Key words. 18 CC-201

Molecular Mechanisms of Tumor Formation

Ghassan Samara, MD, Michael Hurwitz, MD, Mark Sawicki, MD, Edward Passaro, Jr., MD, Los Angeles, California

DNA is the fundamental guide for the cell's processes. The alterations in DNA that can lead to abnormal or absent proteins and the role of chromosomal alterations in changing the function of the DNA are reviewed. The role of viruses in derailing a cell's normal functions, the major mechanisms of oncogene action, and tumor suppressor genes are also discussed.

The parents of the yet unborn fetus asked why a genetic defect that did not affect them would be so detrimental to their child. The fact that only one properly functioning copy of the gene was necessary explained the parents' lack of symptoms since they each had one normal copy. The unborn child, on the other hand, had inherited a defective gene from each parent. Thus, the child would not have the ability to make the proper protein, the cystic fibrosis transmembrane conductance regulator, needed to regulate chloride secretion. This would lead to overly viscous secretions, multiple pulmonary infections, and nutrient malabsorption.

The underlying derangement in cancer is an alteration in the cells' DNA, the fundamental blueprint for cellular processes. When the cells' DNA is damaged or altered (mutated), the normal functions go awry, leading to cell death or one of the steps in tumor formation. The latter is especially true if the damage is to the cells' growth control systems. As these defects accumulate, they lead progressively to malignant changes. The alterations in the DNA, however, do not affect cell function directly. Rather, the deranged DNA leads to abnormal RNA and, in turn, abnormal proteins. Herein, we briefly review the changes leading to tumor formation by first discussing the alteration of the DNA blueprint and then how the product of these alterations (altered proteins) affects cell function.

There are several differing types of mutations (**Figure 1**). Mutations that affect a single nucleotide base are called point mutations. Point mutations can cause termination of the peptide chain if the RNA copy of the mutated DNA contains a stop codon, a sequence of three bases instructing the process of translation to stop. These mutations can completely destroy a protein's function and are called nonsense mutations.

More common are the missense point mutations. These may not result in any significant alteration in protein function since the codon that is mutated may just be changed to another codon for the same or a similar amino acid without hindering the protein product's function. There are many examples of point mutations, however, that do cause significant alterations in function. For example, the point mutations at codons 12, 13, or 61 in the *ras* oncogene cause activation of the oncogene.

Mutations that are caused by the loss or gain of one or more base pairs in the DNA are called deletion or insertion mutations, respectively. These mutations can vary in size from one to thousands of base pairs in length and can span large areas of the genome. When the insertion or deletion of bases changes the reading frame of the codons in the transcript, many amino acids being incorporated in the subsequent portion of the peptide chain will be different. These frameshift mutations can severely alter protein structure.

From the Surgical Service, Veterans Affairs Medical Center, and the Department of Surgery, UCLA School of Medicine, Los Angeles, California.

Requests for reprints should be addressed to Edward Passaro, Jr., MD, Surgical Service (W112), Veterans Affairs Medical Center, Los Angeles, California 90073.

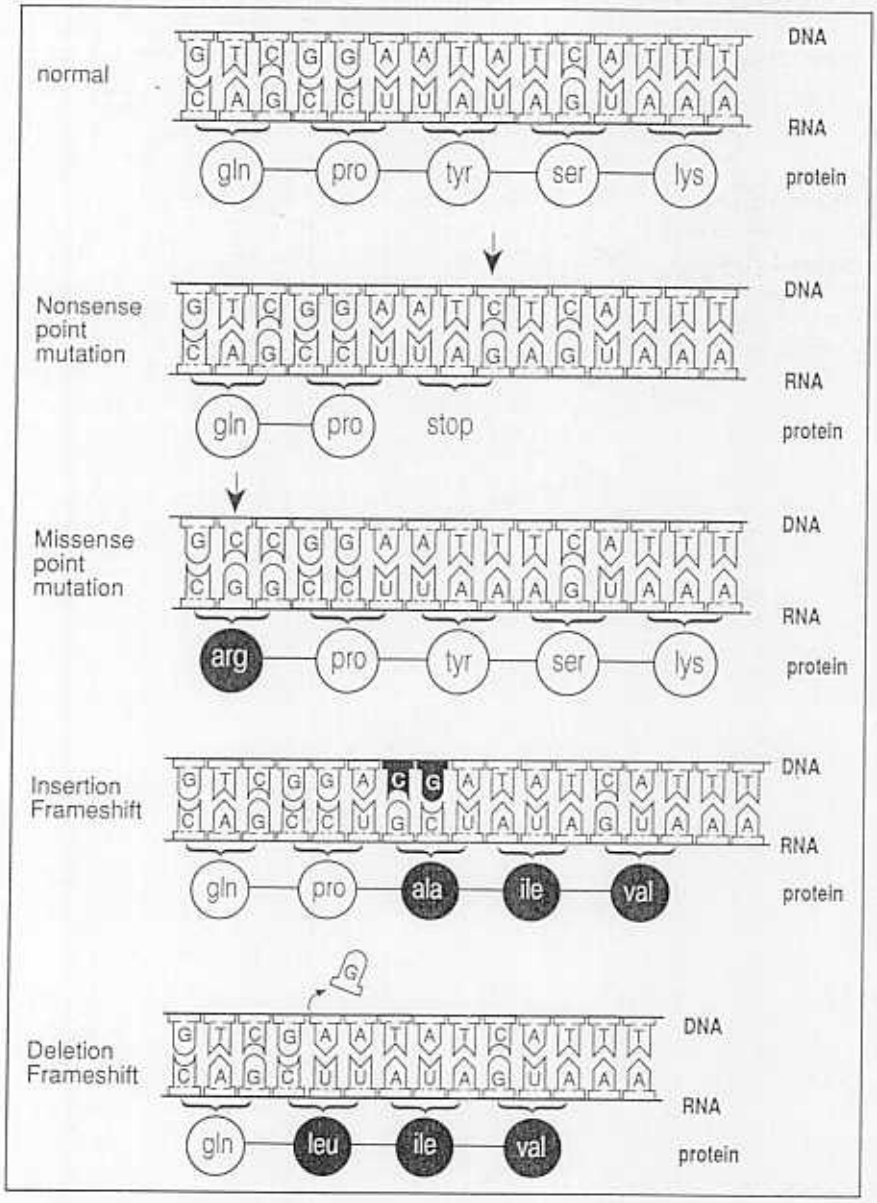

Figure 1. Different types of mutations.

CHROMOSOMAL ALTERATIONS

The largest mutations in size involve entire chromosomes (**Figure 2**). These mutations can be detected through cytogenetic studies in which the cell's chromosomes are stained and then viewed through a light microscope. In the past, this technique was effectively used in hematologic malignancies. Now, however, it is becoming more and more useful for solid tumors. Cytogenetic abnormalities include losses of genetic material such as deletions or monosomy, additions of DNA such as insertions or trisomy, or rearrangements like inversions or translocations.

The majority of rearrangements activate oncogenes by one of two methods. The first is the formation of active fusion proteins. A fusion protein results when the genetic codes of two different proteins are physically combined, leading to a protein product with portions of each parent represented. One example is a translocation between chromosomes 9 and 22, leading to a fusion between a guanosine triphosphatase-activating protein, BCR, and a tyrosine kinase, ABL. The BCR-ABL fusion protein acts as an oncogene, causing a transformation of the cell. This translocation gives rise to the Philadelphia chromosome associated with chronic myelogenous leukemia.

The alteration of a chromosome can also lead to oncogenesis if a normal protein product is present in increased quantity due to changes in a regulatory portion of the adjacent DNA. One solid tumor example is a parathyroid adenoma in which an inversion on chromosome 11 causes the regulatory elements of parathyroid hormone (PTH) to enhance the production of mRNA for the PRAD1 gene, a probable oncogene.

Deletions within chromosomes can remove genes that normally act to limit cell growth (tumor suppressor

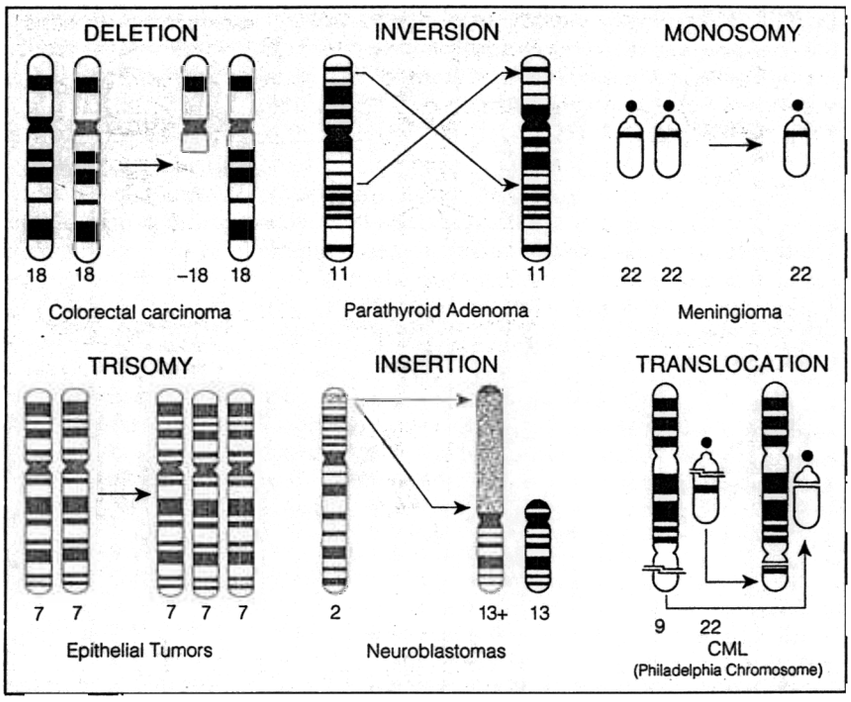

Figure 2. Chromosomal abnormalities.

genes). If the deletion involves an entire chromosome, monosomy results, e.g., monosomy 22 associated with meningioma. Deletion of a portion of chromosome 18 on the long arm removes a gene known as DCC (deleted in colorectal carcinoma), the protein product of which is believed to play a role in contact inhibition. This is one of the later changes that occur in the multistep process of colon tumor formation.

Insertions behave similarly to inversions and translocations but may also act by amplification, an increase in the number of DNA copies coding for a particular oncogene. This is in contrast to the previous PTH/PRAD1 example, in which the increase occurs at the level of transcription into RNA. As with deletions, insertions can also involve an entire chromosome. This results in trisomy, e.g., trisomy 7 associated with epithelial tumors. Usually trisomy results only in a twofold or so amplification of a set of genes. However, other mechanisms can produce greater amplification, on the order of 100 times. Amplification of a gene called c-*erb*B2 (also called HER-2/neu) occurs in about one third of breast cancers, and detection of this amplification may soon be useful as a prognostic marker.

VIRAL AGENTS

It is estimated that viruses cause 15% to 20% of all human tumors worldwide. Hepatocellular and cervical carcinoma together account for 80% of virally linked cancers. There are several mechanisms of tumor formation that are mediated by viruses (**Figure 3**). Human immunodeficiency virus infection is associated with increased neoplasia due to loss of the normal immune response to tumors, i.e., T cells. Some viruses act as nonspecific inducers of cell growth and multiplication, leading to increased chances of mutation, e.g., viral hepatitis. In contrast to these indirect effects, viruses can directly cause oncogene activation by three methods.

To understand how retroviruses cause cancer, it is important to review how retroviruses propagate. A retrovirus attaches to a cell and discharges its RNA genome into the cell. This RNA transcript is accompanied by reverse transcriptase, an enzyme that makes DNA based on an RNA template. This new DNA copy is inserted into the host genome, and RNA copies of it are made and packaged into new viral particles.

Occasionally, recombination between the viral and host genomes results in a virus that contains a cellular gene, usually in a mutated state. Such a virus can then introduce this gene into a new cell (transduction). If the gene is an oncogene, the newly infected cell may proceed to neoplasia. This explains the phenomenon of viral oncogenes corresponding to normal cellular genes. *Ras* genes were first identified in viruses (viral oncogenes) but were later also found in normal cells. The insertion of a gene coding for an oncogene that results in transformation is termed the *trans* effect.

There are examples of viral genes that act via the *trans* effect that may not be of cellular origin. The human papilloma viruses (HPV) are associated with cancers of the anogenital tract. When the DNA of these viruses is translated to protein, two products, E6 and E7, are among those produced. It has been found that these bind to two important tumor suppressor genes, p53 and the retinoblastoma (RB) gene. This binding may be part of the mechanism whereby HPV cause carcinoma.

The second direct mechanism of oncogene activation by viruses is termed the *cis* effect. In this pathway, a viral promoter is inserted within the human DNA such that it causes increased transcription of a cellular gene. This is analogous to the example of the PTH enhancer and

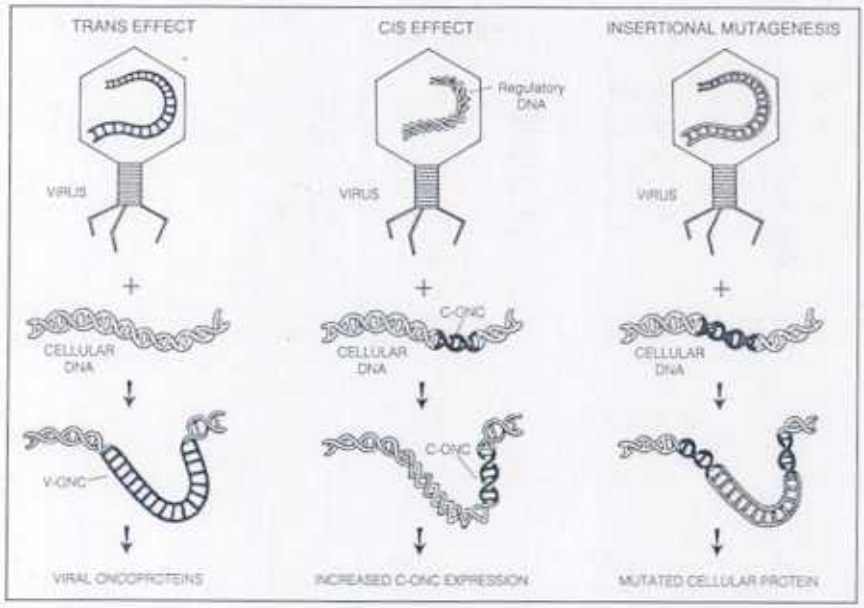

Figure 3. Viral mechanisms of tumor formation.

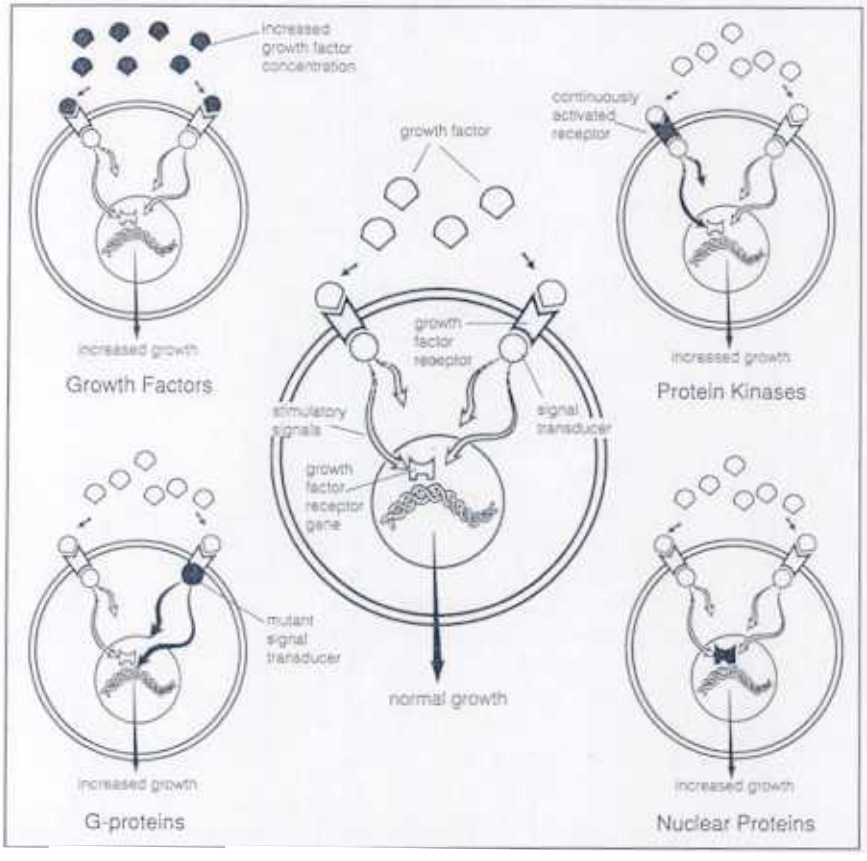

Figure 4. Selected mechanisms of oncogene action.

PRAD1 cited above. Although there is evidence for this mechanism in animal models of hepatocellular carcinoma, a clear example of this has yet to be found in a human tumor type. In these models, there is increased transcription of several oncogenes (c-*myc*, *erb*-A) resulting from viral integration near the respective gene.

Finally, an oncogene may be activated (or a tumor suppressor inactivated) by insertional mutagenesis. Here, a virus inserts into the genome within the gene and alters the function of that gene. Both retroviruses and DNA viruses are known to integrate into host genomes, often but not always at specific sites.

ONCOGENES

A gene whose protein product induces tumor formation is called an oncogene. An oncogene can be a mutated form of a normal cellular gene, a viral gene introduced into a cell, or even a normal cellular gene that is present in an abnormal number of copies. The normal (or cellular) counterpart of an oncogene is called a proto-oncogene or cellular oncogene. These cellular oncogenes are signified by a c- preceding the name of the oncogene, e.g., c-*myc*. Similarly, an oncogene with a viral origin is called a viral oncogene and is signified by a v- as in v-*myc*.

More than 60 oncogenes have been discovered. Some were discovered as viral oncogenes; others were found by transfection experiments. Oncogenes are often homologous to elements of the cell's signaling pathways. They act in a dominant manner, i.e., they act as dominant genes. One activated oncogene in a cell can change the phenotype of the cell despite the normal remaining copy. Oncogenes can be arranged into four functional groups: growth factors, protein kinases, guanosine triphosphate (GTP)-binding proteins, and nuclear proteins (**Figure 4**).

The growth factor oncogenes act as hormonal signals to the cell to multiply. Normally, a quiescent cell must receive two signals for mitosis to proceed: first, an initiating signal, followed by a progression signal, which must follow within a short period of time. The initiating signal is a growth factor that renders the cell competent to leave the resting state and begin the cell cycle, culminating in mitosis. Several such factors have been found and include epidermal growth factor (EGF), platelet-derived growth factor (PDGF), and fibroblast growth factor (FGF). Next, a progression factor is necessary for a cell to continue the process toward mitosis. Insulin and insulin-like growth factor 1 (IGF-1) are two examples of such growth factors. These progression factors are antagonized by the action of cytokines such as interferon and tumor necrosis factor (TNF). Each class of cells reacts uniquely to these different factors because of the signal transduction machinery inherent in the cell type, i.e., one cell type may differentiate under the same stimulus that causes another to multiply wildly.

When a growth factor gene is inappropriately expressed, inappropriate growth may follow. If a cell simultaneously produces and responds to the same growth factor (autocrine), a positive feedback loop results. This autocrine loop probably accounts for the increased growth in tumor cells with activated growth factor oncogenes. The oncogene *int*-2 is one example. This oncogene is found in breast cancer.

Protein kinases are a class of enzymes that add a phosphate group to other proteins. This process usually activates or inactivates the substrates of the kinase and is one of the major paths of cellular regulation. The most significant group of oncogenic protein kinases are the receptor tyrosine kinases. These are a class of membrane-bound receptors that phosphorylate proteins on tyrosine residues when a specific ligand binds the extracellular portion of the receptor. This phosphorylation results in differing effects based on the receptor and the cell type but, in general, is a signal for growth.

Most growth factor receptors such as PDGF, FGF, EGF, and insulin receptors are of this class. Many oncogenes code for these tyrosine kinases. A prominent example is *erb*B2, a receptor tyrosine kinase that is homologous to the EGF family of receptors. When present (often in breast carcinoma) as an oncogene, *erb*B2 exhibits increased tyrosine kinase activity in the cell. This increased activity is often the result of increased amounts of the protein (due to increased transcription or gene amplification), but inherently activated forms of the protein that result from a point mutation are known. These exhibit tyrosine kinase activity in the absence of growth factors. The usual signals for cell growth are thus bypassed, and uncontrolled cell division takes place.

Not all signals of cell growth are transduced by protein kinases. Some cell surface receptors act via cyclic adenosine monophosphate (cAMP), a common second messenger in cells. Normally, a ligand binds a receptor that is in proximity to a GTP-binding protein. The GTP-binding protein (G-protein) then binds GTP and becomes active. The active G-protein then stimulates adenylate cyclase, which, in turn, converts adenosine triphosphate to cAMP. High cAMP levels in the cell can be stimulatory. The G-protein then cleaves GTP to guanosine diphosphate and becomes inactive. This inactivation of the G-protein is facilitated by yet another protein called a GAP protein.

The family of oncogenes called the *ras* oncogenes are G-proteins. They are found in colon, lung, pancreatic, and other carcinomas. The point mutations that occur in these oncogenes inactivate that portion of the protein that cleaves GTP, thus a mutated *ras* cannot become inactive. Again, this leads to stimulation of the cell without regard to the normal homeostatic signals.

All the signals the cell receives to grow must ultimately reach the nucleus to take effect. The proteins that directly interact with the DNA and regulate cell processes are located there. These include proteins that transcribe DNA to RNA and copy DNA and those that enhance or inhibit both transcription and replication. The oncogenes that have protein products that localize to the nucleus of the cell are also a diverse group. Many have regions within them that represent DNA-binding protein structures.

One commonly found example of a nuclear oncogene is *myc*. Like many other oncogenes, it was first discovered as a viral oncogene (v-*myc*), although several cellular forms have subsequently been found (c-*myc*, N-*myc*, and L-*myc*). The protein has been found to have regions that bind a specific DNA sequence, and it also has structural similarity to known transcription factors. This implies a role for *myc* in transcriptional regulation. Research is ongoing to find the exact mechanisms of *myc* action.

TUMOR SUPPRESSOR GENES

The tumor suppressor genes are the most recently discovered class of genes important to the study of malignancy. A gene that normally suppresses tumor formation but allows tumors to form when deleted or mutated is called an anti-oncogene or a tumor suppressor gene. Due

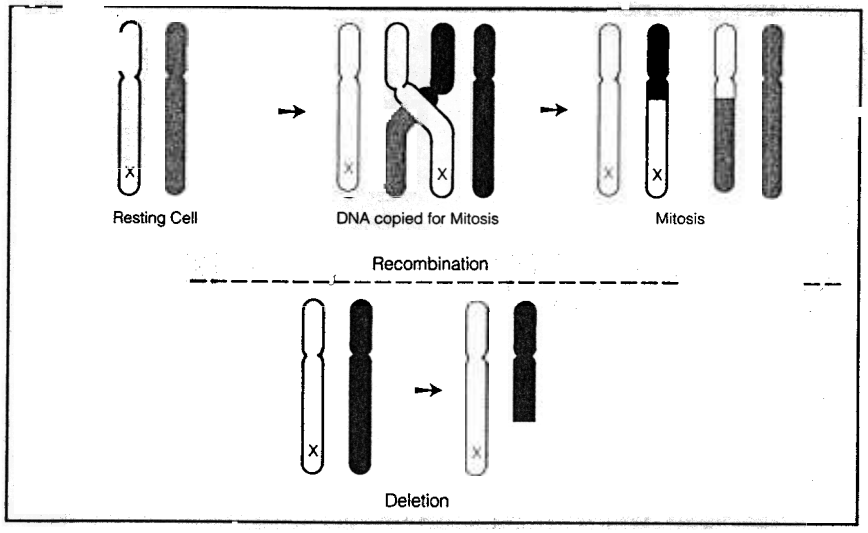

Figure 5. Loss of heterozygosity.

to the difficulty of isolating tumor suppressor genes, less than 10 have been cloned, despite the recognition of many genomic regions that contain them. They generally act in a recessive manner, i.e., both copies of the gene must be inactivated to affect the cell. They are also distinguished in that many familial neoplastic conditions are linked to tumor suppressor genes. In fact, familial conditions such as familial RB were important in the discovery of tumor suppressor genes. These familial conditions result when a tumor suppressor gene is mutated in the germline. This gives each cell in the affected individual only one functional gene instead of two. Such a cell is thus much more prone to inactivation of the remaining tumor suppressor gene, with progression toward tumor formation. The function of tumor suppressor gene products is, in general, to limit cell growth. They do this by several mechanisms.

Each cell normally has two copies of the genome, a maternal and a paternal set of chromosomes. If one gene is activated by a mutation, it usually changes the phenotype of the cell, but, if a normally active gene is inactivated, the other copy can make up the difference. A phenomenon called loss of heterozygosity (LOH) illustrates an important mechanism of inactivation of tumor suppressor genes (**Figure 5**). LOH refers to the loss of alleles that are different (heterozygous) on the two genomic copies. This can result from recombination during mitosis. A portion of the paternal chromosome is exchanged for a portion of the maternal chromosome so that when mitosis proceeds, one cell receives both maternal copies and the other cell receives the paternal ones for the region exchanged. Deletion of a portion of one chromosome would also have the same effect. If one of the genes was mutated, the cell receiving both mutated copies may proceed to tumor formation (assuming the mutation inactivates a gene that usually serves to regulate growth). Due to the rarity of mutations, it is 100 times more likely for recombination to inactivate the second allele than for a separate mutation to do so.

The gene associated with RB is thought to limit cell growth by its interactions with transcription factors (including c-*myc*). The protein product of the gene acts in the nucleus of the cell and is regulated by phosphorylation. It is thought that the unphosphorylated form of the protein is active in inhibiting transcription and, in turn, slowing cell growth. The RB gene is deleted or mutated in the majority of small cell lung carcinomas.

The gene responsible for neurofibromatosis (NF-1) is also a tumor suppressor gene. The absence of the gene causes the formation of tumors in neural crest-derivative cells. The protein product of this gene is related to the GTPase-activating proteins (GAP) that inhibit G-proteins (like *ras*). The GAP proteins inhibit the G-proteins by increasing the hydrolysis of GTP bound to the G-protein, thereby converting the G-protein to its inactive state.

Another important tumor suppressor gene is the DCC gene. This gene was found because LOH in the region containing the gene was noted for many colon carcinomas. This gene codes for a large membrane-bound protein that is homologous to other cell-adhesion molecules. It is thought to play a role in cell-to-cell communication.

Finally, p53 is a tumor suppressor gene that can also act as an oncogene. This gene is mutated in patients with the Li Fraumeni multiple carcinoma syndrome. It is also the most frequently mutated gene in human carcinomas. This gene codes for a nuclear protein that was initially thought to be an oncogene since tumor cells were found to have increased levels of the protein, and the protein was capable of transforming cells in culture. However, it has been learned that the increased levels of protein result from the abnormal p53 protein binding with the normal proteins and preventing the normal proteins' movement into the nucleus. Mutant forms of the protein can transform cells by inactivating the nonmutant forms of p53 by binding to them. Normal p53 inhibits cell growth and thus is a tumor suppressor. A mutant allele inactivating its normal counterpart is termed dominant negative.

Progression from a normal cell to a malignant tumor is now recognized as a multistep process. Mutations of

oncogenes and tumor suppressors accumulate in a population of cells, and those with the highest growth rates are most likely to exhibit further mutations. Currently, the best model of this process occurs in colon carcinoma, in which many (60%) large polyps are found to have a mutated *ras* gene, later lesions are more likely to contain deletions of DCC, and p53 is then often deleted or mutated. Other oncogenes and tumor suppressors also play a role, but each tumor may acquire mutations in a different order or in different genes.

In summary, changes in DNA result from spontaneous mutations, chromosomal aberrations, or viral effects. The altered DNA ultimately results in altered proteins. The oncogene protein products act in a dominant manner and generally stimulate cell growth by activating the normal cell machinery for growth without regard to upstream signals. The tumor suppressor genes act in a recessive manner and limit cell growth until they are inactivated. Mutations in the oncogenes and tumor suppressor genes act synergistically to propel a clone of cells toward neoplasia. The cells that are growing rapidly are most prone to further mutations and a continuation of the process of neoplasia.

ANNOTATED BIBLIOGRAPHY

1. Science 1991; 254 (5035).

This issue of *Science* has seven excellent reviews of topics such as tumor suppressors, viruses, growth factors, and chromosomal aberrations. They are readable and comprehensive. If at all possible, read this issue.

2. Dalgleish AG. Viruses and cancer. Br Med Bull 1991; 47: 21–46.

This is a rather comprehensive review of viruses and cancer. It discusses in much greater detail how viruses are implicated in causing cancer. It is clearly written and is targeted for the clinician.

3. Marshall CJ. Tumor suppressor genes. Cell 1991; 64: 313–26.

This article goes into further detail regarding interactions and characteristics of the known tumor suppressor genes, with emphasis on the techniques used to identify them and on their mechanisms of action.

GLOSSARY

Allele: One member of a pair of genes for a given trait. Each member is of maternal or paternal origin.

Amplify: To increase the quantity of a specific gene by a variety of techniques.

Aneuploid: A cell containing extra chromosomes.

Band: A pattern of light and dark regions by Giemsa staining that can serve as landmarks on chromosomes.

Base pairing: The pairing of specific nitrogenous bases between complementary strands of DNA. For example, adenine is always paired with thymine and guanine with cytosine.

cDNA (complementary DNA): Single- or double-stranded DNA made from an RNA template using the enzyme reverse transcriptase.

Centimorgan (cM): A measure of the statistical probability of recombination between alleles. One cM represents a 1% chance of recombination per meiotic event.

Centromere: The constricted region of a chromosome separating the short and long arms from one another.

Chromosome: A single, linear, highly condensed DNA molecule.

Clone (noun): One of a collection of cells or vectors containing identical genetic material.

Clone (verb): The act of duplicating genetic material within a vector.

Codon: A group of three consecutive nucleotides within messenger RNA (mRNA) that encodes one of 20 amino acids or encodes a message to stop translation (see **Translation**).

Cosmid: A vector that incorporates components of plasmids and phage to carry larger clones (up to 40 kilobases).

Diploid: Cells containing copies of both the maternal and paternal chromosomes.

DNA (deoxyribonucleic acid): The molecule responsible for storing and transmitting genetic information; composed of two strands of nucleotides twisted around each other in the shape of a double helix.

Double helix: The twisted double-strand shape assumed by DNA.

Exon: A contiguous segment of genomic DNA that is translated into polypeptide (see **Intron**).

Familial: An inherited trait.

Flow cytometry: Method used to measure nuclear DNA quantity in order to determine ploidy status.

Gene: A segment of DNA within a chromosome encoding a single protein.

Genome: The entire complement of genetic material in the form of DNA for a given organism.

Genetic map: The ordering of genes by the statistical determination of recombination events between them. Genes separated by greater distances are more likely to recombine.

Heterozygous: An individual containing dissimilar alleles for a given gene or locus.

Homozygous: An individual containing identical alleles for a given gene or locus.

Hybridization: The alignment of complementary strands of DNA (or RNA) via base pairing, widely used to identify portions of DNA on a Southern (or Northern) blot, using labeled probes.

Intron: A noncoding sequence of DNA within the gene (see **Exon**).

Karyotype: The physical appearance of the full complement of stained chromosomes for an individual.

Locus: A site on a segment of DNA.

Long arm (q): One of the two prominent segments of a chromosome; the short or "p" arm is the other. The arms of a given chromosome join at its centromere.

Library: A collection of recombinant genes cloned into a vector.

Linkage: A measure of proximity between two alleles determined by recombination events. If they are not linked, they are on separate chromosomes; if loosely

linked, they are distant to each other on the same chromosome. The closer they are to one another, the more tightly they are linked.

mRNA (messenger RNA): The single-stranded edited copy of a gene ultimately translated into protein.

Northern blot: The transfer of size-separated RNA fragments to a synthetic membrane for further studies.

Nucleotide: One of the four building blocks of DNA (dATP, dGTP, dCTP, or dTTP) or RNA (ATP, CTP, GTP, or ATP) that are combined to form the nucleic acids.

Oncogene: A cancer-inducing gene.

Phage: A virus that infects a bacterial host, used in the laboratory as a cloning vector.

Physical map: Analysis of the distance, in base pairs, between loci.

Plasmid: Circularized DNA fragment, distinct from genomic DNA, found within bacteria, used as a cloning vector or to alter characteristics of the bacteria.

Polymerase chain reaction (PCR): An efficient, simple, and rapid technique to multiply a length of DNA in a test tube.

Promoter site: Region of a DNA molecule found in front of a gene that controls the expression of the gene.

Proto-oncogene: Normal gene that may become an oncogene; also called cellular oncogene.

Recombinant DNA: The combination of foreign DNA inserts with vector DNA (e.g., plasmid, phage, or cosmid) to produce a clone within a host.

Recombination: The rearrangement of DNA by breaking and re-ligations of the DNA strands; also called crossovers.

Replicon: A sequence in the DNA that initiates replication.

Restriction mapping: The creation of a physical map by ordering enzymatically cut DNA fragments.

Restriction endonuclease: Enzyme, isolated from bacteria, that recognizes specific base-pair sequences to cut DNA. The sites vary from frequent to rare cutting, depending upon the length of the restriction site.

Restriction fragment length polymorphism (RFLP): Variation in the distance between restriction enzyme cleavage sites that exist within a population producing unique DNA fingerprint patterns.

RNA (ribonucleic acid): Single-strand nucleic acid found mainly in the nucleolus and ribosomes; contains ribose sugar and uracil, whereas DNA contains thymine.

Short arm (p): One of the two prominent segments of a chromosome; the long or "q" arm is the other. The arms of a given chromosome join at its centromere.

Southern blot: The transfer of size-separated DNA fragments to a synthetic membrane for further studies, initially described by E.N. Southern.

Sporadic: Not familial.

Sticky ends: The cut pieces of a single strand of DNA, which, because of base pairing, can be made to rejoin again at complementary base pairs, using the enzyme DNA ligase.

Telomere: The distal ends of the chromosome.

Transcription: The copying of DNA into messenger RNA.

Transduction: The incorporation of a cellular gene in a viral genome that can then be introduced into other cells.

Transfection: The process of placing foreign DNA into mammalian cells.

Transformation: The cancerous alteration of mammalian cells; also the act of putting foreign DNA into bacteria.

Translation: The process of converting the genetic code into polypeptides. mRNA codons are recognized by tRNA anti-codons. Each tRNA codes for a single amino acid.

Tumor suppressor: A gene that prevents tumor formation until deleted or mutated.

Vector: A construct used to propagate DNA in a host (bacteria, yeast, or cultured cells) (see **Plasmid**, **Phage**, and/or **Yeast Artificial Chromosome**).

Virion: A replication virus particle.

Western analysis: Protein electrophoresis characterizing size of unknown protein.

Yeast artificial chromosome (YAC): A vector used in yeast that can propagate large fragments of DNA.