destruction of cancer cells, a process that is known as cancer immunosurveillance. Disseminating cancer cells avoid immunosurveillance through many mechanims that can classified into two broad categories: the outgrowth of poorly immunogenic tumour-cell variants (immunoselection) and the subversion of the immune system (immunosubversion). The former category include a series of mechanisms by which disseminating cancer cells conceal or down-regulate antigens and recognition molecule complexes at their surface. The latter category include the production of sets of cytokines that down-regulate immune responses and the stimulation of regulatory T-cells that induce a form of immune tolerance towards cancer cells. Metastasis appears to correlate with changes in the immunogenic properties of tumour cells.

Metastatic colonisation. The process by which disseminating cancer cells leave the bloodstream to enter the parenchyma of another organ is

termed extravasation. Metastatic cells extravasate by breaching the capillaries in which they are embedded, either by vascular-remodelling events that allow migration across the capillary wall or as a result of mechanical disruption of capillaries by expanding tumour emboli. On entry into another organ, tumour cells are confronted with a different microenvironment in which they must survive, develop, and eventually expand in the same way as they did in their organ of origin. To help them in the process of establishing a new "home" in their adoptive tissue, cancer cells recruit bone-marrow-derived progenitor cells and other local cells that provide a permissive "niche" for metastasis [22]. Once metastatic cells are established, active colonisation proceeds through the recruitment of organ-specific components of the tumour microenvironment, such as the activation of bone-resorbing osteoclasts by breast cancer cells during osteolytic metastasis [23]. Full metastatic colonisation can occur by immediate growth of cancer cells upon their extravasation, or after a prolonged period of micrometastatic dormancy.

Epidermal-mesenchyme transition and the concept of metastatic cancer stem cells

Most solid tumours start with an epithelial phenotype. However, during tumour progression, this phenotype becomes altered and some cells undergo a transition to assume a more mesenchymal phenotype. These mesenchymal-like cancer cells acquire a high migratory capacity and may represent one of the main forms into which cancer cells can disseminate into the organism. Conversely, at the time of extravasation, these cells undergo a reverse mesenchymal-epithelial transition which regenerates high proliferative status and allows formation of a metastasis with a morphology that resembles the primary tumour. It has emerged that this process closely resembles Epithelial-Mesenchyme Transition (EMT), a mechanism that is vital for morphogenesis during embryonic development [20,24]. During gastrulation in mammals, cells migrate from primitive epithelial-like structures to spatially reorganise and form one of the three main embryonic layers, the

mesoderm. In this process, epithelial cells acquire fibroblast-like properties, show reduced adhesion to ECM and increased mobility, exactly like metastatic cancer cells.

EMT is essential for many morphogenetic events such as organogenesis, wound healing, tissue remodelling and heart development. A landmark of EMT is the loss of E-Cadherin expression, a phenomenon that is common in many epithelial tumours. However, E-Cadherin expression remains detectable in many invasive tumours, raising questions about the whether EMT is a general phenomenon in advanced cancer, or a property assumed only by a limited number of cells. This paradox has been largely resolved by the observation that, in cancer, EMT could generate cells with properties of stem cells, including in particular self-renewal through asymetric division. In normal tissues, stem cells are present only in proliferative areas such as the basal layer of squamous mucosa or crypts of glandular mucosa. Such stem cells become embedded within small, early cancer masses as static cancer stem cells (SCSC) [2]. These SCS are, to a large extent, responsible for sustained production of daughter cancer cells which assume an epithelial phenotype and constitute the bulk of the tumour mass. In certain conditions, SCSC can undergo EMT and become mobile, migrating cells while retaining their capacity for self renewal. The signals that trigger this EMT may correspond to a form of disturbed wound healing response generated by the breakdown of basal membrane and the increased severity of the tissue lesion caused by the tumour.

Mobile, migrating cancer stem cells (MCSC) may actually represent only a small fraction of the cells that are shed in the bloostream. However, their stem status endows them with the capacity to survive during migration as well as to re-differentiate into epithelial-like cells upon extravasation and colony formation into distant organs. Upon entry into the stroma of a target organ, MCSC may locally recruit normal fibroblasts and other cell types to constitute an appropriate niche for undergoing mesenchyme-epithelial transition and giving rise to rapidly growing metastases. It

follows from this model that two critical characteristics of mobile cancer stem cells are the cyclical activation and inhibition of expression of genes involved in EMT, as well as the capacity to recruit normal, non-cancer cells to become essential partners in the metastatic process.

Treatment of metastatic cancer

The presence, number and organ location of metastases are critical parameters in selecting appropriate therapeutic methods. In most instances treatment will consist of a combination of local therapy aimed at removing or neutralising the metastases, and systemic therapy aimed at destroying micrometastases as well as preventing the formation of additional ones [25] [26]. The main local treatments are surgery and radiosurgery (that is, the use of 3-dimensional radiation treatment to deliver high radiation doses in a very delimited area of the body). Systemic treatments include chemotherapy and radiation therapy (which are active against both primary and metastatic cancers) as well as biological and, when appropriate, hormone therapy. Biological therapies may use monoclonal antibodies that target cancer cells, or factors that block processes involved in metastasis such as angiogenesis. Current approaches for drug development are focusing on the neutralisation of specific factors involved in invasion and metastasis, such as metalloproteinases or integrins.

The choice of treatment depends upon many factors, principally the type of cancer, the size, number and localisation of metastases, the general condition and age of the patients, and the treatments the patient has already received in the case of secondary metastatic cancer. In many instances, available treatments are not capable to provide a complete cure for metastatic cancer, although they can induce remission, improve quality of life, and significantly increase survival after diagnosis. Finding new, efficient and better tolerated treatments for metastatic cancer is a major challenge in current cancer research and clinical trials.

WEBSITES

How cancer grows and spreads: An interactive, animated presentation that shows how cancer progresses through the 14 stages of a typical cancer. http://www.childrenshospital.org/research/_cancer/index.html

Metacancer: resources and support for metastatic cancer survivors and their caregivers. http://www.metacancer.org/index.php

REFERENCES

1. Li F, Tiede B, Massague J, et al. (2007). Beyond tumorigenesis: cancer stem cells in metastasis. *Cell Res* 17: 3-14.

2. Brabletz T, Jung A, Spaderna S, et al. (2005). Opinion: migrating cancer stem cells - an integrated concept of malignant tumour progression. *Nat Rev Cancer* 5: 744-749.

3. Freudenberg LS, Rosenbaum-Krumme SJ, Bockisch A, et al. (2008). Cancer of unknown primary. *Recent Results Cancer Res* 170: 193-202.

4. Chambers AF, Groom AC, MacDonald IC (2002). Dissemination and growth of cancer cells in metastatic sites. *Nat Rev Cancer* 2: 563-572.

5. Jodele S, Blavier L, Yoon JM, et al. (2006). Modifying the soil to affect the seed: role of stromal-derived matrix metalloproteinases in cancer progression. *Cancer Metastasis Rev* 25: 35-43.

6. Riethdorf S, Wikman H, Pantel K (2008). Review: Biological relevance of disseminated tumor cells in cancer patients. *Int J Cancer* 123: 1991-2006.

7. Yokota J (2000). Tumor progression and metastasis. *Carcinogenesis* 21: 497-503.

8. Bogenrieder T and Herlyn M (2003). Axis of evil: molecular mechanisms of cancer metastasis. *Oncogene* 22: 6524-6536.

9. Nguyen DX and Massague J (2007). Genetic determinants of cancer metastasis. *Nat Rev Genet* 8: 341-352.

10. Hanahan D and Weinberg RA (2000). The hallmarks of cancer. *Cell* 100: 57-70.

11. Fidler IJ, Singh RK, Yoneda J, et al. (2000). Critical determinants of neoplastic angiogenesis. *Cancer J* 6 Suppl 3: S225-S236.

12. Engbring JA and Kleinman HK (2003). The basement membrane matrix in malignancy. *J Pathol* 200: 465-470.

13. Glukhova M, Deugnier MA, Thiery JP (1995). Tumor progression: the role of cadherins and integrins. *Mol Med Today* 1: 84-89.

14. Moschos SJ, Drogowski LM, Reppert SL, et al. (2007). Integrins and cancer. *Oncology (Williston Park)* 21: 13-20.

15. Jothy S (2003). CD44 and its partners in metastasis. *Clin Exp Metastasis* 20: 195-201.

16. Wai PY and Kuo PC (2008). Osteopontin: regulation in tumor metastasis. *Cancer Metastasis Rev* 27: 103-118.

17. Kazerounian S, Yee KO, Lawler J (2008). Thrombospondins in cancer. *Cell Mol Life Sci* 65: 700-712.

18. Duffy MJ, McGowan PM, Gallagher WM (2008). Cancer invasion and metastasis: changing views. *J Pathol* 214: 283-293.

19. Christiansen JJ and Rajasekaran AK (2006). Reassessing epithelial to mesenchymal transition as a prerequisite for carcinoma invasion and metastasis. *Cancer Res* 66: 8319-8326.

20. Guarino M, Rubino B, Ballabio G (2007). The role of epithelial-mesenchymal transition in cancer pathology. *Pathology* 39: 305-318.

21. Simpson CD, Anyiwe K, Schimmer AD (2008). Anoikis resistance and tumor metastasis. *Cancer Lett* 272: 177-185.

22. Psaila B, Kaplan RN, Port ER, et al. (2006). Priming the 'soil' for breast cancer metastasis: the pre-metastatic niche. *Breast Dis* 26: 65-74.

23. Kaplan RN, Psaila B, Lyden D (2006). Bone marrow cells in the 'pre-metastatic niche': within bone and beyond. *Cancer Metastasis Rev* 25: 521-529.

24. Moustakas A and Heldin CH (2007). Signaling networks guiding epithelial-mesenchymal transitions during embryogenesis and cancer progression. *Cancer Sci* 98: 1512-1520.

25. Fidler IJ (2002). Critical determinants of metastasis. *Semin Cancer Biol* 12: 89-96.

26. Fidler IJ, Kim SJ, Langley RR (2007). The role of the organ microenvironment in the biology and therapy of cancer metastasis. *J Cell Biochem* 101: 927-936.

Emerging Technologies

Summary

> Proteomics is a general term that covers a variety of conceptually related technologies targeting protein expression and function, often in high-throughput formats. Proteomics provides a window to biological function and regulation that is strongly complementary to those provided by genomics and expression profiling

> Micro-RNAs (miRNAs) and small interfering RNAs (siRNAs) are small, double-stranded RNAs that bind to messenger RNAs (mRNAs) and regulate their expression and/or their degradation. si-RNAs have become an extremely powerful research tool for understanding gene function and are thought to have important therapeutic potential

> High density array technologies have been steadily increasing in sophistication and assay density since the mid-1990s. Currently, the three main applications of these technologies are expression profiling (used to measure the expression of many genes simultaneously), array CGH (used to search for DNA duplications or deletions in tumor samples at high resolution), and high-density SNP genotyping (used for genome-wide SNP association studies)

Proteomics

Proteomics is the study of the proteome, the protein complement of the genome in a biological system at a given point in time. The terms proteomics and proteome were first coined in the early 1990s [1], and today this rapidly developing field is employed to study protein localisation, differential expression,

post-translational modifications, protein interactions, protein structure and splice variants. Proteins are the main effectors of biological functions, and the knowledge gained from proteomics studies is invaluable for biomarker and drug discovery.

The proteome is dynamic; protein levels depend not only on the corresponding transcript levels but also on a multitude of translational controls and on regulation of protein degradation [2]. Proteomics, in contrast to genomics, also has the potential to explore large-scale measurements of protein modifications and their quantitative changes following cell perturbations, which are often just as important for protein activity as protein expression levels [3].

Completion of the Human Genome Project was critical for the large-scale development of proteomics not only because the genome sequence provides a full list of possible protein coding sequences, but also because the Genome Project changed the paradigm for large-scale biology projects. Just as the human genome project spawned many daughter projects, the Human Proteome Organization (HUPO) has launched initiatives on human organs and cell systems, established standards, and created antibody (Human Protein Atlas) and mouse models to overcome current limitations of proteomics.

The rapid development of proteomics as a field has depended upon substantial technological advances in many specific areas including gel-based or gel-free protein separation and sequencing techniques (shotgun sequencing), protein chips (SELDI-MS, protein-, tissue-, and antibody arrays), mass spectrometry (MS) (sensitivity, resolution, speed and throughput), and bioinformatics [4,5]. Currently there is no technique that can cover the large number of single proteins contained in a complex sample and the wide dynamic range in the abundance of the individual protein species [6], but large-scale studies of protein complexes are emerging that show how the cell organises to deliver function at the molecular level.

Proteomics projects generate enormous quantities of data, and public domain databases have been developed to manage and assemble this information. Examples of existing databases include PRIDE (http://www.ebi.ac.uk/pride/), PeptideAtlas (http://www.peptideatlas.org/), and IMEx (http://imex.sourceforge.net/index.html). A critical consequence of the use of these public databases is that raw data are released after publication of manuscripts so that other investigators can re-analyse original data or incorporate the original data into new studies.

Proteomics provides an attractive approach to study complex diseases including cancer. Clinical proteomics has focused on the discovery of diagnostic, prognostic and predictive disease biomarkers with a particular focus on biomarkers that can be assayed from easily available samples such as blood or urine. However, disease-driven marker discovery (or marker validation) studies are in many respects more difficult than basic science studies. The main difference is that basic science studies can proceed from relatively small numbers of samples gathered under carefully controlled laboratory conditions, while marker validation studies require large numbers of samples gathered from human subjects under clinical conditions. For example, many studies have reported serum peptide signatures revealed by mass spectrometry that appear to distinguish between cancer free controls and individuals with prostate cancer, breast cancer, lung cancer, colon cancer, etc. (for examples, see [7-11]). However, none of these signatures yet enjoy the level of replication that has been achieved by, for example, mRNA expression profiling of breast tumours. The difficulty in extracting replicable disease-associated serum proteomics profiles can in part be attributed to the intrinsic difficulty of the research goal. However, that difficulty also raises challenges, including creation of biological resource centres that contain large numbers of well-documented biosamples of the correct type and in the correct state of physical preservation to support specific proteomics studies, validation of candidate biomarkers in large well-characterised cohorts (dependent on biological resource centres),

and reduction of validated proteomics-based bioassays to robust and efficient procedures that will work in the clinic.

micro-RNAs

In 1993, Lee and Ambros discovered that lin-4, a gene known to control the timing of C. elegans' larval development does not code for a protein but instead produces a pair of small RNAs [12]. They demonstrated that these small lin-4 RNAs base pair with the 3' untranslated region of the lin-14 mRNA and result in translational repression of this mRNA. The importance of this discovery did not become evident until several years later, when other small RNA molecules with regulatory functions were found [13]. Since then, about 4000 small regulatory RNAs, termed microRNAs (miRNAs), have been identified in a variety of animals, plants,

and viruses and have been deposited in publicly available databases, such as miRBase (http://microrna.sanger.ac.uk/).

It is clear now that miRNAs together with small interfering RNAs (siRNAs) are members of a widespread class of small, evolutionarily conserved, non-coding, double-stranded RNAs (dsRNAs) with regulatory functions. miRNA and siRNA differ in terms of their origin and processing. Once in their single-strand form, either can regulate the expression of downstream genes by binding to a target messenger RNA (mRNA) at a specific complementary target sequence and guide the targeted mRNA to the double stranded RNA-induced silencing complex (RISC), responsible for its cleavage or its translational inhibition (Figure 1). Up to 30% of protein-coding genes may be regulated by miRNA [14], including tran-

scription factors, oncogenes and tumour suppressor genes. Therefore miRNAs play an essential role in multiple biological processes. Moreover, miRNA expression has been shown to be deregulated in a number of cancers [15-17].

siRNAs are easily designed, synthesised in vitro, and built into expression cassettes that can be used to regulate the expression of experimenter-selected target genes in a laboratory setting. Consequently, siRNA have been used as powerful experimental tool to explore gene function. However, it is becoming more and more evident that the potential applications of RNA interference go much further. Transcriptional profiling using genomic microarrays and beads have enabled the discovery of numerous miRNAs that are differentially expressed in normal tissues compared with tumours and are associated with cancer development, diagnosis, and prognosis [17]. miRNAs have also become targets for development of anticancer gene therapy; antisense molecules that can inhibit miRNA activity are currently being tested for their anti-tumour activity [18].

The discovery of viral encoded microRNAs indicates that viruses also use this mode of gene regulation. Viral miRNAs seem to play an important role in regulating both the viral life cycle and the interaction between viruses and their hosts [19,20]; therefore microRNAs may act as critical modulators of viral mediated oncogenesis. In addition, viral gene-specific siRNAs are theoretically very promising antiviral inhibitors and have been examined in a broad range of medically important viruses.

Without a doubt, the phenomenon of RNA interference has been harnessed to create enormously powerful research tools. RNA interference has clear potential to become an important partner in fighting cancer. Future applications of miRNA-related technologies will become even more powerful as new miRNA targets are identified and miRNA-related regulatory mechanisms better understood.
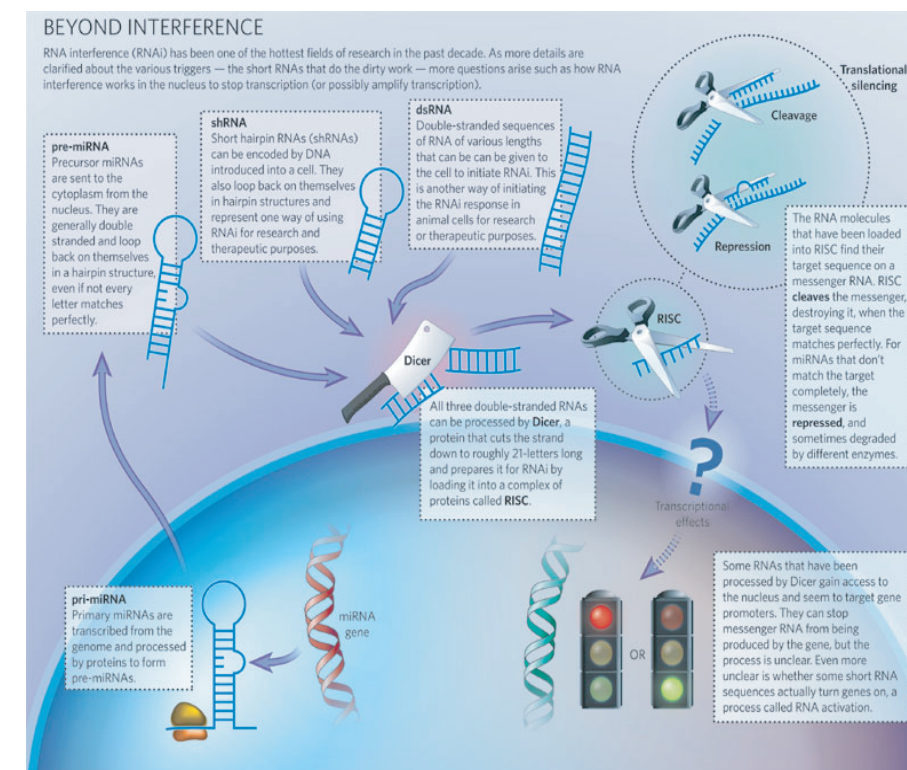

Fig. 3.6.1 Erika Check (2007) Nature 448, 855-858

High density array technologies

Natural deoxyribonucleic acid (DNA) molecules are mixed polymers of the 4 deoxyribonucleotides deoxyadenosine (dA), deoxycytidine (dC), deoxyguanosine (dG) and thymidine (dT). Single-stranded DNA molecules have the very useful characteristic that they hybridise with DNA molecules of complementary sequence to form stable double stranded DNA duplexes. In these duplexes, dA pairs with dT and dC pairs with dG; the hybridisation, or base pairing, is sequence-specific. Therefore, if you know the sequence of one DNA strand, you know the sequence of its complement.

Complementary base pairing is fundamental to the biological processes of DNA replication, transcription of DNA into RNA, and translation of RNA to make protein. In addition, complementary base pairing has been harnessed to create an enormous number of molecular biology protocols. One family of these protocols is high-density array hybridisation, which arose from the earlier protocol of filter hybridisation. The key concept behind filter hybridisation is that a mixed population of DNA molecules can be fractionated, denatured to make single-stranded DNA, and then affixed to a solid support such as a nitrocellulose filter in a way that preserves positional information from the fractionation. If a substantially pure preparation of labelled probe (radioactive or fluorescent or biotinylated, etc.) is then hybridised in solution to the filter, the probe will hybridise to and be positionally concentrated by its complement on the filter. The signal from the label will then reveal the position of its complement on the filter; under suitable conditions, the signal may also reveal the quantity of its complement on the filter. This is the idea behind such venerable procedures as the Southern blot, northern blot, and dot blot.

More recently, the logic behind these protocols has been reversed. That is, a substantially pure population of single-stranded DNA molecules of a specific sequence can be affixed to a specific position or address on a solid support,

constituting "a feature", and then allowed to hybridise in solution to a mixed population of many labelled single-stranded DNA molecules. During the hybridisation, the DNA sequence at the feature will find its complement (its target) in the mixed population of labelled DNAs and positionally concentrate the labelled target at its address on the solid support. The signal from the label will then reveal the presence or absence of the target in the mixed population and, under suitable conditions, will also reveal the relative quantity of the target present in the population. Because the sequence of each feature on the solid support is known, the address of the feature encodes the identity (i.e. gene, allele, etc.) of the target to which it hybridises. Packing large numbers of distinct features onto a small solid support is the idea behind high-density array hybridisation.

Over the last 15 years, a number of technical development-oriented labs and biotechnology companies have competed to develop useful high-density array hybridisation applications. Most of the popular applications have fallen into one of four categories: expression profiling, array comparative genome hybridisation (array CGH), genome-wide SNP genotyping and viral serotyping.

For expression profiling, the DNAs affixed at the features on the array are either parts of specific cDNA clones or DNA oligonucleotides that correspond to the sequence of specific RNA transcripts. RNA or cDNA prepared from a particular sample (cell line, tumour, tissue type, etc.) is labelled and then hybridised to the array. After hybridisation, the signal present at each feature of the array provides information about the quantity of the corresponding transcript in the original sample. In principle, multi-gene patterns of gene expression can then be extracted from the expression data and used to group tumours into classes or predict differential treatment responses. One of the early successes of the expression profiling approach was the demonstration that acute myeloid leukemia can be distinguished from acute lymphoblastic leukemia by expres-

sion profile alone [21]. A more subtle success has been the stratification of breast cancer into at least four distinct sub-types based on gene expression profiles [22,23].

Comparative genome hybridisation is a procedure for finding DNA copy number changes—amplifications or deletions—in tumour DNA samples. The original CGH strategy was to fluorescently label tumour DNA one colour, label reference genomic DNA a second colour, and then hybridise equal quantities of the two labelled DNAs to a metaphase chromosomal spread. Chromosomal regions that are amplified in the tumour will disproportionately fluoresce in the colour that the tumour DNA was labelled, whereas chromosomal regions that are heterozygously or homozygously deleted in the tumour will disproportionately fluoresce in the colour that the reference DNA was labelled. The drawback of this approach is that the positional resolution provided by hybridisation to metaphase chromosomes is limited to about 20 Mb. To overcome this limitation, various investigators have produced hybridisation arrays where the individual features are fragmented BAC clones, cDNA clones or even long synthetic oligonucleotides. These arrays improve the resolution of array CGH to the level of individual genes or, if oligos are used as probes, even individual exons [24]. Many of the experimental results achieved by expression profiling have been recapitulated by array CGH. For example, Bergamaschi et al recently demonstrated that the breast cancer subtypes identified by expression profiling can also be identified by array CGH [25]. Moreover, array CGH has the advantage over expression profiling that its substrate DNA is much more stable than the RNA required for expression profiling—potentially an important advantage for clinical applications.

Genome-wide SNP association studies (GWA studies) seek associations between common SNPs and risk of one or another disease without having to rely on prior hypotheses of which genes or genetic pathways are involved in the disease. Three trends have merged to

make such studies possible: (1) the human SNP haplotype mapping project has mapped and measured disequilibrium between more than 3 000 000 common human SNPs and in so doing showed that genotyping about 500 000 well-selected SNPs captures much of the information present in the full set of 3 000 000 [26]; (2) the total number of features that could be packed on to a high-density DNA hybridisation array increased from the low thousands to several hundred thousand (as of early 2008, more than 1 million), making it possible to genotype a genome-wide representative set of SNPs in a single experiment; and (3) epidemiology research groups have joined into consortia that can assemble series of 5000 or more cases and controls, sufficient to overcome the statistical multiple testing problems inherent in association studies that seek to test hundreds of thousands of independent hypotheses.

The combined result has been that in 2007 and 2008, international research consortia announced results from GWA studies in breast cancer, colon cancer, lung cancer and prostate cancer [27-36]. It is also very likely that GWA studies of some of the less common cancers will be completed over the next few years.

Beyond providing irrefutable proof that inheritance of common SNPs does influence the risk of common cancers, what are some of the most interesting results that have emerged from these studies? Perhaps the most intriguing new information has been that SNPs arrayed across a small segment of chromosome 8q, not far from the oncogene MYCC (c-myc), influence the risk of breast cancer, colon cancer, and prostate cancer. A second result has been the finding that that very few SNPs with frequencies of >10% confer odds ratios >1.5. Most of

the common risk-SNPs detectable at the scale of GWA study conducted so far confer odds ratios in the range of 1.1–1.2. On one hand, this means that, individually, the risk that they confer is not substantial enough to be medically useful. On the other hand, it may eventually be possible to create polygenic SNP profiles that provide sufficiently informative risk prediction to be clinically useful.

Around the world, scientists and engineers continuously invent new technologies or improve on old technologies. A small fraction of these inventions open whole new avenues of research, leading to important advances in medical science. Another overlapping fraction lead to important improvements in clinical medical practice. Thus today's relatively new technologies become commonplace as we look forward to tomorrow's new technologies.
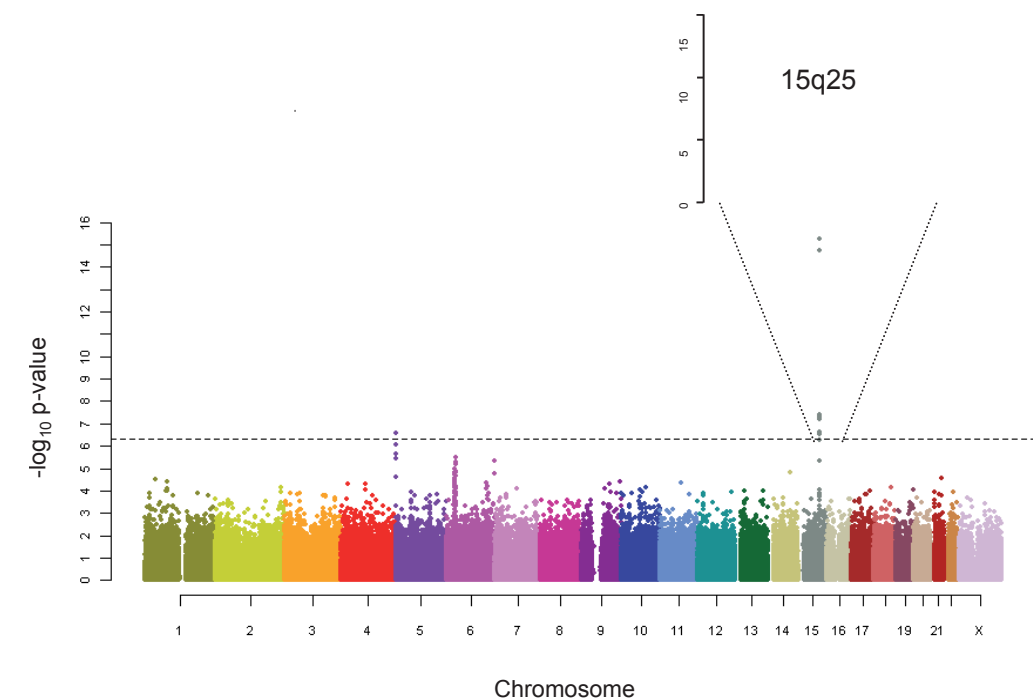

Fig. 3.6.2 *Genomewide SNP scan for lung cancer susceptibility loci. This scatter plot of p-values in −log scale from the trend test for 315 956 variants comparing 2971 lung cancer cases and 3745 controls provides strong evidence for a locus at chromosome 15q25*

REFERENCES

1. Wilkins MR, Sanchez JC, Gooley AA, et al. (1996). Progress with proteome projects: why all proteins expressed by a genome should be identified and how to do it. *Biotechnol Genet Eng Rev* 13: 19-50.

2. Gygi SP, Rochon Y, Franza BR, et al. (1999). Correlation between protein and mRNA abundance in yeast. *Mol Cell Biol* 19: 1720-1730.

3. Cox J and Mann M (2007). Is proteomics the new genomics? *Cell* 130: 395-398.

4. Aebersold R and Mann M (2003). Mass spectrometry-based proteomics. *Nature* 422: 198-207.

5. Washburn MP, Wolters D, Yates JR. (2001). Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nat Biotechnol* 19: 242-247.

6. Anderson NL and Anderson NG (2002). The human plasma proteome: history, character, and diagnostic prospects. *Mol Cell Proteomics* 1: 845-867.

7. Ornstein DK, Rayford W, Fusaro VA, et al. (2004). Serum proteomic profiling can discriminate prostate cancer from benign prostates in men with total prostate specific antigen levels between 2.5 and 15.0 ng/ml. *J Urol* 172: 1302-1305.

8. Villanueva J, Shaffer DR, Philip J, et al. (2006). Differential exoprotease activities confer tumor-specific serum peptidome patterns. *J Clin Invest* 116: 271-284.

9. Liu XP, Shen J, Li ZF, et al. (2006). A serum proteomic pattern for the detection of colorectal adenocarcinoma using surface enhanced laser desorption and ionization mass spectrometry. *Cancer Invest* 24: 747-753.

10. Belluco C, Petricoin EF, Mammano E, et al. (2007). Serum proteomic analysis identifies a highly sensitive and specific discriminatory pattern in stage 1 breast cancer. *Ann Surg Oncol* 14: 2470-2476.

11. Patz EF, Jr., Campa MJ, Gottlin EB, et al. (2007). Panel of serum biomarkers for the diagnosis of lung cancer. *J Clin Oncol* 25: 5578-5583.

12. Lee RC, Feinbaum RL, Ambros V (1993). The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. *Cell* 75: 843-854.

13. Lee RC and Ambros V (2001). An extensive class of small RNAs in Caenorhabditis elegans. *Science* 294: 862-864.

14. Lewis BP, Burge CB, Bartel DP (2005). Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell* 120: 15-20.

15. Schetter AJ, Leung SY, Sohn JJ, et al. (2008). MicroRNA expression profiles associated with prognosis and therapeutic outcome in colon adenocarcinoma. *JAMA* 299: 425-436.

16. Blenkiron C, Goldstein LD, Thorne NP, et al. (2007). MicroRNA expression profiling of human breast cancer identifies new markers of tumor subtype. *Genome Biol* 8: R214.

17. Lu J, Getz G, Miska EA, et al. (2005). MicroRNA expression profiles classify human cancers. *Nature* 435: 834-838.

18. Melnikova I (2007). Therapies for Alzheimer's disease. *Nat Rev Drug Discov* 6: 341-342.

19. Stern-Ginossar N, Elefant N, Zimmermann A, et al. (2007). Host immune system gene targeting by a viral miRNA. *Science* 317: 376-381.

20. Murphy E, Vanicek J, Robins H, et al. (2008). Suppression of immediate-early viral gene expression by herpesvirus-coded microRNAs: implications for latency. *Proc Natl Acad Sci U S A* 105: 5453-5458.

21. Golub TR, Slonim DK, Tamayo P, et al. (1999). Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science* 286: 531-537.

22. Perou CM, Sorlie T, Eisen MB, et al. (2000). Molecular portraits of human breast tumours. *Nature* 406: 747-752.

23. Sorlie T, Perou CM, Tibshirani R, et al. (2001). Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc Natl Acad Sci U S A* 98: 10869-10874.

24. Dhami P, Coffey AJ, Abbs S, et al. (2005). Exon array CGH: detection of copy-number changes at the resolution of individual exons in the human genome. *Am J Hum Genet* 76: 750-762.

25. Bergamaschi A, Kim YH, Wang P, et al. (2006). Distinct patterns of DNA copy number alteration are associated with different clinicopathological features and gene-expression subtypes of breast cancer. *Genes Chromosomes Cancer* 45: 1033-1040.

26. Frazer KA, Ballinger DG, Cox DR, et al. (2007). A second generation human haplotype map of over 3.1 million SNPs. *Nature* 449: 851-861.

27. Gudmundsson J, Sulem P, Rafnar T, et al. (2008). Common sequence variants on 2p15 and Xp11.22 confer susceptibility to prostate cancer. *Nat Genet* 40: 281-283.

28. Eeles RA, Kote-Jarai Z, Giles GG, et al. (2008). Multiple newly identified loci associated with prostate cancer susceptibility. *Nat Genet* 40: 316-321.

29. Thomas G, Jacobs KB, Yeager M, et al. (2008). Multiple loci identified in a genome-wide association study of prostate cancer. *Nat Genet* 40: 310-315.

30. Duggan D, Zheng SL, Knowlton M, et al. (2007). Two genome-wide association studies of aggressive prostate cancer implicate putative prostate tumor suppressor gene DAB2IP. *J Natl Cancer Inst* 99: 1836-1844.

31. Hunter DJ, Kraft P, Jacobs KB, et al. (2007). A genome-wide association study identifies alleles in FGFR2 associated with risk of sporadic postmenopausal breast cancer. *Nat Genet* 39: 870-874.

32. Easton DF, Pooley KA, Dunning AM, et al. (2007). Genome-wide association study identifies novel breast cancer susceptibility loci. *Nature* 447: 1087-1093.

33. Broderick P, Carvajal-Carmona L, Pittman AM, et al. (2007). A genome-wide association study shows that common alleles of SMAD7 influence colorectal cancer risk. *Nat Genet* 39: 1315-1317.

34. Tomlinson I, Webb E, Carvajal-Carmona L, et al. (2007). A genome-wide association scan of tag SNPs identifies a susceptibility variant for colorectal cancer at 8q24.21. *Nat Genet* 39: 984-988.

35. Zanke BW, Greenwood CM, Rangrej J, et al. (2007). Genome-wide association scan identifies a colorectal cancer susceptibility locus on chromosome 8q24. *Nat Genet* 39: 989-994.

36. Hung RJ, McKay JD, Gaborieau V, et al. (2008). A susceptibility locus for lung cancer maps to nicotinic acetylcholine receptor subunit genes on 15q25. *Nature* 452: 633-637.

Biomarkers at a Crossroads:
Implications for Early Cancer Detection and Diagnosis

Summary

> Biomarkers are the signposts during the process of carcinogenesis

> If identified carefully and selected rationally, biomarkers could provide a window of opportunity for risk assessment and cancer diagnosis and hasten the move towards personalised prevention and treatment

> The key to discovery of the right biomarkers lies in the selection of clinically annotated, well characterized specimens, appropriate study designs to address the clinical questions, and associated assays and data analysis

> Validation of biomarkers remains a key bottleneck in bringing biomarkers to clinical fruition. However, the National Cancer Institute's Early Detection Research Network is addressing this critical step by launching several validation trials, an example of which is presented in this chapter

> The future of biomarker research depends on the successful demonstration of biomarker-based diagnosis, prevention and treatment to earn it acceptance in clinical practice

Carcinogenesis is a complex process requiring the coordinated interactions of numerous genes, proteins, signalling pathways and cell types. As a result of extensive studies on the molecular carcinogenesis of cancer, a number of regulatory pathways and networks have been identified. These pathways have revealed several unique events, marked by structural modifications to cells and the expression of genes and proteins that accompany oncogenic transformation. Thus, both cellular morphology and molecular signatures change during cancer development. By discerning these changes accurately with the help of biomarkers, we can improve the early detection and diagnosis of individual cancers. Biomarkers are the molecular signposts that indicate how far the process of carcinogenesis has travelled across the network or pathway leading to the development of a tumour. Biomarkers are the major measures by which future medicine will be personalised for individuals, and prevention or treatment will be based on unique target-specific molecules, as opposed to standard systemic infusions of toxic chemotherapy agents [1].

Definition of biomarkers

There is no standard definition for "biomarker" that is universally used. In 1999, the US National Institutes of Health/Food and Drug Administration Working Group drafted a definition of a biomarker as a characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes or pharmacological response to a therapeutic intervention [2]. However, this definition is broad and may not be easily understood by the general public. Biomarkers may be defined as quantifiable molecules, including DNA, RNA, proteins and metabolites, that are found in body fluids or tissues at an abnormal level that signal a pathologic condition, such as cancer. A biomarker might be a molecule secreted by a malignancy, or it can be a specific response of the body to the presence of cancer. Alterations in gene sequence or expression and in protein structure and function have been associated with every type of cancer and with their progression through the various stages of development. Changes in gene expression and in protein expression or modification can be used to detect cancer, determine prognosis and monitor disease progression and therapeutic response [3].

Search for biomarkers

Tremendous progress has been made in developing high-throughput technologies that accelerate the discovery of genes and pro-teins. But the challenge now is to identify the related biomarkers that provide an earlier indication of disease and are more reliable and precise in predictive ability than current clinical methods. This remains a daunting task and continues to challenge researchers with finding the "needle in a haystack." Yet, the technologies that provide the means to inventory components within the "haystack" at an unprecedented rate have exponentially expanded knowledge of the different types of proteins within serum, and opened the way for novel technologies for diagnosing cancers.

The coupling of high-throughput technologies enables samples from hundreds of patients to be rapidly compared. These technologies have greatly advanced the fields of: proteomics (the study of the structure and function of proteins including the way they work and interact with each other inside cells); genomics (the study of the organisation of genomes and the nucleotide sequences of the component genes); and transcriptomics (the study of genes transcribed from DNA within living cells to molecules of messenger RNA as the first step in protein synthesis). As a result, a number of candidate biomarkers have been identified for various cancer types. The next challenge is how to pick the right biomarkers from among the hundreds of promising candidates.

Selecting the right biomarker

The US National Cancer Institute's Early Detection Research Network [4], a consortium of more than 40 laboratories and 300 investigators, has established guiding principles, commonly known as the five-phase approach, for developing, evaluating and validating biomarkers. These guidelines are used to facilitate the transition of biomarkers toward clinical applications. The five phases provide the principles and study design foundations for validating biomarkers headed for clinical use in risk assessment and early detection of cancer. Phase 1, the discovery phase, includes exploratory studies to identify potentially useful biomarkers. In Phase 2, the validation phase, biomarkers are thor-oughly analysed and verified to determine their capacity for distinguishing between people with cancer and those without. Phase 3 focuses on the capacity of a biomarker to detect preclinical disease by testing the marker against tissues collected longitudinally over time from various research cohorts. Phase 4 comprises prospective screening studies. In Phase 5, large-scale population studies that evaluate both the role of the biomarker for detection of cancer, and the overall impact of screening on the population are conducted [5]. Examples of some biomarkers are provided in Table 3.7.1.

In the context of cancer biomarker testing, the sensitivity of a biomarker refers to the proportion of case subjects (individuals with confirmed disease) who test positive for the biomarker. Specificity refers to the proportion of control subjects (individuals without disease) who test negative for the biomarker. An ideal biomarker test would have 100% sensitivity and specificity; that is, everyone with cancer would have a positive test, and everyone without cancer would have a negative test. The lower the sensitivity, the more often individuals with cancer will not be detected. The lower the specificity, the more often someone without cancer will test positive. None of the currently available biomarkers achieve 100% sensitivity and specificity. For example, prostate specific antigen (PSA), currently the best overall serum biomarker for identifying prostate cancer, has high sensitivity (greater than 90%) but low specificity (about 25%), which results in many men having biopsies when they do not have detectable prostate cancer [6,7]. The serum tumour biomarker for breast cancer, CA15.3, has only 23% sensitivity and 69% specificity, and is only useful in monitoring therapy for advanced breast cancer or recurrence [8]. Other frequently used terms are: positive predictive value, which is the possibility that a person with a positive test has cancer, and negative predictive value, which is the possibility that a person with a negative test does not have cancer. Positive predictive value and, to a lesser degree, negative predictive value are affected by the prevalence of disease in the screened population. For a given sensitivity and specificity, the higher the prevalence, the higher the positive predictive value. The Early Detection Research Network creates the biomarker pipeline for a specific cancer type based on the diagnostic performance criteria discussed above and additional considerations, such as incremental benefits over existing practice of care, cost and acceptance by the patients and caregivers. Once the candidate biomarkers are identified, they are subjected to specimen reference sets collected by the Early Detection Research Network [4] to verify the biomarker achieves the intended clinical goal, such as early detection, diagnosis or prognosis. If the preset criteria for the intended goal are met, the markers move to the next testing stage. An example is provided in Figure 3.7.1.

Use of five-phase guidelines: a case study

In the USA, bladder cancer is the fourth most common malignancy in men and the seventh most common malignancy in women. Bladder cancer occurs in two clinically significant forms: (1) superficial (TNM: Ta, TiS, T1) and (2) invasive (TNM: >T2). Seventy-five percent of individuals with bladder cancer have superficial disease and only a minority of those (approximately 15%) are at risk for disease progression. Most individuals (approximately 70%) with superficial disease will experience relapse during a 10-year period. The majority of recurrences occur within the first 2 years after the diagnosis. Therefore, these individuals require frequent surveillance for recurrence that includes cystoscopy and urine cytology every 3 months for 2 years and then annually, and radiographic evaluation of the upper urinary tract every year. Although urine cytology and cystoscopy are considered standard of care, they are less than optimal in detecting all forms of bladder cancer. The sensitivity and specificity of urinary cytology are 25–50% and 90–100%, respectively. The sensitivity and specificity of cystoscopy is 90–100% and 75%, respectively. Consequently, there is a need to improve the current practice of bladder cancer surveillance [9,10].

Microsatellite analysis is a promising new technique for the surveillance of bladder cancer. The technology, which permits the separation by electrophoresis of polymerase chain reaction-amplified DNA sequences from non-malig-

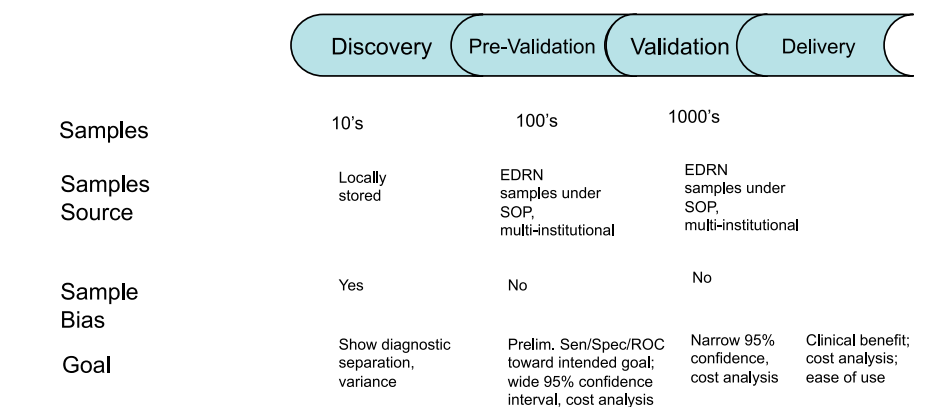|	Discovery	Pre-Validation	Validation	Delivery
Samples	10's	100's	1000's	
Samples Source	Locally stored	EDRN samples under SOP, multi-institutional	EDRN samples under SOP, multi-institutional	
Sample Bias	Yes	No	No	
Goal	Show diagnostic separation, variance	Prelim. Sen/Spec/ROC toward intended goal; wide 95% confidence interval, cost analysis	Narrow 95% confidence, cost analysis	Clinical benefit; cost analysis; ease of use

Fig. 3.7.1 *Evolution of the biomarker pipeline from discovery to development to validation to clinical application in the Early Detection Research Network*

nant and malignant sources, has been applied to the diagnosis of solid tumours arising in the colon, lung, oropharynx, kidney and bladder. Microsatellite analysis can detect genetic changes indicative of carcinoma from urothelial cells obtained in voided urine specimens. The genetic profile of DNA purified from urine is compared to that of DNA purified from peripheral lymphocytes that are considered "normal." Once the DNA from uroepithelial cells has been obtained, polymerase chain reaction is performed with specific oligonucleotide primers for each chromosomal locus. The polymerase chain reaction products are then examined for evidence of microsatellite instability (a change that occurs in the DNA of certain cells, such as tumour cells, in which the number of repeats of microsatellites—short, repeated sequences of DNA—is different than the number of repeats in the DNA when it was inherited) and loss of heterozygosity, which are genetic characteristics of epithelial tumours. Preliminary work shows that MSA detects 95% of cancers [9,10]. However, the effectiveness of microsatellite analysis testing must be validated using a prospective collection of samples from geographically diverse populations. The Early Detection Research Network has selected this promising marker for a Phase III clinical trial since this has met all the requirements of Phase I and Phase II clinical research.

The goals of the study are:

– To determine sensitivity and specificity of microsatellite analysis of urine sediment, using a panel of 15 microsatellite markers, in detecting bladder cancer in participants requiring cystoscopy. This technique will be compared to the diagnostic standard of cystoscopy, as well as to urine cytology.
– To determine the temporal performance characteristics of microsatellite analysis of urine sediment.
– To determine which of the 15 individual markers or combination of markers that make up the microsatellite analysis test are most predictive of the presence of bladder cancer.

Three populations will be included in this study. Two of the populations will include 200 participants (100 each) without bladder cancer (controls). The control population will include two cohorts: (1) a cohort of 100 participants without a history of or current urologic diseases or devices and with a normal urinalysis and urine cytology examination, referred to as Control Group 1; and, (2) a cohort of 100 participants with one of four urologic processes requiring cystoscopy, which in the past have had confounding results with urinary tumour detection assays, referred to as Control Group 2. Both control groups will undergo urinalysis and cytology at baseline along with microsatellite analysis, but only the second group (potentially confounding gastrointestinal conditions) will undergo cystoscopy. The rationale is that the first cohort includes truly "healthy" participants with no gastrointestinal complaints; they have no medical reason to undergo a cystoscopy. Requiring cystoscopy would severely limit recruitment, potentially induce selection bias, and expose otherwise healthy participants to the small but known risks associated with cystoscopy. However, the second control cohort is presenting with gastrointestinal complaints that would ordinarily indicate the need for cystoscopy.

The third population to be enrolled will include 300 participants with bladder cancer, both incident and recurrent. This group of participants will undergo urine cytology and cystoscopy at baseline, and quarterly follow-up surveillance cystoscopy and cytology determinations with synchronous microsatellite analysis determinations.

The accrual is now complete and follow-ups are underway. This first-ever study will provide evidence as to whether or not microsatellite analysis should be used alone or in combination with urine cytology and cystoscopy to monitor the progression of bladder cancer.

A number of other validation studies are underway for markers for pancreatic, lung, mesothelioma, prostate and bladder cancers (Table 3.7.1).

Future directions

Because a single biomarker may not have sufficient sensitivity and specificity to be useful for early detection, there is interest in multiplexing biomarkers (that is, developing a panel of them for concurrent use) that would probably perform better than a single diagnostic marker. Flexible technology platforms are being developed by diagnostic companies that allow for the analysis of a number of biomarkers on a single platform. These multiplexed platforms are designed to simultaneously analyse a panel of protein or nucleic acid biomarkers or more than one kind of biomarker. The multiplexing approach can eliminate time-consuming manual processing of samples, making it faster, efficient and more convenient and allowing for real-time data acquisition and efficient sample comparison. Another important innovation in biochip technology is the microfluidic chip-based immunoassay, which can analyse the expression of serum proteins comparable to commercial enzyme-linked immunosorbent assays, a method using antibodies to quantify levels of a biological marker. However, multiplexing can be a confounding task when optimising the assay conditions, and there is still need for the development of efficient tools for analysing such high dimensional and high throughput data.

With continued attention, support and open cross-disciplinary, multi-institutional collaborations, the challenges of finding and developing accurate and useful biomarkers for early cancer detection and cancer risk will fade and new, long-awaited, less-invasive tools brought into clinical use.

Candidate/Panel	Organ	Status	Part of Multiplex	Reference Number
Annexin 1, LAMR 1, 14-3-3theta	Lung Adenocarcinoma	Phase II	yes	[11]
LCN2, TIMP1, REG1A, REG3 and IGFBP4	Pancreatic	Phase II	yes	[12]
SPINK1, PCA3, GOLPH2, TMPRSS2-ERG	Prostate	Phase II	yes	[13]
CA-125, MIF-1, prolactin, osteopontin, IGF-2 and leptin	Ovary	Phase II	yes	[14]

Table 3.7.1 List of clinical biomarker candidates transitioning through the five-phase approach established by the Early Detection Research Network

REFERENCES

1. Srivastava S and Wagner P (2007). Risk-based and diagnostics-linked personalized medicine for cancer. Personal Medicine 4: 33-43.

2. Biomarkers and surrogate endpoints: preferred definitions and conceptual framework (2001). Clin Pharmacol Ther 69: 89-95.

3. Srivastava S (2007). Cancer biomarker discovery and development in gastrointestinal cancers: Early Detection Research Network-a collaborative approach. Gastrointest Cancer Res 1: s60-s63.

4. National Cancer Institute (2008). Early Detection Research Network. Available from: http://www.cancer.gov/edrn Date Accessed: November 19,2008.

5. Pepe MS, Etzioni R, Feng Z, et al. (2001). Phases of biomarker development for early detection of cancer. J Natl Cancer Inst 93: 1054-1061.

6. Ornstein DK and Kang J (2001). How to improve prostate biopsy detection of prostate cancer. Curr Urol Rep 2: 218-223.

7. Ito K, Yamamoto T, Ohi M, et al. (2003). Natural history of PSA increase with and without prostate cancer. Urology 62: 64-69.

8. Li J, Zhang Z, Rosenzweig J, et al. (2002). Proteomics and bioinformatics approaches for identification of serum biomarkers to detect breast cancer. Clin Chem 48: 1296-1304.

9. Mao L, Schoenberg MP, Scicchitano M, et al. (1996). Molecular detection of primary bladder cancer by microsatellite analysis. Science 271: 659-662.

10. Steiner G, Schoenberg MP, Linn JF, et al. (1997). Detection of bladder cancer recurrence by microsatellite analysis of urine. Nat Med 3: 621-624.

11. Qiu J, Choi G, Li L, et al. (2008). Occurrence of autoantibodies to annexin I, 14-3-3 theta and LAMR1 in prediagnostic lung cancer sera. J Clin Oncol 26: 5060-5066.

12. Faca VM, Song KS, Wang H, et al. (2008). A mouse to human search for plasma proteome changes associated with pancreatic tumor development. PLoS Med 5: e123 DOI:10.1371/journal.pmed.0050123.

13. Laxman B, Morris DS, Yu J, et al. (2008). A first-generation multiplex biomarker analysis of urine for the early detection of prostate cancer. Cancer Res 68: 645-649.

14. Visintin I, Feng Z, Longton G, et al. (2008). Diagnostic markers for early detection of ovarian cancer. Clin Cancer Res 14: 1065-1072.

Stem Cells and Cancer Stem Cells

Summary

> Stem cells constitute a distinct population of cells characterised by the ability to renew themselves indefinitely through mitotic division and to differentiate into a diverse range of specialised cell types

> The two broad types of mammalian stem cells are: the embryonic stem cells that are found in early embryos, and the adult stem cells that are found in many adult tissues

> The main properties of stem cells are self-renewal, essential for maintenance of the stem cells pool, and the ability to differentiate in different lineage required for the integrity and function of tissues

> Cancer stem cell (CSC) is an operational term to functionally define a distinct subpopulation of tumour cells with unlimited renewal potential. Cancer stem cells share many key properties with embryonic stem cells, including the infinitive proliferation potential and the capacity to invade tissues and organs

> Research on stem cells and cancer stem cells holds great promise to advance the design of novel strategies in cancer therapy

The unprecedented pace of discovery in the field of stem cells has turned academic, political and public attention to the potential application of stem cells in medicine and biomedical research. Stem cells are found in all multi-cellular organisms and are likely to be present as a discrete population in most tissues. Stem cells can be grown in culture and differentiated into specialised cells with properties specific to various tissues. Recent landmark discoveries by Takahashi and Yamanaka [1] that induced pluripotent stem cells (iPSCs), which share very similar properties with embryonic stem cells, could be derived from differentiated cells (skin fibroblasts), have already been reproduced for a variety of human cell types [2]. Furthermore, two studies have described the generation of iPSC lines from individuals harbouring both simple and complex genetic diseases [3,4]. Therefore, stem cells have been seen as an essential toolbox in cloning and regenerative medicine. Interestingly, recent advances have revealed that stem cells can be a source of cancer cells, and that better understanding of the behaviour and properties of stem cells could be exploited to devise an arsenal of novel tools to fight against cancer. Although technical advances that allow the isolation and manipulation of embryonic stem cells and possibility of human cloning have provoked intense ethical debates, this should not undermine the tremendous potential of stem cells in the treatment of various human diseases such as neuro-degenerative disorders and cancer [5,6].

Embryonic and tissue-specific stem cells

Every cell in the body is a descendent of a single cell (fertilized egg or zygote). The life of an organism starts with fertilisation of an egg, and from this moment until death involves the passage through several developmental stages (Figure 3.8.1). Multiplication of fertilised eggs gives rise to different cell types of the body. This process involves generation of populations of stem cells that can be propagated indefinitely in culture under adequate conditions. These cells are called the embryonic stem (ES) cells that are able to give rise to any cell type and to reconstitute the entire embryo. In addition, many adult tissues contain a discrete population of undifferentiated cells with properties of stem cells. These cells are known as tissue-specific stem cells (somatic stem cells). Hematopoietic stem cells are the best-characterised tissue-specific stem cells that generate all blood lineages and mature blood elements. The adult stem cells are identified in many other tissues such as brain, skin and liver. While only a few tissue-specific stem cells have undergone a rigorous identification and characterisation, it is likely that stem cells are present in any tissue that undergoes renewal.

Tissue-specific stem cells also have the capacity to perpetuate itself through self-renewal and to produce various mature cells of a particular tissue through differentiation [7]. Tissue-specific stem cells constitute a tiny cell population in adult tissues, yet they are essential in the maintenance of tissue homeostasis and their deregulation may trigger diseases, most notably cancer. The two characteristics of stem cells that distinguish them from all other cells are self-renewal and pluripotency. Self-renewal is the capacity of a cell to divide and produce identical daughter cells over long time period. This property is crucial as it allows that stem cells persist for the lifetime of an organism. Pluripotency is the capacity of stem cells to differentiate into many highly-specialised cells such as neurons, muscle fibers, and blood elements. This feature is important for the maintenance of integrity and function of many tissues. Stem cells are essential for the development and integrity and function of an organism and thus can be considered as a precious treasure from the beginning of embryonic life to the death. Given their properties, function and proliferation of stem cells need to be tightly monitored. Deregulation of the surveillance mechanisms for proliferation and differentiation of stem cells may trigger a shift in the balance between self-renewal and differentiation leading to either stem cell loss, associated with degenerative disorders, or abnormal proliferation of stem cells that may be a source of malignant cells.

Cancer stem cells

Stem cells have been discovered a quarter century ago and have been exploited extensively for the generation of genetically-modified animal models (for example, knockout mice), an essential tool in cancer research. However, the identification of the first human stem cells and in particular so-called cancer stem cells triggered unprecedented attention of the cancer research community. It has long been accepted that most tumours are derived from a single cell that has been transformed into a cancer-initiating cell through acquisition of a series of genetic and epigenetic lesions. These initial events allow expansion of transformed cells and formation of a population of altered cells (clone) with capacity to grow and divide in defiance of normal cellular control. Continuing selection of "fitter" and more aggressive cells results in a generation of cancer clones capable of invading and destroying neighbouring tissues and migrating to distant organs to form secondary tumours (metastasis). It is now believed that many human cancers arise from deregulated control of stem cells (Figure 3.8.2). Moreover, recent studies indicated that many genetic and epigenetic changes underlying aggressive and destructive behaviour of cancer are orchestrated by discrete population of cancer cells with stem cell properties. These cells are known as cancer stem cells. The idea that cancer develops from stem cells was suggested as early as 1875 when Cohnheim hypothesised that stem cancer displaced during embryonic development may be the origin of malignant cells later in the life [8]. However, it was not until the identification of the first cancer stem cells that the cancer stem cell concept received considerable attention.

The cancer stem cell hypothesis suggests that cancer clones are maintained exclusively by a rare fraction of cells with stem cell properties. Many cancers are found to contain cells with properties of stem cells. However, in most cases the existence of cancer stem cells has been documented functionally. This means that the presence of cancer stem cells in the bulk of cancer cells is discerned by their capacity to form tumours after transplantation into an immunocompromised animal host (usually mice). These assays revealed that only a small fraction of cancer cells were capable of forming new tumours in the host. Importantly, these cells are shown to be able not only to form tumours upon transplantation but also to recapitulate tumour heterogeneity [10]. However, until very recently it proved extremely difficult to isolate cancer stem cell population using molecular signatures, cell-surface markers or mutation profiles [5,11].

Cancer stem cells share many key properties with embryonic stem cells. These include the infinitive proliferation potential and the capacity to invade tissues and organs, and promote formation of blood vessels for their own supply. While we have seen important progress on the identification of cancer stem cells, the origin of cancer stem cells remains mysterious. It is believed that cancer stem cells may arise in different ways. First, cancer stem cells can be derived from normal tissue-specific stem cells as a result of specific genetic and epigenetic changes that abrogate their proliferation control. Second, differentiated cells that normally have a limited life span can regain stem cell properties (de-differentiate) and become cancer stem cells. Third, it is also possible that normal stem cell fuse with various differentiated cells and resulting hybrid cells may be cancer-initiating cells with stem cell properties [11]. These hypotheses are not mutually exclusive, and the genesis of cancer stem cells may involve more than one mechanism.

Gene wiring that instructs stem cell identity

Although the features that distinguish stem cells from all differentiated cells have been known
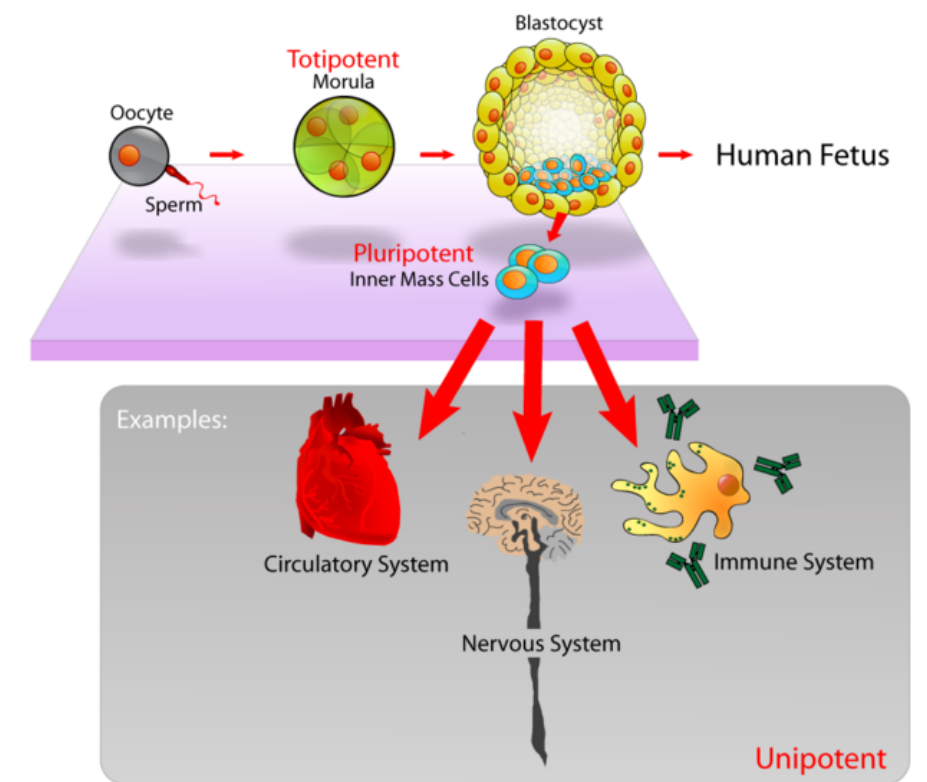

Fig. 3.8.1 Embryonic stem cells can become all cell types of the body because they are pluripotent. Adult stem cells are generally limited to differentiating into different cell types of their tissue of origin

for many decades, it was not until recently that we have begun to understand the genetic basis of stem cell identity. The development of powerful tools in genomics for genome-wide screens allowed the identification of genes and gene networks that keep stem cells in a special state. Using these tools, scientists have discovered a handful of genes that are necessary and sufficient to maintain self-renewal and pluripotency, two distinguishing features of stem cells. These genes are known as "masters of stemness". The genes Oct4, Sox2, and Nanog belong to this privileged club [12]. These genes encode for specialised proteins known as transcription factors whose duty is to control the transcription of other genes (Figure 3.8.3). This forms a kind of gene wiring that instructs stem cell behaviour and identity. When these genes are inactivated or mutated, stem cells may differentiate into specialized cells and stem cell pool may be rapidly depleted. This can impede regeneration and integrity of normal tissues leading to degenerative diseases [13].

The discovery of stem cell master genes enabled another even more tantalising adventure: the reversal of specialised (differentiated) cells into immature pluripotent (stem) cells, the process known as de-differentiation—in other words, making specialised cells (such as neurons or muscle fibres) become cells with stem cell properties that would allow the generation of just about any type of cell. This would solve important ethical issues associated with the use of embryos as a source of stem cells. Recent studies demonstrated just that [1,14,15]. Several laboratories showed that the introduction of as few as 4 of master genes into differentiated cells of either humans of mice into stem cells. This remarkable phenomenon argues that differentiated cells can be reprogrammed and that differentiation clock can be reversed. An important implication of these findings is that cancer stem cells may also arise by non-genetic changes that confer certain features to differentiated cells. In support of this hypothesis is the fact that all cells including stem cells in any given organisms share an identical genome (the sum of genetic codes).
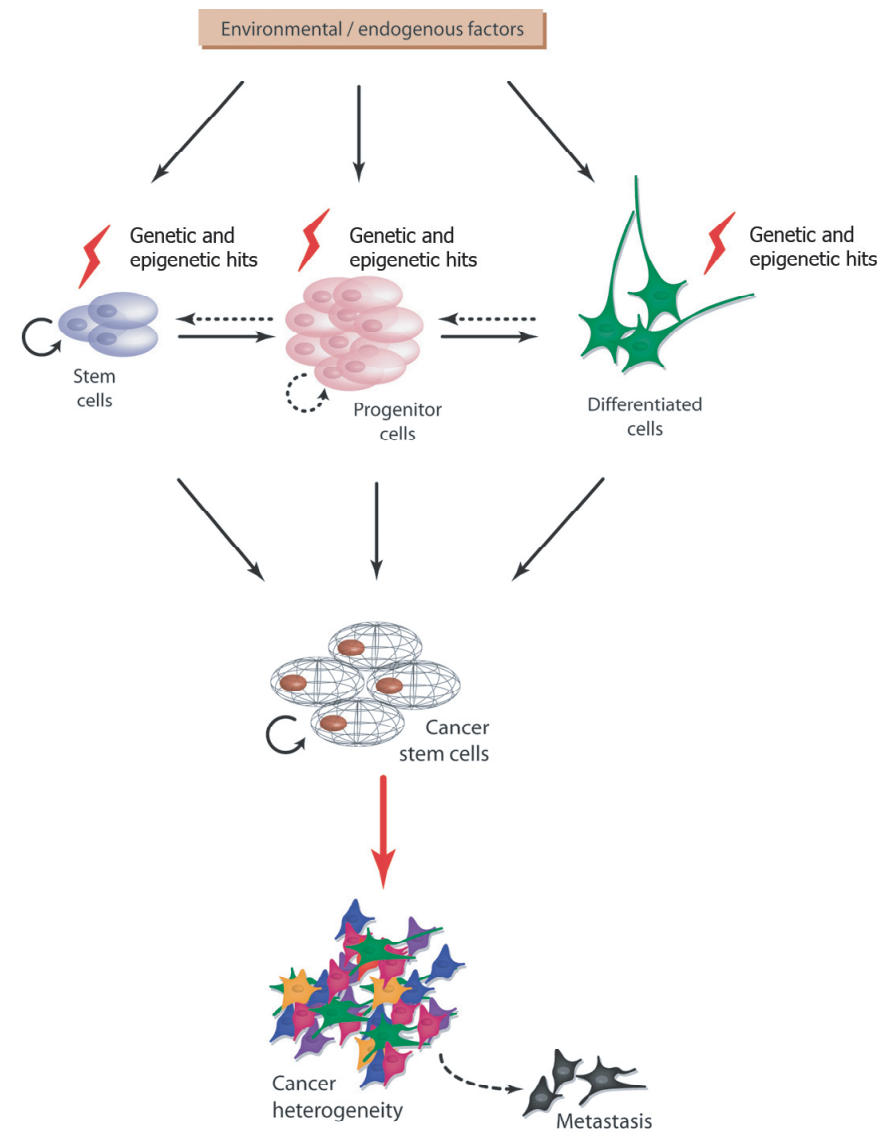

Environmental / endogenous factors

Genetic and epigenetic hits

Genetic and epigenetic hits

Genetic and epigenetic hits

Stem cells

Progenitor cells

Differentiated cells

Cancer stem cells

Cancer heterogeneity

Metastasis

Fig. 3.8.2 Genetic and epigenetic changes in stem/progenitor cells may be an early event in the development of cancer and give rise to cancer stem cells and contribute to tumour heterogeneity. Abnormal expression and function of a set of genes in more differentiated cells may contribute to reprogramming into a pluripotent state and also cancer stem cells. [9]

These findings suggest that special features of stem cells are due to non-genetic (epigenetic) events. In other words, gene expression patterns of stem cells but also differentiated cells are controlled by epigenetic mechanisms. Self-renewal and pluripotency represent opposing demands on genome of stem cells. Self-renewal potential requires a long-term memory system for stable maintenance of transcriptional patterns without changes in the genomic code.
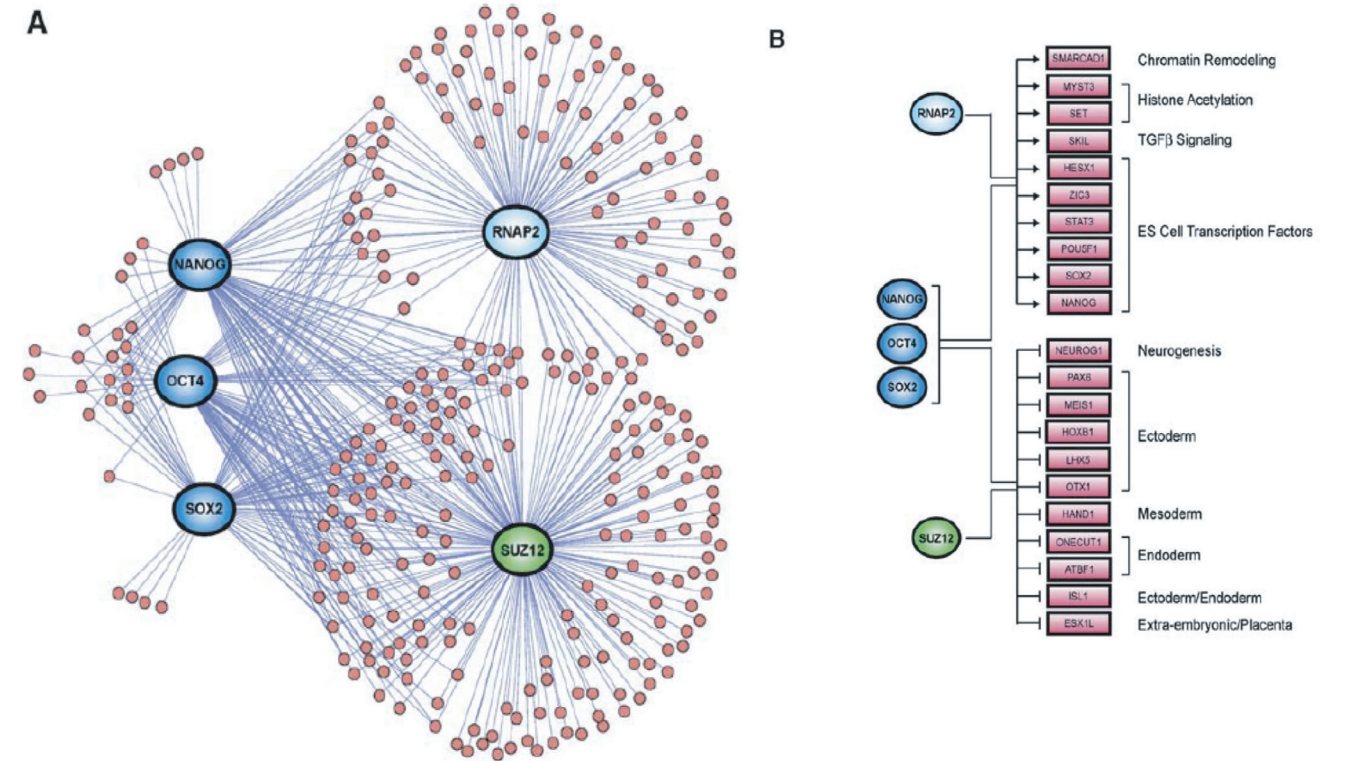

Fig. 3.8.3 Transcriptional regulatory network governed by several key regulators in human embryonic stem cells.

In contrast, the potential for multi-lineage differentiation requires plasticity of the genome allowing multiple differentiation decisions. This apparent dichotomy of stem cells is reflected by the presence of specific patterns in DNA methylation (epigenetic modification of DNA molecule which does not involve changes in genetic code) and histone modifications (markings of special proteins which ensure protection and compaction of DNA chain)[9]. Therefore, it is quite plausible that deregulation of epigenetic mechanisms may lead to an altered potential of stem cell self-renewal and expansion of epigenetically modified stem cell pools. Stem cells modified in this manner exhibit no genetic changes, yet they may represent a precursor pool susceptible to acquisition of mutations and further epigenetic alterations. In this scenario,

disruption of epigenetic states is the first step of tumorigenesis and is a contributing factor to polyclonal tumour phenotype.

Stem cell and cancer therapy

Much of the current interest in a stem cells and cancer stem cells comes from the realisation that this tiny yet critical population of cancer cells represents an opportunity to devise novel strategies for cancer therapy [5-7]. It is believed that many current protocols for cancer therapy fail and cancer reappears due to the failure to eradicate cancer stem cells. For example, classical chemotherapy regimes have been developed to rapidly shrink tumours; however these effects are transient and are followed by tumour relapse (Figure 3.8.4). This is explained by the

fact that current drugs can efficiently kill rapidly growing cancer (non-stem) cells, whereas slower growing cancer stem cells may be spared. Thus, the major challenges will be: (i) to discover efficient ways to identify and isolate tissue-specific stem cells and cancer stem cells; (ii) to gain insights into the mechanisms of self-renewal and pluripotency of normal stem cells and cancer stem cells; and (iii) to pinpoint genetic and epigenetic events that are at the heart of cancer-initiating reprogramming and cancer stem cell development. The ever-increasing research efforts in the field of stem cells carry a promise to provide missing pieces of the puzzle of discovering and targeting the Achilles heel of cancer cells, which would make a major impact on the development of novel strategies to conquer cancer.
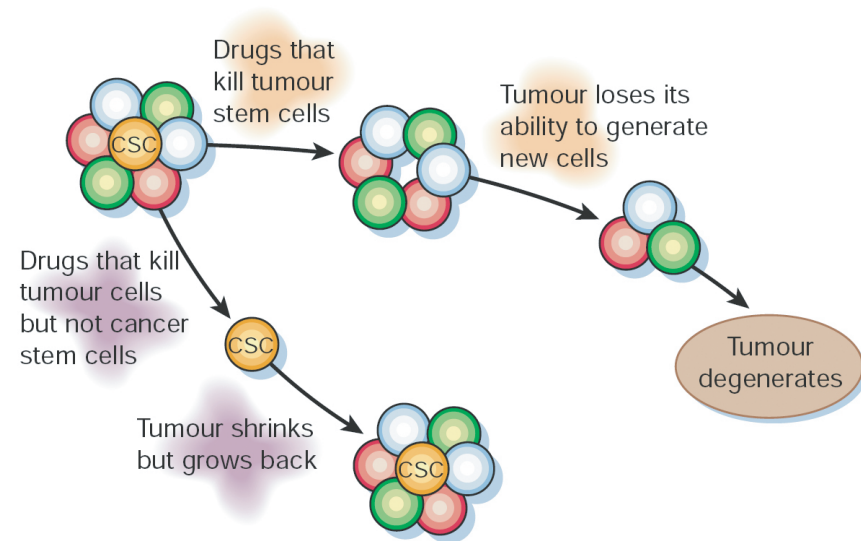
Fig. 3.8.4 *Conventional therapies may reduce tumours by killing mainly differenciated tumour cells. If the putative cancer stem cells are less sensitive to these therapies, then they will remain viable after therapy and re-establish the tumour. By contrast, if therapies can be targeted against cancer stem cells, then they might more effectively kill the cancer stem cells, rendering the tumours unable to maintain themselves or grow. Thus, even if cancer stem cell-directed therapies do not shrink tumours initially, they may eventually lead to cures. [7]*

REFERENCES

1. Takahashi K and Yamanaka S (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* 126: 663-676.

2. Yamanaka S (2007). Strategies and new developments in the generation of patient-specific pluripotent stem cells. *Cell Stem Cell* 1: 39-49.

3. Dimos JT, Rodolfa KT, Niakan KK, et al. (2008). Induced pluripotent stem cells generated from patients with ALS can be differentiated into motor neurons. *Science* 321: 1218-1221.

4. Park IH, Arora N, Huo H, et al. (2008). Disease-specific induced pluripotent stem cells. *Cell* 134: 877-886.

5. Polyak K and Hahn WC (2006). Roots and stems: stem cells in cancer. *Nat Med* 12: 296-300.

6. Baumann M, Krause M, Hill R (2008). Exploring the role of cancer stem cells in radioresistance. *Nat Rev Cancer* 8: 545-554.

7. Reya T, Morrison SJ, Clarke MF, et al. (2001). Stem cells, cancer, and cancer stem cells. *Nature* 414: 105-111.

8. Cohnheim V (1875). Congenitales, quergestreiftes muskelsarkom der nieren. *Virchows Arch Pathol Anat Physiol Klin Med* 65: 64-69.

9. Shukla V, Vaissiere T, Herceg Z (2008). Histone acetylation and chromatin signature in stem cell identity and cancer. *Mutat Res* 637: 1-15.

10. Visvader JE and Lindeman GJ (2008). Cancer stem cells in solid tumours: accumulating evidence and unresolved questions. *Nat Rev Cancer* 8: 755-768.

11. Bjerkvig R, Tysnes BB, Aboody KS, et al. (2005). Opinion: the origin of the cancer stem cell: current controversies and new insights. *Nat Rev Cancer* 5: 899-904.

12. Boiani M and Scholer HR (2005). Regulatory networks in embryo-derived pluripotent stem cells. *Nat Rev Mol Cell Biol* 6: 872-884.

13. Erlandsson A and Morshead CM (2006). Exploiting the properties of adult stem cells for the treatment of disease. *Curr Opin Mol Ther* 8: 331-337.

14. Cowan CA, Atienza J, Melton DA, et al. (2005). Nuclear reprogramming of somatic cells after fusion with human embryonic stem cells. *Science* 309: 1369-1373.

15. Ivanova N, Dobrin R, Lu R, et al. (2006). Dissecting self-renewal in stem cells with RNA interference. *Nature* 442: 533-538.

Biobanks and Biological Resource Centres

Summary

> Collections of human biological specimens are the cornerstone of research on gene-environment interactions. They are also strategic resources for translational medicine and for the discovery of biomarkers useful in cancer management

> Biobanks are collections of biological samples and associated data that are structured in a way that makes them accessible for research. Such biobanks must abide by extremely strict technical, legal and ethical standards

> Biological Resource Centres regroup biobanks and technical services required to handle all aspects of specimen processing from collection to long-term storage and shipping to laboratories that perform biomarker analyses. BRCs are now at the centre of molecular epidemiology and evidence-based medicine

> Many studies on biomarkers require the analysis of large number of specimens, making it necessary to gather specimens in multiple centres involving several BRCs. Therefore, it is critical to promote the adoption of common minimal standards and protocols that are widely applicable in different field, laboratory or hospital contexts

> IARC, as a research centre with global outreach, is involved in collecting biospecimens in many parts of the world. This chapter summarises the recommendations and guidelines IARC gives to its collaborators worldwide for developing biobanks

Biobanks are at the centre of recent advances in cancer research. A "biobank" is an infrastructure to store biospecimens, e.g. blood samples or tissues, and to make them available for laboratory analyses. The term "Biological Resource Centre" (BRC) is used to identify specialised units that handle the acquisition, quality control, storage, processing and distribution of biospecimens to laboratories where they will be analysed. Thus this is more than just a "bank of samples": it encompasses the whole chain from the study participant (cancer patient or healthy volunteer) to the laboratory that performs biological analyses. Operating a BRC requires compliance with specific regulations and recommendations setting standards for collection, labeling, annotation, processing, storage, retrieval and analysis of the biospecimens, while ensuring biological safety and protection of personal data. The keyword in these tasks is traceability, allowing for the tracking of the biospecimens at all steps from collection to laboratory analyses. These complex operations involve bioethical aspects (How to inform study participants and ask them for consent? Who has the right to decide about how the specimens given in this way should be used? What is the proper action when specimen analysis reveals something unexpected that may be important for the participant? What are the procedures after death of the sample donor?), technical problems (How to make sure that the specimens collected are of consistent high quality? How to best preserve them in the long-term?), development of databases (What information on the participant and on the specimen should be stored? How to protect the confidentiality of this information? How to follow-up patients during and after treatment?), and huge logistical problems (Where to store specimens? How to manipulate them? How to handle, process, transport thousands of specimens?). Finally running a BRC requires highly trained, dedicated staff, defined standard operating procedures, and quality assurance/quality control procedures. This has serious economical implications in particular for the long-term sustainability of large biobanks. This chapter addresses the roles of BRC in cancer research and explains how the development of such BRC will shape not only tomorrow's cancer science but also medical practice.

The need for large biobanks and biobank networks

BRCs play a number of critical roles in all aspects of biological research. The role of BRCs in biological research in general, and their impact on medical, societal and economical issues has been extensively discussed in a recent report of the Organisation for Economic Co-operation and Development [1]. Storing specimens is not a new activity: biologists, natural scientists and doctors, including in particular pathologists, have always preserved samples for diverse scientific and medical uses. In the 19th century, these samples often were tissues or whole organs immersed in glass jars filled with formalin or alcohol. More recently, and indeed as is still the case today, hospital pathologists preserve tissues for diagnosis in paraffin wax. But with the advent of modern genetics and molecular biology, these questions have taken a new, unprecedented twist: it is now possible to analyse hundreds of thousands of markers in the DNA of a single subject, or to measure the whole set of genes expressed by a healthy or diseased tissue. Thus, scientists become capable of performing very detailed and precise molecular characterisation of biological samples.

Research on the molecular mechanisms of cancer has identified many molecules that can be measured and used as indicators of the effects of environmental exposures, of genetic susceptibility, of early steps of cell transformation, of early cancer disease, and of cancer progression towards invasiveness. Such biomarker molecules are critical not only to better classifying and understanding cancer diseases, but also to better detect and diagnose them, prognosticate their evolution, predict their responses to therapy, and provide targets for new drugs aimed at improving cancer treatment. A new area of biomedicine, often referred to as the "-omics", is in full development to combine large-scale biological analysis with bio-computing, generating large amounts of data on the status and level of multiple molecular biomarkers (see chapter on "Biomarkers"). The suffix "-omics", as for example in "genomics", conveys the notion that these approaches are aimed at assessing molecular makers in their wholeness. This distinguishes genomics from genetics, which usually concentrates on the study of a small number of related genes. Similarly, new words have been coined to identify the global study of gene transcription ("transcriptomics"), of proteins ("proteomics") or metabolic activities ("metabolomics"). Studies on human specimens are also becoming critical in the process of discovering new mechanisms involved in causing cancer or in determining its progression, resistance/response to treatment and clinical outcome.

Biobanks and personalised medicine

Collecting and analysing biological specimens is a necessary procedure for pathology-based diagnosis and is also a mechanism for allowing patients to benefit from the applications of molecular cancer research. Today, BRCs are the foundation of three rapidly expanding domains of biomedical sciences: molecular and genetic epidemiology (aimed at assessing the genetic and environmental basis of cancer causation in the general population as well as in families), molecular pathology (aimed at developing molecular-based classification and diagnosis procedures for cancer diseases), and pharmacogenomics/ pharmacoproteomics (understanding the correlation between an individual patient's genotype/phenotype and response to drug treatment). Close involvement of the pathology department at collecting centres is essential to facilitate the use of banked fresh frozen samples in diagnostic procedures. In the future, correct assessment of patient status and therapeutic needs may require the determination of a number of molecular parameters and will require systematic preservation of frozen biospecimens or derived biomaterial. With the continuing improvement of survival after therapy, performing such molecular-based evaluations may become a systematic requirement not only at diagnosis but also at different stages of patient follow-up. While the present chapter specifically deals with biorepositories for research, it is recognised that developing BRCs may rapidly become part of recommended, if not mandatory, medical practice. Thus, gathering know-how and procedures for collecting, storing and analysing human specimens is a major contribution to the development of biomedical practice worldwide. In the chain from laboratory discovery to medical application, biobanks have made a key contribution to life science research and development (R&D). Progress in medicine is dependent upon innovation, development and translation of laboratory findings into clinical practice. Access to human biological specimens is often a prerequisite for such R&D advances. Thus, development of high-quality BRCs has the potential to accelerate and facilitate this translational process.

Importance of networking and exchanges between BRCs

Cancer is a global disease, the understanding and management of which requires comparisons between disease patterns in different parts of the world. In addition, studies on many rare forms of cancer are limited by the difficulty in recruiting a sufficient number of cases within any single collection centre. Furthermore, large numbers of human biospecimens must be available to harness the full potential of novel, large scale technologies in genomics. Due to the molecular diversity of cancer, it is not possible to make accurate distinctions between sets of several thousand of biomarkers by analysing only a few hundred specimens. Statistical power is the key problem. Numbers of specimens analysed must be commensurate with the number of biomarkers simultaneously assessed, as well as with the prevalence of these biomarkers. To achieve this, molecular epidemiological studies often have to include thousands (if not hundreds of thousands) of subjects, recruited in many different locations. Thus, biobanks should be developed not as single autonomous "units" in a given hospital or research institutes, but as parts of a network of biobanks capable of sharing specimens. This leads us to the second basic requirement: the specimens analysed must be of constant, controlled quality, independently of their origin. These large studies must be made strictly comparable. Any lack of due care in specimen collection, processing, transport, storage or distribution to the laboratories may ruin the work of large consortiums of epidemiologists, doctors, nurses and statisticians—not mentioning the ethical responsibility towards subjects who have volunteered to participate in such large studies. Therefore, all laboratories involved in biobanking should adopt common technical standards for specimen collection, storage, annotation and data management. BRCs have an important role in facilitating such exchanges and in providing logistics and infrastructure for multi-centre research projects (epidemiological studies as well as clinical trials). It is recommended that the institution develop tools to enable up-to-date, anonymous information retrieval of clinical annotation on individuals, and set up communications between departments of (e.g.) oncology, surgery, pathology and clinical chemistry.

Legal and ethical implications

Developing and using BRCs requires the active involvement of many actors at different levels (national policy makers, institutional administrators, epidemiologists, pathologists, surgeons, clinicians, bioinformaticians, laboratory scientists) and has complex ethical and legal implications. The perception of these issues and the way they are regulated and managed varies according to legislative, cultural and economical contexts. This paragraph does not intend to provide general answers to these questions, but to put into perspective some important challenges associated to the collection, storage and use in research of human biospecimens. The Helsinki Declaration provides the general framework in which these questions should be addressed [2].

Firstly, the rights of the individuals whose tissues or biological specimens are to be included in the BRC should be strictly considered and protected. Crucial aspects in this process are the development of appropriate methods to obtain informed consent according to the local standards where the definition of protocols are fully

compatible with the three basic ethical requirements of autonomy, beneficence (nonmalevolence) and justice [3]. In this process, it is critical that individuals receive accurate information regarding the potential use of their specimens in large national and international studies as well as in collaborative studies involving third parties such as industrial or commercial partners. Secondly, the institutions that organize and oversee BRCs have the responsibility of protecting individual information and data, to guarantee safe and adequate long-term preservation of banked specimens, to inform, train and protect staff involved in specimen management, to ensure biological and environmental safety, and to make collections accessible and available under defined conditions for research purposes. Thirdly, the scientists who wish to use banked biospecimens for research purposes must submit their research proposals and protocols to appropriate scientific and ethical review. As custodians of the biospecimens, institutions have the duty to take into account the non-renewable nature of the specimens in making priorities for scientific use. In requesting specimens from a biobank, scientists should develop detailed power calculations and provide pilot data to ensure the optimal use of biological resources. Distribution of specimens for research should be done within clear transfer agreements. Such agreements may include return of data and leftover materials to the BRCs. They should also make provisions for users to contribute to the economical sustainability of the BRCs (see below), and should also acknowledge the rights of the BRCs and its scientific contributors to intellectual property derived from research performed using specimens made available by the BRC.

Principles for sustainable Biological Resource Centres

Developing and sustaining a BRC has a high initial cost as well as running financial cost, and can strain economically underprivileged institutions. These constraints are a significant obstacle to developing BRCs in middle or low-resource countries. Lifting these obstacles in these coun-

tries requires a significant, international solidarity effort. The public sector (local, national governments, international bodies and organisations) has a responsibility for contributing to the funding of the baseline infrastructure of BRCs. On the other hand, the responsibility for development and maintenance of sustainable and useable specimen collections lies primarily with the clinical and scientific institutions. These institutions should make provisions towards maintenance of infrastructure, equipment, and running costs as well as data management systems. In addition, users of BRCs should contribute to the general financial and structural sustainability of BRCs. Thus, access to biorepositories of human specimens should entail a contribution from researchers, either in the public or private sector, to the costs of collecting, annotating, storing, retrieving and processing of biospecimens. However, human biospecimens should not be sold in any circumstances. Regardless of the role of industry in core funding of BRCs, which is a matter of debate with serious implications, the responsibility for specimen collection and storage must remain within institutions. In defining mechanisms for BRC sustainability, there is a need to develop safeguards against exploitation and improper use of human biospecimens.

Recommendations for BRCs

Protection of persons

The first, basic requirement of a BRC is Safety. This includes protection of persons and of the environment against biological and chemical hazards, as well as protection of the data and information associated to the specimen collected. The management of these risks should be based on a general implementation of a principled, precautionary approach similar to those used in laboratories and clinical settings, and should be embodied in a general safety management plan.

Biological hazards

All biological specimens should be considered as potentially infectious. Their collection

and processing represents a source of hazard both for the subject who is the source of the specimens and for the staff involved in these processes. Immunization of BRC staff is recommended when appropriate vaccines are available. In particular, immunization against the Hepatitis B Virus (HBV) is mandatory for staff involved in collecting and processing human blood or tissues. Other significant risks are posed by exposure to the Hepatitis C Virus (HCV), the Human Immunodeficiency Virus (HIV) as well as to the prion that causes Creutzfeld-Jacobs diseases. Further sources of biological risk are identified by Grizzle and Fredenburgh [4] [Picture from http://www.foto-search.fr/photos-images/symbole-biohazard.html (RF Libres de Droits)]

General laboratory safety

In addition to biosafety, BRCs need to follow strict general safety regulations and procedures in relation to chemical, physical and electrical safety. The use of liquid gases such as liquid nitrogen for cryopreservation is a serious source of hazard. There are also risks associated with the use of chemical fixatives and solvents used in tissue processing. Electrical safety is an important concern. Deep-freezers must be properly wired to adequate sources of electrical supply, and grounded. Work in a BRC also entails a number of occupational hazards typical of the laboratory environment. These risks must be taken into account before setting up a BRC, and their prevention must be integrated in all aspects of the Standard Operating Procedures of the BRC.

Data management and informatics safety

The protection of personal information and individual data associated with specimen collection is a fundamental requirement of a BRC. This should be achieved through the use of safe, structured bioinformatics systems. The mechanisms of access to these systems, as well as the permissions, should be clearly defined. Backups should be made on a regular basis to avoid data loss. The communication to third parties

or authorities of data files containing personal information and identifiers should be strictly prohibited. Personal data archived in the BRC management system should be protected with the same stringency as patient clinical files.

General considerations for setting up a BRC

A number of factors must be taken into account in setting-up and running a BRC. A detailed description of these requirements can be found in the "best practices for biological repositories" developed by the International Society of Biological and Environmental repositories" [5]. The paragraph below underlines aspects of particular importance in setting up a BRC for cancer research.

Institutional commitment

Many factors contribute to the decision to develop a BRC. In practice, the process often starts from the willingness of medical doctors and scientists to develop a resource useful for diagnosis, prognosis and research purposes. However, initiating a BRC must not only rely on individual action but also requires a clear commitment by the institution to ensure that collections are developed within appropriate legal, ethical, clinical, scientific and technical guidelines, to provide historical continuity in specimen and record keeping, and to ensure that the materials stored by the BRC can be made available for research.

The purpose of the BRC must be clearly formulated and documented. BRCs that contract with third parties for laboratory service should keep detailed records of the nature of the contract, the identity of the contractor, and the inclusive dates of the contract period. In the case of loss of funding or other adverse events that may prevent the institution from maintaining its commitment, it is the responsibility of the institution to take necessary steps to transfer collected specimens and data to another institution that will take over the commitment to the long-term maintenance of the collection.

BRC management and staff

BRCs should be adequately staffed, and the personnel assigned to these tasks must have an appropriate level of training. The BRC should be placed under the overall supervision of a manager with sufficient training, experience and seniority to fulfil the scope of the activities of the BRC. The manager is responsible for operations, including compliance with current regulations. The manager has also a critical role in receiving, processing and answering requests for access to stored specimens.

Running a BRC requires dedicated staff for specimen processing and storage and for data management. The job description, tasks and reporting to authority of each supervisory and technical staff contributing to the BRC must be documented. This is of particular importance in the many instances where the staff contributing to the BRC also performs other tasks within the institution (e.g. pathology service or service activities in molecular biology). Staff must have adequate educational background, experience and training to ensure that assigned tasks are performed in accordance with the BRC's established procedures.

Infrastructure and facilities

The BRC's infrastructure depends upon the types of material being stored, the required storage conditions, the projected retention periods, and the projected use of the materials.

BRCs should have dedicated facilities that are not shared with other activities. Sufficient air conditioning must be provided for air circulation and to maintain ambient temperature equal or less than 22°C at the level of the freezers/refrigerators in order to prevent excess freezer wear and early failure. Rooms that contain liquid nitrogen tanks should be equipped with appropriate air flow systems to avoid the accumulation of N_2 in case of leakage. Storage facilities and instruments should be monitored by appropriate alarm systems (Figure 3.9.1). Response systems must be in place to respond to an alarm in a

time frame that prevents or minimises loss or damage to the collection materials.

BRCs should be equipped with a system that adequately limits access to appropriate staff and protects against physical intrusion. In principle, only persons assigned to the BRC operation should have access to the material, and all materials added to or withdrawn are documented (Figure 3.9.2).

BRCs require a constant source of electrical power. Given that all commercial power will fail at some time, a backup power system is required. The most common type of backup power is the motor generator. Such a system should have the capacity to run for sufficient time to allow the restoration of power supply (typically 48 to 72 hrs) and should be regularly tested.

Adequate back-up capacity for low temperature units must be maintained. The total amount of back-up storage required for large repositor-

Fig. 3.9.1 *Monitors of LN2 level*

Fig. 3.9.2 *Controlled access to the BRC*

ies must be determined empirically, but will typically be 1.5–3% of the total freezer capacity.

Every repository should employ basic security systems. The systems must be monitored and alarms responded to 24h per day, 7 days per week. Response systems must be in place such that a responsible individual can take the necessary action to respond to an alarm in a time frame that prevents or minimises loss or damage to the collection materials. Systems should allow for calls to other key staff from a list of staff phone numbers when the first individual fails to acknowledge the alarm.

Storage conditions

Biospecimens should be stored in a stabilised state. In selecting the biospecimen storage temperature, consider the biospecimen type, the anticipated length of storage, the biomolecules of interest, and whether goals include preserving viable cells.

Cryopreservation

Cryopreservation is the recommended standard for preservation of human biological samples for a wide range of research applications. Cryopreservation, is a process where

cells or whole tissues are preserved by cooling to low sub-zero temperatures, such as (typically) -80°C or -196°C (the boiling point of liquid nitrogen). At these low temperatures, any biological activity, including the biochemical reactions that would lead to cell death is effectively stopped. However, due to the particular physical properties of water, cryopreservation may damage cells and tissue by thermal stress, dehydration and increase in salt concentration, and formation of water crystals. Table 3.9.1 lists the most commonly accepted cryopreservation standards for human tissue and body fluids. It should be noted that specific applications (e.g. proteomics or development of primary cultures) may require more complex cryopreservation procedures. General information on the principles of cryopreservation may be found at http://www.cryobiosystem-imv.com/CBS/Cryobiology/cons_cbs.asp (Figure 3.9.3)

Specimen freezing is generally performed by placing the specimen in a sealed container, and by immersing the specimen into a rapid freezing medium. The ideal medium for rapid freezing is isopentane that has been cooled to its freezing point (-160°C). To achieve this, the vessel containing the isopentane must be introduced into another container of liquid nitrogen. The freezing point approximately corresponds to the

moment when opaque drops begin to appear in the isopentane. Direct contact of the specimen with liquid nitrogen should be avoided, as this damages tissue structures.

Other fixation and preservation methods

Formalin or alcohol fixation and paraffin embedding may be used as an alternative method to preserve tissues at relatively low cost when adequate freezing procedures and storage facilities are not available. Fixed paraffin blocks may be stored in the dark at 22°C in a correctly ventilated cupboard (Figures 3.9.4 and 3.9.5).

Tissues fixed according to strict protocols may be used for DNA extraction. The DNA

Fig. 3.9.3 *Storage at -196°C in liquid nitrogen*

is usually fragmented but remains suitable for PCR-based analysis of short DNA fragments (up to 1-2 kbp). However, fixed tissues are of limited usefulness for RNA extraction. RNAlater is a commercial aqueous, non-toxic tissue storage reagent that rapidly permeates tissues to stabilize and protect cellular RNA. RNAlater eliminates the need to immediately process tissue samples or to freeze samples in liquid nitrogen for later processing. Tissue pieces can be harvested and submerged in RNAlater for storage without jeopardizing the quality or quantity of RNA obtained after subsequent RNA isolation. However, specimens placed in RNAlater cannot be further used for pathological analysis.

Working with Liquid Nitrogen

Where liquid nitrogen (LN_2) refrigeration is employed, an adequate supply of refrigerant must be maintained. The supply maintained on hand should be at least 20% more than the normal re-fill usage to allow for emergency situations.

Handling liquid nitrogen has serious safety implication. Flesh freezes instantly at -196°C, causing severe burns. Because nitrogen displaces oxygen, care must be taken when LN_2 freezers are employed. The risk is inversely correlated with the size of the room. Oxygen level sensors should always be employed when LN_2 freezers are used in a repository.

When bulk storage and piping systems are used, blockage of relief valves and/or overpressure may lead to simultaneous leakage of N_2 from a number of relief valves, causing a "white-out" condition in a matter of a few seconds. Visibility drops to near zero and the oxygen level in the area is below that necessary to sustain life. Personnel must evacuate immediately.

Liquid nitrogen expands to 800 times its original volume at room temperature, causing a form of explosion hazard. Plastic and glass containers can easily explode if liquid is trapped when the container is removed from the freezer. Heavy

gloves, a face shield, and a protective garment should always be used under these conditions.

Liquid Nitrogen (LN_2) tanks

The critical temperature for sensitive tissues, organisms and cells is generally considered to be -132°C, the glass transition temperature (Tg). Vapour phase storage is preferred over liquid phase storage. Design of the tank/freezer is critical to allow maintaining a sufficient amount of LN2 in the vapour phase. Use of vapour phase avoids the safety hazards inherent in liquid phase storage, including the risk of transmission of infectious agents (Figure 3.9.6).

Mechanical Freezers

Mechanical freezers are employed in a variety of storage temperature ranges, including -20°C, -40°C, -70 to -80°C, and occasionally -140°C. Freezers should be equipped with alarms set at
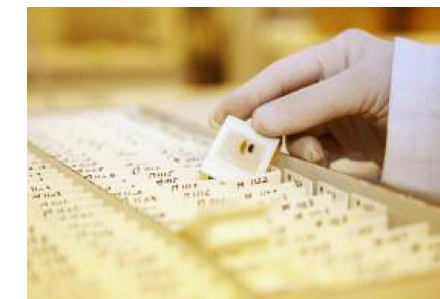
about 10°C warmer than the nominal operating temperature of the unit (Figure 3.9.7).

Dry Ice

Dry ice or solid-phase carbon dioxide is frequently used as a refrigerant for shipping and emergency backup for mechanical freezers. Handling precautions should be employed when handling this material, which exists at a nominal -70°C. As dry ice sublimates, the CO_2 level in the surroundings can increase. In confined areas the carbon dioxide can displace oxygen, presenting an asphyxiation hazard.

Standard Operating Procedures

BRCs should develop, document and regularly update policies and procedures in a standardized written format incorporated into a Standard Operating Procedures (SOP) manual that is

Fig. 3.9.4 *Fixed formalin paraffin embedded tissues*

Fig. 3.9.5 *Histological slides*

Fig. 3.9.6 *Liquid nitrogen room*

Fig. 3.9.7 *Deep freezers room*

Temperature	Properties of water/Liquid Nitrogen	Cryopreservation method	Biological relevance
0°C- +4°C	Ice melting	Fridge	
-0.5°C to -27°C	Ice fusion area	Freezer	
-27°C to -40°C	Ice	Freezer	Limit of protein mobility/ DNA stability
-40°C to -80°C	Limit of water molecules mobility	Freezer	RNA stability
-80°C to -130°C	Ice transition	Freezer/Liquid nitrogen	No metabolic activity Recommended storage for blood and urine
-130°C to -150°C	Liquid nitrogen vapour	Liquid nitrogen	Recommended storage for tissue
-150°C to -196°C	Liquid nitrogen liquid	Liquid nitrogen	Possible micro-fractures Recommended storage for living cells

Table 3.9.1 *Basic standards of crypreservation and applications to biological specimens*

readily available to all laboratory personnel. The SOP manual should specifically include:

- Specimen handling policies and procedures including supplies, methods and equipment
- Laboratory procedures for tests and any aliquoting or other specimen processing
- Policies and procedures for shipping and receiving specimens
- Records management policies
- Quality assurance and quality control policies and procedures for supplies, equipment, instruments, reagents, labels, and processes employed in sample retrieval and processing
- Emergency and safety policies and procedures, including reporting of staff injuries and exposure to potential pathogens
- Policies and procedures for the investigation, documentation and reporting of accidents, errors, complaints and adverse outcomes
- Policies and procedures and schedules for equipment inspection, maintenance, repair and calibration
- Procedures for disposal of medical and other hazardous waste
- Policies and procedures describing requirements of training programs for BRC staff.

BRCs should have an appropriate QA and QC programs regarding equipment maintenance and repair, staff training, data management and recordkeeping, and adherence to Good Laboratory Practice. All BRC operations must be subjected to regular audits. The timing, scope and outcome of these audits should be documented. QA is an integrated system of management activities involving planning, implementation, documentation, assessment, and improvement to ensure that a process or item is of the type and quality needed for the project. QC is the system of technical activities that measures the attributes and performances of a process, or item, against defined standards, to verify that the stated requirements are fully met.

Records Management

BRCs must develop a record management system that permits detailed records to be made concurrently with the performance of each step in the collection, processing and distribution of specimens. This may include but is not limited to: informed consent, procurement, processing, preservation, quarantining, testing, record review, releasing, labelling, storage, distribution and quality control of specimens. Records shall be created and maintained in a manner that allows steps to be clearly traced. Record security systems shall be adequate to ensure confidentiality and safety. Record management should be regularly audited. Records should be kept for at least 10 years after expiration of specimen storage or specimen distribution. Electronic records should be adequately protected (regular back-ups on appropriate media, intrusion-proof management systems).

The BRC should be inventoried at regular intervals (e.g. every two years) to assess the concordance between stored specimens and records. The specific position of every stored aliquot should be tracked. Each freezer, refrigerator or room temperature storage cabinet should have a unique identifier. A convention should be established for numbering shelves, racks, boxes, as well as, each location within the container. The biorepository database should be updated each time a biospecimen is moved within or out of the biorepository.

Specimen labelling

Each specimen should be labelled in such a manner that the labelling will survive all potential storage conditions, in particular dry ice and liquid nitrogen.

- Ink used on the label should be resistant to all common laboratory solvents.
- Labels should be printed with a linear bar-code if possible, thus providing a direct link to database software. However, it is also impor-

tant to include human-readable indications of contents.
- Suggested information for the label is the tissue bank's unique identifier number, sample type and date of collection/banking, plus a bar-code if available (Figures 3.9.8 and 3.9.9).

Specimen collection, processing, storage

The methods used to collect biospecimens will vary depending on how the specimens will be processed and what is intended to be the end use. This paragraph provides general recommendations for collection of blood, solid tissues, urine and wide blood cells. These recommendations are derived from those described in the Biorepository Protocols developed by the Australasian Biospecimen Network. [6]
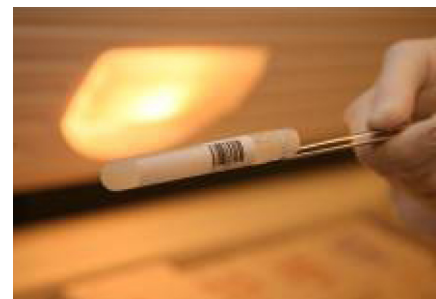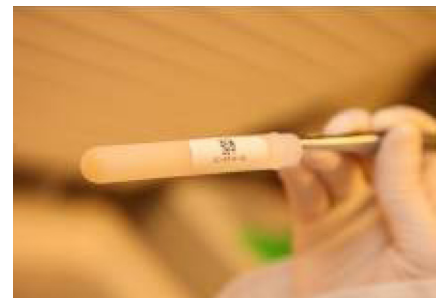

Fig. 3.9.8 Linear bar-code

Fig. 3.9.9 2D Bar-code

Collection of Blood

Detailed instructions and protocols for collection of blood specimens are given in the Protocols section. The following general guidelines should be considered.

- All blood should be treated as potentially infectious. It is recommended to take tissue bank blood samples concurrently with routine clinical blood samples, so as to limit discomfort to patients.
- Blood may be collected into EDTA, ACD (Acid Citrate Dextrose), lithium heparin, or into a clotted tube containing separating gel. Either EDTA and ACD tubes can be used if DNA is to be extracted or lymphocyte cell lines to be made; however, ACD is more appropriate if there is to be an extended time lapse between blood collection and processing. Lithium heparin is generally only used if cytology studies are being performed. If DNA is to be extracted from the blood or lymphocyte cells lines made, collecting into lithium heparin is not recommended [7].
- Tubes should be clearly labelled (Figure 3.9.10).
- The amount of blood usually collected varies for different diseases. In most cases, 2 tubes (18-20 ml) blood is an ideal collection amount. The volume collected is guided by ethics clearance. Reduced volume of blood in a tube containing additives should be noted so as to avoid confounding of results by variation in additive concentration
- Time of bleed and time of freezing should be recorded, as well as any variations to the processing protocol.
- Blood should be transported at room temperature, unless otherwise specified for particular applications (for some proteomic applications require transport on dry ice).
- All blood should be processed within 24 hours of collection. Cell viability decreases rapidly after 24 hours, resulting in poor cell structure in slide preparations, or degradation of proteins and nucleic acids.
- Serum and plasma should be stored within 2 hours.

- Blood spot collection should be considered as alternative to whole blood when protocols call for easier collection and cheap room-temperature storage [8]. Guthrie cards are made from pure cotton and can be used for the extraction of DNA (Figure 3.9.11).

Collection of Solid Tissues

Solid tissues are collected by biopsy or during surgical procedures. Collection should be carefully planned with surgeons and clinical staff, and all materials and instruments should be prepared in advance.

- The collection of samples for research should never compromise the diagnostic integrity of a specimen. Only tissue exceeding diagnostic needs should be banked.
- All tissue should be treated as potentially infectious; the collection process should be carried out in the most aseptic conditions possible.
- The intact operative specimen should be sent as soon as possible to pathology.
- It is recommended to process specimens within 30min of excision [9]. Transfer of specimens must be carried out as quickly as possible in order to minimise the effect of hypoxia upon genetic expression, and degradation of RNA and other tissue constituents. A record of the timing of events from excision to fixation or freezing should be kept.
- Each specimen receptacle must be clearly labelled (Figure 3.9.12).
- For transport from surgery to pathology, or to the repository, specimens should remain fresh (not fixed) and be placed in a closed, sterile container on wet ice.
- A pathologist should supervise the procurement of the tissue. The pathologists will examine the sample and, allowing adequate tissue for histological diagnosis and assessment of margins, will remove a portion of the tumour and adjacent normal tissue for specimen banking. When selecting specimens, areas with massive ischemia and/or necrosis should be avoided.

- The anatomical site from which the tissue is taken must be recorded and documented by a picture.
- Tissue bank staff must be present in pathology to freeze or fix the tissue as quickly as possible. Tissues should be placed in appropriate containers before freezing. Direct contact of the tissue with the liquid nitrogen should be avoided. Samples requiring snap freezing

Fig. 3.9.10 Blood sample identification

Fig. 3.9.11 Blood spots storage

Fig. 3.9.12 Collection of solid tissues

can be frozen in a Dewar of liquid nitrogen or on dry ice at the time of collection.

– When dry ice / liquid nitrogen is not readily available, tissue collections into RNA*later* may be a good alternative provided that this tissue is not required for diagnosis and clearance is given by the pathologist.

Collection of other specimens

Urine. Urine is easy to collect and is suitable source of protein, hormones, metabolites and DNA from exfoliated bladder cells. However, storage of urine specimens is space-consuming. Urine should be stored at -80°C or in liquid nitrogen vapour.

Buccal cells. The collection of buccal cells is not difficult and does not require highly trained staff. Buccal cell collection should therefore be considered when non-invasive, self-administered or mailed collection protocols are required [8]. Donors who do not give blood may also be asked to donate a buccal cell specimen; however, buccal cell collection will yield only limited amounts of DNA in comparison to blood. A collection kit (containing mouthwash, 50 ml plastic tube, plastic biohazard bottle, and courier packaging) may mailed or given to the participant, along with an instruction sheet.

Specimen annotations, data collection

It is recommended that BRC adopt standardised systems for annotating the characteristics of collected specimens as well as data on the patients or subjects who are the source of these specimens. The nature and extent of data collection may vary depending upon the project in which the specimens are collected as well as depending upon on the type of cancer and nature of specimen collected. The paragraphs below provide a brief outline of the structure of minimal annotation datasets.

Annotations on patients/subjects

– Local Patient Case Code

– Tumour topography and morphology according to the International Classification of Disease–Oncology (ICD-O 10) Histopathological (Figure 3.9.13)
– TNM staging
– Tumour grade
– Age at time of specimen collection (in years)
– Gender
– Place of residence (city/region/country)
– Ethnicity
– History of previous cancer disease
– Evidence for familial history of cancer
– Involvement into clinical trial/cohort study
– If appropriate, information on medical history, treatment and response to therapy, concomitant disease, secondary tumours/laboratory data

Annotations on stored specimen

– Local BRC inventory code
– Tissue condition (tumour/non-tumour/interface)
– Preservation protocol
– Time (in min) elapsed between tissue removal and fixation/freezing
– Duration of storage and record of storage incidents
– History of freezing/thawing
– Amount of tissue collected and amount left over in storage

Specimen shipping and sending

Human biospecimens are considered as "dangerous goods", that is, "articles or substances which are capable of posing a risk to health, safety, property of the environment". According to UN regulations, dangerous goods meet the criteria of one or more of nine UN hazard classes (see links to references below). The relevant class for biological specimens is Class 6, division 6.2: Infectious substances.

The shipping and sending of biospecimens is subject to international regulations. These regulations, applicable to any mode of transport are based upon the Recommendations by the Committee of Experts on the Transport of Dangerous Goods (UNCETDG), a committee of the United Nations Economic and Social Council.

Technical Instructions for the Safe Transport of Dangerous Goods by Air published by the International Civil Aviation Organization (ICAO) are the legally binding international regulations. The International Air Transport Association (IATA) publishes Dangerous Goods Regulations (DGR) that incorporates the ICAO provisions and may add further restrictions. The ICAO rules apply on all international flights. For national flights, i.e. flights within one country, national civil aviation authorities apply national legislation. This is normally based on the ICAO provisions, but may incorporate variations. State and operator variations are published in the ICAO Technical Instructions and in the IATA Dangerous Goods Regulations.

The following links refer to these regulations:

– UNECE (United Nations Economic Commission for Europe)
UN Recommendations on the Transport of Dangerous Goods. Model Regulations.
http://www.unece.org/trans/danger/publi/unrec/rev13/13files_e.html

– IATA (International Air Transport Association)
Dangerous Goods Regulations 2005.
http://www.iata.org/ps/publications/9065.htm

– ICAO (International Civil Aviation Organization)
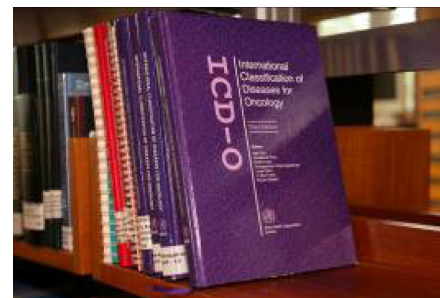*http://www.icao.int/icao/en/m_publications.html*

Fig. 3.9.13 *ICD-O 10 Book*

– WHO (World Health Organization)
Transport of infectious substances 2005
http://www.who.int/csr/resources/publications/biosafety/WHO_CDS_CSR_LYO_2005_22r%20.pdf

When preparing to transport biospecimens, it is important to consider shipping time, distance, climate, season, method of transport, and regulations as well as the type and number of biospecimens to be sent and their intended use. Below are some general guidelines:

Regulations

Infectious substances fall into two categories.

Category A comprises substances which are transported in a form that, when exposure to them occur, are capable of posing permanent disability or life-threatening or fatal disease to humans or animals. Category A specimens include, but are not restricted to, specimens contaminated by highly pathogenic viruses (Ebola, Hantaan, Marburg, Lassa, etc.) or cultures of viruses such as Dengue, Human Immunodeficiency Virus (HIV), Hepatitis B Virus (HBV). The proper shipping name for such substances is UN2814: "Infectious substances affecting humans" or UN2900: "Infectious substances affecting animal only".

Category B comprises substances that do not meet the above criteria. Most human specimens such as blood samples, tissues, exfoliated cells or urine, not contaminated by highly pathogenic viruses, will fall into Category B. The proper shipping name for such substances is UN3373: "Biological Substance, Category B".

Biospecimens or derived products that have been specifically treated to neutralize infectious agents, or for which there is a minimal likelihood that pathogens are present, are not subject to these regulations. The proper shipping name for such substances is "Exempt Human (or Animal) Specimens".

Packaging

The basic triple packaging system applies to all substances. It consists of three layers, as follows:

– *Primary receptacle*: a primary watertight, leak-proof receptacle containing the specimen, packaged with enough absorbent material to absorb all fluids in case of breakage;
– *Secondary packaging*: a second, durable watertight, leak-proof packaging to enclose and protect the primary receptacle. Several primary receptacles may be placed in one secondary packaging but additional absorbent material sufficient to absorb all fluid should be used in case of breakage;
– *Outer packaging*: an outer shipping packaging of suitable cushioning material, protecting the contents from outside influences while in transit (Figures 3.9.14 and 3.9.15).

Use appropriate insulation, e.g. for +8°C to -20°C use gel packs, for -80°C use dry ice, and if samples need to be kept at -150°C, transport them in a dry shipper containing liquid nitrogen. Ensure enough refrigerant is included to allow for a 24-hour delay in shipping.

The triple packaging system also applies to Exempt Human Specimens such as paraffin-embedded samples (that should be shipped at room temperature in insulated packaging to protect from extreme fluctuations in temperature), Guthrie cards (that should be transported in watertight plastic bags) or histopathological slides (that need to be cushioned to prevent breakage). In all cases, desiccants should be used for samples sensitive to humidity.

Labelling

All outer packages must bear United Nations packaging specification marking according to the category in which the specimens fall. For category A, the packaging instruction PI602 applies. For category B, the relevant packaging instruction is PI650. Detailed instructions are described in the IATA "Infectious Substances

and Diagnostic Specimens Shipping Guidelines 2005" (www.iata.org).

When shipping biospecimens overseas, be aware of the receiver country's requirements prior to the initiation of the shipment, and ensure that the consignment adheres to these regulations.

Access to stored materials and data for research purposes

Access to human biological specimens for research purposes is crucial for most fields of

Fig. 3.9.14 *Preparation of the basic triple packaging system*

Fig. 3.9.15 *Preparation of the basic triple packaging system*

cancer research and in particular to genomics, proteomics, metabolomics or molecular imaging. Each BRC should establish clear guidelines for distribution and sharing of biospecimens and data, compatible with local, national and international prevailing laws, ethical principles, and protection of Intellectual Property. However, BRCs should not serve exclusively to satisfy individual needs or research projects and all efforts should be made to make specimens and data available to the wider scientific community. So far little has been done internationally to standardize access to biospecimens. The following paragraph, based on the recommendations developed by the NCI [10], develops general principles to guide the procedure for access to specimens for research purposes.

– Although BRCs have the right to establish priorities for access to specimens, in principle, BRCs should commit themselves to providing equal right of access to researchers.
– A mechanism of rapid peer and/or stakeholder review should be in place to set up priorities as to how collected specimens should be allocated to qualified recipient investigators.
– The proposed research project and use of specimen should be consistent with participants' consent, research purpose, and allowable use of specimens.
– Within the above principles, the main criteria for approving request for access should be the scientific validity of the research proposal, the investigator and institutional research qualifications, the investigator written agreement covering confidentiality, use, disposition, and security of specimens and associated data, the investigator's written agreement in a Material Transfer Agreement covering publication, sharing of research results, and ownership of future intellectual property, the Ethical approval of the proposed research, and the funding level for the project.

Constructing and running a large BRC: The example of the EPIC biobank

There are many types of BRC. Tumour banks, for examples, often are hospital-based collections of cancer tissues which are "leftovers" from diagnostic or surgical procedures. Many collections are also developed in the context of clinical trials: the patients recruited in these trials donate blood or tissue specimens, the analysis of which often enrich the results of the trial by allowing a better understanding of the parameters that determine good or bad response to a treatment. But the largest and the more systematic collections are those associated with large cohort studies developed in molecular epidemiological contexts. Typically, in such cohorts, healthy subjects are recruited, donate specimens at the time of recruitment (for example, blood, urine, saliva, or exfoliated cells from the buccal cavity) and are then followed up for a period of time that can extend over several decades. With time, a proportion of these subjects develop chronic diseases, including cancer (and also diabetes, heart diseases and other common conditions). It then becomes possible to compare subjects who developed a particular disease with those who did not, and to carry molecular studies using the specimens collected at recruitment, to identify biomarkers that predict or explain why these individuals have developed the disease under study.

EPIC (European Prospective Investigation into Cancer) is a typical example of such a large cohort study. It was developed by the IARC as a long-term, multi-centric prospective study in Western Europe to investigate the relationships between nutrition and cancer, taking advantage both of the contrast in cancer rates and dietary habits between centres and countries. As a rule, healthy subjects were invited to participate either by mail or in person. Individuals who agreed to participate signed an informed consent agreement and were mailed a questionnaire on diet and a questionnaire on lifestyle. Most participants completed these questionnaires at home and were then invited to a study centre for an examination that included collection of the completed questionnaires, blood donation, anthropometry and measurement of blood pressure. The enrolment of subjects took place between 1992 and 2000. The cohort participants are now followed over time for the occurrence of cancer and other diseases, as well as for overall mortality. The study has recruited 519 978 participants in 23 centres located in 10 European countries.

Blood was obtained by venipuncture and separated into plasma, serum, white blood cells and erythrocytes. They were collected from 385 747 of the 519 978 EPIC study participants. To make storage easier, blood samples were aliquoted into 28 plastic straws containing 0.5 ml each. The samples were then split into two mirror halves of 14 aliquots each. One set was stored locally, and the other one was transported to IARC to be stored in liquid nitrogen in a central biorepository located at IARC, where the specimens are kept under N2 liquid phase at -196°C. The biobank contains about 3.8 million straws, labelled with the participant's ID and colour-coded to indicate its contents.

The EPIC provides a framework for addressing a wide range of questions relevant to cancer. When biological samples are involved, studies mostly use the nested case–control approach. Typically, cases are subjects who developed a particular pathology after they were recruited in the cohort (incident cases) and had not been diagnosed with cancer before or at the time of recruitment. Controls are usually chosen at random among all cohort members who were alive without cancer at the time of recruitment of the case. The logistical tasks related to specimen management, retrieval from the biobank and distribution are handled by the team of Laboratory Infrastructure and Resources (LIR) at IARC. Based on lists of specimens and on their known position in the biorepository, the LIR technicians develop an ordered retrieval plan that minimizes the time of opening of each LN2 tank. Specimen retrieval is performed manually. It takes about 5 minutes to access one specific storage position and to retrieve either one or several straws of materials from the same subject. Standard operating procedures include double checking of 10% of all retrieved specimens to minimise the risk of individual error. On average, a trained technician can retrieve specimens for about 150 subjects over one normal working day. Specimen retrieval is a limiting factor in the pre-analytical processing of EPIC biospecimens and its demand in terms of manpower entails important costs, in particular for studies in which several thousand specimens are included. The LIR team offers a range of biobanking services including automated DNA extraction, quantification, aliquoting in various tube or microplate format, and specimen shipping. Currently, the EPIC biobank is providing support for research by over 250 scientists in Europe and beyond, and has provided the basis to several hundred international scientific publications. Constructing and maintaining such a biobank is a major effort that can be estimated, overall, to over 10 million US dollars.

The future of biobanking

The EPIC example shows the importance of developing large biobanks by networking the efforts of scientists in different countries. There is indeed a huge benefit in networking. Cancer diseases are very diverse and have complex relationships with both the genetic makeup of individuals and their lifestyles, so comparisons across different countries, ethnic groups and cultural backgrounds are extremely informative. Thus, tomorrow's BRC will be made of networks and hubs, interconnecting many collection centres and making it possible to access large series of specimens for research. By networking, it is possible to share the burden of investing into large biobanks as well as the benefits of research. Today's cancer research is a vast, collective endeavour, in which scientists and doctors have to team up in powerful networks, capable of delivering the best of human's mind creativity to the bedside of cancer patients.

REFERENCES

1. OECD Directorate for Science, Technology and Industry (2007). *OECD Best Practice Guidelines for Biological Resource Centres*. Available from: http://www.oecd.org/dataoecd/7/13/38777417.pdf Date Accessed: November 15, 2008.

2. 18th World Medical Assembly (1964). Declaration of Helsinki. Recommendations guiding physicians in biomedical research involving human subjects. Helsinki: Finland.

3. National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research (1979). *The Belmont Report. Ethical principles and guidelines for the protection of human subjects of research*. Bethesda, MD: US Department of Health, Education, and Welfare.

4. Grizzle WE and Fredenburgh J (2001). Avoiding biohazards in medical, veterinary and research laboratories. *Biotech Histochem* 76: 183-206.

5. International Society for Biological and Environmental Repositories (2008). *ISBER Best Practices For Repositories. Collection, storage, retrieval and distribution of biological materials for research*. Available from: http://www.isber.org/Pubs/BestPractices2008.pdf Date Accessed: November 15,2008.

6. Australian Biospecimen Network (2007). *Biorepository Protocols. Revision 4*. East Melbourne, Australia: ABN.

7. Yokota M, Tatsumi N, Nathalang O, et al. (1999). Effects of heparin on polymerase chain reaction for blood white cells. *J Clin Lab Anal* 13: 133-140.

8. Steinberg K, Beck J, Nickerson D, et al. (2002). DNA banking for epidemiologic studies: a review of current practices. *Epidemiology* 13: 246-254.

9. Eiseman E, Bloom G, Brower J, et al. (2003). *Case studies of existing human tissue repositories : "Best practices" for a biospecimen resource for the genomic and proteomic era*. Santa Monica, CA: Rand Corporation.

10. National Cancer Institute. Office of Biorepositories and Biospecimen Research (2007). *NCI Best Practices for Biospecimen Resources*. Available from: http://biospecimens.cancer.gov/practices/ Date Accessed: November 15, 2008.

CANCER IN THE EASTERN MEDITERRANEAN AND NORTH AFRICAN REGION (EMRO)

According to WHO 2002 mortality estimates, cancer is the fourth-ranked cause of death in the Eastern Mediterranean and North African Region (EM Region), after cardiovascular diseases, infectious/parasitic diseases and injuries. It is estimated that cancer kills 272 000 people each year in the EM Region, more than HIV/AIDS, tuberculosis and malaria combined (241 000 deaths per year). Although cancer incidence in the EM region is still much lower than in other parts of the world, the largest increase in cancer incidence among the WHO regions in the next 15 years is likely to be in the EM region, in which projection modelling predicts an increase of between 100% and 180% [Rastogi *et al.* 2004]. At present, resources for cancer control in the EM region as a whole are not only inadequate but directed almost exclusively to treatment. However, the impact of preventive measures on incidence is not fully exploited, while the lack of approaches to earlier diagnosis reduce the value of therapy; the curability of cancer is directly related to its stage at the time of diagnosis, and in the majority of EM countries, cancer is generally diagnosed when at a relatively advanced stage (Table 1).

In response to the above situation, WHO/EMRO has developed a regional strategy for the prevention and control of cancer in its Member States, a draft of which was presented and discussed in a consultative meeting in Marrakesh, Morocco, November 2007 and will be finalised and formally launched in a meeting planned in April 2008. A Regional Alliance Against Cancer,

bringing together various NGOs and interested parties working in EM Member States, was formally created during the Marrakesh meeting under the leadership of WHO and in collaboration with the Lalla Salma Association Against Cancer and H.R.H. Princess Lalla Salma, Patroness of Prevention and Care for cancer in the Eastern Mediterranean Region.

The strategy lays a foundation for the development of a coordinated approach that seeks to take advantage of the strengths of some of the regional resources to overcome some of the weaknesses that exist in the Region. An important function of the strategy resides in its twin goals of sensitising authorities in Member States to the pressing need to control cancer more effectively, while at the same time providing technical guidance and a foundational formula for regional cooperation in this endeavour. The strategy encourages countries to develop their National Cancer Control Programmes (NCCP), an essential first step towards more effective cancer control.

The EM regional strategy is in keeping with the WHO Global Action Plan against Cancer (GAPAC) and pursues the same goals, which are to:

– Prevent Preventable Cancers (through avoiding or reducing exposure to risk factors, i.e. prevention strategies);
– Cure Curable Cancers (early detection, diagnostic and treatment strategies);
– Relieve Pain and Improve Quality of Life (Palliative care strategies); and

– Manage for Success (strengthening health care systems; management, monitoring and evaluation of interventions).

WHO/EMRO continues to assist countries to develop their NCCPs, and in 2007 participated in two missions to Yemen and Syria in collaboration with IAEA's PACT Programme.

The pattern of cancer in EM Region is shown in Table 2. Data are obtained from the GLOBOCAN database and updated, for many countries, directly by national focal points based on latest information from their cancer registries. Breast cancer has the highest incidence rate in most countries, while cervical cancer is the leading type of cancer in Djibouti and Somalia.

website: www.emro.who.int

	Breast Cancer			Cervical Cancer		
Stage	USA	Saudi Arabia	Egypt	USA	Saudi Arabia	Egypt
Localised	65%	29%	25.5%	58%	35%	35.9%
Regional	30%	55%	58%	33%	51%	53.2%
Distant	5%	16%	16.5%	9%	14%	10.9%

Table 1. *Stage at diagnosis in breast and cervical cancer as reported by a population based registry in Saudi Arabia [National cancer registry report, 2002], Tanta Cancer registry (Gharbiah 2000-2002, Egypt), and US [SEER, 9 Registries 1988-2003].*

Country	1st cancer	2nd cancer	3rd cancer	4th cancer	5th cancer
Afghanistan	Breast	Stomach	Esophagus	Lung	Oral Cavity
Bahrain	Breast	Lung	Colon	Bladder	Leukaemia
Djibouti	Cervix	Liver	Breast	Esophagus	Kaposi
Egypt	Breast	bladder	NHL	Liver	Lung
Iran	Breast	Stomach	Colon	Bladder	Esophagus
Iraq	Breast	Leukaemia	Lung	Brain and CNS; Larynx	Bladder
Jordan	Breast	Colon	Lung	Bladder	NHL
Kuwait	Breast	Lung	Colon	NHL	Leukaemia
Lebanon	Breast	Lung	Bladder	Cervix	Larynx
Libya	Breast	Lung	Colon	Head & neck; Cervix	Bladder
Morocco	Breast	Lung	Cervix	Prostate	lymphoma
Oman	Stomach	Breast	Lung	NHL	Liver
Pakistan	Breast	Oral Cavity	Lung	Esophagus	Bladder
Qatar	Lung	Breast	Colon	Bladder	Liver
Saudi Arabia	Breast	NHL	Liver	Colon	Thyroid
Somalia	Cervix	Liver	Esophagus	Breast	NHL
Sudan	Breast	Cervix	Oral Cavity	Esophagus	Colon
Syria	Breast	lung	NHL	CNS	Bladder
Tunisia	Lung	Breast	Bladder	Colon	NHL
UAE	Breast	Colon	Blood Leukaemia	Lymphomas	Thyroid
Yemen	Breast	NHL	Colon	NHL	Esophagus

Table 2. *Commonest 5 cancers in EM countries*