18

Altering the Genetic Message

Concept Outline

18.1 Mutations are changes in the genetic message.

Mutations Are Rare But Important. Changes in genes provide the raw material for evolution.
Kinds of Mutation. Some mutations alter genes themselves, others alter the positions of genes.
Point Mutations. Radiation damage or chemical modification can change one or a few nucleotides.
Changes in Gene Position. Chromosomal rearrangement and insertional inactivation reflect changes in gene position.

18.2 Cancer results from mutation of growth-regulating genes.

What Is Cancer? Cancer is a growth disorder of cells.
Kinds of Cancer. Cancer occurs in almost all tissues, but more in some than others.
Some Tumors Are Caused by Chemicals. Chemicals that mutate DNA cause cancer.
Other Tumors Result from Viral Infection. Viruses carrying growth-regulating genes can cause cancer.
Cancer and the Cell Cycle. Cancer results from mutation of genes regulating cell proliferation
Smoking and Cancer. Smoking causes lung cancer.
Curing Cancer. New approaches offer promise of a cure.

18.3 Recombination alters gene location.

An Overview of Recombination. Recombination is produced by gene transfer and by reciprocal recombination.
Gene Transfer. Many genes move within small circles of DNA called plasmids. Plasmids can move between bacterial cells and carry bacterial genes. Some gene sequences move from one location to another on a chromosome.
Reciprocal Recombination. Reciprocal recombination can alter genes in several ways.
Trinucleotide Repeats. Increases in the number of repeated triplets can produce gene disorders.

18.4 Genomes are continually evolving.

Classes of Eukaryotic DNA. Unequal crossing over expands eukaryotic genomes.

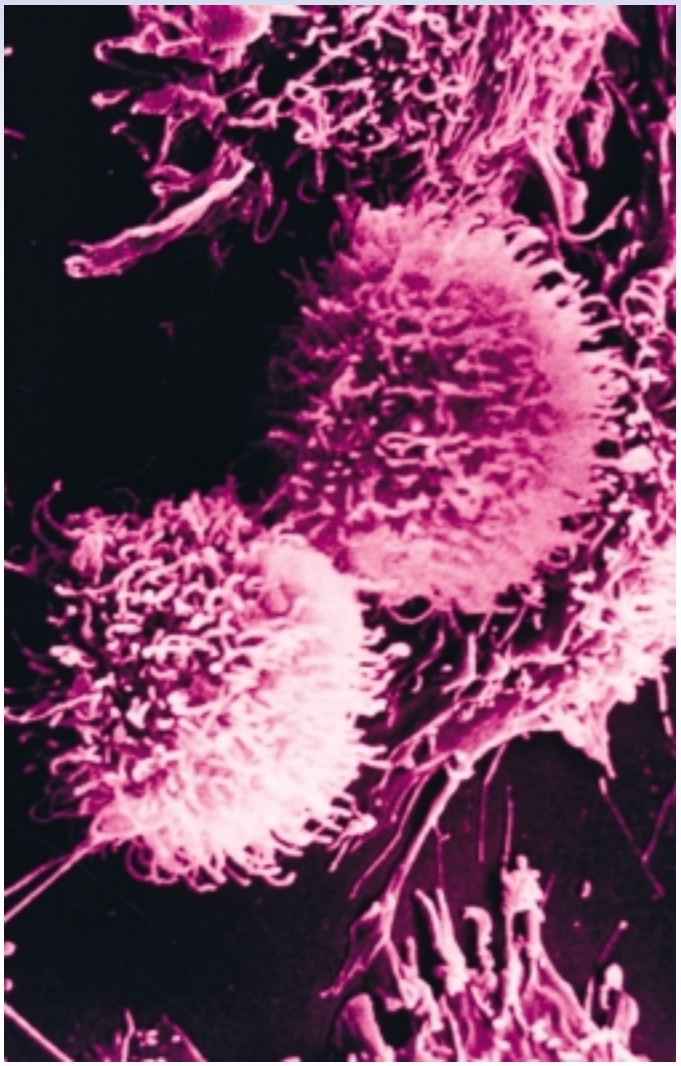


FIGURE 18.1
Cancer. A scanning electron micrograph of deadly cancer cells (8000×).

In general, the genetic message can be altered in two broad ways: mutation and recombination. A change in the content of the genetic message—the base sequence of one or more genes—is referred to as a mutation. Some mutations alter the identity of a particular nucleotide, while others remove or add nucleotides to a gene. A change in the position of a portion of the genetic message is referred to as recombination. Some recombination events move a gene to a different chromosome; others alter the location of only part of a gene. In this chapter, we will first consider gene mutation, using cancer as a focus for our inquiry (figure 18.1). Then we will turn to recombination, focusing on how it has affected the organization of the eukaryotic genome.

18.1 Mutations are changes in the genetic message.

Mutations Are Rare But Important

The cells of eukaryotes contain an enormous amount of DNA. If the DNA in all of the cells of an adult human were lined up end-to-end, it would stretch nearly 100 billion kilometers—60 times the distance from Earth to Jupiter! The DNA in any multicellular organism is the final result of a long series of replications, starting with the DNA of a single cell, the fertilized egg. Organisms have evolved many different mechanisms to avoid errors during DNA replication and to preserve the DNA from damage. Some of these mechanisms "proofread" the replicated DNA strands for accuracy and correct any mistakes. The proofreading is not perfect, however. If it were, no variation in the nucleotide sequences of genes would be generated.

FIGURE 18.2
Mutation. Normal fruit flies have one pair of wings extending from the thorax. This fly is a mutant because of changes in *bithorax*, a gene regulating a critical stage of development; it possesses two thoracic segments and thus two sets of wings.

Mistakes Happen

In fact, cells do make mistakes during replication, and damage to the genetic message also occurs, causing mutation (figure 18.2). However, change is rare. Typically, a particular gene is altered in only one of a million gametes. If changes were common, the genetic instructions encoded in DNA would soon degrade into meaningless gibberish. Limited as it might seem, the steady trickle of change that does occur is the very stuff of evolution. Every difference in the genetic messages that specify different organisms arose as the result of genetic change.

The Importance of Genetic Change

All evolution begins with alterations in the genetic message: mutation creates new alleles, gene transfer and transposition alter gene location, reciprocal recombination shuffles and sorts these changes, and chromosomal rearrangement alters the organization of entire chromosomes. Some changes in germ-line tissue produce alterations that enable an organism to leave more offspring, and those changes tend to be preserved as the genetic endowment of future generations. Other changes reduce the ability of an organism to leave offspring. Those changes tend to be lost, as the organisms that carry them contribute fewer members to future generations.

Evolution can be viewed as the selection of particular combinations of alleles from a pool of alternatives. The rate of evolution is ultimately limited by the rate at which these alternatives are generated. Genetic change through mutation and recombination provides the raw material for evolution.

Genetic changes in somatic cells do not pass on to offspring, and so have less evolutionary consequence than germ-line change. However, changes in the genes of somatic cells can have an important immediate impact, particularly if the gene affects development or is involved with regulation of cell proliferation.

Rare changes in genes, called mutations, can have significant effects on the individual when they occur in somatic tissue, but are only inherited if they occur in germ-line tissue. Inherited changes provide the raw material for evolution.

Kinds of Mutation

Because mutations can occur randomly anywhere in a cell's DNA, mutations can be detrimental, just as making a random change in a computer program or a musical score usually worsens performance. The consequences of a detrimental mutation may be minor or catastrophic, depending on the function of the altered gene.

Mutations in Germ-Line Tissues

The effect of a mutation depends critically on the identity of the cell in which the mutation occurs. During the embryonic development of all multicellular organisms, there comes a point when cells destined to form gametes **(germline cells)** are segregated from those that will form the other cells of the body **(somatic cells).** Only when a mutation occurs within a germ-line cell is it passed to subsequent generations as part of the hereditary endowment of the gametes derived from that cell.

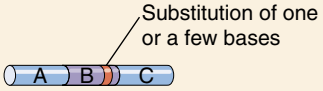
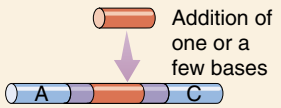
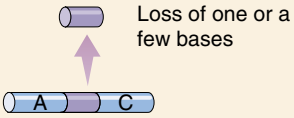
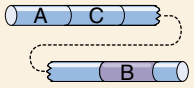
Mutations in Somatic Tissues

Mutations in germ-line tissue are of enormous biological importance because they provide the raw material from which natural selection produces evolutionary change. Change can occur only if there are new, different allele combinations available to replace the old. Mutation produces new alleles, and recombination puts the alleles together in different combinations. In animals, it is the occurrence of these two processes in germ-line tissue that is important to evolution, as mutations in somatic cells **(somatic mutations)** are not passed from one generation to the next. However, a somatic mutation may have drastic effects on the individual organism in which it occurs, as it *is* passed on to all of the cells that are descended from the original mutant cell. Thus, if a mutant lung cell divides, all cells derived from it will carry the mutation. Somatic mutations of lung cells are, as we shall see, the principal cause of lung cancer in humans.

Point Mutations

One category of mutational changes affects the message itself, producing alterations in the sequence of DNA nucleotides (table 18.1 summarizes the sources and types of mutations). If alterations involve only one or a few base-pairs in the coding sequence, they are called **point mutations.** While some point mutations arise due to spontaneous pairing errors that occur during DNA replication, others result from damage to the DNA caused by **mutagens,** usually radiation or chemicals. The latter class of mutations is of particular practical importance because modern industrial societies often release many chemical mutagens into the environment.

Table 18.1 Types of Mutation

Mutation	Example result
NO MUTATION	
	Normal B protein is produced by the *B* gene.
POINT MUTATION	
Base substitution 	B protein is inactive because changed amino acid disrupts function.
Insertion 	B protein is inactive because inserted material disrupts proper shape.
Deletion 	B protein is inactive because portion of protein is missing.
CHANGES IN GENE POSITION	
Transposition 	*B* gene or B protein may be regulated differently because of change in gene position.
Chromosomal rearrangement 	*B* gene may be inactivated or regulated differently in its new location on chromosome.

Changes in Gene Position

Another category of mutations affects the way the genetic message is organized. In both bacteria and eukaryotes, individual genes may move from one place in the genome to another by **transposition.** When a particular gene moves to a different location, its expression or the expression of neighboring genes may be altered. In addition, large segments of chromosomes in eukaryotes may change their relative locations or undergo duplication. Such **chromosomal rearrangements** often have drastic effects on the expression of the genetic message.

Point mutations are changes in the hereditary message of an organism. They may result from spontaneous errors during DNA replication or from damage to the DNA due to radiation or chemicals.

Point Mutations

Physical Damage to DNA

Ionizing Radiation. High-energy forms of radiation, such as X rays and gamma rays, are highly mutagenic. When such radiation reaches a cell, it is absorbed by the atoms it encounters, imparting energy to the electrons in their outer shells. These energized electrons are ejected from the atoms, leaving behind free radicals, ionized atoms with unpaired electrons. Free radicals react violently with other molecules, including DNA.

When a free radical breaks *both* phosphodiester bonds of a DNA helix, causing a **double-strand break,** the cell's usual mutational repair enzymes cannot fix the damage. The two fragments created by the break must be aligned while the phosphodiester bonds between them form again. Bacteria have no mechanism to achieve this alignment, and double-strand breaks are lethal to their descendants. In eukaryotes, which almost all possess multiple copies of their chromosomes, the synaptonemal complex assembled in meiosis is used to pair the fragmented chromosome with its homologue. In fact, it is speculated that meiosis may have evolved initially as a mechanism to repair double-strand breaks in DNA (see chapter 12).

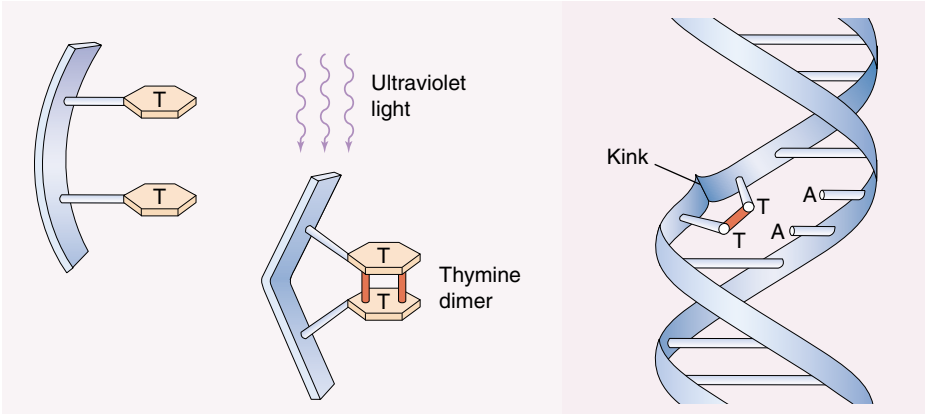


FIGURE 18.3
Making a pyrimidine dimer. When two pyrimidines, such as two thymines, are adjacent to each other in a DNA strand, the absorption of UV radiation can cause covalent bonds to form between them—creating a pyrimidine dimer. The dimer introduces a "kink" into the double helix that prevents replication of the duplex by DNA polymerase.

Ultraviolet Radiation. Ultraviolet (UV) radiation, the component of sunlight that tans (and burns), contains much less energy than ionizing radiation. It does not induce atoms to eject electrons, and thus it does not produce free radicals. The only molecules capable of absorbing UV radiation are certain organic ring compounds, whose outer-shell electrons become reactive when they absorb UV energy.

DNA strongly absorbs UV radiation in the pyrimidine bases, thymine and cytosine. If one of the nucleotides on either side of the absorbing pyrimidine is also a pyrimidine, a double covalent bond forms between them. The resulting cross-link between adjacent pyrimidines is called a **pyrimidine dimer** (figure 18.3). In most cases, cellular UV repair systems either cleave the bonds that link the adjacent pyrimidines or excise the entire pyrimidine dimer from the strand and fill in the gap, using the other strand as a template (figure 18.4). In those rare instances in which a pyrimidine dimer goes unrepaired, DNA polymerase may fail to replicate the portion of the strand that includes the dimer, skipping ahead and leaving the problem area to be filled in later. This filling-in process is often error-prone, however, and it may create mutational changes in the base sequence of the gap region. Some unrepaired pyrimidine dimers block DNA replication altogether, which is lethal to the cell.

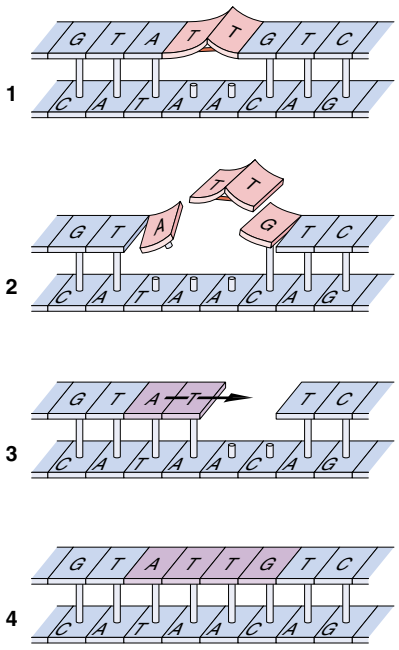


FIGURE 18.4
Repair of a pyrimidine dimer. Some pyrimidine dimers are repaired by excising the dimer, as well as a short run of nucleotides on either side of it, and then filling in the gap using the other strand as a template.

Sunlight can wreak havoc on the cells of the skin because its UV light causes mutations. Indeed, a strong and direct association exists between exposure to bright sunlight, UV-induced DNA damage, and skin cancer. A deep tan is *not* healthy! A rare hereditary disorder among humans called **xeroderma pigmentosum** causes these problems after a lesser exposure to UV. Individuals with this disorder develop extensive skin tumors after exposure to sunlight because they lack a mechanism for repairing the DNA damage UV radiation causes. Because of the many different proteins involved in excision and repair of pyrimidine dimers, mutations in as many as eight different genes cause the disease.

Chemical Modification of DNA

Many mutations result from direct chemical modification of the DNA. The chemicals that act on DNA fall into three classes: (1) chemicals that resemble DNA nucleotides but pair incorrectly when they are incorporated into DNA (figure 18.5). Some of the new AIDS chemotherapeutic drugs are analogues of nitrogenous bases that are inserted into the viral or infected cell DNA. This DNA cannot be properly transcribed, so viral growth slows; (2) chemicals that remove the amino group from adenine or cytosine, causing them to mispair; and (3) chemicals that add hydrocarbon groups to nucleotide bases, also causing them to mispair. This last group includes many particularly potent mutagens commonly used in laboratories, as well as compounds sometimes released into the environment, such as mustard gas.

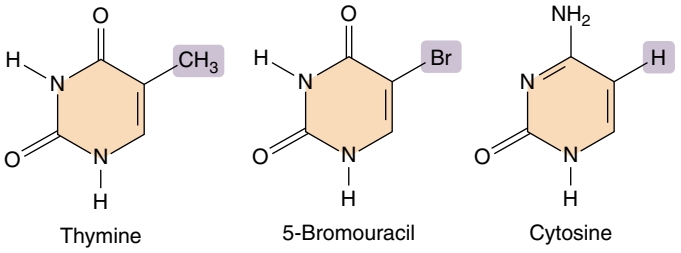


FIGURE 18.5
Chemicals that resemble DNA bases can cause mutations. For example, DNA polymerase cannot distinguish between thymine and 5-bromouracil, which are similar in shape. Once incorporated into a DNA molecule, however, 5-bromouracil tends to rearrange to a form that resembles cytosine and pairs with guanine. When this happens, what was originally an A-T base-pair becomes a G-C base-pair.

Spontaneous Mutations

Many point mutations occur spontaneously, without exposure to radiation or mutagenic chemicals. Sometimes nucleotide bases spontaneously shift to alternative conformations, or isomers, which form different kinds of hydrogen bonds than the normal conformations. During replication, DNA polymerase pairs a different nucleotide with the isomer than it would have otherwise selected. Unrepaired spontaneous errors occur in fewer than one in a billion nucleotides per generation, but they are still an important source of mutation.

Sequences sometimes misalign when homologous chromosomes pair, causing a portion of one strand to loop out. These misalignments, called **slipped mispairing,** are usually only transitory, and the chromosomes quickly revert to the normal arrangement (figure 18.6). If the error-correcting system of the cell encounters a slipped mispairing before it reverts, however, the system will attempt to "correct" it, usually by excising the loop. This may result in a **deletion** of several hundred nucleotides from one of the chromosomes. Many of these deletions start or end in the middle of a codon, thereby shifting the reading frame by one or two bases. These so-called **frame-shift mutations** cause the gene to be read in the wrong three-base groupings, distorting the genetic message, just as the deletion of the letter F from the sentence, THE FAT CAT ATE THE RAT shifts the reading frame of the sentence, producing the meaningless message, THE ATC ATA TET HER AT. Some chemicals specifically promote deletions and frame-shift mutations by stabilizing the loops produced during slipped mispairing, thus increasing the time the loops are vulnerable to excision.

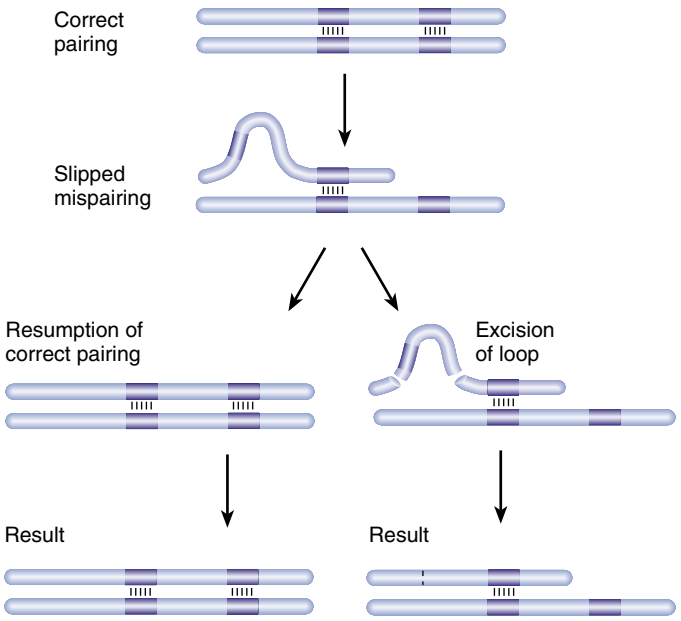


FIGURE 18.6
Slipped mispairing. Slipped mispairing occurs when a sequence is present in more than one copy on a chromosome and the copies on homologous chromosomes pair out of register, like a shirt buttoned wrong. The loop this mistake produces is sometimes excised by the cell's repair enzymes, producing a short deletion and often altering the reading frame. Any chemical that stabilizes the loop increases the chance it will be excised.

The major sources of physical damage to DNA are ionizing radiation, which breaks the DNA strands; ultraviolet radiation, which creates nucleotide cross-links whose removal often leads to errors in base selection; and chemicals that modify DNA bases and alter their base-pairing behavior. Unrepaired spontaneous errors in DNA replication occur rarely.

Changes in Gene Position

Chromosome location is an important factor in determining whether genes are transcribed. Some genes cannot be transcribed if they are adjacent to a tightly coiled region of the chromosome, even though the same gene can be transcribed normally in any other location. Transcription of many chromosomal regions appears to be regulated in this manner; the binding of specific proteins regulates the degree of coiling in local regions of the chromosome, determining the accessibility RNA polymerase has to genes located within those regions.

Chromosomal Rearrangements

Chromosomes undergo several different kinds of gross physical alterations that have significant effects on the locations of their genes. The two most important are **translocations,** in which a segment of one chromosome becomes part of another chromosome, and **inversions,** in which the orientation of a portion of a chromosome is reversed. Translocations often have significant effects on gene expression. Inversions, on the other hand, usually do not alter gene expression, but they are nonetheless important. Recombination within a region that is inverted on one homologue but not the other (figure 18.7) leads to serious problems: none of the gametes that contain chromatids produced following such a crossover event will have a complete set of genes.

Other chromosomal alterations change the number of gene copies an individual possesses. Particular genes or segments of chromosomes may be deleted or duplicated, whole chromosomes may be lost or gained *(aneuploidy)*, and entire sets of chromosomes may be added *(polyploidy)*. Most deletions are harmful because they halve the number of gene copies within a diploid genome and thus seriously affect the level of transcription. Duplications cause gene imbalance and are also usually harmful.

Insertional Inactivation

Many small segments of DNA are capable of moving from one location to another in the genome, using an enzyme to cut and paste themselves into new genetic neighborhoods. We call these mobile bits of DNA transposable elements, or **transposons.** Transposons select their new locations at random, and are as likely to enter one segment of a chromosome as another. Inevitably, some transposons end up inserted into genes, and this almost always inactivates the gene. The encoded protein now has a large meaningless chunk inserted within it, disrupting its structure. This form of mutation, called **insertional inactivation,** is common in nature. Indeed, it seems to be one of the most significant causes of mutation. The original white-eye mutant of *Drosophila* discovered by Morgan (see chapter 13) is the result of a transposition event, a transposon nested within a gene encoding a pigment-producing enzyme.

As you might expect, a variety of human gene disorders are the result of transposition. The human transposon called *Alu*, for example, is responsible for an X-linked hemophilia, inserting into clotting factor IX and placing a premature stop codon there. It also causes inherited high levels of cholesterol (hypercholesterolemia), *Alu* elements inserting into the gene encoding the low density lipoprotein (LDL) receptor. In one very interesting case, a *Drosophila* transposon called *Mariner* proves responsible for a rare human neurological disorder called Charcot-Marie-Tooth disease, in which the muscles and nerves of the legs and feet gradually wither away. The Mariner transposon is inserted into a key gene called *CMT* on chromosome 17, creating a weak site where the chromosome can break. No one knows how the *Drosophila* transposon got into the human genome.

Many mutations result from changes in gene location or from insertional inactivation.

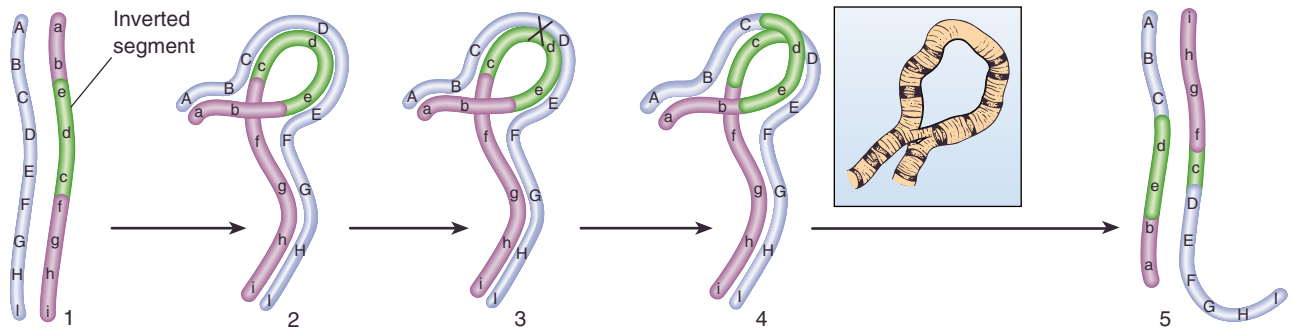


FIGURE 18.7
The consequence of inversion. (*1*) When a segment of a chromosome is inverted, (*2*) it can pair in meiosis only by forming an internal loop. (*3*) Any crossing over that occurs within the inverted segment during meiosis will result in nonviable gametes; some genes are lost from each chromosome, while others are duplicated (*4* and *5*). For clarity, only two strands are shown, although crossing over occurs in the four-strand stage. The pairing that occurs between inverted segments is sometimes visible under the microscope as a characteristic loop (*inset*).

18.2 Cancer results from mutation of growth-regulating genes.

What Is Cancer?

Cancer is a growth disorder of cells. It starts when an apparently normal cell begins to grow in an uncontrolled and invasive way (figure 18.8). The result is a cluster of cells, called a **tumor,** that constantly expands in size. Cells that leave the tumor and spread throughout the body, forming new tumors at distant sites, are called **metastases** (figure 18.9). Cancer is perhaps the most pernicious disease. Of the children born in 1999, one-third will contract cancer at some time during their lives; one-fourth of the male children and one-third of the female children will someday die of cancer. Most of us have had family or friends affected by the disease. In 1997, 560,000 Americans died of cancer.

Not surprisingly, researchers are expending a great deal of effort to learn the cause of this disease. Scientists have made a great deal of progress in the last 20 years using molecular biological techniques, and the rough outlines of understanding are now emerging. We now know that cancer is a gene disorder of somatic tissue, in which damaged genes fail to properly control cell proliferation. The cell division cycle is regulated by a sophisticated group of proteins described in chapter 11. Cancer results from the mutation of the genes encoding these proteins.

Cancer can be caused by chemicals that mutate DNA or in some instances by viruses that circumvent the cell's normal proliferation controls. Whatever the immediate cause, however, all cancers are characterized by unrestrained growth and division. Cell division never stops in a cancerous line of cells. Cancer cells are virtually immortal—until the body in which they reside dies.

Cancer is unrestrained cell proliferation caused by damage to genes regulating the cell division cycle.

FIGURE 18.8
Lung cancer cells (530×). These cells are from a tumor located in the alveolus (air sac) of a lung.

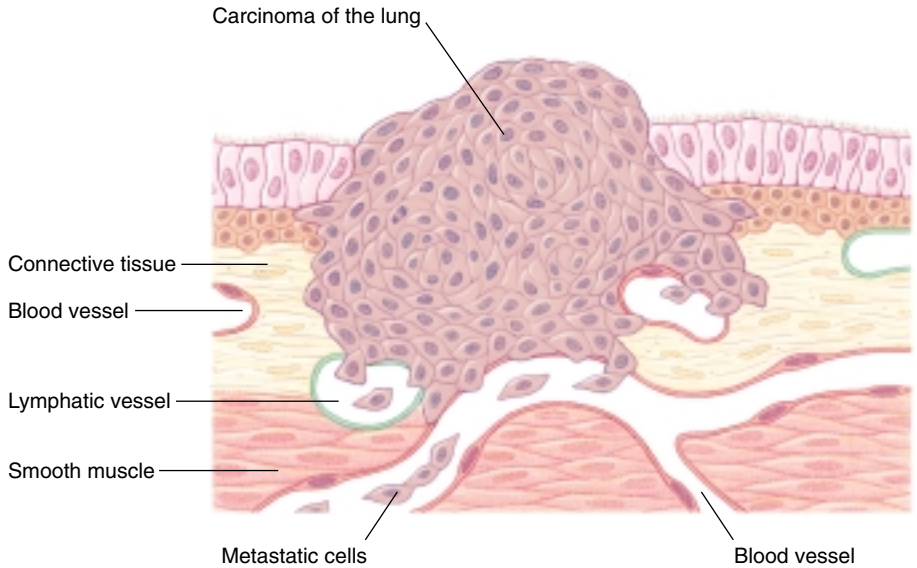


FIGURE 18.9
Portrait of a cancer. This ball of cells is a carcinoma (cancer tumor) developing from epithelial cells that line the interior surface of a human lung. As the mass of cells grows, it invades surrounding tissues, eventually penetrating lymphatic and blood vessels, both plentiful within the lung. These vessels carry metastatic cancer cells throughout the body, where they lodge and grow, forming new masses of cancerous tissue.

Kinds of Cancer

Cancer can occur in almost any tissue, so a bewildering number of different cancers occur. Tumors arising from cells in connective tissue, bone, or muscle are known as **sarcomas,** while those that originate in epithelial tissue such as skin are called **carcinomas.** In the United States, the three deadliest human cancers are lung cancer, cancer of the colon and rectum, and breast cancer (table 18.2). Lung cancer, responsible for the most cancer deaths, is largely preventable; most cases result from smoking cigarettes. Colorectal cancers appear to be fostered by the high-meat diets so favored in the United States. The cause of breast cancer is still a mystery, although in 1994 and 1995 researchers isolated two genes responsible for hereditary susceptibility to breast cancer, *BRCA1* and *BRCA2* (Breast Cancer genes #1 and #2 located on human chromosomes 17 and 13); their discovery offers hope that researchers will soon be able to unravel the fundamental mechanism leading to hereditary breast cancer, about one-third of all breast cancers.

Table 18.2 Incidence of Cancer in the United States in 1999

Type of Cancer	New Cases	Deaths	% of Cancer Deaths
Lung	171,600	158,900	28
Colon and rectum	129,400	56,600	10
Leukemia/lymphoma	94,200	49,100	9
Breast	176,300	43,700	8
Prostate	179,300	37,000	7
Pancreas	28,600	28,600	5
Ovary	25,200	14,500	3
Stomach	21,900	13,500	2
Liver	14,500	13,600	2
Nervous system/eye	19,000	13,300	2
Bladder	54,200	12,100	2
Oral cavity	29,800	8,100	2
Kidney	30,000	11,900	2
Cervix/uterus	50,200	11,200	2
Malignant melanoma	44,200	7,300	1
Sarcoma (connective tissue)	10,400	5,800	1
All other cancers	143,000	77,900	14

In the United States in 1999 there were 1,221,800 reported cases of new cancers and 563,100 cancer deaths, indicating that roughly half the people who develop cancer die from it.

Source: Data from the American Cancer Society, Inc., 1999.

The association of particular chemicals with cancer, particularly chemicals that are potent mutagens, led researchers early on to the suspicion that cancer might be caused, at least in part, by chemicals, the so-called **chemical carcinogenesis theory.** Agents thought to cause cancer are called **carcinogens.** A simple and effective way to test if a chemical is mutagenic is the Ames test (figure 18.10), named for its developer, Bruce Ames. The test uses a strain of *Salmonella* bacteria that has a defective histidine-synthesizing gene. Because these bacteria cannot make histidine, they cannot grow on media without it. Only a back-mutation that restores the ability to manufacture histidine will permit growth. Thus the number of colonies of these bacteria that grow on histidine-free medium is a measure of the frequency of back-mutation. A majority of chemicals that cause back-mutations in this test are carcinogenic, and vice versa. To increase the sensitivity of the test, the strains of bacteria are altered to disable their DNA repair machinery. The search for the cause of cancer has focused in part on chemical carcinogens and other environmental factors, including ionizing radiation such as X rays (figure 18.11).

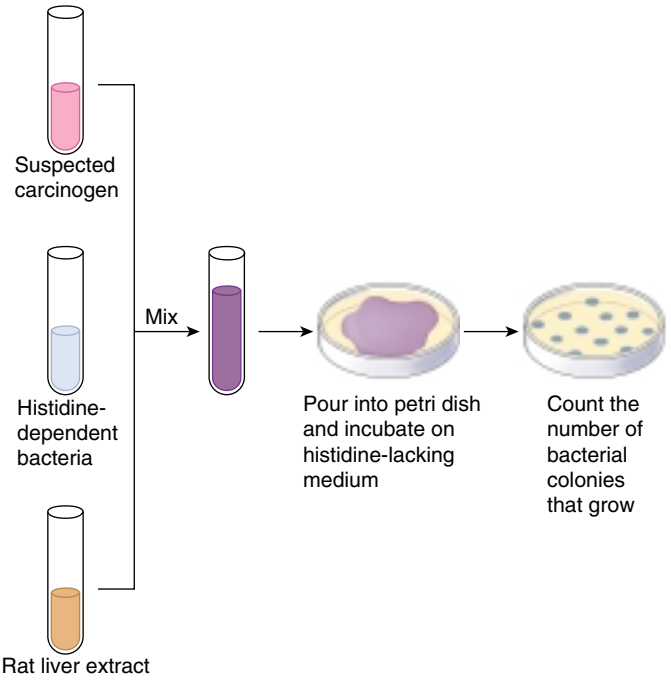


FIGURE 18.10
The Ames test. This test is uses a strain of *Salmonella* bacteria that requires histidine in the growth medium due to a mutated gene. If a suspected carcinogen is mutagenic, it can reverse this mutation. Rat liver extract is added because it contains enzymes that can convert carcinogens into mutagens. The mutagenicity of the carcinogen can be quantified by counting the number of bacterial colonies that grow on a medium lacking histidine.

Cancers occur in all tissues, but are more common in some than others.

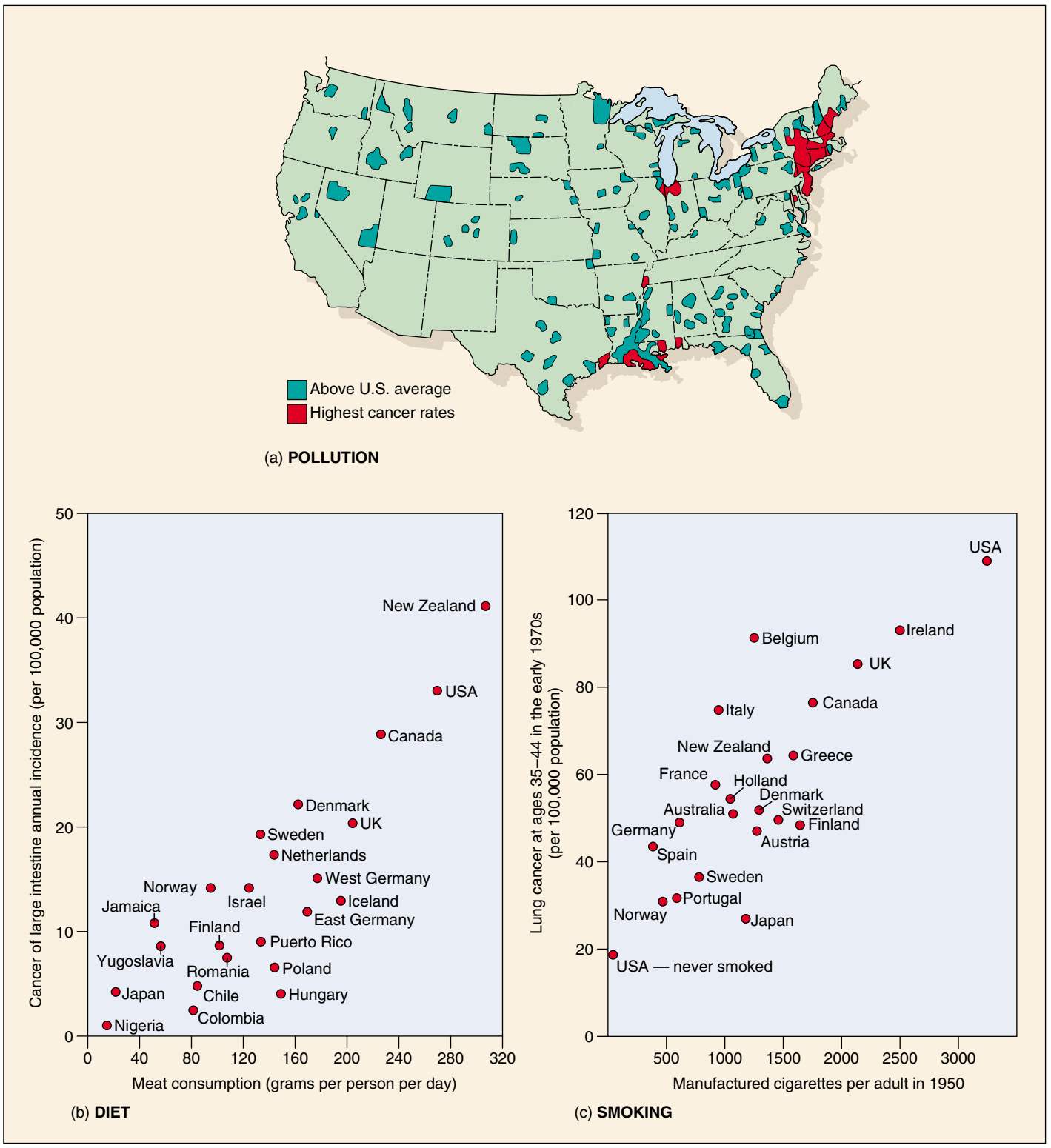


FIGURE 18.11
Potential cancer-causing agents. (*a*) The incidence of cancer per 1000 people is not uniform throughout the United States. The incidence is higher in cities and in the Mississippi Delta, suggesting that pollution and pesticide runoff may contribute to the development of cancer. (*b*) One of the deadliest cancers in the United States, cancer of the large intestine, is uncommon in many other countries. Its incidence appears to be related to the amount of meat a person consumes: a high-meat diet slows the passage of food through the intestine, prolonging exposure of the intestinal wall to digestive waste. (*c*) The biggest killer among cancers is lung cancer, and the most deadly environmental agent producing lung cancer is cigarette smoke. The incidence of lung cancer among men 35 to 44 years of age in various countries strongly correlates with the cigarette consumption in that country 20 years earlier.

Some Tumors Are Caused by Chemicals

Early Ideas

The chemical carcinogenesis theory was first advanced over 200 years ago in 1761 by Dr. John Hill, an English physician, who noted unusual tumors of the nose in heavy snuff users and suggested tobacco had produced these cancers. In 1775, a London surgeon, Sir Percivall Pott, made a similar observation, noting that men who had been chimney sweeps exhibited frequent cancer of the scrotum, and suggesting that soot and tars might be responsible. British sweeps washed themselves infrequently and always seemed covered with soot. Chimney sweeps on the continent, who washed daily, had much less of this scrotal cancer. These and many other observations led to the hypothesis that cancer results from the action of chemicals on the body.

Demonstrating That Chemicals Can Cause Cancer

It was over a century before this hypothesis was directly tested. In 1915, Japanese doctor Katsusaburo Yamagiwa applied extracts of coal tar to the skin of 137 rabbits every 2 or 3 days for 3 months. Then he waited to see what would happen. After a year, cancers appeared at the site of application in seven of the rabbits. Yamagiwa had induced cancer with the coal tar, the first direct demonstration of chemical carcinogenesis. In the decades that followed, this approach demonstrated that many chemicals were capable of causing cancer. Importantly, most of them were potent mutagens.

Because these were lab studies, many people did not accept that the results applied to real people. Do tars in fact induce cancer in humans? In 1949, the American physician Ernst Winder and the British epidemiologist Richard Doll independently reported that lung cancer showed a strong link to the smoking of cigarettes, which introduces tars into the lungs. Winder interviewed 684 lung cancer patients and 600 normal controls, asking whether each had ever smoked. Cancer rates were 40 times higher in heavy smokers than in nonsmokers. Doll's study was even more convincing. He interviewed a large number of British physicians, noting which ones smoked, then waited to see which would develop lung cancer. Many did. Overwhelmingly, those who did were smokers. From these studies, it seemed likely as long as 50 years ago that tars and other chemicals in cigarette smoke induce cancer in the lungs of persistent smokers. While this suggestion was (and is) resisted by the tobacco industry, the evidence that has accumulated since these pioneering studies makes a clear case, and there is no longer any real doubt. Chemicals in cigarette smoke cause cancer.

Carcinogens Are Common

In ongoing investigations over the last 50 years, many hundreds of synthetic chemicals have been shown capable of causing cancer in laboratory animals. Among them are trichloroethylene, asbestos, benzene, vinyl chloride, arsenic, arylamide, and a host of complex petroleum products with chemical structures resembling chicken wire. People in the workplace encounter chemicals daily (table 18.3).

Table 18.3 Chemical Carcinogens in the Workplace

Chemical	Cancer	Workers at Risk for Exposure
COMMON EXPOSURE		
Benzene	Myelogenous leukemia	Painters; dye users; furniture finishers
Diesel exhaust	Lung	Railroad and bus-garage workers; truckers; miners
Mineral oils	Skin	Metal machinists
Pesticides	Lung	Sprayers
Cigarette tar	Lung	Smokers
UNCOMMON EXPOSURE		
Asbestos	Mesothelioma, lung	Brake-lining, insulation workers
Synthetic mineral fibers	Lung	Wall and pipe insulation and duct wrapping users
Hair dyes	Bladder	Hairdressers and barbers
Paint	Lung	Painters
Polychlorinated biphenyls	Liver, skin	Users of hydraulic fluids and lubricants, inks, adhesives, insecticides
Soot	Skin	Chimney sweeps; bricklayers; firefighters; heating-unit service workers
RARE EXPOSURE		
Arsenic	Lung, skin	Insecticide/herbicide sprayers; tanners; oil refiners
Formaldehyde	Nose	Wood product, paper, textiles, and metal product workers

In addition to identifying potentially dangerous substances, what have the studies of potential carcinogens told us about the nature of cancer? What do these cancer-causing chemicals have in common? *They are all mutagens, each capable of inducing changes in DNA.*

Chemicals that produce mutations in DNA are often potent carcinogens. Tars in cigarette smoke, for example, are the direct cause of most lung cancers.

Other Tumors Result from Viral Infection

Chemical mutagens are not the only carcinogens, however. Some tumors seem almost certainly to result from viral infection. Viruses can be isolated from certain tumors, and these viruses cause virus-containing tumors to develop in other individuals. About 15% of human cancers are associated with viruses.

A Virus That Causes Cancer

In 1911, American medical researcher Peyton Rous reported that a virus, subsequently named **Rous avian sarcoma virus (RSV),** was associated with chicken sarcomas. He found that RSV could infect and initiate cancer in chicken fibroblast (connective tissue) cells growing in culture; from those cancerous cells, more viruses could be isolated. Rous was awarded the 1966 Nobel Prize in Physiology or Medicine for this discovery. RSV proved to be a kind of RNA virus called a **retrovirus.** When retroviruses infect a cell, they make a DNA copy of their RNA genome and insert that copy into the host cell's DNA.

How RSV Causes Cancer

How does RSV initiate cancer? When RSV was compared to a closely related virus, RAV-O, which is unable to transform normal chicken cells into cancerous cells, the two viruses proved to be identical except for one gene that was present in RSV but absent from RAV-O. That gene was called the *src* gene, short for sarcoma.

How do viral genes cause cancer? An essential clue came in 1970, when temperature-sensitive RSV mutants were isolated. These mutants would transform tissue culture cells into cancer cells at 35°C, but not at 41°C. Temperature sensitivity of this kind is almost always associated with proteins. It seemed likely, therefore, that the *src* gene was actively transcribed by the cell, rather than serving as a recognition site for some sort of regulatory protein. This was an exciting result, suggesting that the protein specified by this cancer-causing gene, or **oncogene,** could be isolated and its properties studied.

The *src* protein was first isolated in 1977 by J. Michael Bishop and Harold Varmus, who won the Nobel Prize for their efforts. It turned out to be an enzyme of moderate size that phosphorylates (adds a phosphate group to) the tyrosine amino acids of proteins. Such enzymes, called **tyrosine kinases,** are quite common in animal cells. One example is an enzyme that also serves as a plasma membrane receptor for **epidermal growth factor,** a protein that signals the initiation of cell division. This finding raised the exciting possibility that RSV may cause cancer by introducing into cells an active form of a normally quiescent growth-promoting enzyme. Later experiments showed this is indeed the case.

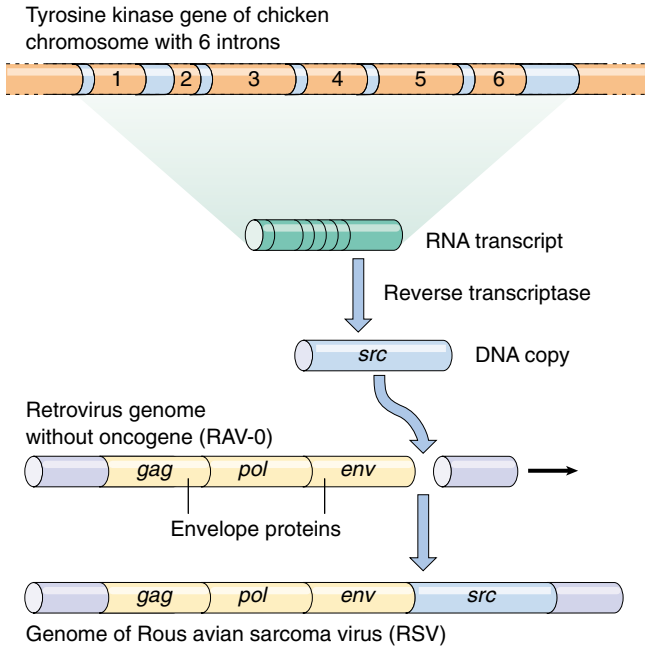


FIGURE 18.12
How a chicken gene got into the RSV genome. RSV contains only a few genes: *gag* and *env*, which encode the viral protein coat and envelope proteins, and *pol*, which encodes reverse transcriptase. It also contains the *src* gene that causes sarcomas, which the RAV-O virus lacks. RSV originally obtained its *src* gene from chickens, where a copy of the gene occurs normally and is controlled by the chicken's regulatory genes.

Origin of the *src* Gene

Does the *src* gene actually integrate into the host cell's chromosome along with the rest of the RSV genome? One way to answer this question is to prepare a radioactive version of the gene, allow it to bind to complementary sequences on the chicken chromosomes, and examine where the chromosomes become radioactive. The result of this experiment is that radioactive *src* DNA does in fact bind to the site where RSV DNA is inserted into the chicken genome—but it also binds to a second site where there is no part of the RSV genome!

The explanation for the second binding site is that the *src* gene is not exclusively a viral gene. It is also a growth-promoting gene that evolved in and occurs normally in chickens. This normal chicken gene is the second site where *src* binds to chicken DNA. Somehow, an ancestor of RSV picked up a copy of the normal chicken gene in some past infection. Now part of the virus, the gene is transcribed under the control of viral promoters rather than under the regulatory system of the chicken genome (figure 18.12).

Studies of RSV reveal that cancer results from the inappropriate activity of growth-promoting genes that are less active or completely inactive in normal cells.

Cancer and the Cell Cycle

An important technique used to study tumors is called **transfection.** In this procedure, the nuclear DNA from tumor cells is isolated and cleaved into random fragments. Each fragment is then tested individually for its ability to induce cancer in the cells that assimilate it.

Using transfection, researchers have discovered that most human tumors appear to result from the mutation of genes that regulate the cell cycle. Sometimes the mutation of only one or two gene is all that is needed to transform normally dividing cells into cancerous cells in tissue culture (table 18.4).

Point Mutations Can Lead to Cancer

The difference between a normal gene encoding a protein that regulates the cell cycle and a cancer-inducing version can be a single point mutation in the DNA. In one case of *ras*-induced bladder cancer, for example, a single DNA base change from guanine to thymine converts a glycine in the normal *ras* protein into a valine in the cancer-causing version. Several other *ras*-induced human carcinomas have been shown to also involve single nucleotide substitutions.

Telomerase and Cancer

Telomeres are short sequences of nucleotides repeated thousands of times on the ends of chromosomes. Because DNA polymerase is unable to copy chromosomes all the way to the tip (there is no place for the primer necessary to copy the last Okazaki fragment), telomeric segments are lost every time a cell divides.

In healthy cells a tumor suppressor inhibits production of a special enzyme called telomerase that adds the lost telomere material back to the tip. Without this enzyme, a cell's chromosomes lose material from their telomeres with each replication. Every time a chromosome is copied as the cell prepares to divide, more of the tip is lost. After some 30 divisions, so much is lost that copying is no longer possible. Cells in the tissues of an adult human have typically undergone 25 or more divisions. Cancer can't get very far with only the 5 remaining cell divisions. Were cancer to start, it would grind to a halt after only a few divisions for lack of telomere.

Thus, we see that the cell's inhibition of telomerase in somatic cells is a very effective natural brake on the cancer process. Any mutation that destroys the telomerase inhibitor releases that brake, making cancer possible. Thus, when researchers looked for telomerase in human ovarian tumor cells, they found it. These cells contained mutations that had inactivated the cell control that blocks the transcription of the telomerase gene. Telomerase produced in these cells reversed normal telomere shortening, allowing the cells to proliferate and gain the immortality of cancer cells.

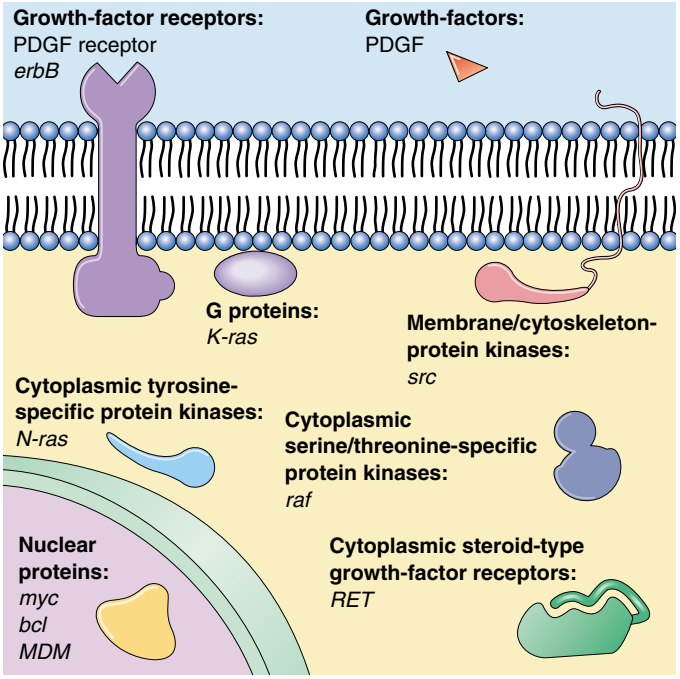


FIGURE 18.13
The main classes of oncogenes. Before they are altered by mutation to their cancer-causing condition, oncogenes are called proto-oncogenes (that is, genes able to become oncogenes). Illustrated here are the principal classes of proto-oncogenes, with some typical representatives indicated.

Mutations in Proto-Oncogenes: Accelerating the Cell Cycle

Most cancers are the direct result of mutations in growth-regulating genes. There are two general classes of cancer-inducing mutations: mutations of proto-oncogenes and mutations of tumor-suppressor genes.

Genes known as **proto-oncogenes** encode proteins that stimulate cell division. Mutations that overactivate these stimulatory proteins cause the cells that contain them to proliferate excessively. Mutated proto-oncogenes become cancer-causing genes called **oncogenes** (Greek *onco-*, "tumor") (figure 18.13). Often the induction of these cancers involves changes in the activity of intracellular signalling molecules associated with receptors on the surface of the plasma membrane. In a normal cell, the signalling pathways activated by these receptors trigger passage of the G_1 checkpoint of cell proliferation (see figure 11.17).

The mutated alleles of these oncogenes are genetically dominant. Among the most widely studied are *myc* and *ras*. Expression of *myc* stimulates the production of cyclins and cyclin-dependent protein kinases (Cdks), key elements in regulating the checkpoints of cell division.

Table 18.4 Some Genes Implicated in Human Cancers

Gene		Product	Cancer
ONCOGENES			
Genes Encoding Growth Factors or Their Receptors			
erb-B	Receptor for epidermal growth factor		Glioblastoma (a brain cancer); breast cancer
erb-B2	A growth factor receptor (gene also called *neu*)		Breast cancer; ovarian cancer; salivary gland cancer
PDGF	Platelet-derived growth factor		Glioma (a brain cancer)
RET	A growth factor receptor		Thyroid cancer
Genes Encoding Cytoplasmic Relays in Intracellular Signaling Pathways			
K-ras	Protein kinase		Lung cancer; colon cancer; ovarian cancer; pancreatic cancer
N-ras	Protein kinase		Leukemias
Genes Encoding Transcription Factors That Activate Transcription of Growth-Promoting Genes			
c-myc	Transcription factor		Lung cancer; breast cancer; stomach cancer; leukemias
L-myc	Transcription factor		Lung cancer
N-myc	Transcription factor		Neuroblastoma (a nerve cell cancer)
Genes Encoding Other Kinds of Proteins			
bcl-2	Protein that blocks cell suicide		Follicular B cell lymphoma
bcl-1	Cyclin D1, which stimulates the cell cycle clock (gene also called *PRAD1*)		Breast cancer; head and neck cancers
MDM2	Protein antagonist of p53 tumor-supressor protein		Wide variety of sarcomas (connective tissue cancers)
TUMOR-SUPRESSOR GENES			
Genes Encoding Cytoplasmic Proteins			
APC	Step in a signaling pathway		Colon cancer; stomach cancer
DPC4	A relay in signaling pathway that inhibits cell division		Pancreatic cancer
NF-1	Inhibitor of ras, a protein that stimulates cell division		Neurofibroma; myeloid leukemia
NF-2	Inhibitor of ras		Meningioma (brain cancer); schwannoma (cancer of cells supporting peripheral nerves)
Genes Encoding Nuclear Proteins			
MTS1	p16 protein, which slows the cell cycle clock		A wide range of cancers
p53	p53 protein, which halts cell division at the G_1 checkpoint		A wide range of cancers
Rb	Rb protein, which acts as a master brake of the cell cycle		Retinoblastoma; breast cancer; bone cancer; bladder cancer
Genes Encoding Proteins of Unknown Cellular Locations			
BRCA1	?		Breast cancer; ovarian cancer
BRCA2	?		Breast cancer
VHL	?		Renal cell cancer

The *ras* gene product is involved in the cellular response to a variety of growth factors such as EGF, an intercellular signal that normally initiates cell proliferation. When EGF binds to a specific receptor protein on the plasma membrane of epithelial cells, the portion of the receptor that protrudes into the cytoplasm stimulates the *ras* protein to bind to GTP. The *ras* protein/GTP complex in turn recruits and activates a protein called Raf to the inner surface of the plasma membrane, which in turn activates cytoplasmic kinases and so triggers an intracellular signaling system (see chapter 7). The final step in the pathway is the activation of transcription factors that trigger cell proliferation. Cancer-causing mutations in *ras* greatly reduce the amount of EGF necessary to initiate cell proliferation.

Mutations in Tumor-Suppressor Genes: Inactivating the Cell's Inhibitors of Proliferation

If the first class of cancer-inducing mutations "steps on the accelerator" of cell division, the second class of cancer-inducing mutations "removes the brakes." Cell division is normally turned off in healthy cells by proteins that prevent cyclins from binding to Cdks. The genes that encode these proteins are called **tumor-suppressor genes.** Their mutant alleles are genetically recessive.

Among the most widely studied tumor-suppressor genes are *Rb*, *p16*, *p21*, and *p53*. The unphosphorylated product of the *Rb* gene ties up transcription factor E2F, which transcribes several genes required for passage through the G_1 checkpoint into S phase of the cell cycle (figure 18.14). The proteins encoded by *p16* and *p21* reinforce the tumor-suppressing role of the Rb protein, preventing its phosphorylation by binding to the appropriate Cdk/cyclin complex and inhibiting its kinase activity. The p53 protein senses the integrity of the DNA and is activated if the DNA is damaged (figure 18.15). It appears to act by inducing the transcription of *p21*, which binds to cyclins and Cdk and prevents them from interacting. One of the reasons repeated smoking leads inexorably to lung cancer is that it induces *p53* mutations. Indeed, almost half of all cancers involve mutations of the *p53* gene.

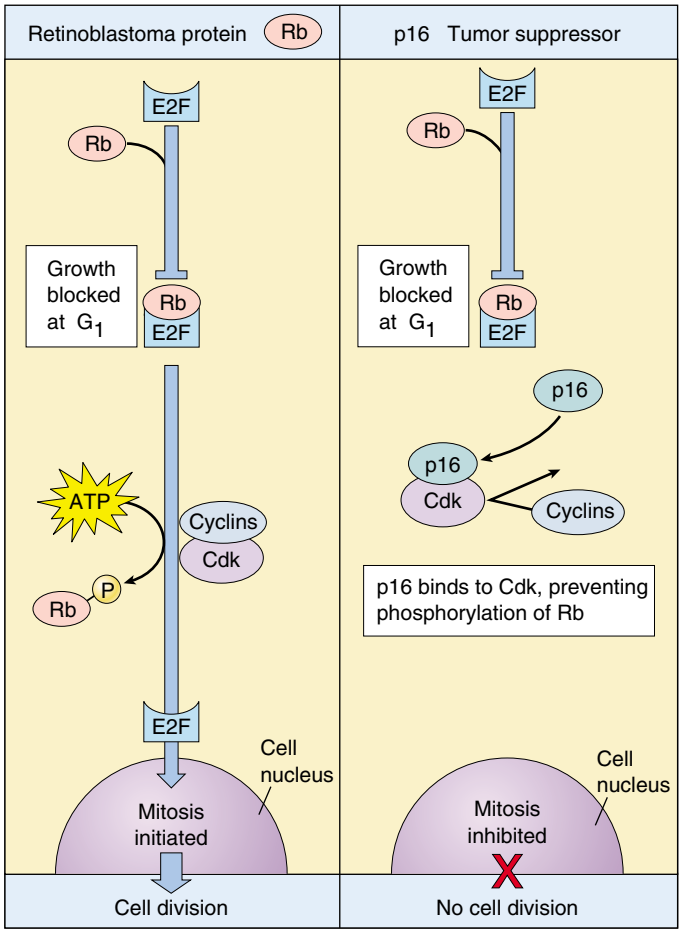


FIGURE 18.14
How the tumor-suppressor genes *Rb* and *p16* interact to block cell division. The retinoblastoma protein (Rb) binds to the transcription factor (E2F) that activates genes in the nucleus, preventing this factor from initiating mitosis. The G_1 checkpoint is passed when Cdk interacts with cyclins to phosphorylate Rb, releasing E2F. The p16 tumor-suppressor protein reinforces Rb's inhibitory action by binding to Cdk so that Cdk is not available to phosphorylate Rb.

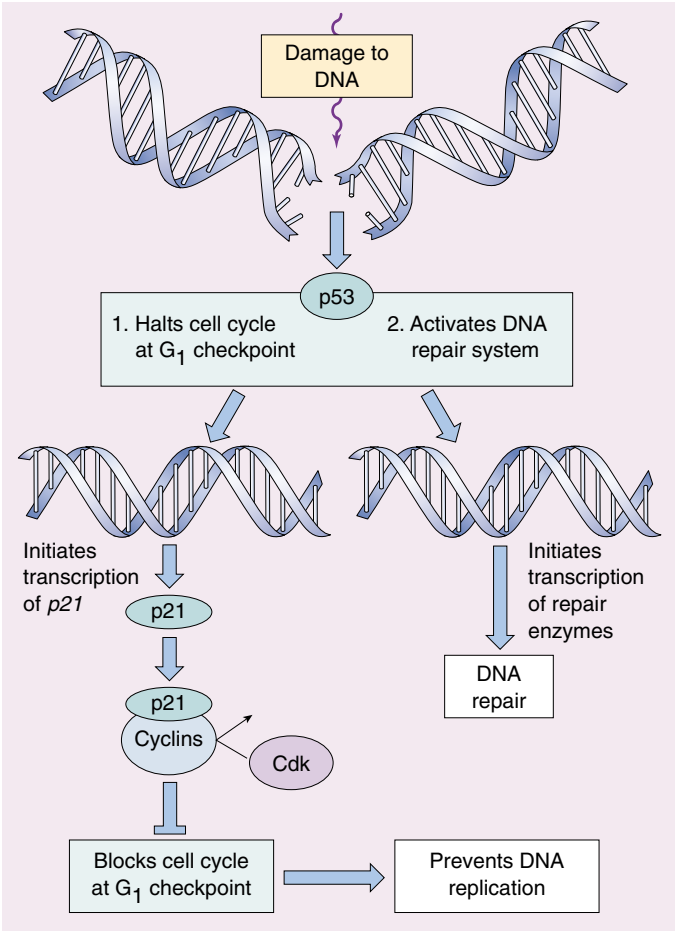


FIGURE 18.15
The role of tumor-suppressor *p53* in regulating the cell cycle. The p53 protein works at the G_1 checkpoint to check for DNA damage. If the DNA is damaged, p53 activates the DNA repair system and stops the cell cycle at the G_1 checkpoint (before DNA replication). This allows time for the damage to be repaired. p53 stops the cell cycle by inducing the transcription of *p21*. The p21 protein then binds to cyclins and prevents them from complexing with Cdk.

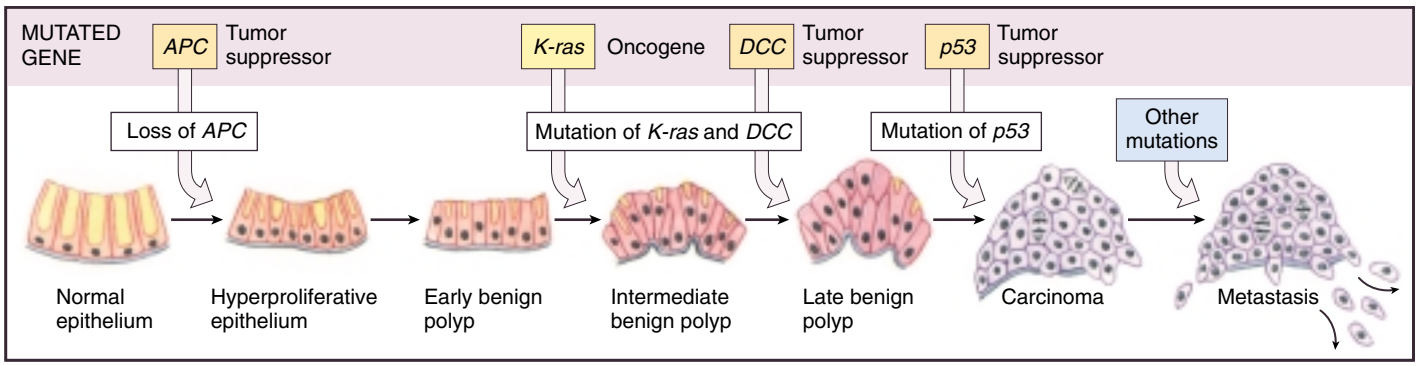


FIGURE 18.16
The progression of mutations that commonly lead to colorectal cancer. The fatal metastasis is the last of six serial changes that the epithelial cells lining the rectum undergo. One of these changes is brought about by mutation of a proto-oncogene, and three of them involve mutations that inactivate tumor-suppressor genes.

Cancer-Causing Mutations Accumulate over Time

Cells control proliferation at several checkpoints, and all of these controls must be inactivated for cancer to be initiated. Therefore, the induction of most cancers involves the mutation of multiple genes; four is a typical number (figure 18.16). In many of the tissue culture cell lines used to study cancer, most of the controls are already inactivated, so that mutations in only one or a few genes transform the line into cancerous growth. The need to inactivate several regulatory genes almost certainly explains why most cancers occur in people over 40 years old (figure 18.17); in older persons, there has been more time for individual cells to accumulate multiple mutations. It is now clear that mutations, including those in potentially cancer-causing genes, do accumulate over time. Using the polymerase chain reaction (PCR), researchers in 1994 searched for a certain cancer-associated gene mutation in the blood cells of 63 cancer-free people. They found that the mutation occurred 13 times more often in people over 60 years old than in people under 20.

Cancer is a disease in which the controls that normally restrict cell proliferation do not operate. In some cases, cancerous growth is initiated by the inappropriate activation of proteins that regulate the cell cycle; in other cases, it is initiated by the inactivation of proteins that normally suppress cell division.

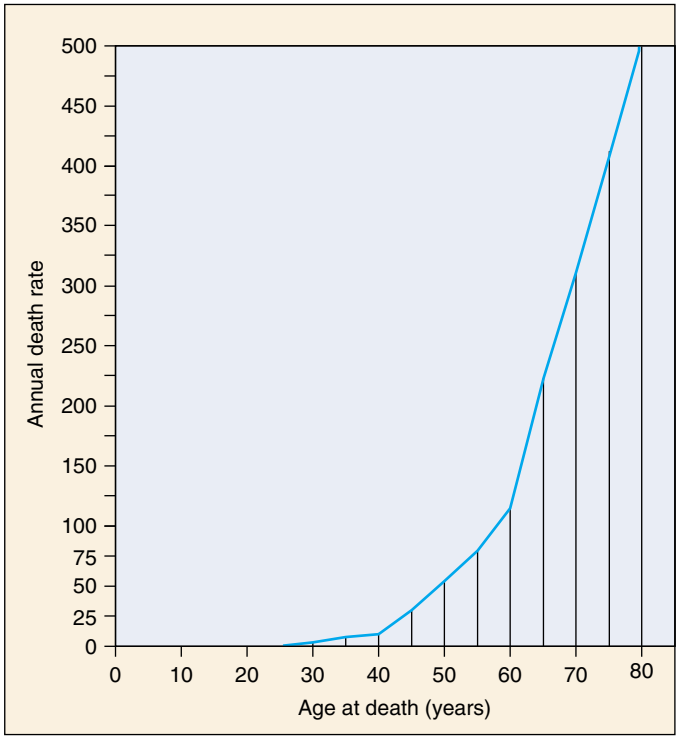


FIGURE 18.17
The annual death rate from cancer climbs with age. The rate of cancer deaths increases steeply after age 40 and even more steeply after age 60, suggesting that several independent mutations must accumulate to give rise to cancer.

Smoking and Cancer

How can we prevent cancer? The most obvious strategy is to minimize mutational insult. Anything that decreases exposure to mutagens can decrease the incidence of cancer because exposure has the potential to mutate a normal gene into an oncogene. It is no accident that the most reliable tests for the carcinogenicity of a substance are tests that measure the substance's mutagenicity.

The Association between Smoking and Cancer

About a third of all cases of cancer in the United States are directly attributable to cigarette smoking. The association between smoking and cancer is particularly striking for lung cancer (figure 18.18). Studies of male smokers show a highly positive correlation between the number of cigarettes smoked per day and the incidence of lung cancer (figure 18.19). For individuals who smoke two or more packs a day, the risk of contracting lung cancer is at least 40 times greater than it is for nonsmokers, whose risk level approaches zero. Clearly, an effective way to avoid lung cancer is not to smoke. Other studies have shown a clear relationship between cigarette smoking and reduced life expectancy (figure 18.20). Life insurance companies have calculated that smoking a single cigarette lowers one's life expectancy by 10.7 minutes (longer than it takes to smoke the cigarette)! Every pack of 20 cigarettes bears an unwritten label:

"The price of smoking this pack of cigarettes is 3½ hours of your life."

Smoking Introduces Mutagens to the Lungs

Over half a million people died of cancer in the United States in 1999; about 28% of them died of lung cancer. About 140,000 persons were diagnosed with lung cancer each year in the 1980s. Around 90% of them died within three years after diagnosis; 96% of them were cigarette smokers.

Smoking is a popular pastime. In the United States, 24% of the population smokes, and U.S. smokers consumed over 450 billion cigarettes in 1999. The smoke emitted from these cigarettes contains some 3000 chemical components, including vinyl chloride, benzo[*a*]pyrenes, and nitroso-*nor*-nicotine, all potent mutagens. Smoking places these mutagens into direct contact with the tissues of the lungs.

Mutagens in the Lung Cause Cancer

Introducing powerful mutagens to the lungs causes considerable damage to the genes of the epithelial cells that line the lungs and are directly exposed to the chemicals. Among the genes that are mutated as a result are some whose normal function is to regulate cell proliferation. When these genes are damaged, lung cancer results.

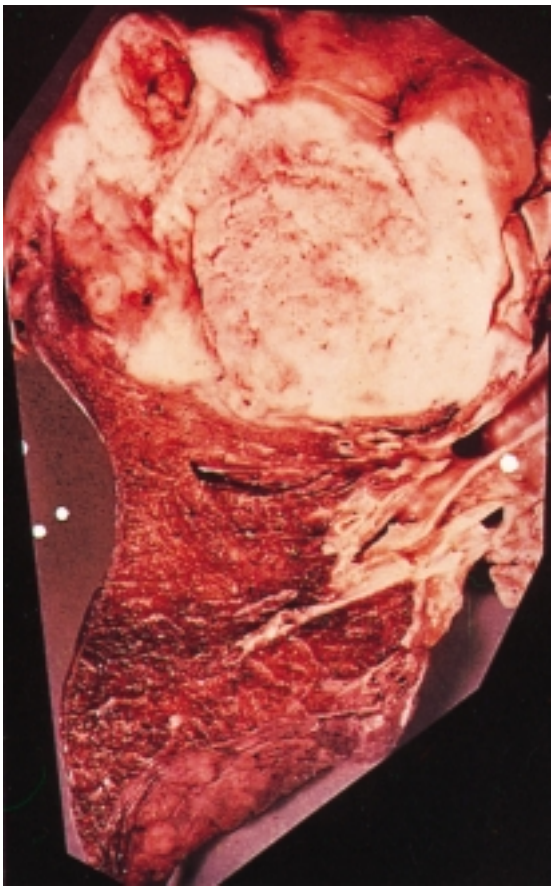


FIGURE 18.18
Photo of a cancerous human lung. The bottom half of the lung is normal, while a cancerous tumor has completely taken over the top half. The cancer cells will eventually break through into the lymph and blood vessels and spread through the body.

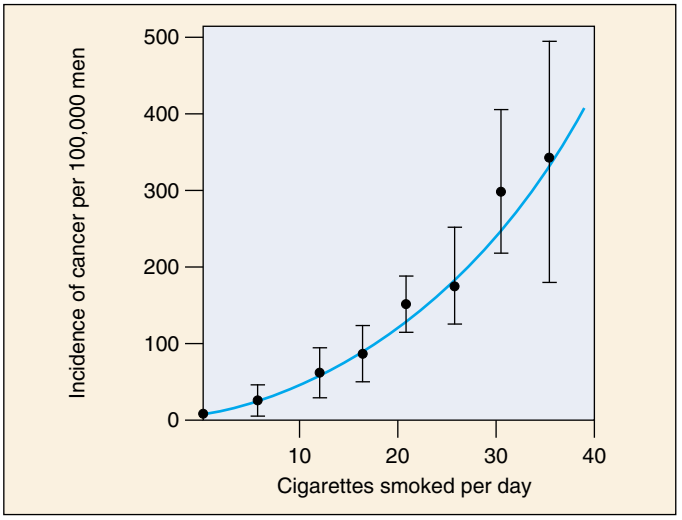


FIGURE 18.19
Smoking causes cancer. The annual incidence of lung cancer per 100,000 men clearly increases with the number of cigarettes smoked per day.

This process has been clearly demonstrated for benzo[*a*]pyrene (BP), one of the potent mutagens released into cigarette smoke from tars in the tobacco. The epithelial cells of the lung absorb BP from tobacco smoke and chemically alter it to a derivative form. This derivative form, benzo[*a*]pyrene-diolepoxide (BPDE), binds directly to the tumor-suppressor gene *p53* and mutates it to an inactive form. The protein encoded by *p53* oversees the G_1 cell cycle checkpoint described in chapter 11 and is one of the body's key mechanisms for preventing uncontrolled cell proliferation. The destruction of *p53* in lung epithelial cells greatly hastens the onset of lung cancer—*p53* is mutated to an inactive form in over 70% of lung cancers. When examined, the *p53* mutations in cancer cells almost all occur at one of three "hot spots." The key evidence linking smoking and cancer is that when the mutations of *p53* caused by BPDE from cigarettes are examined, they occur at the same three specific "hot spots!"

The Incidence of Cancer Reflects Smoking

Cigarette manufacturers argue that the causal connection between smoking and cancer has not been proved, and that somehow the relationship is coincidental. Look carefully at the data presented in figure 18.21 and see if you agree. The upper graph, compiled from data on American men, shows the incidence of smoking from 1900 to 1990 and the incidence of lung cancer over the same period. Note that as late as 1920, lung cancer was a rare disease. About 20 years after the incidence of smoking began to increase among men, lung cancer also started to become more common.

Now look at the lower graph, which presents data on American women. Because of social mores, significant numbers of American women did not smoke until after World War II, when many social conventions changed. As late as 1963, when lung cancer among males was near current levels, this disease was still rare in women. In the United States that year, only 6588 women died of lung cancer. But as more women smoked, more developed lung cancer, again with a lag of about 20 years. American women today have achieved equality with men in the numbers of cigarettes they smoke, and their lung cancer death rates are today approaching those for men. In 1990, more than 49,000 women died of lung cancer in the United States. The current annual rate of deaths from lung cancer in male and female smokers is 180 per 100,000, or about 2 out of every 1000 smokers *each year*.

The easiest way to avoid cancer is to avoid exposure to mutagens. The single greatest contribution one can make to a longer life is not to smoke.

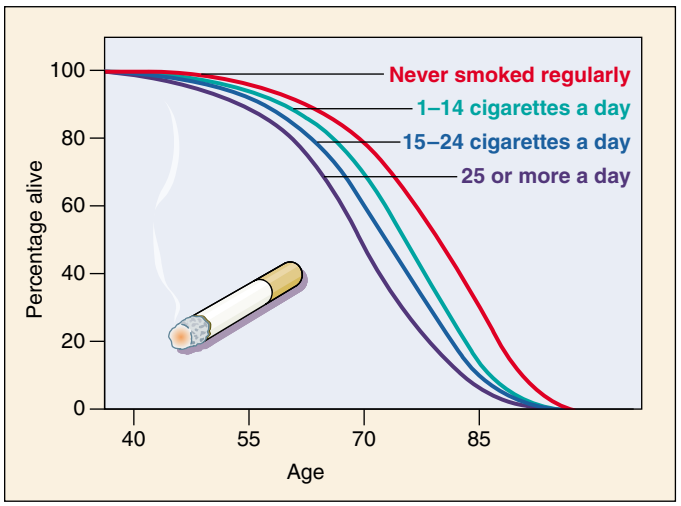


FIGURE 18.20
Tobacco reduces life expectancy. The world's longest-running survey of smoking, begun in 1951 in Britain, revealed that by 1994 the death rate for smokers had climbed to three times the rate for nonsmokers among men 35 to 69 years of age.
Source: Data from *New Scientist*, October 15, 1994.

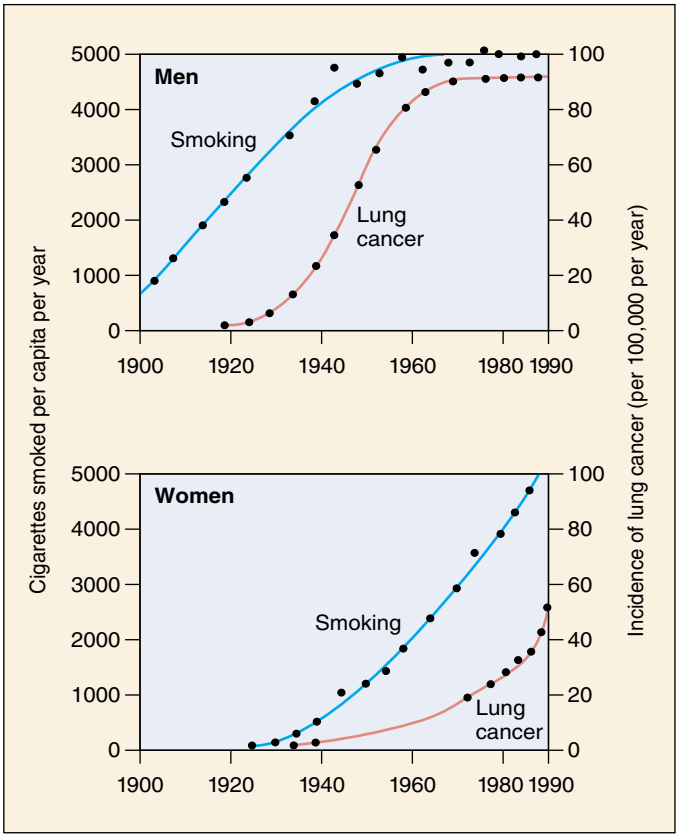


FIGURE 18.21
The incidence of lung cancer in men and women. What do these graphs indicate about the connection between smoking and lung cancer?

Curing Cancer

Potential cancer therapies are being developed on many fronts (figure 18.22). Some act to prevent the start of cancer within cells. Others act outside cancer cells, preventing tumors from growing and spreading.

Preventing the Start of Cancer

Many promising cancer therapies act within potential cancer cells, focusing on different stages of the cell's "Shall I divide?" decision-making process.

1. Receiving the Signal to Divide. The first step in the decision process is the reception of a "divide" signal, usually a small protein called a growth factor released from a neighboring cell. The growth factor is received by a protein receptor on the cell surface. Mutations that increase the number of receptors on the cell surface amplify the division signal and so lead to cancer. Over 20% of breast cancer tumors prove to overproduce a protein called HER2 associated with the receptor for epidermal growth factor.

Therapies directed at this stage of the decision process utilize the human immune system to attack cancer cells. Special protein molecules called "monoclonal antibodies," created by genetic engineering, are the therapeutic agents. These monoclonal antibodies are designed to seek out and stick to HER2. Like waving a red flag, the presence of the monoclonal antibody calls down attack by the immune system on the HER2 cell. Because breast cancer cells overproduce HER2, they are killed preferentially. Genentech's recently approved monoclonal antibody, called "herceptin," has given promising results in clinical tests. In other tests, the monoclonal antibody C225, directed against epidermal growth factor receptors, has succeeded in curing advanced colon cancer. Clinical trials of C225 have begun.

2. The Relay Switch. The second step in the decision process is the passage of the signal into the cell's interior, the cytoplasm. This is carried out in normal cells by a protein called Ras that acts as a relay switch. When growth factor binds to a receptor like EGF, the adjacent Ras protein acts like it has been "goosed," contorting into a new shape. This new shape is chemically active, and initiates a chain of reactions that passes the "divide" signal inward toward the nucleus. Mutated forms of the Ras protein behave like a relay switch stuck in the "ON" position, continually instructing the cell to divide when it should not. 30% of all cancers have a mutant form of Ras.

Therapies directed at this stage of the decision process take advantage of the fact that normal Ras proteins are inactive when made. Only after it has been modified by the special enzyme *farnesyl transferase* does Ras protein become able to function as a relay switch. In tests on animals, farnesyl transferase inhibitors induce the regression of tumors and prevent the formation of new ones.

3. Amplifying the Signal. The third step in the decision process is the amplification of the signal within the cytoplasm. Just as a TV signal needs to be amplified in order to be received at a distance, so a "divide" signal must be amplified if it is to reach the nucleus at the interior of the cell, a very long journey at a molecular scale. Cells use an ingenious trick to amplify the signal. Ras, when "ON," activates an enzyme, a protein kinase. This protein kinase activates other protein kinases that in their turn activate still others. The trick is that once a protein kinase enzyme is activated, it goes to work like a demon, activating hoards of others every second! And each and every one it activates behaves the same way too, activating still more, in a cascade of ever-widening effect. At each stage of the relay, the signal is amplified a thousand-fold. Mutations stimulating any of the protein kinases can dangerously increase the already amplified signal and lead to cancer. Five percent of all cancers, for example, have a mutant hyperactive form of the protein kinase Src.

Therapies directed at this stage of the decision process employ so-called "anti-sense RNA" directed specifically against Src or other cancer-inducing kinase mutations. The idea is that the *src* gene uses a complementary copy of itself to manufacture the Src protein (the "sense" RNA or messenger RNA), and a mirror image complementary copy of the sense RNA ("anti-sense RNA") will stick to it, gumming it up so it can't be used to make Src protein. The approach appears promising. In tissue culture, anti-sense RNAs inhibit the growth of cancer cells, and some also appear to block the growth of human tumors implanted in laboratory animals. Human clinical trials are underway.

4. Releasing the Brake. The fourth step in the decision process is the removal of the "brake" the cell uses to restrain cell division. In healthy cells this brake, a tumor suppressor protein called Rb, blocks the activity of a transcription factor protein called E2F. When free, E2F enables the cell to copy its DNA. Normal cell division is triggered to begin when Rb is inhibited, unleashing E2F. Mutations which destroy Rb release E2F from its control completely, leading to ceaseless cell division. Forty percent of all cancers have a defective form of Rb.

Therapies directed at this stage of the decision process are only now being attempted. They focus on drugs able to inhibit E2F, which should halt the growth of tumors arising from inactive Rb. Experiments in mice in which the E2F genes have been destroyed provide a model system to study such drugs, which are being actively investigated.

5. Checking That Everything Is Ready. The fifth step in the decision process is the mechanism used by the cell to ensure that its DNA is undamaged and ready to divide. This job is carried out in healthy cells by the tumor-suppressor protein p53, which inspects the integrity of the DNA. When it detects damaged or foreign DNA, p53 stops cell division and activates the cell's DNA repair systems. If the damage doesn't

get repaired in a reasonable time, p53 pulls the plug, triggering events that kill the cell. In this way, mutations such as those that cause cancer are either repaired or the cells containing them eliminated. If p53 is itself destroyed by mutation, future damage accumulates unrepaired. Among this damage are mutations that lead to cancer. Fifty percent of all cancers have a disabled p53. Fully 70 to 80% of lung cancers have a mutant inactive p53—the chemical benzo[*a*]pyrene in cigarette smoke is a potent mutagen of p53.

A promising new therapy using adenovirus (responsible for mild colds) is being targeted at cancers with a mutant p53. To grow in a host cell, adenovirus must use the product of its gene *E1B* to block the host cell's p53, thereby enabling replication of the adenovirus DNA. This means that while mutant adenovirus without *E1B* cannot grow in healthy cells, the mutants should be able to grow in, and destroy, cancer cells with defective p53. When human colon and lung cancer cells are introduced into mice lacking an immune system and allowed to produce substantial tumors, 60% of the tumors simply disappear when treated with E1B-deficient adenovirus, and do not reappear later. Initial clinical trials are less encouraging, as many people possess antibodies to adenovirus.

6. Stepping on the Gas. Cell division starts with replication of the DNA. In healthy cells, another tumor suppressor "keeps the gas tank nearly empty" for the DNA replication process by inhibiting production of an enzyme called telomerase. Without this enzyme, a cell's chromosomes lose material from their tips, called telomeres. Every time a chromosome is copied, more tip material is lost. After some thirty divisions, so much is lost that copying is no longer possible. Cells in the tissues of an adult human have typically undergone twenty five or more divisions. Cancer can't get very far with only the five remaining cell divisions, so inhibiting telomerase is a very effective natural break on the cancer process. It is thought that almost all cancers involve a mutation that destroys the telomerase inhibitor, releasing this break and making cancer possible. It should be possible to block cancer by reapplying this inhibition. Cancer therapies that inhibit telomerase are just beginning clinical trials.

Preventing the Spread of Cancer

7. Tumor Growth. Once a cell begins cancerous growth, it forms an expanding tumor. As the tumor grows ever-larger, it requires an increasing supply of food and nutrients, obtained from the body's blood supply. To facilitate this necessary grocery shopping, tumors leak out substances into the surrounding tissues that encourage angiogenesis, the formation of small blood vessels. Chemicals that inhibit this process are called angiogenesis inhibitors. In mice, two such angiogenesis inhibitors, angiostatin and endostatin, caused tumors to regress to microscopic size. This very exciting result has proven controversial, but initial human trials seem promising.

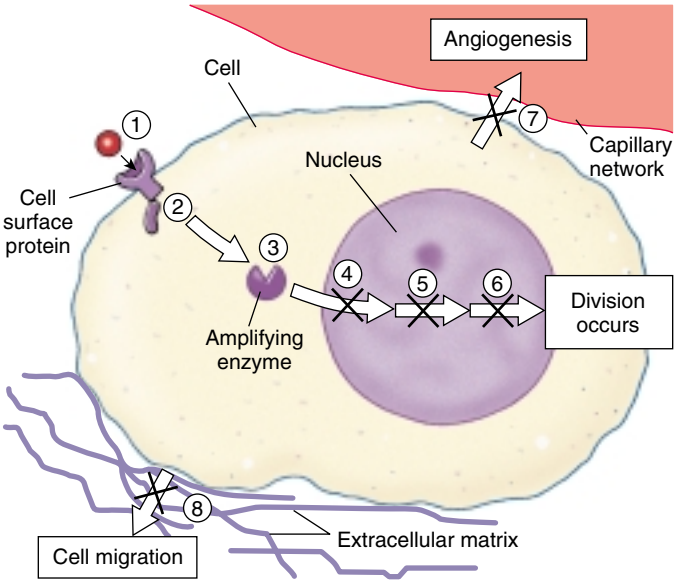


FIGURE 18.22
New molecular therapies for cancer target eight different stages in the cancer process. (*1*) On the cell surface, a growth factor signals the cell to divide. (*2*) Just inside the cell, a protein relay switch passes on the divide signal. (*3*) In the cytoplasm, enzymes amplify the signal. In the nucleus, (*4*) a "brake" preventing DNA replication is released, (*5*) proteins check that the replicated DNA is not damaged, and (*6*) other proteins rebuild chromosome tips so DNA can replicate. (*7*) The new tumor promotes angiogenesis, the formation of growth-promoting blood vessels. (*8*) Some cancer cells break away from the extracellular matrix and invade other parts of the body.

8. Metastasis. If cancerous tumors simply continued to grow where they form, many could be surgically removed, and far fewer would prove fatal. Unfortunately, many cancerous tumors eventually metastasize, individual cancer cells breaking their moorings to the extracellular matrix and spreading to other locations in the body where they initiate formation of secondary tumors. This process involves metal-requiring protease enzymes that cleave the cell-matrix linkage, components of the extracellular matrix such as fibronectin that also promote the migration of several non-cancerous cell types, and RhoC, a GTP-hydrolyzing enzyme that promotes cell migration by providing needed GTP. All of these components offer promising targets for future anti-cancer therapy.

Therapies such as those described here are only part of a wave of potential treatments under development and clinical trial. The clinical trials will take years to complete, but in the coming decade we can expect cancer to become a curable disease.

Understanding of how mutations produce cancer has progressed to the point where promising potential therapies can be tested.

18.3 Recombination alters gene location.

An Overview of Recombination

Mutation is a change in the *content* of an organism's genetic message, but it is not the only source of genetic diversity. Diversity is also generated when existing elements of the genetic message move around within the genome. As an analogy, consider the pages of this book. A point mutation would correspond to a change in one or more of the letters on the pages. For example, " . . . in one or more of the letters *of* the pages" is a mutation of the previous sentence, in which an "n" is changed to an "f." A significant alteration is also achieved, however, when we move the position of words, as in " . . . in one or more of the pages on the letters." The change alters (and destroys) the meaning of the sentence by exchanging the position of the words "letters" and "pages." This second kind of change, which represents an alteration in the genomic *location* of a gene or a fragment of a gene, demonstrates **genetic recombination.**

FIGURE 18.23
A Nobel Prize for discovering gene transfer by transposition.
Barbara McClintock receiving her Nobel Prize in 1983.

Gene Transfer

Viewed broadly, genetic recombination can occur by two mechanisms (table 18.5). In **gene transfer,** one chromosome or genome donates a segment to another chromosome or genome. The transfer of genes from the human immunodeficiency virus (HIV) to a human chromosome is an example of gene transfer. Because gene transfer occurs in both prokaryotes and eukaryotes, it is thought to be the more primitive of the two mechanisms.

Reciprocal Recombination

Reciprocal recombination is when two chromosomes trade segments. It is exemplified by the crossing over that occurs between homologous chromosomes during meiosis. Independent assortment during meiosis is another form of reciprocal recombination. Discussed in chapters 12 and 13, it is responsible for the 9:3:3:1 ratio of phenotypes in a dihybrid cross and occurs only in eukaryotes.

Genetic recombination is a change in the genomic association among genes. It often involves a change in the position of a gene or portion of a gene. Recombination of this sort may result from one-way gene transfer or reciprocal gene exchange.

Table 18.5 Classes of Genetic Recombination

Class	Occurrence
GENE TRANSFERS	
Conjugation	Occurs predominantly but not exclusively in bacteria and is targeted to specific locations in the genome
Transposition	Common in both bacteria and eukaryotes; genes move to new genomic locations, apparently at random
RECIPROCAL RECOMBINATIONS	
Crossing over	Requires the pairing of homologous chromosomes and may occur anywhere along their length
Unequal crossing over	The result of crossing over between mismatched segments; leads to gene duplication and deletion
Gene conversion	Occurs when homologous chromosomes pair and one is "corrected" to resemble the other
Independent assortment	Haploid cells produced by meiosis contain only one randomly selected member of each pair of homologous chromosomes

Gene Transfer

Genes are not fixed in their locations on chromosomes or the circular DNA molecules of bacteria; they can move around. Some genes move because they are part of small, circular, extrachromosomal DNA segments called **plasmids.** Plasmids enter and leave the main genome at specific places where a nucleotide sequence matches one present on the plasmid. Plasmids occur primarily in bacteria, in which the main genomic DNA can interact readily with other DNA fragments. About 5% of the DNA that occurs in a bacterium is plasmid DNA. Some plasmids are very small, containing only one or a few genes, while others are quite complex and contain many genes. Other genes move within **transposons,** which jump from one genomic position to another at random in both bacteria and eukaryotes.

Gene transfer by plasmid movement was discovered by Joshua Lederberg and Edward Tatum in 1947. Three years later, transposons were discovered by Barbara McClintock. However, her work implied that the position of genes in a genome need not be constant. Researchers accustomed to viewing genes as fixed entities, like beads on a string, did not readily accept the idea of transposons. Therefore, while Lederberg and Tatum were awarded a Nobel Prize for their discovery in 1958, McClintock did not receive the same recognition for hers until 1983 (figure 18.23).

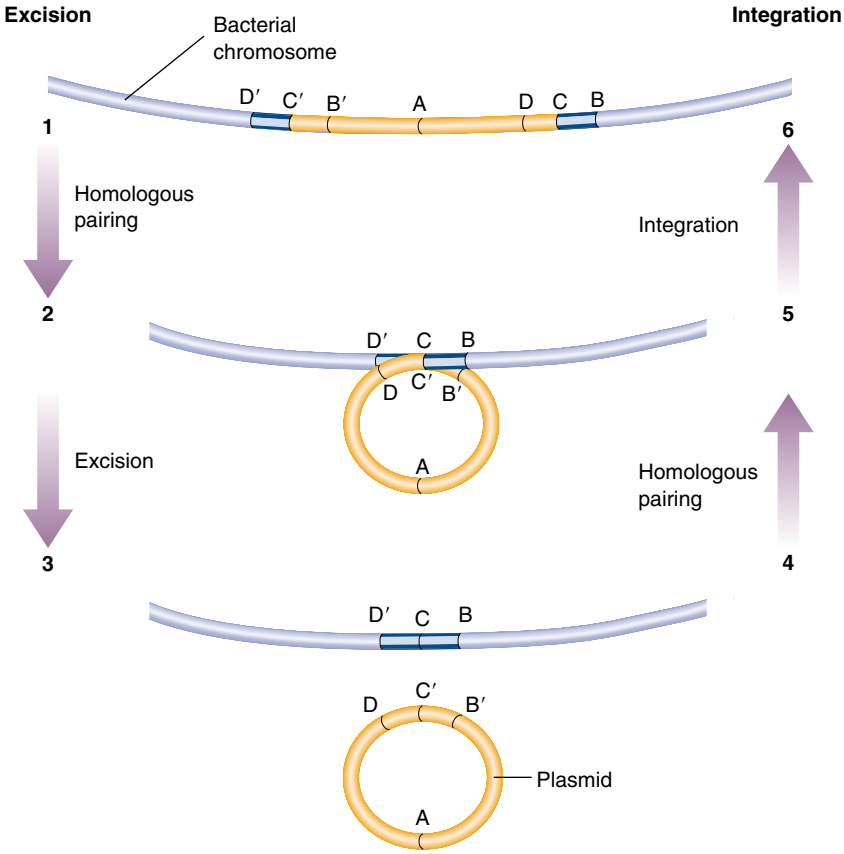


FIGURE 18.24
Integration and excision of a plasmid. Because the ends of the two sequences in the bacterial genome are the same (D′, C′, B′, and D, C, B), it is possible for the two ends to pair. Steps 1–3 show the sequence of events if the strands exchange during the pairing. The result is excision of the loop and a free circle of DNA—a plasmid. Steps 4–6 show the sequence when a plasmid integrates itself into a bacterial genome.

Plasmid Creation

To understand how plasmids arise, consider a hypothetical stretch of bacterial DNA that contains two copies of the same nucleotide sequence. It is possible for the two copies to base-pair with each other and create a transient "loop," or double duplex. All cells have recombination enzymes that can cause such double duplexes to undergo a **reciprocal exchange,** in which they exchange strands. As a result of the exchange, the loop is freed from the rest of the DNA molecule and becomes a plasmid (figure 18.24, steps 1–3). Any genes between the duplicated sequences (such as gene A in figure 18.24) are transferred to the plasmid.

Once a plasmid has been created by reciprocal exchange, DNA polymerase will replicate it if it contains a replication origin, often without the controls that restrict the main genome to one replication per cell division. Consequently, some plasmids may be present in multiple copies, others in just a few copies, in a given cell.

Integration

A plasmid created by recombination can reenter the main genome the same way it left. Sometimes the region of the plasmid DNA that was involved in the original exchange, called the **recognition site,** aligns with a matching sequence on the main genome. If a recombination event occurs anywhere in the region of alignment, the plasmid will integrate into the genome (figure 18.24, steps 4–6). Integration can occur wherever any shared sequences exist, so plasmids may be integrated into the main genome at positions other than the one from which they arose. If a plasmid is integrated at a new position, it transfers its genes to that new position.

Transposons and plasmids transfer genes to new locations on chromosomes. Plasmids can arise from and integrate back into a genome wherever DNA sequences in the genome and in the plasmid match.

Gene Transfer by Conjugation

One of the startling discoveries Lederberg and Tatum made was that plasmids can pass from one bacterium to another. The plasmid they studied was part of the genome of *Escherichia coli*. It was given the name F for fertility factor because only cells which had that plasmid integrated into their DNA could act as plasmid donors. These cells are called Hfr cells (for "high-frequency recombination"). The F plasmid contains a DNA replication origin and several genes that promote its transfer to other cells. These genes encode protein subunits that assemble on the surface of the bacterial cell, forming a hollow tube called a **pilus.**

When the pilus of one cell (F^+) contacts the surface of another cell that lacks a pilus, and therefore does not contain an F plasmid (F^-), the pilus draws the two cells close together so that DNA can be exchanged (figure 18.25). First, the F plasmid binds to a site on the interior of the F^+ cell just beneath the pilus. Then, by a process called **rolling-circle replication,** the F plasmid begins to copy its DNA at the binding point. As it is replicated, the single-stranded copy of the plasmid passes into the other cell. There a complementary strand is added, creating a new, stable F plasmid (figure 18.26). In this way, genes are passed from one bacterium to another. This transfer of genes between bacteria is called **conjugation.**

In an Hfr cell, with the F plasmid integrated into the main bacterial genome rather than free in the cytoplasm, the F plasmid can still organize the transfer of genes. In this case, the integrated F region binds beneath the pilus and initiates the *replication of the bacterial genome*, transferring the newly replicated portion to the recipient cell. Transfer proceeds as if the bacterial genome were simply a part of the F plasmid. By studying this phenomenon, researchers have been able to locate the positions of different genes in bacterial genomes (figure 18.27).

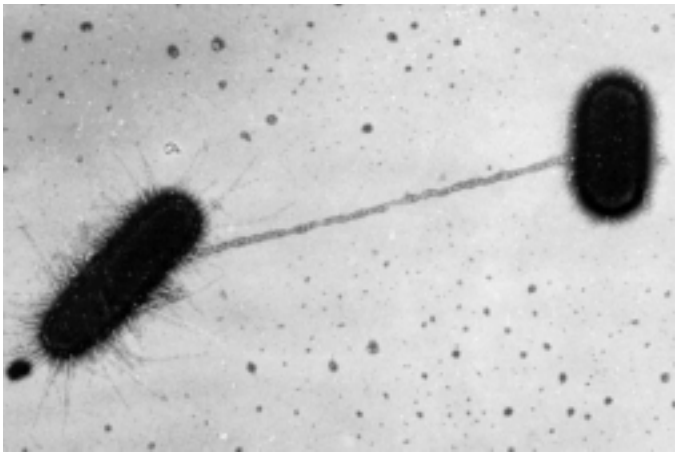


FIGURE 18.25
Contact by a pilus. The pilus of an F^+ cell connects to an F^- cell and draws the two cells close together so that DNA transfer can occur.

Gene Transfer by Transposition

Like plasmids, transposons (figure 18.28) move from one genomic location to another. After spending many generations in one position, a transposon may abruptly move to a new position in the genome, carrying various genes along with it. Transposons encode an enzyme called **transposase,** that inserts the transposon into the genome (figure 18.29). Because this enzyme usually does not recognize any particular sequence on the genome, transposons appear to move to random destinations.

The movement of any given transposon is relatively rare: it may occur perhaps once in 100,000 cell generations. Although low, this rate is still about 10 times as frequent as

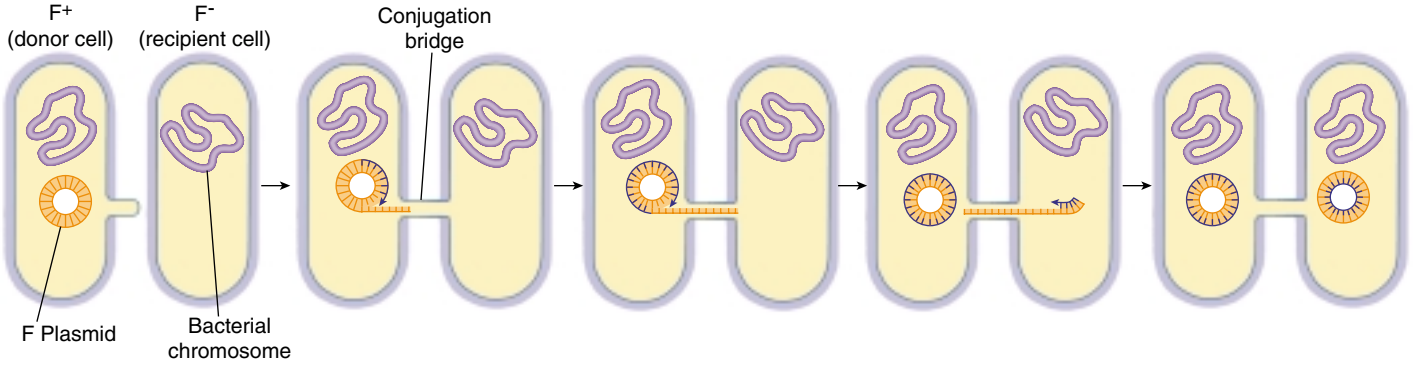
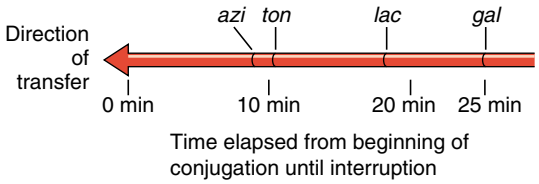
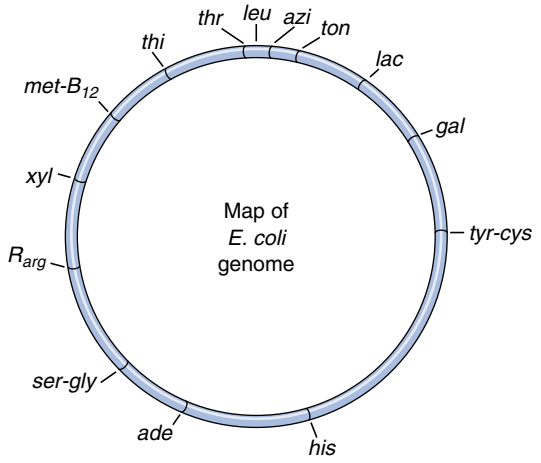


FIGURE 18.26
Gene transfer between bacteria. Donor cells (F^+) contain an F plasmid that recipient cells (F^-) lack. The F plasmid replicates itself and transfers the copy across a conjugation bridge. The remaining strand of the plasmid serves as a template to build a replacement. When the single strand enters the recipient cell, it serves as a template to assemble a double-stranded plasmid. When the process is complete, both cells contain a complete copy of the plasmid.



(a)



(b)

FIGURE 18.27
A conjugation map of the *E. coli* chromosome. Scientists have been able to break the *Escherichia coli* conjugation bridges by agitating the cell suspension rapidly in a blender. By agitating at different intervals after the start of conjugation, investigators can locate the positions of various genes along the bacterial genome. (*a*) The closer the genes are to the origin of replication, the sooner one has to turn on the blender to block their transfer.
(*b*) Map of the *E. coli* genome developed using this method.

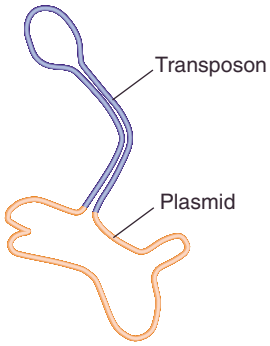
the rate at which random mutational changes occur. Furthermore, there are many transposons in most cells. Hence, over long periods of time, transposition can have an enormous evolutionary impact.

One way this impact can be felt is through mutation. The insertion of a transposon into a gene often destroys the gene's function, resulting in what is termed **insertional inactivation.** This phenomenon is thought to be the cause of a significant number of the spontaneous mutations observed in nature.

Transposition can also facilitate **gene mobilization,** the bringing together in one place of genes that are usually located at different positions in the genome. In bacteria, for example, a number of genes encode enzymes that make the bacteria resistant to antibiotics such as penicillin, and many of these genes are located on plasmids. The simultaneous exposure of bacteria to multiple antibiotics, a common medical practice some years ago, favors the persistence of plasmids that have managed to acquire several resistance genes. Transposition can rapidly generate such composite plasmids, called **resistance transfer factors** (RTFs), by moving antibiotic resistance genes from several plasmids to one. Bacteria possessing RTFs are thus able to survive treatment with a wide variety of antibiotics. RTFs are thought to be responsible for much of the recent difficulty in treating hospital-engendered *Staphylococcus aureus* infections and the new drug-resistant strains of tuberculosis.

Plasmids transfer copies of bacterial genes (and even entire genomes) from one bacterium to another. Transposition is the one-way transfer of genes to a randomly selected location in the genome. The genes move because they are associated with mobile genetic elements called transposons.

FIGURE 18.28
Transposon. Transposons form characteristic stem-and-loop structures called "lollipops" because their two ends have the same nucleotide sequence as inverted repeats. These ends pair together to form the stem of the lollipop.



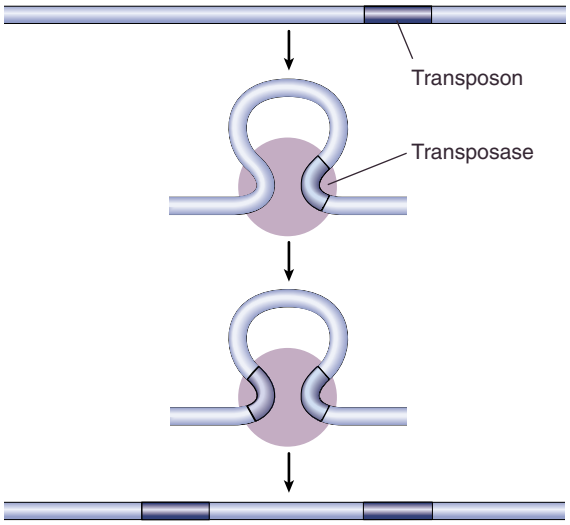


FIGURE 18.29
Transposition. Transposase does not recognize any particular DNA sequence; rather, it selects one at random, moving the transposon to a random location. Some transposons leave a copy of themselves behind when they move.

Reciprocal Recombination

In the second major mechanism for producing genetic recombination, reciprocal recombination, two homologous chromosomes exchange all or part of themselves during the process of meiosis.

Crossing Over

As we saw in chapter 12, crossing over occurs in the first prophase of meiosis, when two homologous chromosomes line up side by side within the synaptonemal complex. At this point, the homologues exchange DNA strands at one or more locations. This exchange of strands can produce chromosomes with new combinations of alleles.

Imagine, for example, that a giraffe has genes encoding neck length and leg length at two different loci on one of its chromosomes. Imagine further that a recessive mutation occurs at the neck length locus, leading after several rounds of independent assortment to some individuals that are homozygous for a variant "long-neck" allele. Similarly, a recessive mutation at the leg length locus leads to homozygous "long-leg" individuals.

It is very unlikely that these two mutations would arise at the same time in the same individual because the probability of two independent events occurring together is the product of their individual probabilities. If the spontaneous occurrence of both mutations in a single individual were the only way to produce a giraffe with both a long neck and long legs, it would be extremely unlikely that such an individual would ever occur. Because of recombination, however, a crossover in the interval between the two genes could in one meiosis produce a chromosome bearing both variant alleles. This ability to reshuffle gene combinations rapidly is what makes recombination so important to the production of natural variation.

Unequal Crossing Over

Reciprocal recombination can occur in any region along two homologous chromosomes with sequences similar enough to permit close pairing. Mistakes in pairing occasionally happen when several copies of a sequence exist in different locations on a chromosome. In such cases, one copy of a sequence may line up with one of the duplicate copies instead of with its homologous copy. Such misalignment causes slipped mispairing, which, as we discussed earlier, can lead to small deletions and frame-shift mutations. If a crossover occurs in the pairing region, it will result in unequal crossing over because the two homologues will exchange segments of unequal length.

In unequal crossing over, one chromosome gains extra copies of the multicopy sequences, while the other chromosome loses them (figure 18.30). This process can generate a chromosome with hundreds of copies of a particular gene, lined up side by side in tandem array.

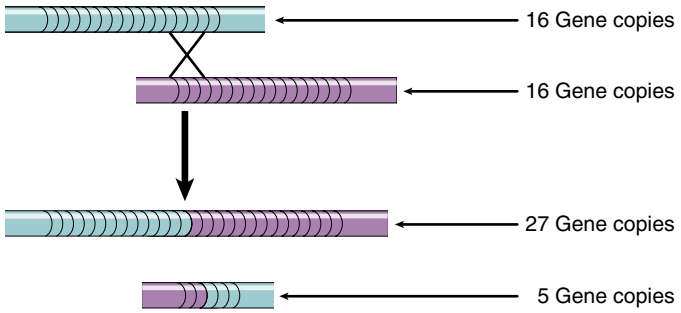


FIGURE 18.30
Unequal crossing over. When a repeated sequence pairs out of register, a crossover within the region will produce one chromosome with fewer gene copies and one with more. Much of the gene duplication that has occurred in eukaryotic evolution may well be the result of unequal crossing over.

Because the genomes of most eukaryotes possess multiple copies of transposons scattered throughout the chromosomes, unequal crossing over between copies of transposons located in different positions has had a profound influence on gene organization in eukaryotes. As we shall see later, most of the genes of eukaryotes appear to have been duplicated one or more times during their evolution.

Gene Conversion

Because the two homologues that pair within a synaptonemal complex are not identical, some nucleotides in one homologue are not complementary to their counterpart in the other homologue with which it is paired. These occasional nonmatching pairs of nucleotides are called **mismatch pairs.**

As you might expect, the cell's error-correcting machinery is able to detect mismatch pairs. If a mismatch is detected during meiosis, the enzymes that "proofread" new DNA strands during DNA replication correct it. The mismatched nucleotide in one of the homologues is excised and replaced with a nucleotide complementary to the one in the other homologue. Its base-pairing partner in the first homologue is then replaced, producing two chromosomes with the same sequence. This error correction causes one of the mismatched sequences to convert into the other, a process called **gene conversion.**

Unequal crossing over is a crossover between chromosomal regions that are similar in nucleotide sequence but are not homologous. Gene conversion is the alteration of one homologue by the cell's error-detection and repair system to make it resemble the other homologue.

Trinucleotide Repeats

In 1991, a new kind of change in the genetic material was reported, one that involved neither changes in the identity of nucleotides (mutation) nor changes in the position of nucleotide sequences (recombination), but rather an increase in the number of copies of repeated trinucleotide sequences. Called **trinucleotide repeats,** these changes appear to be the root cause of a surprisingly large number of inherited human disorders.

The first examples of disorders resulting from the expansion of trinucleotide repeat sequences were reported in individuals with *fragile X syndrome* (the most common form of developmental disorder) and *spinal muscular atrophy*. In both disorders, genes containing runs of repeated nucleotide triplets (CGG in fragile X syndrome and CAG in spinal muscular atrophy) exhibit large increases in copy number. In individuals with fragile X syndrome, for example, the CGG sequence is repeated hundreds of times (figure 18.31), whereas in normal individuals it repeats only about 30 times.

Ten additional human genes are now known to have alleles with expanded trinucleotide repeats (figure 18.32). Many (but not all) of these alleles are GC-rich. A few of the alleles appear benign, but most are associated with heritable disorders, including Huntington's disease, myotonic dystrophy, and a variety of neurological ataxias. In each case, the expansion transmits as a dominant trait. Often the repeats are found within the exons of their genes, but sometimes, as in the case of fragile X syndrome, they are located outside the coding segment. Furthermore, although the repeat number is stably transmitted in normal families, it shows marked instability once it has abnormally expanded. Siblings often exhibit unique repeat lengths.

As the repeat number increases, disease severity tends to increase in step. In fragile X syndrome, the CGG triplet number first increases from the normal stable range of 5 to 55 times (the most common allele has 29 repeats) to an unstable number of repeats ranging from 50 to 200, with no detectable effect. In offspring, the number increases markedly, with copy numbers ranging from 200 to 1300, with significant mental retardation (see figure 18.31). Similarly, the normal allele for myotonic dystrophy has 5 GTC repeats. Mildly affected individuals have about 50, and severely affected individuals have up to 1000.

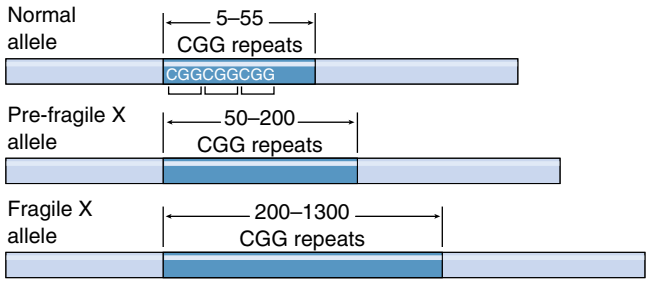


FIGURE 18.31
CGG repeats in fragile X alleles. The CGG triplet is repeated approximately 30 times in normal alleles. Individuals with pre-fragile X alleles show no detectable signs of the syndrome but do have increased numbers of CGG repeats. In fragile X alleles, the CGG triplet repeats hundreds of times.

Trinucleotide repeats appear common in human genes, but their function is unknown. Nor do we know the mechanism behind trinucleotide repeat expansion. It may involve unequal crossing over, which can readily produce copy-number expansion, or perhaps some sort of stutter in the DNA polymerase when it encounters a run of triplets. The fact that di- and tetranucleotide repeat expansions are not found seems an important clue. Undoubtedly, further examples of this remarkable class of genetic change will be reported in the future. Considerable research is currently focused on this extremely interesting area.

Many human genes contain runs of a trinucleotide sequence. Their function is unknown, but if the copy number expands, hereditary disorders often result.

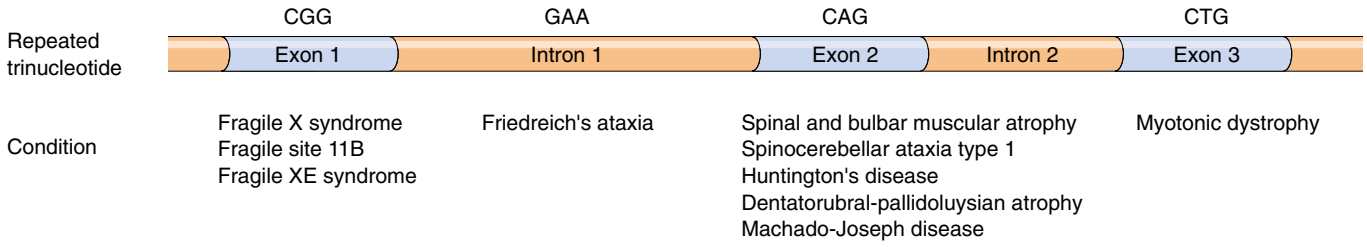


FIGURE 18.32
A hypothetical gene showing the locations and types of trinucleotide repeats associated with various human diseases. The CGG repeats of fragile X syndrome, fragile XE syndrome, and fragile site 11B occur in the first exon of their respective genes. GAA repeats characteristic of Friedreich's ataxia exist in the first intron of its gene. The genes for five different diseases, including Huntington's disease, have CAG repeats within their second exons. Lastly, the myotonic dystrophy gene contains CTG repeats within the third exon.

18.4 Genomes are continually evolving.

Classes of Eukaryotic DNA

The two main mechanisms of genetic recombination, gene transfer and reciprocal recombination, are directly responsible for the architecture of the eukaryotic chromosome. They determine where genes are located and how many copies of each exist. To understand how recombination shapes the genome, it is instructive to compare the effects of recombination in bacteria and eukaryotes.

Comparing Bacterial and Eukaryotic DNA Sequences

Bacterial genomes are relatively simple, containing genes that almost always occur as single copies. Unequal crossing over between repeated transposition elements in their circular DNA molecules tends to *delete* material, fostering the maintenance of a minimum genome size (figure 18.33*a*). For this reason, these genomes are very tightly packed, with few or no noncoding nucleotides. Recall the efficient use of space in the organization of the *lac* genes described in chapter 16.

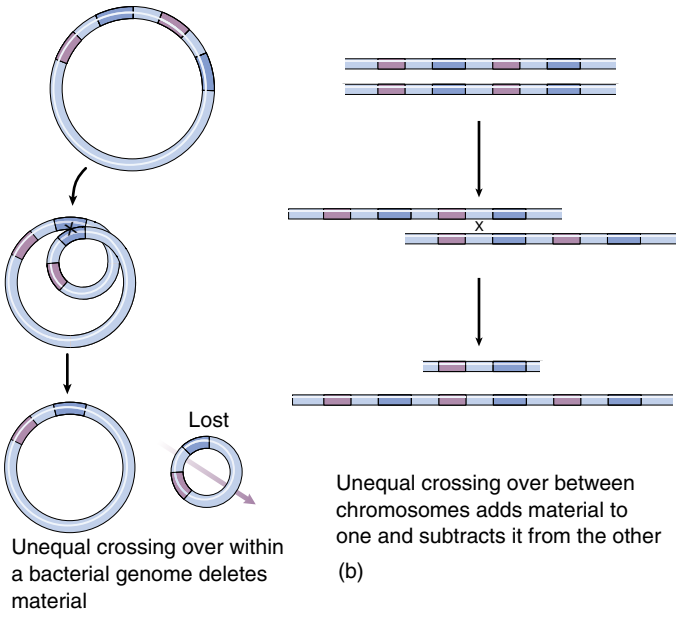


FIGURE 18.33
Unequal crossing over has different consequences in bacteria and eukaryotes. (*a*) Bacteria have a circular DNA molecule, and a crossover between duplicate regions within the molecule deletes the intervening material. (*b*) In eukaryotes, with two versions of each chromosome, crossing over adds material to one chromosome; thus, gene amplification occurs in that chromosome.

In eukaryotes, by contrast, the introduction of *pairs* of homologous chromosomes (presumably because of their importance in repairing breaks in double-stranded DNA) has led to a radically different situation. Unequal crossing over between homologous chromosomes tends to promote the *duplication* of material rather than its reduction (figure 18.33*b*). Consequently, eukaryotic genomes have been in a constant state of flux during the course of their evolution. Multiple copies of genes have evolved, some of them subsequently diverging in sequence to become different genes, which in turn have duplicated and diverged.

Six different classes of eukaryotic DNA sequences are commonly recognized, based on the number of copies of each (table 18.6).

Transposons

Transposons exist in multiple copies scattered about the genome. In *Drosophila*, for example, more than 30 different transposons are known, most of them present at 20 to 40 different sites throughout the genome. In all, the known transposons of *Drosophila* account for perhaps 5% of its DNA. Mammalian genomes contain fewer kinds of transposons than the genomes of many other organisms, although the transposons in mammals are repeated more often. The family of human transposons called *ALU* elements, for example, typically occurs about 300,000 times in each cell. Transposons are transcribed but appear to play no functional role in the life of the cell. As noted earlier in this chapter, many transposition events carry transposons into the exon portions of genes, disrupting the function of the protein specified by the gene transcript. These insertional inactivations are thought to be responsible for many naturally occurring mutations.

Tandem Clusters

A second class consists of DNA sequences that are repeated many times, one copy following another in tandem array. By transcribing all of the copies in these **tandem clusters** simultaneously, a cell can rapidly obtain large amounts of the product they encode. For example, the genes encoding rRNA are present in several hundred copies in most eukaryotic cells. Because these clusters are active sites of rRNA synthesis, they are readily visible in cytological preparations, where they are called **nucleolar organizer regions.** When transcription of the rRNA gene clusters ceases during cell division, the nucleolus disappears from view under the microscope, but it reappears when transcription begins again.

The genes present in a tandem cluster are very similar in sequence but not always identical; some may differ by one

Table 18.6 Classes of DNA Sequences Found in Eukaryotes

Class	Description
Transposons	Thousands of copies scattered around the genome
Tandem clusters	Clusters containing hundreds of nearly identical copies of a gene
Multigene families	Clusters of a few to several hundred copies of related but distinctly different genes
Satellite DNA	Short sequences present in millions of copies per genome
Dispersed pseudogenes	Inactive members of a multigene family separated from other members of the family
Single-copy genes	Genes that exist in only one copy in the genome

or a few nucleotides. Each gene in the cluster is separated from its neighbors by a short "spacer" sequence that is not transcribed. Unlike the genes, the spacers in a cluster vary considerably in sequence and in length.

Multigene Families

As we have learned more about the nucleotide sequences of eukaryotic genomes, it has become apparent that many genes exist as parts of **multigene families,** groups of related but distinctly different genes that often occur together in a cluster. Multigene families differ from tandem clusters in that they contain far fewer genes (from three to several hundred), and those genes differ much more from one another than the genes in tandem clusters. Despite their differences, the genes in a multigene family are clearly related in their sequences, making it likely that they arose from a single ancestral sequence through a series of unequal crossing over events. For example, studies of the evolution of the hemoglobin multigene family indicate that the ancestral globin gene is at least 800 million years old. By the time modern fishes evolved, this ancestral gene had already duplicated, forming the α and β forms. Later, after the evolutionary divergence of amphibians and reptiles, these two globin gene forms moved apart on the chromosome; the mechanism of this movement is not known, but it may have involved transposition. In mammals, two more waves of duplication occurred to produce the array of 11 globin genes found in the human genome. Three of these genes are silent, encoding nonfunctional proteins. Other genes are expressed only during embryonic (ζ and ε) or fetal (γ) development. Only four (δ, β, α_1, and α_2) encode the polypeptides that make up adult human hemoglobin.

Satellite DNA

Some short nucleotide sequences are repeated several million times in eukaryotic genomes. These sequences are collectively called **satellite DNA** and occur outside the main body of DNA. Almost all satellite DNA is either clustered around the centromere or located near the ends of the chromosomes, at the telomeres. These regions of the chromosomes remain highly condensed, tightly coiled, and untranscribed throughout the cell cycle; this suggests that satellite DNA may serve some sort of structural function, such as initiating the pairing of homologous chromosomes in meiosis. About 4% of the human genome consists of satellite DNA.

Dispersed Pseudogenes

Silent copies of a gene, inactivated by mutation, are called **pseudogenes.** Such mutations may affect the gene's promoter (see chapter 16), shift the reading frame of the gene, or produce a small deletion. While some pseudogenes occur within a multigene family cluster, others are widely separated. The latter are called **dispersed pseudogenes** because they are believed to have been dispersed from their original position within a multigene family cluster. No one suspected the existence of dispersed pseudogenes until a few years ago, but they are now thought to be of major evolutionary significance in eukaryotes.

Single-Copy Genes

Ever since eukaryotes appeared, processes such as unequal crossing over between different copies of transposons have repeatedly caused segments of chromosomes to duplicate, and it appears that no portion of the genome has escaped this phenomenon. The duplication of genes, followed by the conversion of some of the copies into pseudogenes, has probably been the major source of "new" genes during the evolution of eukaryotes. As pseudogenes accumulate mutational changes, a fortuitous combination of changes may eventually result in an active gene encoding a protein with different properties. When that new gene first arises, it is a **single-copy gene,** but in time it, too, will be duplicated. Thus, a single-copy gene is but one stage in the cycle of duplication and divergence that has characterized the evolution of the eukaryotic genome.

Gene sequences in eukaryotes vary greatly in copy number, some occurring many thousands of times, others only once. Many protein-encoding eukaryotic genes occur in several nonidentical copies, some of them not transcribed.

Chapter 18

www.mhhe.com

www.biocourse.com

Summary | Questions | Media Resources

18.1 Mutations are changes in the genetic message.

- A mutation is any change in the hereditary message.
- Mutations that change one or a few nucleotides are called point mutations. They may arise as a result of damage from ionizing or ultraviolet radiation, chemical mutagens, or errors in pairing during DNA replication.

1. What are pyrimidine dimers? How do they form? How are they repaired? What may happen if they are not repaired?

2. Explain how slipped mispairing can cause deletions and frame-shift mutations.

- Mutations
- DNA repair

- Experiment: Luria/Delbrück-Mutations Occur in Random

18.2 Cancer results from mutation of growth-regulating genes.

- Cancer is a disease in which the regulatory controls that normally restrain cell division are disrupted.
- A variety of environmental factors, including ionizing radiation, chemical mutagens, and viruses, have been implicated in causing cancer.
- The best way to avoid getting cancer is to avoid exposure to mutagens, especially those in cigarette smoke.

3. What is transfection? What has it revealed about the genetic basis of cancer?

4. About how many genes can be mutated to cause cancer? Why do most cancers require mutations in multiple genes?

- Polymerase Chain Reaction
- Student Research: Age and Breast Cancer

On Science Articles:

- Understanding Cancer
- Evidence Links Cigarette Smoking to Lung Cancer
- Deadly Cancer is Becoming More Common

18.3 Recombination alters gene location.

- Recombination is the creation of new gene combinations. It includes changes in the position of genes or fragments of genes as well as the exchange of entire chromosomes during meiosis.
- Genes may be transferred between bacteria when they are included within small circles of DNA called plasmids.
- Transposition is the random movement of genes within transposons to new locations in the genome. It is responsible for many naturally occurring mutations, as the insertion of a transposon into a gene often inactivates the gene.
- Crossing over involves a physical exchange of genetic material between homologous chromosomes during the close pairing that occurs in meiosis. It may produce chromosomes that have different combinations of alleles.

5. What is genetic recombination? What mechanisms produce it? Which of these mechanisms occurs in prokaryotes, and which occurs in eukaryotes?

6. What is a plasmid? What is a transposon? How are plasmids and transposons similar, and how are they different?

7. What are mismatched pairs? How are they corrected? What effect does this correction have on the genetic message?

- Recombinant DNA/Technology

- Experiments: McClintock/Stern

18.4 Genomes are continually evolving.

- Satellite sequences are short sequences of nucleotides repeated millions of times.
- Tandem clusters are genes that occur in thousands of copies grouped together at one or a few sites on a chromosome. These genes encode products that are required by the cell in large amounts.
- Multigene families consist of copies of genes clustered at one site on a chromosome that diverge in sequence more than the genes in a tandem cluster.

8. What kinds of genes exist in multigene families? How are these families thought to have evolved?

9. What are pseudogenes? How might they have been involved in the evolution of single-copy genes?

- Student research: DNA repair in fish