Review Article:

CANCER GENETICS: FROM CHROMOSOMES TO GENES

MAHIMA BHAKAR MOR, RAVI TANDON and H.D.KHANNA*

Department of Medicine and Biophysics*,
Institute of Medical Sciences,
Banaras Hindu University,
Varanasi-221005

A neoplasm, as defined by Willis, is an abnormal mass of tissue, the growth of which exceeds and is uncoordinated with that of the normal tissue and persists in the same excessive manner after the cessation of the stimuli which evoked the change. Depending upon the neoplasm's potential clinical behaviour, neoplasms have been categorized into – benign and malignant types. The term malignant implies that lesion can invade, destroy adjacent structures and spread to distant sites to cause death.

Cancer is a synonym for *malignant neoplasm*. The malignant phenotype often requires mutations in several different genes that regulate cell proliferation, survival, DNA repair, motility, invasion and angiogenesis. *Thus, cancer is a genetic disease.*

The field that studies the type of mutations, as well as the consequence of these mutations in tumor cells, is known as *cancer genetics*. The aim of cancer genetics is to understand the multi-step mutational and selective pathway that allow a normal somatic cell to form a population of proliferating and invasive cancer cells. This can be deeply confusing as there are so many genes, so many mutations, such an infinity of combinations.....every tumor is unique. But what they have in common is the product of selection for a specific set of definable capabilities, better known as pathways. Thus, it is more important to concentrate on principles of carcinogenesis, rather than cataloging genes and mutations.

The new knowledge in the field of cancer genetics has lead to significant progress in detection, diagnosis and treatment of cancers.

Whenever a cell is subjected to any kind of insult at the genetic level, there are three basic sets of regulatory genes which exert their effect through their ability to control cell division or cell death. These cells are proto-oncogenes, TS genes and DNA repair genes.

1) PROTO-ONCOGENES

These are genes that are normally present in a cell and are stimulated by external signals such as growth factors, which bind to cell surface receptors, stimulates intracellular signaling pathways leading to alterations in gene expression. In the normal cell, the expression of these proto-oncogenes is tightly controlled and they are transcribed at the appropriate stages of growth and development of cells thereby controlling proliferation and differentiation.

However, alterations in these genes or their control sequences lead to inappropriate expression resulting in formation of genes with oncogenic potential i.e. ONCOGENES.

There are three main ways of activation of proto-oncogenes to oncogenes :

a) STRUCTURAL ALTERATION

Proto-oncogenes acquire mutations that lead to increased activity of the gene product. This mutational event typically occurs in a single allele of the oncogene and acts in a dominant fashion. Point mutations which promote receptor dimerization in absence of growth factor binding are a general mechanism of activation of some of the receptor tyrosine kinases. Classically this has been described in the RET gene in the inherited disorders, familial medullary thyroid cancers and multiple endocrine neoplasia types 2A and 2B

.

b) AMPLIFICATION

This results in overproduction of the normal protein by amplification of the proto-oncogene. For example, amplification of the MYC gene is found in some cancers. In breast cancer, amplification of the HER2 gene is a common finding and has been shown in many studies to be an independent prognostic indicator predicting a more aggressive disease behavior in node-positive patients. Over expression of EGFR has been shown to be a strong prognostic indicator in head and neck cancer, ovarian, cervical, bladder and esophageal cancers.

c) DEREGULATED EXPRESSION

This results in loss of appropriate control mechanisms. One of the best studied of these is the t(9,22) found in CML which places the ABL oncogene on chromosome 9 next to the breakpoint cluster region (BCR) of the Philadelphia gene on chromosome 22.The fusion gene produced produces a fusion protein with constitutive protein tyrosine kinase activity . Another important aspect of BCR-ABL signal transduction is the cytoplasmic localization of the chimeric protein circumventing ABL's role in the nucleus to mediate DNA damage-induced apoptosis.
Another example is activation of MYC oncogene in Burkitt's lymphomas. Translocation between chromosomes 8 and 14 involving MYC and the immunoglobulin heavy chain gene is seen in approximately 80% of cases of Burkitt's lymphoma.

There is no single consistent mechanism of activation of any one oncogene. RAS is primarily activated by point mutation but amplification is also found; MYC is amplified in many tumors but abnormal expression is also associated with deregulation following chromosomal translocation. Whatever the mechanism for the activation of these genes, the end result is to produce a protein which can cause abnormal growth.

2) TUMOR SUPPRESSOR GENES (TS GENES)

These are defined as genes involved in control of abnormal cell proliferation and whose loss or inactivation is associated with the development of malignancy. They act by effectively inhibiting or putting the brake on cell growth and cell cycling.
Both copies of the gene have to be mutated before tumors develop. These genes are recessive at the level of the cell although they show dominant inheritance when associated with a familiar cancer syndrome.

MECHANISM OF INACTIVATION OF TS GENES:

The two major types of somatic lesions observed in TS genes during tumor development are *point mutations* and *large deletions*. Point mutations in the coding region of tumor-suppressor genes will frequently lead to truncated

protein products or otherwise nonfunctional proteins. Similarly, deletions lead to the loss of a functional product and sometimes encompass the entire gene or even the entire chromosome arm, leading to loss of heterozygosity (LOH) in the tumor DNA compared to the corresponding normal tissue DNA. LOH in tumor DNA is considered a hallmark for the presence of a tumor-suppressor gene at a particular locus. Gene silencing, which occurs in conjunction with hypermethylation of the promoter, is another mechanism of tumor-suppressor gene inactivation.

FAMILIAL CANCER SYNDROMES

A small fraction of cancers occur in patients with a genetic predisposition. The majority are inherited as autosomal dominant traits, although some of those associated with DNA repair abnormalities (xeroderma pigmentosum, Fanconi's anemia, ataxia telangiectasia) are autosomal recessive. Most forms of cancer do not follow simple patterns of inheritance. In many instances (e.g., lung cancer), a strong environmental contribution is at work. Even in such circumstances some individuals may be more genetically susceptible to developing cancer, given the appropriate exposure, due to the presence of modifier alleles. In these families, the affected individuals have a predisposing loss-of-function mutation in one allele of a TS gene.

Retinoblastoma, a childhood cancer of the retina which occurs in both sporadic and familial forms .40% of cases are hereditary, transmitted as an autosomal dominant trait. Of these cases around 10–15% are transmitted from an affected parent, the remainder arising as de novo germ-line mutations, more usually on the paternal germ-line than the maternal. Tumors frequently arise in both eyes, the average number of tumor foci being three to five. The remaining 60% of cases are sporadic and characteristically tumors are seen in only one eye.

In 1971, Knudson postulated that the disease arose from two sequential mutational events. In the hereditary form of the disease, one mutation is inherited in the germ-line and is phenotypically harmless, confirming the recessive nature of the mutation at the cellular level. A second ‘hit’, occurring in a retinal cell, causes the tumor to develop.

As there are a large number of retinoblasts in the eye ,which are all at risk because they already carry one mutation, a second ‘hit’ will occur frequently enough to cause a high proportion of tumors in at least one eye and often in both. In sporadic form of the disease, both mutations occur in the somatic tissue. The probability of two mutations occurring in the same cell is low, therefore the disease is primarily unilateral.

CELL CYCLE REGULATORS

The cell cycle consists of the DNA synthesis phase (S) and cell division at mitosis (M) separated by two gap intervals G1 and G2. Upon growth factor stimulation, cells move from the resting phase, G0 into G1. Cycling will stop at this stage if mitogenic signals are removed prior to a specific point—the restriction point (R)—at the end of G1 and will return to G0.

The orderly progression of cells through the various phases of the cell cycle is orchestrated by cyclin-dependent kinases (CDKs), which are activated by binding to *cyclins*. The CDK-cyclin complexes phosphorylate crucial target proteins that drive the cell through the cell cycle. On completion of this task, cyclin levels decline rapidly.

There are two main cell cycle checkpoints, one at the G_1/S transition and the other at G_2/M. To function properly, cell cycle checkpoints require sensors of DNA damage, signal transducers and effector molecules.

Before a cell makes the final commitment to replicate, the G_1/S checkpoint checks for DNA damage; if damage is present, the DNA-repair machinery and mechanisms that arrest the cell cycle are put in motion. The delay in cell cycle progression provides the time needed for DNA repair; if the damage is not repairable, apoptotic pathways are activated to kill the cell. Thus, the G_1/S checkpoint prevents the replication of cells that have defects in DNA, which would be perpetuated as mutations or chromosomal breaks in the progeny of the cell.

The G_2/M checkpoint monitors the completion of DNA replication and checks whether the cell can safely initiate mitosis and separate sister chromatids.

In the G_1/S checkpoint, cell cycle arrest is mostly mediated through p53, which induces the cell cycle inhibitor p21. Arrest of the cell cycle by the G_2/M checkpoint involves both p53-dependent and p53-independent mechanisms.

Defects in cell cycle checkpoint components are a major cause of genetic instability in cancer cells. Mutations that dysregulate the activity of cyclins and CDKs favor cell proliferation.

The cyclin D genes are over expressed in many cancers, including those affecting the breast, esophagus, liver, and a subset of lymphomas. Amplification of the *CDK4* gene occurs in melanomas, sarcomas, and glioblastomas. Mutations affecting cyclin B and cyclin E and other CDKs also occur, but they are much less frequent.

RB GENE

RB protein, the product of the *RB* gene, is a ubiquitously expressed nuclear phosphoprotein that plays a key role in regulating the cell cycle. RB exists in an active hypophosphorylated state in quiescent cells and an inactive hyperphosphorylated state in the G_1/S cell cycle transition.

The importance of RB lies in its enforcement of G_1, the most important checkpoint in the cell cycle clock. The initiation of DNA replication requires the activity of cyclin E–CDK2 complexes, and expression of cyclin E is dependent on the E2F family of transcription factors. Early in G_1, RB is in its hypophosphorylated active form and it binds to and inhibits the E2F family of transcription factors, preventing transcription of cyclin E. Hypophosphorylated RB blocks E2F-mediated transcription in at least two ways. First, it sequesters E2F, preventing it from interacting with other transcriptional activators. Second, RB recruits chromatin-remodeling proteins, such as histone deacetylases and histone methyltransferases, which bind to the promoters of E2F-responsive genes such as cyclin E. E2Fs are not the sole effectors of Rb-mediated G_1 arrest. Rb also controls the stability of the cell cycle inhibitor p27.

If RB is absent (due to gene mutations) or its ability to regulate E2F transcription factors is derailed, the molecular brakes on the cell cycle are released and the cells move through the cell cycle.

*The emerging paradigm is that loss of normal cell cycle control is central to malignant transformation and that at least one of four key regulators of the cell cycle (*p16/INK4a, cyclin D, CDK4, RB) *is dysregulated in the vast majority of human cancers*. In cells that harbor mutations in any one of these other genes, the function of RB is disrupted even if the *RB* gene itself is not mutated.

TP53: GUARDIAN OF THE GENOME

The *p53* gene is located on chromosome 17p13.1.It is the most common target for genetic alteration in human tumors. Homozygous loss of *p53* occurs in virtually every type of cancer, including carcinomas of the lung, colon, and breast—the three leading causes of cancer death.

In non-stressed healthy cells, p53 has a short half-life (20 minutes) because of its association with MDM2, a protein that targets it for destruction. When the cell is stressed, p53 undergoes post-transcriptional modifications that release it from MDM2 and increase its half-life. Unshackled from MDM2, p53 also becomes activated as a transcription factor.

In most cases, the inactivating mutations affect both *p53* alleles and are acquired in somatic cells. p53 is a transcription factor that is at the center of a large network of signals that sense cellular stress, such as DNA damage, shortened telomeres and hypoxia. In the majority of tumors without a *p53* mutation, the function of the p53 pathway is blocked by mutation in another gene that regulates p53 function mutated. Indeed, MDM2 is amplified in 33% of human sarcomas, thereby causing functional loss of p53 in these tumors.

p53 has a number of functions in the cell. p53 thwarts neoplastic transformation by three interlocking mechanisms: activation of temporary cell cycle arrest (quiescence), induction of permanent cell cycle arrest (senescence), or triggering of programmed cell death (apoptosis). p53 also stimulates expression of genes which prevent the growth of new blood vessels needed to allow tumor growth . Cells with p53 mutations are therefore more likely to be able to recruit new vessels and give them a growth advantage.

GENOMIC INSTABILITY

Defects in DNA repair pathways leading to either single or oligo-nucleotide mutations (microsatellite instability) or more commonly chromosomal instability leading to aneuploidy. Genomic instability is an universal feature of cancer cells. Instability may be of two types:

1) Chromosomal instability(CIN) is the commonest form. Tumor cells typically have abnormal karyotypes, with extra and missing chromosomes, rearrange-ments, and so on. Tumor cell lines are often chromosomally unstable, acquir-ing new changes during culture.

2) Micosatellite instability(MIN) is a DNA- level instability seen in certain tumors, especially some colon carcinomas.

Instability is necessary to enable a cell to amass enough mutations to complete the microevolution, from a normal somatic cell to an invasive cell.

However, tumors normally show either CIN or MIN but not both, this suggests that instability is not a chance feature but is the result of selection.

There are three main mechanisms of production of genomic instability - defective spindle check point, undamaged DNA repair and limitless replicative potential mediated by telomeres.

TELOMERES

After 60-70 divisions, the cells lose their ability to divide and become senescent. This has been ascribed to progressive shortening of *telomeres* at the ends of chromosomes. These are indeed composed of tandem repeats of a six-nucleotide sequence (TTAGGG) that associate with specialized telomere-binding proteins to form a T-loop structure that protects the ends of chromosomes from being mistakenly recognized as damaged. The loss of telomeric repeats with each cell division cycle causes gradual telomere shortening leading to growth arrest (*replicative senescence*). Cells where the checkpoints are disabled by *p53* or *RB1* mutations, the non-homologous end-joining pathway is activated joining the shortened ends of two chromosomes.

For tumors to grow indefinitely loss of growth restraints is not enough. Tumor cells must also develop ways to avoid both cellular senescence and mitotic catastrophe.

Telomere maintenance is seen in virtually all types of cancers. In 85% to 95% of cancers, this is due to up-regulation of the enzyme telomerase. A few tumors use other mechanisms, termed *alternative lengthening of telomeres*, which probably depend on DNA recombination.

ESSENTIAL ALTERATIONS FOR MALIGNANT TRANSFORMATION

There are six fundamental changes in cell physiology that together determine the malignant phenotype. Each of the cancer-associated genes has a specific function, the dysregulation of which contributes to the origin or progression of malignancy.

1) Self-sufficiency in growth signals

2) *Insensitivity to growth-inhibitory signals*

3) *Evasion of apoptosis*

4) *Limitless replicative potential*

5) *Sustained angiogenesis*

6) *Ability to invade and metastasize*

Mutations in genes that regulate these cellular traits are seen in every cancer. However, the precise genetic pathways that give rise to these attributes differ between individual cancers, even within the same organ. The occurrence of mutations in cancer-related genes is conditioned by the robustness of the DNA-repair machinery and protective mechanisms such as apoptosis that prevent the proliferation of cells with damaged DNA.

THE CLONAL ORIGIN AND MULTISTEP NATURE OF CANCER:

Carcinogenesis is a multistep process at both the phenotypic and the genetic levels, resulting from the accumulation of multiple mutations. No single mutation can fully transform non-immortalized cells but *cells can generally be transformed by combinations of mutations.*

A classic example of this is colon carcinoma which evolve through a series of morphologically identifiable stages: colon epithelial hyperplasia followed by formation of adenomas that progressively enlarge and ultimately undergo malignant transformation. Inactivation of the *APC* TS gene occurs first, followed by activation of *RAS* and, ultimately, loss of a TS gene on 18q and loss of *p53*.

All cancers originate from a single cell, this clonal origin is a critical discriminating feature between neoplasia and hyperplasia. During progression, tumor cells are subjected to immune and non-immune selection pressures. A growing tumor, therefore, tends to be enriched for subclones that adapt at survival, growth, invasion and metastasis. Therefore, *even though most malignant tumors are monoclonal in origin, by the time they become clinically evident their constituent cells are extremely heterogeneous.*

MOLECULAR TECHNIQUES FOR ANALYSIS OF GENES

Genes can be studied at three levels - DNA, RNA or protein. These molecules can be either within the cell in prepared tissue sections (in situ) or in isolation. These two approaches give different information.

Isolated DNA can be examined for qualitative and quantitative abnormalities. This is useful when looking for rearrangements or mutations within a particular gene and also for assessing absolute levels of a gene (gene amplification). Analysis of isolated RNA gives information about the level of transcription (gene expression). Direct analysis of protein allows the determination of protein levels or of changes in protein size.

In situ analysis provides information concerning the spatial distribution of molecules in the cell and therefore, shows which cells are expressing a particular gene, RNA sequence or protein molecule, but cannot easily be used for quantitative analysis. When used on chromosome spreads this technique gives information about the chromosomal localization of a gene or, once a gene has been mapped, can be used to identify gene deletions.

The analysis of genes has been revolutionized by the introduction of new technologies: PCR, RT-PCR, FISH, Microarrays etc. Most notable of these is microarray technology which is a high-throughput method that enables information to be generated about gene expression and function.

Advances in automation and bioinformatics have resulted in a discipline of biology termed genomics, which is defined as the generation and analysis of information about genes and genomes. This has lead to identification and characterization of individual genes and patterns of gene expression that distinguish malignant and premalignant cells from their normal counterparts.

The comprehensive study of proteins and protein systems is now called proteomics. From a biomedical point of view proteomics has great potential as the bulk of phamacological interactions and diagnostic tests are directed at proteins rather than their genes.

THERAPEUTIC APPLICATIONS

Better cancer treatment is the need of the hour. Rationally designed, target-based therapeutic agents, directed against the specific molecular derangements that distinguish malignant from nonmalignant cells, have become possible with advances in the understanding of oncogene and tumor-suppressor pathways.

Therapeutic targets in cancer treatment are -

1) Cell cycle abnormalities

Flavopiridol was the first relatively selective CDK inhibitor. Regulation of cellular transcription elongation by the CDK7/cyclin H and CDK9/cyclin T1 complexes may be the critical target of flavopiridol. Responses have been reported in CLL, phase II clinical trials are in progress.

Development of "oncolytic" viruses that replicate selectively in tumor cells with defined genetic lesions i.e those that lack functional p53 (adenovirus mutant with viral p55 deleted) or have defects in the pRB pathway.

2) Telomerases

The reverse transcriptase activity of telomerase is a prime target for small-molecule pharmaceuticals. The protein component of telomerase can act as a tumor-associated antigen recognized by antigen-specific cytotoxic T lymphocytes that lyse human melanoma, prostate, lung, breast and colon cancer cells in vitro. Clinical trials of telomerase vaccines are underway.

3) Signal transduction pathways

a) Tyrosine kinases

Many tyrosine kinases act on signaling pathways. Receptor Tyrosine Kinases are transmembrane glycoproteins that undergo dimerization upon ligand binding, with activation of their cytoplasmic tyrosine kinase domains by proximity-induced trans-phosphorylation of the activation loop.

The Abl tyrosine kinase inhibitor, Imatinib Mesylate has validated the concept of a molecularly targeted approach to cancer treatment. Clinical studies have demonstrated remarkable activity of this agent in CML.

Relapse during treatment with Imatinib was associated with reactivation of the tyrosine kinase by amplification of the *Bcr-Abl* locus or point mutations within the Bcr-Abl kinase domain that decreased Imatinib binding without loss of Bcr-Abl kinase activity. These data constitute genetic proof that the target of imatinib is the Bcr-Abl tyrosine kinase and that Bcr-Abl kinase activity is still required by imatinib-resistant cells.

Two drugs have been developed (Dasatinib and Nilotinib) that are potent inhibitors against most Imatinib resistant mutants; these compounds have demonstrated significant activity in patients with imatinib-resistant CML.

Imatinib is a low-molecular-weight competitive inhibitor of the ATP binding site of Bcr-Abl, c-Abl, PDGFR and c-Kit; hence it is not absolutely specific for the *Bcr-Abl* oncogene product. It has also demonstrated targeted activity in other diseases, including gastrointestinal stromal tumors, rare mesenchymal tumors of the GI tract (stomach and small intestine). Imatinib has become the paradigm of targeted drug development in other diseases.

GIST are thought to arise from or share a common stem cell with the interstitial cells of Cajal, which give rise to the myenteric plexus of the GI tract . The pathogenic molecular event for most patients with this disease is mutation of the proto-oncogene c-*Kit*, leading to the constitutive activation of this receptor tyrosine kinase without the binding of its physiologic ligand.

Imatinib, which inhibits the c-Kit kinase domain, has demonstrated significant activity (>50% partial responses usually lasting 1–2 years) in this chemotherapy-refractory tumor. Resistance to Imatinib develops due to secondary mutations in c-Kit. About 10% of GIST encode activating mutations of the PDGFR instead of c-*Kit*. Many of these tumors are susceptible to treatment with the multitargeted TK inhibitor Sunitinib that has activity against c-Kit as well as the PDGF and VEGFR. Sunitinib is an FDA approved drug for treatment of patients with imatinib-resistant GIST or who are intolerant of Imatinib.

These examples extend the proof of principle that targeting of signaling pathways in cancer cells can be highly efficacious with minimal toxicity, even when the drug does not have absolute target specificity.

b) Non-tyrosine kinases

EGFR mutations define a novel subset of lung cancers. Clinical studies of two high-affinity competitive inhibitors of the ATP binding site in the EGFR kinase domain, Gefitinib and Erlotinib, have provided important insights into the pathogenesis of different subsets of patients with non-small cell lung cancer.

FDA approval after ~10–20% of advanced-stage patients treated with single-agent Gefitinib or Erlotinib had objective tumor responses. Responders tended to have adenocarcinoma or bronchoalveolar histology and were never-smokers, women, and of Eastern Asian origin. DNA sequence analysis of the *EGFR* gene isolated from the tumors of responding patients demonstrated that most had acquired mutations of the kinase domain that led to increased tyrosine kinase activity.

Mutated K-*Ras*, which occurs to the exclusion of EGFR mutation, appears to define a subset of patients with low likelihood of response to EGFR inhibitors. It has been proposed that the pathogenesis of NSCLC in never-smokers occurs through a novel pathway that is dependent on activated EGFR, and that tumors are *addicted* to this oncogene, rendering them highly susceptible to its inhibition. No EGFR kinase domain mutations have yet been found in tumors other than NSCLC.

Thus, these studies define a novel oncogenic pathway for an important human cancer and provide a mechanism to identify subsets of patients likely to respond to the targeted therapy.

4) Alteration in gene transcription

5) Apoptosis

6) Metastasis

7) Cancer stem cells

8) Tumor angiogenesis

MOLECULARLY TARGETED AGENTS FOR TREATMENT OF CANCER

(FDA APPROVED)

Drug	Molecular Target	Disease	Mechanism of Action
All-*trans* retinoic acid (ATRA)	PML-RARα oncogene	Acute promyelocytic leukemia M3 AML; t(15;17)	Inhibits transcriptional repression by the PML-RARα
Imatinib (Gleevec)	Bcr-Abl, c-Abl, c-Kit, PDGFR-α/β,	Chronic myelogenous leukemia; GIST	Blocks ATP binding to tyrosine kinase active site.
Sunitinib (Sutent)	c-Kit, VEGFR-2, PDGFR-β, Flt-3	GIST; renal cell cancer	Inhibits activated c-Kit and PDGFR in GIST; inhibits VEGFR in RCC.
Sorafinib (Nexavar)	RAF, VEGFR-2, PDGFR-α/β, Flt-3, c-Kit	RCC; may have activity in melanoma when combined with chemotherapy	Targets VEGFR pathways in RCC. Possible activity against BRAF in melanoma, colon cancer, and others.
Erlotinib (Tarceva)	EGFR	Non-small cell lung cancer; pancreatic cancer	Competitive inhibitor of the ATP binding site of the EGFR.
Gefitinb (Iressa)	EGFR	Non-small cell lung cancer	Inhibitor of EGFR tyrosine kinase.
Bortezomib (Velcade)	Proteasome	Multiple myeloma	Inhibits proteolytic degradation of multiple cellular proteins.

THE FUTURE

Identification of cancer genes and genetic alterations has led to better understanding of the molecular pathways and the tumorigenesis process. The oncogenes, TS genes, their RNA transcripts, and their protein products are therefore, potential targets for attack by specific therapeutic agents. It may lead to the development of sensitive strategies for early detection of cancer and novel therapies based on pathophysiology rather than empiricism.

REFERENCES

1) Robbins Basic Pathology, 7 th edition
2) Harrison's principles of internal medicine, 17 th edition
3) Human molecular genetics by Tom Strachan and Andrews Read, 4 th edition
4) Molecular Biology of Cancer by F. Macdonald , C H J Ford and A G Casson , 2 nd edition