The biology of cancer

Kevin J Harrington

Abstract

Cancer is a genetic disease. Most common cancers are caused by acquired mutations in somatic cells. In contrast, specific germline mutations account for rare hereditary cancer syndromes. In general, cancer-associated genes can be divided into two groups: oncogenes and tumour suppressor genes (TSGs). Oncogenes undergo activation and are phenotypically dominant, while TSGs undergo inactivation and are phenotypically recessive. Oncogenic activation can occur by specific point mutations within the gene sequence, amplification of the number of copies of the gene or translocation of DNA to sites where transcription is more active or where a new fusion gene is formed that encodes a protein with enhanced biological activity. TSGs are inactivated by mutations that destroy the function of the protein encoded by the gene, or by silencing of the gene's promoter. The biological behaviour of cancer can be considered in terms of eight specific hallmarks and two additional so-called enabling characteristics. Improved understanding of the mechanistic basis of these processes has revolutionized diagnosis, treatment and prognostication in cancer medicine.

Keywords Angiogenesis; apoptosis; cancer; growth factor; hallmarks; immune evasion; invasion; metastasis; oncogene; telomerase; tumour suppressor gene; Warburg effect

Introduction

Cancer is a genetic disease that occurs when the information in cellular DNA is corrupted, leading to abnormal patterns of gene expression. Consequently, the effects of normal genes that control normal cellular functions, such as growth, survival and invasion/motility, are enhanced and those of genes that suppress these effects are repressed. The main mechanism of alteration is through the accumulation of mutations, although non-mutational (epigenetic) changes are increasingly being seen as central to the process. Aberrant gene expression causes fundamental changes to biological processes within cancer cells – the so-called hallmarks and enabling characteristics of cancer.[1,2] (See also *Genetic Predisposition to Cancer* on pages 65–68 of this issue.)

Structure and function of DNA

The genetic code exists in the form of deoxyribonucleic acid (DNA) that is packaged in chromosomes in the cellular nucleus. DNA molecules consist of a sugar–phosphate backbone (deoxyribose sugars linked by phosphate groups) with each sugar bearing one of four nucleotide bases (the purines adenine (A) and guanine (G), and the pyrimidines thymine (T) and cytosine (C)). Two DNA strands spiral around one another to form a double helix with the bases forming hydrogen bonds with specific partners in the opposite strand: normally A only pairs with T (through two hydrogen bonds), and C only pairs with G (through three hydrogen bonds). The helix is wound around nucleosomes, consisting of histone proteins, and is further condensed to form chromatin. Chromatin is loosely packaged (euchromatin) in regions of active gene transcription (see below), and tightly packed (heterochromatin) in transcriptionally silent areas of the genome. The entire genetic code comprises approximately 3.2×10^9 bases and contains approximately 20,000–25,000 genes, which account for only 1% of the genome. The DNA sequence within a gene comprises both exons and introns. The exons represent the protein-coding regions that are translated at the ribosome, whereas the introns are non-encoding and are edited out (see below).

The function of genes is to make proteins. This process occurs in two steps. First, the DNA code is copied into messenger ribonucleic acid (mRNA) during the process of transcription. RNA differs from DNA in two ways: (1) the sugar backbone contains ribose (not deoxyribose); and (2) thymine is replaced by uracil (U). The process by which a gene is transcribed into mRNA and then translated into a cellular protein is complex and subject to multiple levels of control. Regulation of transcription is the key initiating event and is mediated by interactions between enhancer/promoter elements in the DNA and specific proteins (>100 individual subunits) that bind to them. At the transcription start site, DNA-dependent RNA polymerase II is recruited and begins to synthesize mRNA complementary to the coding strand of DNA, such that the DNA sequence CGTA becomes GCAU in the mRNA (as there is no T in RNA). Post-transcriptional modifications of mRNA include editing out (splicing) of introns and processing to make the mRNA suitable for export from the nucleus. Alternate splicing of transcripts of approximately 70% of human genes allows different proteins to be produced from the same gene. On average, each gene produces about four alternatively spliced mRNAs and this means that the human genome encodes about 100,000 different proteins. The mRNA can be degraded by small non-coding RNA molecules called microRNAs (miRNAs) before translation occurs. Specific overexpression of miRNAs has been identified in a variety of different cancers. If the mRNA reaches the ribosome, its three-base (triplet) code of bases is translated into specific protein molecules. It is these protein products of genes that mediate the phenotypic changes that we recognize as cancer.

Cancer genes

Cancer is driven by two classes of genes – oncogenes and tumour suppressor genes (TSGs) – each of which provides an essential function in normal cells.

Oncogenes are derived from mutated versions of normal cellular genes (called proto-oncogenes) that control cell proliferation, survival and invasion/motility. In normal cells, expression of proto-oncogenes is very carefully regulated to avoid uncontrolled cell growth. In cancer, activating mutations of proto-oncogenes cause uncontrolled cell division, enhanced survival (even following anti-cancer treatment) and dissemination. Oncogenes are described as phenotypically dominant – meaning that a single mutated copy of a proto-oncogene is sufficient to promote cancer – and are generally not associated with

Kevin J Harrington *PhD FRCP FRCR is Professor in Biological Cancer Therapies at The Institute of Cancer Research, Royal Marsden NIHR Biomedical Research Centre, Targeted Therapy Laboratory, Division of Cancer Biology, London, UK. Competing interest: none declared.*

inherited cancer syndromes. Exceptions to this rule include mutations in the *RET* proto-oncogene that are associated with multiple endocrine neoplasia syndromes (types 2A and 2B) and germline mutations in *H-RAS* that can cause Costello's syndrome (high birth weight, cardiomyopathy, predisposition to cancers). Oncogenes are activated in three ways to cause cancers (Figure 1): (1) gene mutations alter the sequence of a gene to give it enhanced biological function (e.g. *RAS* in pancreatic, lung and colorectal cancers); (2) gene amplification occurs when a gene retains its normal sequence but is multiply repeated in the chromosome (e.g. *N-MYC* in neuroblastoma); (3) gene translocation involves movement of a genetic sequence from its normal chromosomal position (locus) to a new position (usually on a different chromosome) where it is controlled by a new, more active promoter element or generates a novel, fusion protein with enhanced biological activity (e.g. *BCR-ABL* in chronic myeloid leukaemia).

TSGs are normal cellular genes that function to inhibit cell proliferation and survival. They are frequently involved in controlling cell cycle progression and programmed cell death/apoptosis. TSGs are phenotypically recessive — meaning that both copies must be functionally altered to promote cancer — and are responsible for inherited cancer syndromes (see Tables 1–3 in *Genetic Predisposition to Cancer* on pages 65–68 of this issue). In familial cancer syndromes, individuals inherit a germline mutation in one copy (allele) of a TSG such that every cell in the body is affected. It is therefore highly likely that at least one cell in the body will suffer complete loss of TSG function since only one copy has to be altered (so-called loss of heterozygosity). As a result, hereditary cancer syndromes often give rise to multiple cancers early in life.

In recent years, The Cancer Genome Atlas (http://cancergenome.nih.gov/) programme has catalogued the specific genetic alterations in an increasing number of tumour types. This information is providing new insights into the events that drive particular cancers and is shaping research into new anti-cancer treatments. In the coming years, it is likely that detailed genetic profiles will be available for cohorts of patients with all common and many rare tumour types.

Hallmarks of cancer and enabling characteristics

In 2000, Hanahan and Weinberg described six fundamental changes in cancers (growth factor independence, evading growth suppressors, avoiding cell death, maintaining replicative potential, angiogenesis, invasion/metastasis) that can largely explain their malignant behaviour.[1] Their description was subsequently updated in 2011 through the addition of two emerging hallmarks (reprogramming energy metabolism, evading immune destruction) and two enabling characteristics (genomic instability, inflammation) (Table 1).[2] The role played by each of these processes will be reviewed below.

Growth factor independence

A general scheme for the function of growth factor receptors and their ligands in promoting cell growth (and other effects) is shown in Figure 2. Binding of the cognate ligand to a specific ligand-binding domain on the extracellular component of the receptor leads to a signal being passed from the membrane to the nucleus via a cascade of intermediary messengers, such that ligand binding on the cell surface alters the pattern of gene expression in the cell. Normally, activation of growth factor receptors is very tightly regulated — as is the synthesis and release of the ligands that stimulate them. Cancer cells almost ubiquitously deregulate normal growth factor signalling pathways and use them to promote unrestrained cell division.[3]

Cancer cells achieve self-sufficiency in growth factors by three main mechanisms: (1) they make and release growth factors to stimulate their own receptors (autocrine signalling) and those of their neighbours (paracrine signalling); (2) they alter the

Oncogenes can be activated via three pathways

a Specific mutations in the sequence of the DNA code lead to alterations in the amino acid sequence of the protein, giving it enhanced activity. **b** Increased numbers of normal copies of the gene (amplification) result in the formation of increased amounts of normal protein; **c** Translocation of part of DNA from one chromosomal location to another can result in the generation of a fusion protein with enhanced biological activity.

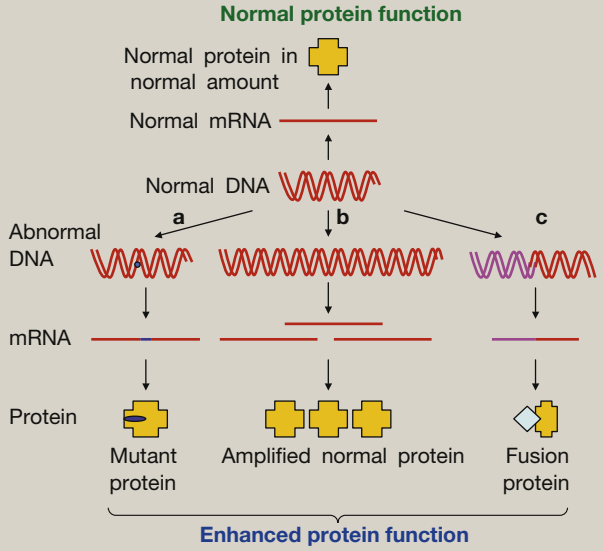

Figure 1

The hallmarks of cancer

- Growth factor independence or self-sufficiency
- Insensitivity to anti-growth signals
- Avoidance of programmed cell death
- Ability to recruit a dedicated blood supply
- Immortalization by reactivation of telomerase
- Ability to invade adjacent normal tissues and metastasize to distant sites
- Reprogrammed energy metabolism
- Evading immune destruction

Enabling characteristics of cancer

- Genomic instability
- Inflammation

Table 1