BASIC CONCEPTS OF CANCER: GENOMIC DETERMINATION

Edith Oláh

Corresponding author's address:
Prof. Edith Olah, Ph.D., D.Sc.
Department of Molecular Genetics,
National Institute of Oncology,
H-1525 Budapest,
Rath Gyorgy u /,
HUNGARY

In the past two decades, cancer researchers have generated a rich and complex body of knowledge, revealing cancer to be a disease involving dynamic changes in the genome. These genomic changes are associated with genetic events that usurp the physiologic function of a normal cell.

Historically, many theoretical models have received temporary favour in efforts empirically to address the problem of cancer etiology, including those founded upon the action of environmental agents, chemical carcinogens, viruses, somatic chromosomal abnormalities, and congenital predisposition. We now know that all of these paradigms are in fact correct by virtue of their convergence into the *genetic paradigm: cancer is the result of an accumulation of mutations in genes that govern the tumour phenotype.*

There is unlikely to exist, now or ever, a more robust biological paradigm than the genetic basis of human cancer development. The genetic foundation of carcinogenesis was implied by some of the earliest practitioners of cancer cell biology and cytogenetics. In the mid-nineteenth century, *Rudolph Virchow* recognized that metastatic cancer cells resemble those of the primary tumour and that all cells of a tumour may arise from a single progenitor cell. Thus, the neoplastic phenotype is heritable from one tumour cell generation to the next. In the early 1900s, *Theodor Boveri* extended this concept to the cytogenetic level, suggesting that gains and losses of specific chromosomes from abnormal segregation might lead to abnormal cell division and other aspects of the cancer phenotype. However, it was not until the discovery of structure of DNA by *James Watson* and *Francis Crick* and the elucidation of the genetic code, that it was possible to begin defining the molecular basis of tumour-genesis in terms of specific mutations in specific genes. In the last quarter of the 20th century since *Bishop* and *Varmus* described the first vertebrate oncogene, the genetic paradigm has been defined in sufficient detail and became generally accepted. Development of recombinant DNA technologies, and establishment of increasingly automatable methods for DNA sequencing set the stage for the *Human Genome Project* to begin in 1990. The completion of a high-quality, comprehensive sequence of the human genome by the fiftieth anniversary of the discovery of the structure of DNA has been a landmark event. The genomic era is now a reality, and allows an unprecedented optimism regarding our understanding of cancer and thus our ability to diagnose it, provide more accurate prognoses, and ultimately, to treat it more effectively.

Over the past decade, our knowledge of the human genome in malignancies has increased enormously. **Genomics** *(the study of the human genome)* and *proteomics (the analysis of the protein complement of the genome)* play a major role in the understanding, diagnosis and potentially also in the treatment of cancer.

It is impossible to summarize the vast literature on cancer molecular genetics and genomics in one chapter. Since genetic mutations are the central aetiologic factor in tumour-genesis, a chapter such as this must include the basic principles of cancer molecular genetics, including evidence for the multistep, multigenic basis of tumour-genesis, and a summary of our current state of knowledge regarding the genes involved in this process. Molecular carcinogenesis is intimately linked to perturbations in cell proliferation and cell death; therefore an overview of the enormous progress recently made in this area will also be presented. The molecular genetics of specific cancer types and hereditary syndromes, and clinical applications will also be given.

1.1 Principles of cancer molecular genetics

All cancers are genetic in origin, in the sense that the driving force of tumour development is genetic *mutation*. A given tumour may arise through the accumulation of mutations that are exclusively somatic in origin, or through the inheritance of a mutation(s) through the germline, followed by the acquisition of additional somatic mutations. These two genetic scenarios distinguish what are colloquially referred to as sporadic and hereditary cancers, respectively. While the neoplastic phenotype is partially derived from epigenetic alterations in gene expression, the sequential mutation of cancer related genes, with their subsequent selection and accumulation in a clonal population of cells, are the determinant factors in regard to whether a tumour develops and the time required for its development and progression. The data to support this multistep, multigenic paradigm are extensive, but perhaps the most compelling evidence is that the age-specific incidence rates for most human epithelial tumours increase at roughly the fourth to eighth power of elapsed time, suggesting that a series of four to eight genetic alterations are rate-limiting for cancer development.

A central aim of cancer research has been to identify the mutated genes that are causally implicated in oncogenesis ('cancer genes'). At the present state of the Cancer Genome Project, close to 300

'cancer genes' are known indicating that mutations in more than 1% of genes of the genome contribute to human cancer.

Genetic alterations in cancer cells have thus far been described in two *major families of 'cancer genes': oncogenes and tumour-suppressor genes*. Proteins encoded by oncogenes may generally be viewed as stimulatory and those encoded by tumour-suppressor genes as inhibitory to the neoplastic phenotype; mutational activation of proto-oncogenes to oncogenes and mutational inactivation of tumour-suppressor genes must both occur for cancer development to take place. Proto-oncogene mutations are nearly always somatic; three known exceptions involve the RET, MET, and CDK4 proto-oncogenes, mutations of which may be inherited through the germline, predisposing to multiple endocrine neoplasia type 2, and papillary renal carcinoma, and hereditary melanoma, respectively. Tumour-suppressor gene mutations may be inherited or acquired somatically. Other than the above-noted exceptions, all hereditary cancer syndromes for which predisposing genes have been identified are linked to tumour-suppressor genes.

1.1.1 Oncogenes

Oncogenes result from gain-of-function mutations in their normal cellular counterpart proto-oncogenes, the normal function of which is to drive cell proliferation in the appropriate contexts. Activated oncogenes behave in a dominant fashion at the cellular level, that is, cell proliferation or development of the neoplastic phenotype is stimulated following the mutation of only one allele. This class of genes was originally discovered through studies of the mechanism of retroviral tumour-genesis, which involves viral transduction of the vertebrate proto-oncogene and re-integration into the host genome under the transcriptional control of viral promoters, such that expression is constitutive and thus oncogenic. The most common mechanisms for mutational activation of human proto-oncogenes are gene amplification, typically resulting in overexpression of an otherwise normal protein product, point mutation, generally leading to constitutive activation of a mutant form of the protein product, and chromosomal translocation, which usually results in juxtaposition of the oncogene with the promoter region of a constitutively expressed gene, thus resulting in over-expression of the oncogene-encoded protein. This latter mechanism is most common in haematopoietic malignancies while the first two are more common in solid cancers. The oncogenes most relevant to human solid malignancies, their mechanism of activation, biochemical function, and the tumour types most often affected by each are summarized in Table 1.

1.1.2 Tumour-suppressor genes

The protein products of tumour-suppressor genes normally function to inhibit cell proliferation and are inactivated through loss-of-function mutations. Knudson's two-hit model established the paradigm for tumour-suppressor gene recessivity at the cellular level, wherein both alleles must typically be inactivated in order for a phenotypic effect to be observed. The most common mutations observed in tumour-suppressor genes are point mutations, missense or nonsense, microdeletions or insertions of one or several nucleotides causing frameshifts, large deletions, and rarely, translocations. A mutation in one allele, whether germline or somatic, is then revealed following somatic inactivation of the homologous wild-type allele. In theory, the same spectrum of mutational events could contribute to inactivation of the second allele, but that is typically observed in tumours is homozygosity or hemizygosity for the first mutation, indicating loss of the wild-type

allele. As originally demonstrated for the retinoblastoma susceptibility gene, loss of the second allele may occur through mitotic non-disjunction or recombination mechanisms, or large deletions. This so-called loss of heterozygosity (LOH) has become recognized as the hallmark of tumour-suppressor gene inactivation at particular genomic loci. Table 2 summarizes the known tumour-suppressor genes, their chromosomal locations, suspected biochemical functions, and the hereditary and sporadic tumours with which they are most commonly associated.

The heterogeneous group of tumour-suppressor genes has been subclassified into gatekeepers, and caretakers. Biallelic inactivation of *gatekeepers* or "classical" tumour suppressors, such as RB1, TP53, APC, or VHL, is rate-limiting for cancer development and is usually tissue-specific. Loss of a *caretaker* gene function is not essential for cancer development, but accelerates the course of other events in the pathogenesis. Caretakers are thus indirect suppressors. The genes involved in various DNA repair systems and cell cycle control mechanisms belong to this group.

1.2 Cancer pathways - genotype to phenotype

A human cancer represents the endpoint of a long and complex process involving multiple changes in genotype and phenotype. Human solid tumours are monoclonal in nature; every cell in a given malignancy may be shown to have arisen from a single progenitor cell. As proposed by *Nowell*, the process through which a cell and its offspring sustain and accumulate multiple mutations, with the stepwise selection of variant sublines, is known as *clonal evolution* or clonal expansion. A long-term goal in studying the molecular genetics of a particular tumour type is to catalogue the specific genes that are affected by mutations, the relative order in which they are affected (in any), and ultimately, to use this molecular blueprint to improve methods of diagnosis, prognostication, and treatment. This task will undoubtedly prove difficult, however, as a defining characteristic of cancer is genetic instability. There are multiple types of such instability, operative at both the chromosomal and molecular levels. Distinguishing the genetic mutations that are simply the byproduct of genetic instability from those are critical to the neoplastic phenotype or, indeed, responsible for increasing genetic instability of one form or another is among the greatest challenges to be faced in cancer research.

The first progress in this context has clearly been achieved for colorectal cancer, and a model has been proposed that applies molecular detail for this particular cancer type to the general paradigm of *multistep tumour-genesis* and clonal evolution (Figure 1). In addition the recent demonstration that most colon cancer cell lines are affected by one of two types of genetic instability, specific molecular genetic alterations have been shown to occur at discreet stages of neoplastic progression in the colon – for example, mutation of the APC tumour-suppressor gene at a very early stage of hyperproliferation, mutation of the K-RAS oncogene in the progression of early to intermediate adenoma, and mutation of the TP53 tumour-suppressor gene in the progression of late adenoma to carcinoma. Mutations in SRC oncogene are found in advanced colorectal cancers sending liver metastasis. The model is limited in applicability to other cancer types, however, as nonmalignant precursor lesions for many solid tumour types are not readily

detectable, and few molecular genetic changes have been described that occur in major fractions of other cancer types.
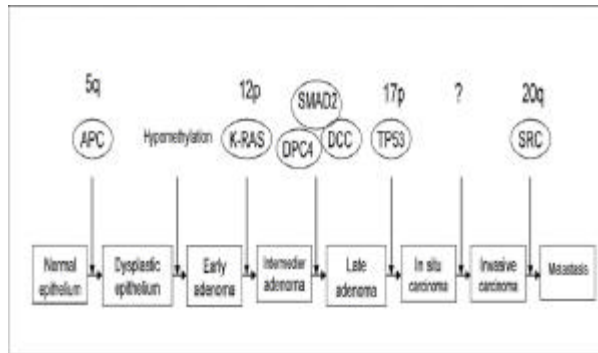

Figure 1. The multistep model of colorectal carcinogenesis ("vogelogram"). (After Fearon and Vogelstein 1990)

Although cancer cells possess many abnormal properties, deregulation of the normal constraints on cell proliferation lies at the heart of malignant transformation. A tumour may increase in size through any one of three mechanisms involving alterations pertaining to the cell cycle: shortening of the time of transit of cells through the cycle, a decrease in the rate of cell death, or the re-entry of quiescent cells into the cycle. In most human cancers, all three mechanisms appear to be important in regulating tumour growth rate, a critical parameter in determining the biological aggressiveness of a tumour.

Hanahan and Weinberg have identified six 'hallmark' features that characterize malignant cells: self-sufficiency in growth signals; insensitivity to growth-inhibitory (antigrowth) signals; evasion of programmed cell death (apoptosis); limitless replication potential; sustained angiogenesis; and tissue invasion and metastasis.

Basically, cancer is a disease of genes that control the proliferation, differentiation, and death of our cells.

Cancer development is driven by the accumulation of DNA changes in about 300 of the approximately 30,000 human chromosomal genes. The genes are the code for the actual players in the cellular processes, the 100,000 – 10 million proteins, which in (pre) malignant cells can also be altered in a variety of ways.

1.2.1 Cell cycle and apoptosis

Homeostasis within a cell population or tissue is a balanced state between cell proliferation and cell death. If this balance is disturbed and the rate of cell proliferation exceeds that of cell death, growing tissue proceeds to form what slowly develops into a tumour.

The fundament of the **cell cycle** is to ensure that the genetic material (i.e. the DNA packed in chromosomes) is faithfully replicated and passed to the next generation of cells. The cell cycle consists of four coordinated phases. Chromosomes replicate during the S(ynthesis) phase. During M(itosis), cellular microtubules form a spindle structure that separates the replicated chromosomes and the nucleus divides. Two G(ap) phases (G1 and G2) intervene between the S and M phases, allowing time for cellular growth and differentiation. Cells in different tissues cycle at different rates: the epithelial cells of the colon or endometrium are in a constantly dividing state, whereas liver cells or fibroblasts are

in a non-dividing state (G0), but can re-enter the G1 phase, for example in response to tissue damage.

If the cellular genome is somehow injured, the cycling cell can be arrested at G1, S, G2 or M checkpoints to allow repair. DNA damage caused by X-rays, oxygen radicals, alkylating agents, UV light, polycyclic aromatic hydrocarbons, anti-tumour agents, spontaneous chemical reactivity or replication errors are sensed and corrected by various nuclear repair mechanism. If the DNA repair mechanisms fail, the cell may be triggered to *apoptosis*, a controlled process of cell death during which the cell shrinks, the nucleus is condensed and cellular DNA is autodigested. Apoptosis does not induce inflammatory response. Cancer cells, and most apparently metastatic cells, do not response adequately to these physiological tissue-specific stimuli, because their ability to undergo apoptosis is lost.

1.3 Inherited predisposition to cancer

The vast majority of mutations in cancer are somatic and are found only in an individual's cancer cells. However, about 1-2 % of all cancers arise in individuals with an unmistakable hereditary cancer syndrome. These individuals carry a particular germline mutation in every cell of their body.

The cardinal feature by which inherited predisposition is recognized clinically is family history. Cancer is common, so families may contain several cases by chance.

It is useful to distinguish between the terms familial and hereditary. The term familial applies to any situation in which family members, either closely or most distantly related to the proband, are also affected – 'cancer runs in the family'. Often the genetic, environmental, or both mechanisms are unknown. Hereditary cancer refers to a situation in which the susceptibility is inherited in a Mendelian manner, suggesting a high-penetrance gene. As a rule, in the rare cancers, the proportion of hereditary cases among all cases is high (e.g. 40% for retinoblastoma) and in the common cancers such as breast- and colon cancer, it is low (usually < 5-7 %).

Identifying germline mutations in high-penetrance cancer predisposition genes performs molecular genetic diagnosis for hereditary cancer. Once a mutation has been identified, the presence or absence of the same mutation can be determined in other members of the family. Genetic testing for cancer susceptibility is used increasingly in cancers in which the results provide information promoting early detection or prevention, or both. This is the case in five childhood or early-onset, relatively rare cancers in which the gene has been isolated and in which many or most cases are caused by mutations in one gene. Among the common cancers, only hereditary breast cancer, breast-ovarian cancer, hereditary nonpolyposis colorectal cancer (HNPCC), and melanoma lend themselves to molecular diagnosis, and only less that 5-10% of all of these cancers presently are caused by germline mutations in known genes. A clinical benefit is clear-cut in breast cancer and in HNPCC and less obvious in melanoma. As false negative, false-positive, erroneous, and uninterpretable molecular findings occur, and because the penetrance of the mutations is

highly variable, it is imperative that genetic testing be performed only in the context of appropriate genetic counseling.

1.4 Concluding remarks

The excitement of genetics, and the perceived medical importance of the human genome sequence, are pegged to the promise of an understanding of cancer as a disease and utilize this new knowledge in the clinical practice.

Molecular genetics and genomics have improved our ability to study genes, proteins and pathways involved in disease and have provided the technology necessary to generate new sets of targets for small-molecule drug design. It has also enabled the creation and production of a new range of biological therapeutics-recombinant proteins, and therapeutic antibodies, which are one of the fastest growing classes of new treatments.

We are undergoing a *revolution in clinical practice* that depends upon a better understanding of disease mechanisms and pathways at a molecular level. Much has already been achieved: an enhanced understanding of cancer-related pathways, new therapies, novel approaches to diagnostics and new tools for identifying those at risk. But more remains to be done before the full impact of genetics on oncology is realized.

References

1. Hanahan D, Weinberg RA: The Hallmarks of cancer. Cell 100: 57-70 (2000)

2. Futreal PA, Coin L, Marshall M, Down T, Hubbard T, Wooster R, Rahman N, Stratton MR: A census of human cancer genes. Nature Reviews Cancer 4: 177-183 (2004)

3. Ponder BAJ: Cancer genetics. Nature 411: 336-341 (2001)

4. Collins FS, Green D, Guttmacher AE, Guyer MS on behalf of the US National Human Genome Research Institute: A vision for the future of genomics research – A blueprint for the genomic era. Nature 422: 835-847 (2003)

5. De la Chapelle A: Genetic predisposition to colorectal cancer. Nature Reviews Cancer 4: 769-780 (2004)

6. Kinzler KW, Vogelstein B: Gatekeepers and caretakers. Nature 386:761-763 (1997)

Table 1. Representative oncogenes mutated in human tumours

Oncogene	Chromosomal location	Mutation	Function	Tumours
ABL	9q34.1	translocation	signal transduction (tyrosine kinase)	CML, other leukaemia
BCL-2	18q21.3	translocation	anti-apoptosis	B-cell lymphoma
CDK4*	12q14	amplification, point mutation	cell cycle-regulation	sarcomas, familial melanoma
cyclin D1	11q13	amplification, translocation	cell cycle-regulation	breast- and other carcinomas, B-cell lymphoma
C-MYC	8q24	amplification, translocation	transcription factor	lymphomas, carcinomas
ERB B		amplification	growth factor (EGF) receptor	glioma, squamous cell carcinoma, carcinomas
FOS	14q24.3	amplification	AP-1 transcription factor	osteosarcoma
Ha-RAS	11q15.5	point mutation	signal transduction (G-protein)	bladder carcinoma, thyroid cancer
HST	11q13.3	amplification	growth factor (FGF-like)	stomach
INT2	11p13	amplification, insertion	growth factor family (FGF-like)	esophagus, breast, glioblastoma
Ki-RAS	12p12.1	point mutation	signal transduction (G-protein)	pancreas-, colon-, lung-adenocarcinoma, endometrium-, other carcinoma, melanoma
KIT*	4q11-21	constitutive activation, point mutation	receptor tyrosine kinase (stem cell factor receptor	(hereditary)GIST, ANLL, testis
MET*	7q31	point mutation	receptor tyrosine kinase	(hereditary) papillary kidney tumour
NEU/ ERB B2	17q11.2-12	amplification	growth factor (EGF) receptor	breast-, ovary-, stomach-, other carcinomas
N-MYC	2p24	amplification	transcription factor	neuroblastoma, SCLC
N-RAS	1p13.2	Point mutation	signal transduction (G-protein)	acute leukaemia, thyroid cancer, melanoma
PIM-1	6p21	constitutive activation	signal transduction (serine-threonine kinase)	T-cell lymphoma
RAF-1	3p25	translocation	signal	stomach

Table 2. Tumour suppressor genes inactivated in hereditary cancer syndromes

Syndrome	Gene (chromosomal location)	Function (Kinzler-Vogelstein category)	Tumours
Familial stomach cancer	E-cadherin /CDH1 (16q22)	regulator of cell adhesion (gatekeeper)	stomach
Familial adenomatous polyposis (FAP) and its variants (Gardner, Turcot, attenuated FAP)	APC (5q21-q22)	β-catenin binding. Wnt signal pathway, transcription factor (gatekeeper)	gastrointestinal tumours, brain, thyroid carcinoma, retina lesions
Gorlin syndrome/ Nevoid basal cell carcinoma	PTCH (9q22)	sonic hedgehog pathway receptor (gatekeeper)	skin basal cell tumours, medulloblastoma
Juvenile polyposis (JPS)	SMAD4/ DPC4/ MADH4 (18q21), BMPR1A (10q22)	SMAD4: TGFβ signalization - intracellular mediator (gatekeeper ?) BMPR1A: (gatekeeper)	gastrointestinal tumours
Li-Fraumeni	TP53 (17q13.1)	transcription factor, cell cycle, apoptosis, DNA-repair, angiogenesis etc. regulator (gatekeeper)	sarcomas, breast-, brain tumours, leukaemia, adenocorticoid tumours
Multiplex endocrine neoplasia 1.	MEN1 (11q13)	inhibition of transcription with JunD transcription factor (?)	insulinoma, gastrinoma, parathyroid tumours, pituitary adenomas
Neurofibromatosis 1. (von Recklingshausen)	NF1 (17q11)	GTP-ase activator (gatekeeper)	neurofibroma, neurofibrosarcoma, brain tumours
Neurofibromatosis 2.	NF2 (22q12)	role in cell adhesion (gatekeeper)	schwannoma, meningioma
Hereditary breast cancer	BRCA1 (17q21) BRCA2 (13q13)	transcription factor, DNA-repair (caretaker), transcription factor, DNA-repair (caretaker)	breast-, ovary-, pancreas-, prostate-cancer female and male breast-, pancreas-, prostate-cancer
Hereditary melanoma	CDKN2/p14/ p16/p19ARF (9p21)	cell cycle regulator (gatekeeper)	melanoma, pancreas-, prostate-, bladder-, esophagus- cancer, leukaemia
Hereditary nonpolyposis colorectal cancer (HNPCC)	MLH1 (3p21), MSH2 (2p22-p21), MSH6 (2p16),	MLH1-PMS2: DNA mismatch repair (caretakers) TGFβR2: TGFβ pathway receptor (gatekeeper?)	gastrointestinal tumours, endometrial, ovarian, hepatobiliary and urinary tract cancer, glioblastoma (Turcot syndrome)