# Genetic Control of Cell Function

G enetic information is stored in the structure of *deoxyribonucleic acid* (DNA). DNA is an extremely stable macromolecule found in the nucleus of each cell. Because of the stable structure of DNA, the genetic information can survive the many processes of reduction division, in which the gametes (*i.e.*, ovum and sperm) are formed, and the fertilization process. This stability is also maintained throughout the many mitotic cell divisions involved in the formation of a new organism from the single-celled fertilized ovum called the *zygote*.

The term *gene* is used to describe a part of the DNA molecule that contains the information needed to code for the types of proteins and enzymes needed for the day-to-day function of the cells in the body. In addition, a gene is the unit of heredity passed from generation to generation. For example, genes control the type and quantity of hormones that a cell produces, the antigens and receptors that are present on the cell membrane, and the synthesis of enzymes needed for metabolism. Of the estimated 35,000 to 140,000 genes that humans possess, more than 5000 have been identified and more than 2300 have been localized to a particular chromosome. With few exceptions, each gene provides the instructions for the synthesis of a single protein. This chapter includes discussions of genetic regulation of cell function, chromosomal structure, patterns of inheritance, and gene technology.

GENETIC CONTROL OF CELL FUNCTION

The genetic information needed for protein synthesis is encoded in the DNA contained in the cell nucleus. A second type of nucleic acid, *ribonucleic acid* (RNA), is involved in the actual synthesis of cellular enzymes and proteins. Cells contain several types of RNA: messenger RNA, transfer RNA, and ribosomal RNA. *Messenger RNA* (mRNA) contains the transcribed instructions for protein synthesis obtained from the DNA molecule and carries them into the cytoplasm. Transcription is followed by translation, the synthesis of proteins according to the instructions carried by mRNA. *Ribosomal RNA* (rRNA) provides the machinery needed for protein synthesis. *Transfer RNA* (tRNA) reads the instructions and delivers the appropriate amino acids to the ribosome, where they are incorporated into the protein being synthesized.

The mechanism for genetic control of cell function is illustrated in Figure 3-1. The nuclei of all the cells in an organism contain the same accumulation of genes derived from the gametes of the two parents. This means that liver cells contain the same genetic information as skin and muscle cells. For this to be true, the molecular code must be duplicated before each succeeding cell division, or mitosis. In theory, although this has not yet been achieved in humans, any of the highly differentiated cells of an organism could be used to produce a complete, genetically identical organism, or *clone.* Each particular cell type

in a tissue uses only part of the information stored in the genetic code. Although information required for the development and differentiation of the other cell types is still present, it is repressed.

Besides nuclear DNA, part of the DNA of a cell resides in the mitochondria. Mitochondrial DNA is inherited from the mother by her offspring (*i.e.*, matrilineal inheritance). Several
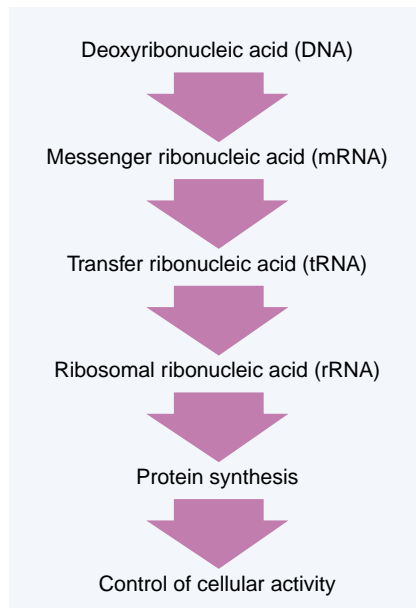
■ FIGURE 3-1 ■ DNA-directed control of cellular activity through synthesis of cellular proteins. Messenger RNA carries the transcribed message, which directs protein synthesis, from the nucleus to the cytoplasm. Transfer RNA selects the appropriate amino acids and carries them to ribosomal RNA where assembly of the proteins takes place.

genetic disorders are attributed to defects in mitochondrial DNA. Leber's hereditary optic neuropathy was the first human disease attributed to mutation in mitochondrial DNA.

Gene Structure

The structure that stores the genetic information in the nucleus is a long, double-stranded, helical molecule of DNA. DNA is composed of *nucleotides*, which consist of phosphoric acid, a five-carbon sugar called *deoxyribose*, and one of four nitrogenous bases. These nitrogenous bases carry the genetic information and are divided into two groups: the *purine bases*, adenine and guanine, which have two nitrogen ring structures, and the *pyrimidine bases*, thymine and cytosine, which have one ring. The backbone of DNA consists of alternating groups of sugar and phosphoric acid; the paired bases project inward from the sides of the sugar molecule. DNA resembles a spiral staircase, with the paired bases representing the steps (Fig. 3-2). A precise complementary pairing of purine and pyrimidine bases occurs in the double-stranded DNA molecule. Adenine is paired with thymine, and guanine is paired with cytosine. Each nucleotide in a pair is on one strand of the DNA molecule, with the bases on opposite DNA strands bound together by hydrogen bonds that are extremely stable under normal conditions. Enzymes called *DNA helicases* separate the two strands so that the genetic information can be duplicated or transcribed.

Several hundred to almost one million base pairs can represent a gene; the size is proportional to the protein product it encodes. Of the two DNA strands, only one is used in transcribing the information for the cell's polypeptide-building machinery. The genetic information of one strand is meaningful and is used as a template for transcription; the complementary code of the other strand does not make sense and is ignored. However, both strands are involved in DNA duplication. Before cell division, the two strands of the helix separate and a complementary molecule is duplicated next to each original strand. Two strands become four strands. During cell division, the newly duplicated double-stranded molecules are separated and placed in each daughter cell by the mechanics of mitosis. As a result, each of the daughter cells again contains the meaningful strand and the complementary strand joined in the form of a double helix. This type of DNA replication has been termed *semiconservative* as opposed to conservative (Fig. 3-3).

The DNA molecule is combined with several types of protein and small amounts of RNA into a complex known as *chromatin*. Chromatin is the readily stainable portion of the cell nucleus. Some DNA proteins form binding sites for repressor molecules and hormones that regulate genetic transcription; others may block genetic transcription by preventing access of nucleotides to the surface of the DNA molecule. A specific group of proteins called *histones* are thought to control the folding of the DNA strands.
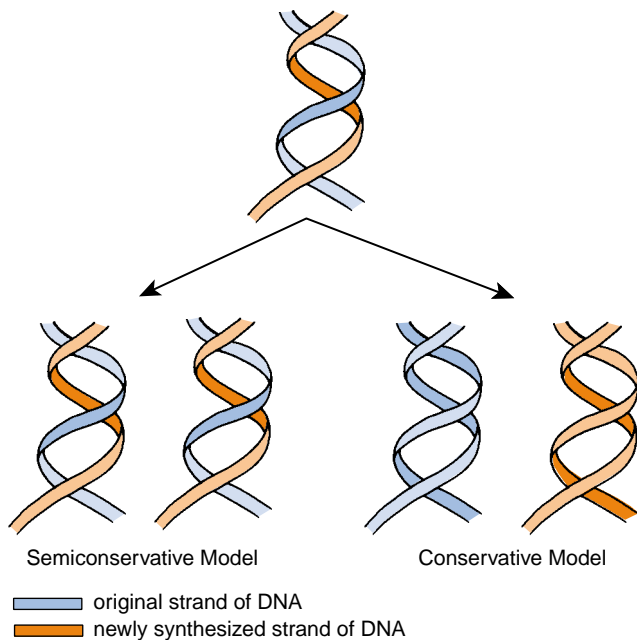
Genetic Code

The four bases—guanine, adenine, cytosine, and thymine (uracil is substituted for thymine in RNA)—make up the alphabet of the genetic code. A sequence of three of these bases forms the fundamental triplet code used in transmitting the genetic information needed for protein synthesis. This triplet code is called
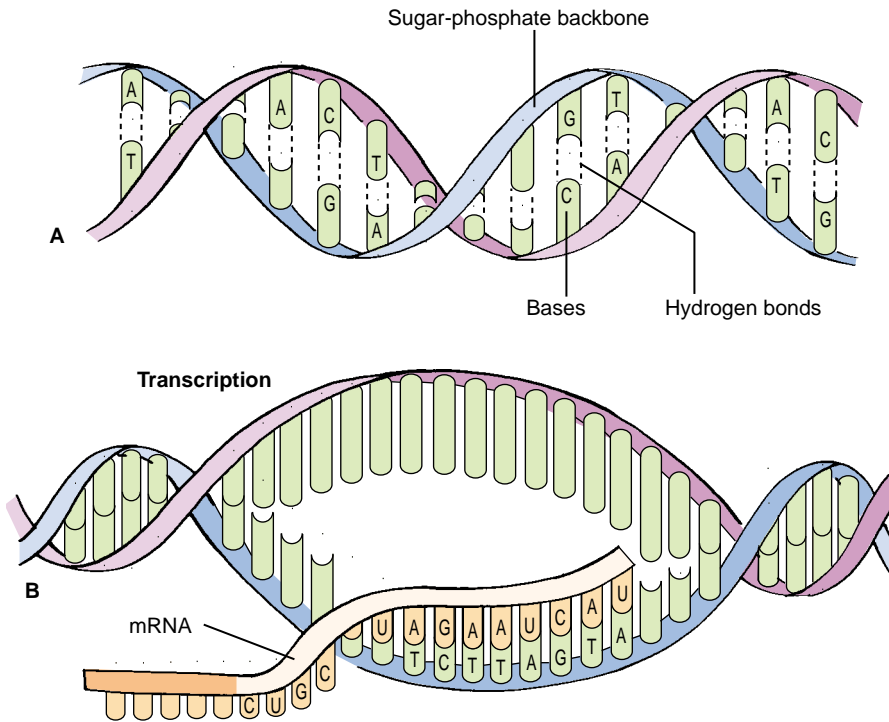
■ FIGURE 3-2 ■ The DNA double helix and transcription of messenger RNA (mRNA). The top panel (**A**) shows the sequence of four bases (adenine [A], cytosine [C], guanine [G], and thymine [T]), which determines the specificity of genetic information. The bases face inward from the sugar-phosphate backbone and form pairs (*dashed lines*) with complementary bases on the opposing strand. In the bottom panel (**B**), transcription creates a complementary mRNA copy from one of the DNA strands in the double helix.

a *codon* (Table 3-1). An example is the nucleotide sequence GCU (guanine, cytosine, and uracil), which is the triplet RNA code for the amino acid alanine. The genetic code is a universal language used by most living cells (*i.e.,* the code for the amino acid tryptophan is the same in a bacterium, a plant, and a human being). *Stop codes,* which signal the end of a protein molecule, are also present. Mathematically, the four bases can be arranged in 64 different combinations. Sixty-one of the triplets correspond to particular amino acids, and three are stop signals. Only 20 amino acids are used in protein synthesis in humans. Several triplets code for the same amino acid; therefore, the genetic code is said to be redundant or *degenerate.* For example, AUG is a part of the initiation or start signal and the codon for the amino acid methionine. Codons that specify the same amino acid are called *synonyms.* Synonyms usually have the same first two bases but differ in the third base.

Protein Synthesis

Although DNA determines the type of biochemical product that the cell synthesizes, the transmission and decoding of information needed for protein synthesis are carried out by RNA, the formation of which is directed by DNA. The general structure of RNA differs from DNA in three respects: RNA is a single-stranded, rather than a double-stranded, molecule; the sugar in each nucleotide of RNA is ribose, instead of deoxyribose; and the pyrimidine base thymine in DNA is replaced by uracil in RNA. Three types of RNA are known: mRNA, tRNA, and rRNA. All three types are synthesized in the nucleus by RNA polymerase enzymes that take directions from DNA. Because the ribose sugars found in RNA are more susceptible to degradation than the sugars in DNA, the types of RNA molecules in the cytoplasm can be altered rapidly in response to extracellular signals.

Messenger RNA

Messenger RNA is the template for protein synthesis. It is a long molecule containing several hundred to several thousand nucleotides. Each group of three nucleotides forms a codon that

■ FIGURE 3-3 ■ Semiconservative vs conservative models of DNA replication as proposed by Meselson and Stahl in 1958. In semiconservative DNA replication, the two original strands of DNA unwind and a complementary strand is formed along each original strand.

Amino Acid	RNA Codons					
Alanine	GCU	GCC	GCA	GCG		
Arginine	CGU	CGC	CGA	CGG	AGA	AGG
Asparagine	AAU	AAC				
Aspartic acid	GAU	GAC				
Cysteine	UGU	UGC				
Glutamic acid	GAA	GAG				
Glutamine	CAA	CAG				
Glycine	GGU	GGC	GGA	GGG		
Histidine	CAU	CAC				
Isoleucine	AUU	AUC	AUA			
Leucine	CUU	CUC	CUA	CUG	UUA	UUG
Lysine	AAA	AAG				
Methionine	AUG					
Phenylalanine	UUU	UUC				
Proline	CCU	CCC	CCA	CCG		
Serine	UCU	UCC	UCA	UCG	AGC	AGU
Threonine	ACU	ACC	ACA	ACG		
Tryptophan	UGG					
Tyrosine	UAU	UAC				
Valine	GUU	GUC	GUA	GUG		
Start (CI)	AUG					
Stop (CT)	UAA	UAG	UGA			

TABLE 3-1 Triplet Codes for Amino Acids

(Guyton A. [2000]. *Textbook of medical physiology* [10th ed., p. 28]. Philadelphia: W.B. Saunders)

is exactly complementary to the triplet of nucleotides of the DNA molecule. Messenger RNA is formed by a process called *transcription,* in which the weak hydrogen bonds of the DNA are broken so that free RNA nucleotides can pair with their exposed DNA counterparts on the meaningful strand of the DNA molecule (see Fig. 3-2). As with the base pairing of the DNA strands, complementary RNA bases pair with the DNA bases. In RNA, uracil replaces thymine and pairs with adenine.

During transcription, a specialized nuclear enzyme, called *RNA polymerase,* recognizes the beginning or start sequence of a gene. The RNA polymerase attaches to the double-stranded DNA and proceeds to copy the meaningful strand into a single strand of RNA as it travels along the length of the gene. On reaching the stop signal, the enzyme leaves the gene and releases the RNA strand. The RNA strand then is processed. Processing involves the addition of certain nucleic acids at the ends of the RNA strand and cutting and splicing of certain internal sequences. Splicing often involves the removal of stretches of RNA. Because of the splicing process, the final mRNA sequence is different from the original DNA template. RNA sequences that are retained are called *exons,* and those excised are called *introns.* The functions of the introns are unknown. They are thought to be involved in the activation or deactivation of genes during various stages of development.

Splicing permits a cell to produce a variety of mRNA molecules from a single gene. By varying the splicing segments of the initial mRNA, different mRNA molecules are formed. For example, in a muscle cell, the original tropomyosin mRNA is spliced in as many as 10 different ways, yielding distinctly different protein products. This permits different proteins to be expressed from a single gene and reduces how much DNA must be contained in the genome.

Transfer RNA

The clover-shaped tRNA molecule contains only 80 nucleotides, making it the smallest RNA molecule. Its function is to deliver the activated form of amino acids to protein molecules in the ribosomes. At least 20 different types of tRNA are known, each of which recognizes and binds to only one type of amino acid. Each tRNA molecule has two recognition sites: the first is complementary for the mRNA codon and the second is for the amino acid itself. Each type of tRNA carries its own specific amino acid to the ribosomes, where protein synthesis is taking place; there it recognizes the appropriate codon on the mRNA and delivers the amino acid to the newly forming protein molecule.
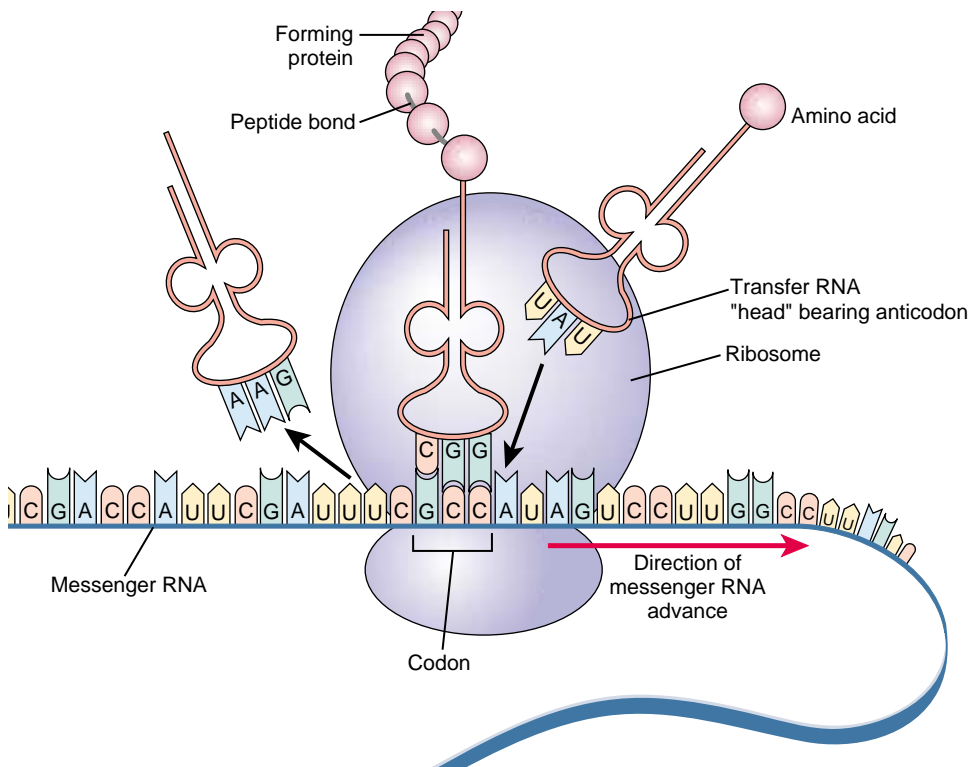
Ribosomal RNA

The ribosome is the physical structure in the cytoplasm where protein synthesis takes place. Ribosomal RNA forms 60% of the ribosome, with the remainder of the ribosome composed of the structural proteins and enzymes needed for protein synthesis. As with the other types of RNA, rRNA is synthesized in the nucleus. Unlike other RNAs, ribosomal RNA is produced in a specialized nuclear structure called the *nucleolus.* The formed rRNA combines with ribosomal proteins in the nucleus to produce the ribosome, which is then transported into the cytoplasm. On reaching the cytoplasm, most ribosomes become attached to the endoplasmic reticulum and begin the task of protein synthesis.

Proteins are made from a standard set of amino acids, which are joined end to end to form the long polypeptide chains of protein molecules. Each polypeptide chain may have as many as 100 to more than 300 amino acids in it. The process of protein synthesis is called *translation* because the genetic code is translated into the production language needed for protein assembly. Besides rRNA, translation requires the coordinated actions of mRNA and tRNA. Each of the 20 different tRNA molecules transports its specific amino acid to the ribosome for incorporation into the developing protein molecule. Messenger RNA provides the information needed for placing the amino acids in their proper order for each specific type of protein. During protein synthesis, mRNA contacts and passes through the ribosome, which "reads" the directions for protein synthesis in much the same way that a tape is read as it passes through a tape player (Fig. 3-4). As mRNA passes through the ribosome, tRNA delivers the appropriate amino acids for attachment to the growing polypeptide chain. The long mRNA molecule usually travels through and directs protein synthesis in more than one ribosome at a time. After the first part of the mRNA is read by the first ribosome, it moves onto a second and a third. As a result, ribosomes that are actively involved in protein synthesis are often found in clusters called *polyribosomes.*

Regulation of Gene Expression

Although all cells contain the same genes, not all genes are active all of the time, nor are the same genes active in all cell types. On the contrary, only a small, select group of genes is active in directing protein synthesis in the cell, and this group varies from one cell type to another. For the differentiation

■ **FIGURE 3-4** ■ Protein synthesis. A messenger RNA strand is moving through a small ribosomal RNA subunit in the cytoplasm. Transfer RNA transports amino acids to the mRNA strand and recognizes the RNA codon calling for the amino acid by base pairing (through its anticodon). The ribosome then adds the amino acid to the growing polypeptide chain. The ribosome moves along the mRNA strand and is read sequently. As each amino acid is bound to the next by a peptide bond, its tRNA is released.

process of cells to occur in the various organs and tissues of the body, protein synthesis in some cells must be different from that in others. To adapt to an ever-changing environment, certain cells may need to produce varying amounts and types of proteins. Certain enzymes, such as carbonic anhydrase, are synthesized by all cells for the fundamental metabolic processes on which life depends.

The degree to which a gene or particular group of genes is active is called *gene expression*. A phenomenon termed *induction* is an important process by which gene expression is increased. Except in early embryonic development, induction is promoted by some external influence. *Gene repression* is a process whereby a regulatory gene acts to reduce or prevent gene expression. Some genes are normally dormant but can be activated by inducer substances; other genes are naturally active and can be inhibited by repressor substances. Genetic mechanisms for the control of protein synthesis are better understood in microorganisms than in humans. However, it can be assumed that the same general principles apply.

The mechanism that has been most extensively studied is the one by which the synthesis of particular proteins can be turned on and off. For example, in the bacterium *Escherichia coli* grown in a nutrient medium containing the disaccharide lactose, the enzyme galactosidase can be isolated. The galactosidase catalyzes the splitting of lactose into a molecule of glucose and a molecule of galactose. This is necessary if lactose is to be metabolized by *E. coli*. However, if the *E. coli* is grown in a medium that does not contain lactose, very little of the enzyme is produced. From these and other studies, it is theorized that the synthesis of a particular protein, such as galactosidase, requires a series of reactions, each of which is catalyzed by a specific enzyme.

At least two types of genes control protein synthesis: *structural genes* that specify the amino acid sequence of a polypeptide chain and *regulator genes* that serve a regulatory function without stipulating the structure of protein molecules. The regulation of protein synthesis is controlled by a sequence of genes, called an *operon*, on adjacent sites on the same chromosome (Fig. 3-5). An operon consists of a set of structural genes that code for enzymes used in the synthesis of a particular product and a promoter site that binds RNA polymerase and initiates transcription of the structural genes. The function of the operon is further regulated by activator and repressor opera-
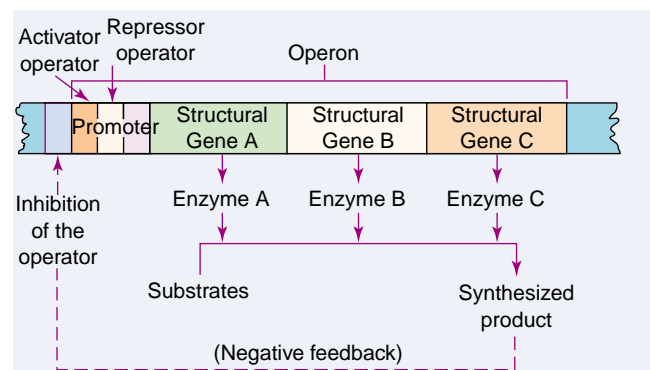
■ **FIGURE 3-5** ■ Function of the operon to control biosynthesis. The synthesized product exerts negative feedback to inhibit function of the operon, in this way automatically controlling the concentration of the product itself. (Guyton A., Hall J.E. [2000]. *Textbook of medical physiology* [10th ed., p. 31]. Philadelphia: W.B. Saunders with permission from Elsevier Science)

tors, which induce or repress the function of the promoter. The activator and repressor sites commonly monitor levels of the synthesized product and regulate the activity of the operon through a negative feedback mechanism. When product levels decrease, the function of the operon is activated, and when levels increase, its function is repressed. Regulatory genes found elsewhere in the genetic complex can exert control over an operon through activator or repressor substances. Not all genes are subject to induction and repression.

Gene Mutations

Rarely, accidental errors in duplication of DNA occur. These errors are called *mutations.* Mutations result from the substitution of one base pair for another, the loss or addition of one or more base pairs, or rearrangements of base pairs. Many of these mutations occur spontaneously; others occur because of environmental agents, chemicals, and radiation. Mutations may arise in somatic cells or in germ cells (ovum or sperm). Only those DNA changes that occur in germ cells can be inherited. A somatic mutation affects a cell line that differentiates into one or more of the many tissues of the body and is not transmissible to the next generation. Somatic mutations that do not have an impact on the health or functioning of a person are called *polymorphisms.* Occasionally, a person is born with one brown eye and one blue eye because of a somatic mutation. The change or loss of gene information is just as likely to affect the fundamental processes of cell function or organ differentiation. Such somatic mutations in the early embryonic period can result in embryonic death or congenital malformations. Somatic mutations are important causes of cancer and other tumors in which cell differentiation and growth get out of control. Each year, hundreds of thousands of random changes occur in the DNA molecule because of environmental events or metabolic accidents. Fortunately, fewer than 1 in 1000 base pair changes result in serious mutations. Most of these defects are corrected by DNA repair mechanisms. Several mechanisms exist, and each depends on specific enzymes such as DNA repair nucleases. Fishermen, farmers, and others who are excessively exposed to the ultraviolet radiation of sunlight have an increased risk for development of skin cancer as a result of potential radiation damage to the genetic structure of the skin-forming cells.

In summary, genes are the fundamental unit of information storage in the cell. They determine the types of proteins and enzymes made by the cell and therefore control inheritance and day-to-day cell function. Genes store information in a stable macromolecule called *DNA.* Genes transmit information contained in the DNA molecule as a triplet code. The arrangement of the nitrogenous bases of the four nucleotides (*i.e.,* adenine, guanine, thymine [or uracil in RNA], and cytosine) forms the code. The transfer of stored information into production of cell products is accomplished through a second type of macromolecule called *RNA.* Messenger RNA transcribes the instructions for product synthesis from the DNA molecule and carries it into the cell's cytoplasm, where ribosomal RNA uses the information to direct product synthesis. Transfer RNA acts as a carrier system for delivering the appropriate amino acids to the ribosomes, where the synthesis of cell products occurs. Although all cells contain the same genes, only a small, select group of genes is active in a given cell type. In all cells, some genetic information is repressed, whereas other information is expressed. Gene mutations represent accidental errors in duplication, rearrangement, or deletion of parts of the genetic code. Fortunately, most mutations are corrected by DNA repair mechanisms in the cell.

CHROMOSOMES

Most genetic information of a cell is organized, stored, and retrieved in small intracellular structures called *chromosomes.* Although the chromosomes are visible only in dividing cells, they retain their integrity between cell divisions. The chromosomes are arranged in pairs; one member of the pair is inherited from the father, the other from the mother. Each species has a characteristic number of chromosomes. In the human, 46 single or 23 pairs of chromosomes are present. Of the 23 pairs of human chromosomes, there are 22 pairs called *autosomes* that are alike in males and females. Each of the 22 pairs of autosomes has the same appearance in all individuals, and each has been given a numeric designation for classification purposes (Fig. 3-6).

The sex chromosomes make up the 23rd pair of chromosomes. Two sex chromosomes determine the sex of a person. All males have an X and Y chromosome (*i.e.,* an X chromosome from the mother and a Y chromosome from the father); all females have two X chromosomes (*i.e.,* one from each parent). Only one X chromosome in the female is active in controlling the expression of genetic traits; however, both X chromosomes are activated during gametogenesis (*i.e.,* formation of the germ cell or ovum). In the female, the active X chromosome is invisible, but the inactive X chromosome can be demonstrated with appropriate nuclear staining. This inactive chromatin mass is seen as the *Barr body* in epithelial cells or as the drumstick body in the chromatin of neutrophils. The genetic sex of a child can be determined by microscopic study of cell or tissue samples. The total number of X chromosomes is equal to the number of Barr bodies plus one (*i.e.,* an inactive plus an active X chromosome). For example, the cells of a normal female have one Barr body and therefore a total of two X chromosomes. A normal male has no Barr bodies. Males with Klinefelter's syndrome (one Y, an inactive X, plus an active X chromosome) exhibit one Barr body. In the female, whether the X chromosome derived from the mother or that derived from the father is active is determined within a few days after conception; the selection is random for each postmitotic cell line. This is called the *Lyon principle,* after Mary Lyon, the British geneticist who described it.

Cell Division

There are two types of cell division: mitosis and meiosis. *Meiosis* is limited to replicating germ cells and takes place only once in a cell line. It results in the formation of gametes or reproductive cells (*i.e.,* ovum and sperm), each of which has only a single set of 23 chromosomes. Meiosis is typically divided into two distinct phases, meiotic divisions I and II. Similar to mitosis, cells
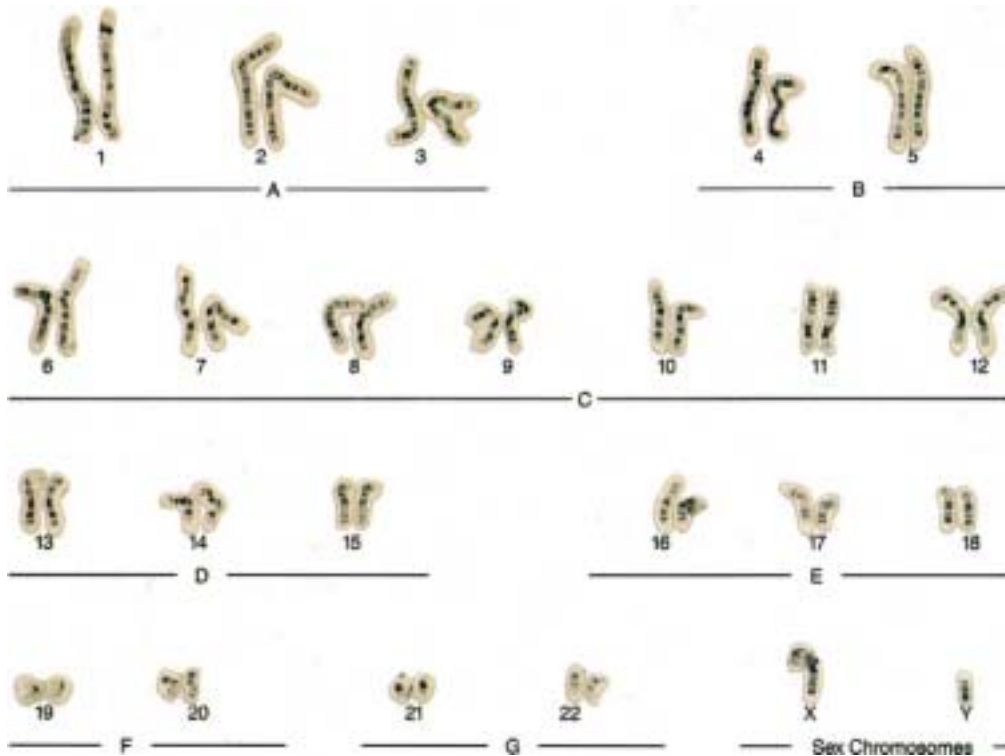
■ **FIGURE 3-6** ■ Karyotype of normal human boy. (Courtesy of the Prenatal Diagnostic and Imaging Center, Sacramento, CA. Frederick W. Hansen, MD, Medical Director)

about to undergo the first meiotic division replicate their DNA during interphase. During metaphase I, homologous chromosomes pair up, forming a synapsis or tetrad (two chromatids per chromosome). They are sometimes called *bivalents*. The X and Y chromosomes are not homologs and do not form bivalents. While in metaphase I, an interchange of chromatid segments can occur. This process is called *crossing over* (Fig. 3-7). Crossing over allows for new combinations of genes, increasing genetic variability. After telophase I, each of the two daughter cells contains one member of each homologous pair of chromosomes and a sex chromosome (23 double-stranded chromosomes).

No DNA synthesis occurs before meiotic division II. During anaphase II, the 23 double-stranded chromosomes (two chromatids) of each of the two daughter cells from meiosis I divide at their centromeres. Each subsequent daughter cell receives 23 single-stranded chromatids. Thus, a total of four daughter cells are formed by a meiotic division of one cell (Fig. 3-8).

Meiosis, which occurs only in the gamete-producing cells found in either testes or ovaries, has a different outcome in males and females. In males, meiosis (spermatogenesis) results in four viable daughter cells called *spermatids* that differentiate into sperm cells. In females, gamete formation or oogenesis is quite different. After the first meiotic division of a primary oocyte, a secondary oocyte and another structure called a *polar body* are formed. This small polar body contains little cyto-

KEY CONCEPTS

CHROMOSOME STRUCTURE

■ The DNA that stores genetic material is organized into 23 pairs of chromosomes. There are 22 pairs of autosomes, which are alike for males and females, and one pair of sex chromosomes, with XX pairing in females and XY pairing in males.

■ Mitosis refers to the duplication of chromosomes in somatic cell lines, in which each daughter cell receives a pair of 23 chromosomes.

■ Meiosis is limited to replicating germ cells and results in the formation of a single set of 23 chromosomes.
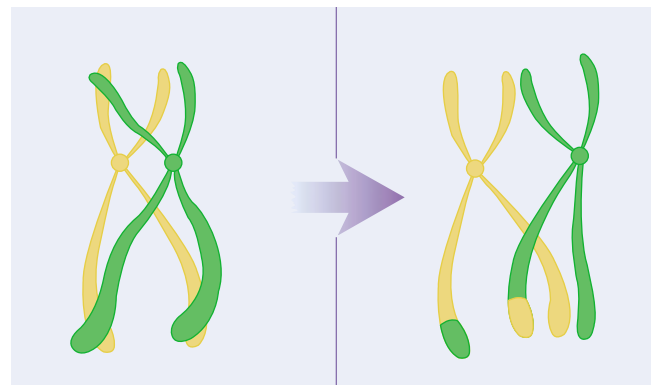
■ **FIGURE 3-7** ■ Crossing over of DNA at the time of meiosis.
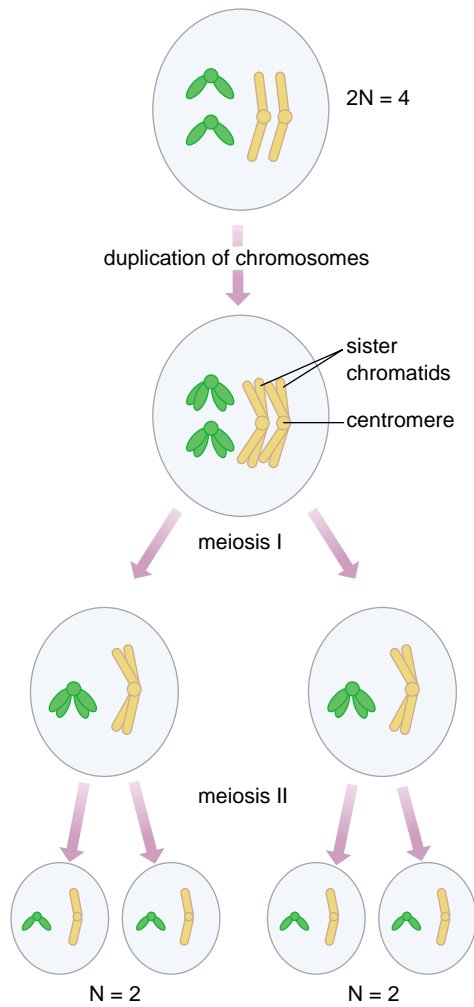
■ **FIGURE 3-8** ■ Separation of chromosomes at the time of meiosis.

■ **FIGURE 3-9** ■ Essential stages of meiosis in a female, with discarding of the first and second polar bodies and formation of the ovum with the haploid number of chromosomes. The *d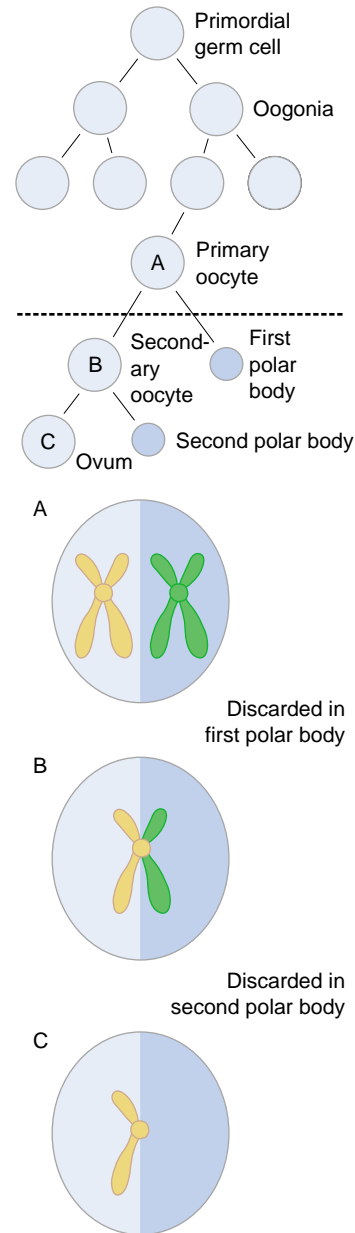otted line* indicates reduction division. (Adapted from Cormack D.H. [1993]. *Essential histology.* Philadelphia: J.B. Lippincott)

plasm, but it may undergo a second meiotic division, resulting in two polar bodies (Fig. 3-9). The secondary oocyte undergoes its second meiotic division, producing one mature oocyte and another polar body. Four viable sperm cells are produced during spermatogenesis, but only one ovum from oogenesis.

Chromosome Structure

Cytogenetics is the study of the structure and numeric characteristics of the cell's chromosomes. Chromosome studies can be done on any tissue or cell that grows and divides in culture. Lymphocytes from venous blood are frequently used for this purpose. After the cells have been cultured, a drug called *colchicine* is used to arrest mitosis in metaphase. A chromosome spread is prepared by fixing and spreading the chromosomes on a slide. Subsequently, appropriate staining techniques show the chromosomal banding patterns so they can be identified. The chromosomes are photographed, and the photomicrograph of each chromosome is cut out and arranged in pairs according to a standard classification system. The completed picture is called a *karyotype,* and the procedure for preparing the picture is called *karyotyping.* A uniform system of chromosome classification was originally formulated at the 1971 Paris Chro-

mosome Conference and was later revised to describe the chromosomes as seen in more elongated prophase and prometaphase preparations.

In the metaphase spread, each chromosome takes the form of chromatids to form an "X" or "wishbone" pattern. Human chromosomes are divided into three types according to the position of the centromere (*i.e.,* central constriction; Fig. 3-10). If the centromere is in the center and the arms are of approximately the same length, the chromosome is said to be metacentric; if it is off center and the arms are of clearly different lengths, it is *submetacentric*; and if it is near one end, it is *acrocentric.* The short arm of the chromosome is designated
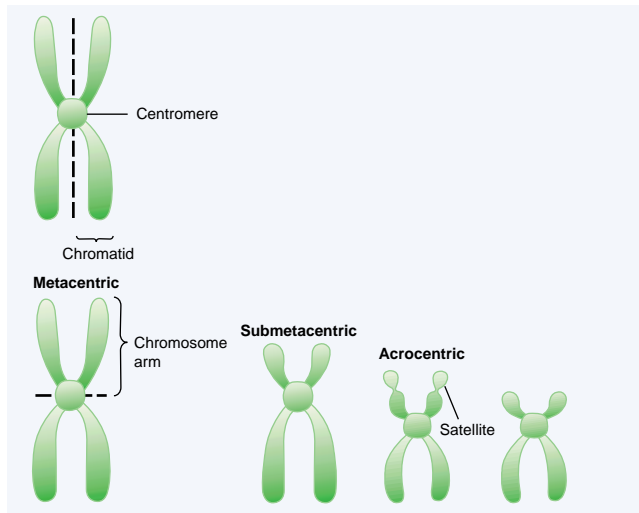
■ **FIGURE 3-10** ■ Three basic shapes and the component parts of human metaphase chromosomes. The relative size of the satellite on the acrocentric is exaggerated for visibility. (Adapted from Cormack D.H. [1993]. *Essential histology.* Philadelphia: J.B. Lippincott)

■ **FIGURE 3-11** ■ The localization of representative inherited diseases on the X chromosome. Notice the nomenclature of arms (P,Q), regions (1,2), bands (*e.g.,* 11,22). (Rubin E., Farber J.L. [1999]. *Pathology* [3rd ed., p. 260]. Philadelphia: Lippincott Williams & Wilkins)

as "p" for "petite," and the long arm is designated as "q" for no other reason than it is the next letter of the alphabet. Arms of the chromosome are indicated by the chromosome number followed by the p or q designation. Chromosomes 13, 14, 15, 21, and 22 have small masses of chromatin called *satellites* attached to their short arms by narrow stalks. At the ends of each chromosome are special DNA sequences called *telomeres.* Telomeres allow the end of the DNA molecule to be replicated completely.

The banding patterns of a chromosome are used in describing the position of a gene. Regions on the chromosomes are numbered from the centromere outward. The regions are further divided into bands and subbands, which are also numbered (Fig. 3-11). These numbers are used in designating the position of a gene on a chromosome. For example, Xp22.2 refers to subband 2, band 2, region 2 of the short arm (p) of the X chromosome.

In summary, the genetic information in a cell is organized, stored, and retrieved as small cellular structures called *chromosomes.* Forty-six chromosomes arranged in 23 pairs are present in the human being. Twenty-two of these pairs are autosomes. The 23rd pair is the sex chromosomes, which determine the sex of a person. Two types of cell division occur, meiosis and mitosis. Meiosis is limited to replicating germ cells and results in the formation of gametes or reproductive cells (ovum and sperm), each of which has only a single set of 23 chromosomes. Mitotic division occurs in somatic cells and results in the formation of 23 pairs of chromosomes. A karyotype is a photograph of a person's chromosomes. It is prepared by special laboratory techniques in which body cells are cultured, fixed, and then stained to demonstrate identifiable banding patterns. A photomicrograph is then made. Often the individual chromosomes are cut out and regrouped according to chromosome number.
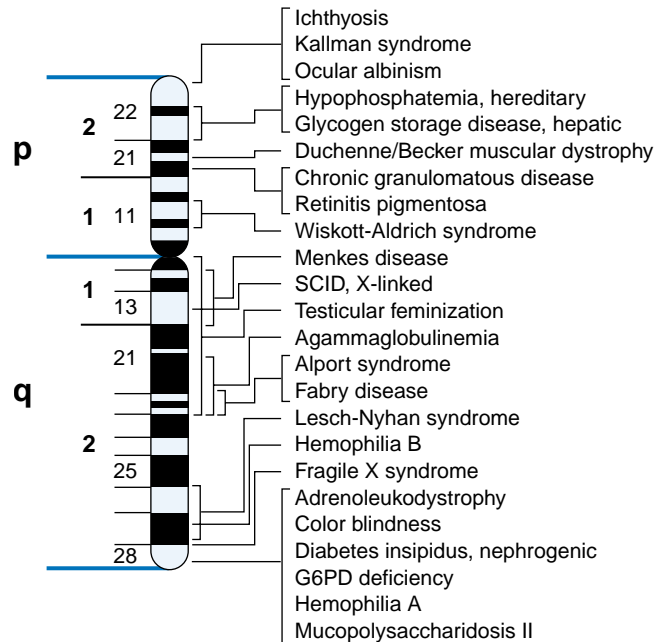
PATTERNS OF INHERITANCE

The characteristics inherited from a person's parents are inscribed in gene pairs located along the length of the chromosomes. Alternate forms of the same gene are possible (*i.e.,* one inherited from the mother and the other from the father), and each may produce a different aspect of a trait.

KEY CONCEPTS

TRANSMISSION OF GENETIC INFORMATION

- The transmission of information from one generation to the next is vested in genetic material transferred from each parent at the time of conception.

- Alleles are the alternate forms of a gene (one from each parent), and the locus is the position that they occupy on the chromosome.

- The genotype of a person represents the sum total of the genetic information in the cells and the phenotype the physical manifestations of that information.

- Penetrance is the percentage in a population with a particular genotype in which that genotype is phenotypically manifested, whereas expressivity is the manner in which the gene is expressed.

Definitions

Genetics has its own set of definitions. The *genotype* of a person is the genetic information stored in the base sequence triplet code. The *phenotype* refers to the recognizable traits, physical or biochemical, associated with a specific genotype. Often, the genotype is not evident by available detection methods. More than one genotype may have the same phenotype. Some brown-eyed persons are carriers of the code for blue eyes, and other brown-eyed persons are not. Phenotypically, these two types of brown-eyed persons are the same, but genotypically they are different.

When it comes to a genetic disorder, not all persons with a mutant gene are affected to the same extent. *Expressivity* refers to the manner in which the gene is expressed in the phenotype, which can range from mild to severe. *Penetrance* represents the ability of a gene to express its function. Seventy-five percent penetrance means 75% of persons of a particular genotype demonstrate a recognizable phenotype. Syndactyly (webbed fingers) and blue sclera are genetic mutations that often do not exhibit 100% penetrance.

The position of a gene on a chromosome is called its *locus,* and alternate forms of a gene at the same locus are called *alleles.* When only one pair of genes is involved in the transmission of information, the term *single-gene trait* is used. Single-gene traits follow the Mendelian laws of inheritance.

Polygenic inheritance involves multiple genes at different loci, with each gene exerting a small additive effect in determining a trait. Most human traits are determined by multiple pairs of genes, many with alternate codes, accounting for some of the dissimilar forms that occur with certain genetic disorders. Polygenic traits are predictable but less so than single-gene traits. *Multifactorial* inheritance is similar to polygenic inheritance in that multiple alleles at different loci affect the outcome; the difference is that multifactorial inheritance includes environmental effects on the genes.

Many other gene–gene interactions are known. These include *epistasis,* in which one gene masks the phenotypic effects of another nonallelic gene; *multiple alleles,* in which more than one allele affects the same trait (*e.g.,* ABO blood types); *complementary genes,* in which each gene is mutually dependent on the other; and *collaborative genes,* in which two different genes influencing the same trait interact to produce a phenotype neither gene alone could produce.

Genetic Imprinting

Besides autosomal and sex-linked genes and mitochondrial inheritance, it was found that certain genes exhibit a "parent of origin" type of transmission, in which the parental genomes do not always contribute equally in the development of an individual (Fig. 3-12). Although rare, it is estimated that approximately 100 genes exhibit genomic, or genetic, imprinting. Evidence suggests there is a genetic conflict over the developing embryo: the male genome attempts to establish larger offspring, whereas the female prefers smaller offspring to conserve her energy for the current and subsequent pregnancies.

An example of a disorder involving parenteral imprinting is *uniparental disomy.* This occurs when two chromosomes of the same number are inherited from one parent. Normally, this is not a problem, except in cases where a chromosome has been imprinted by a parent. If an allele is inactivated by imprinting, the offspring will have only one working copy of the chromosome, resulting in possible problems.

Mendel's Laws

The main feature of inheritance is predictability: given certain conditions, the likelihood of the occurrence or recurrence of a specific trait is remarkably predictable. The units of inheritance are the genes, and the pattern of single-gene expression often

■ **FIGURE 3-12** ■ Pedigree of genetic imprinting. In generation I, male **A** has inherited a mutant allele from his affected mother (not shown); the gene is "turned off" during spermatogenesis, and therefore none of his offspring (generation II) will express the mutant allele, regardless if they are carriers. However, the gene will be "turned on" again during oogenesis in any of his daughters (**B**) who inherit the allele. All offspring (generation III) who inherit the mutant allele will be affected. All offspring of normal children (**C**) will produce normal offspring. Children of female **D** will all express the mutation if they inherit the allele.
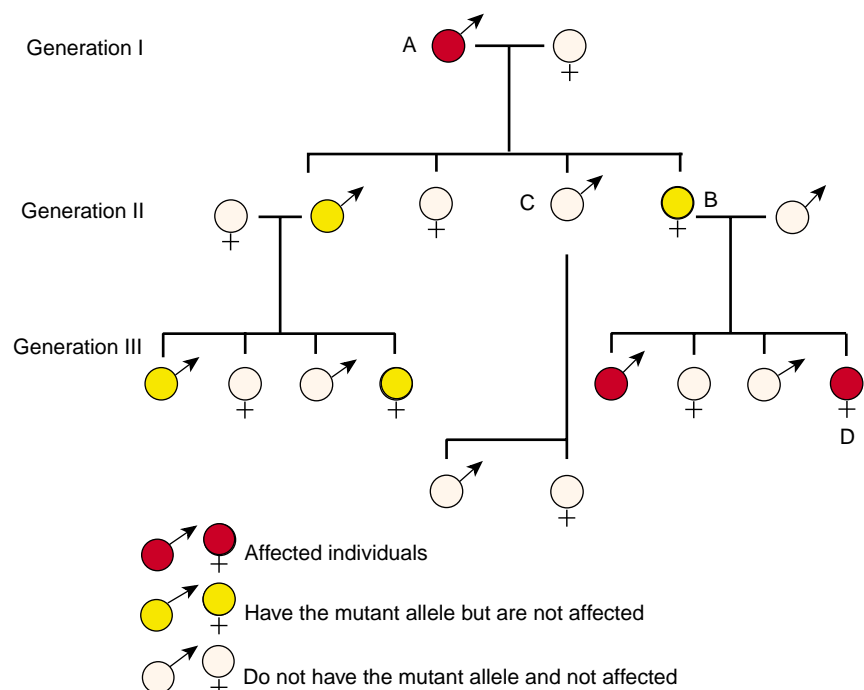
can be predicted using Mendel's laws of genetic transmission. Techniques and discoveries since Gregor Mendel's original work was published in 1865 have led to some modification of his original laws.

Mendel discovered the basic pattern of inheritance by conducting carefully planned experiments with simple garden peas. Experimenting with several phenotypic traits in peas, Mendel proposed that inherited traits are transmitted from parents to offspring by means of independently inherited factors—now known as genes—and that these factors are transmitted as recessive and dominant traits. Mendel labeled dominant factors (his round peas) "A" and recessive factors (his wrinkled peas) "a." Geneticists continue to use capital letters to designate dominant traits and lowercase letters to identify recessive traits. The possible combinations that can occur with transmission of single-gene dominant and recessive traits can be described by constructing a figure called a *Punnett square* using capital and lowercase letters (Fig. 3-13).
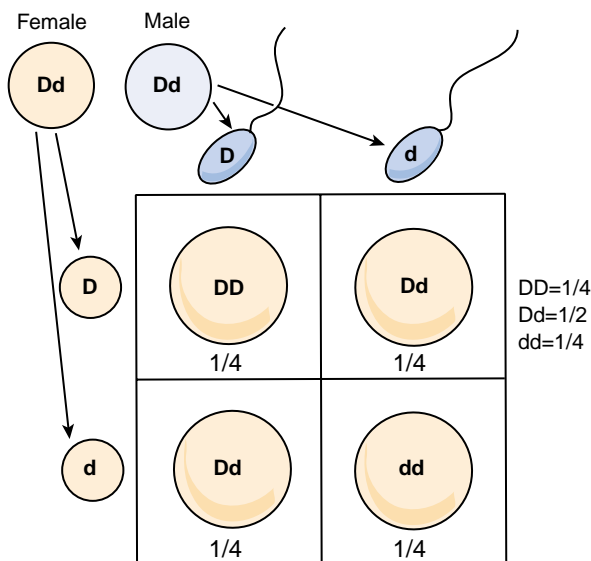
The observable traits of single-gene inheritance are inherited by the offspring from the parents. During maturation, the primordial germ cells (*i.e.,* sperm and ovum) of both parents undergo meiosis, or reduction division, in which the number of chromosomes is divided in half (from 46 to 23). At this time, the two alleles from a gene locus separate so that each germ cell receives only one allele from each pair (*i.e.,* Mendel's first law). According to Mendel's second law, the alleles from the different gene loci segregate independently and recombine randomly in the zygote. Persons in whom the two alleles of a given pair are the same (AA or aa) are called *homozygotes. Heterozygotes* have different alleles (Aa) at a gene locus. A *recessive trait* is one expressed only in a homozygous pairing; a *dominant trait* is one expressed in either a homozygous or a heterozygous pairing. All persons with a dominant allele (depending on the penetrance of the genes) manifest that trait. A *carrier* is a person who is heterozygous for a recessive trait and does not manifest the trait. For example, the genes for blond hair are recessive and those for brown hair are dominant. Therefore, only persons with a genotype having two alleles for blond hair would be blond; persons with either one or two brown alleles would have dark hair.

Pedigree

A pedigree is a graphic method for portraying a family history of an inherited trait. It is constructed from a carefully obtained family history and is useful for tracing the pattern of inheritance for a particular trait.

In summary, inheritance represents the likelihood of the occurrence or recurrence of a specific genetic trait. The genotype refers to information stored in the genetic code of a person, whereas the phenotype represents the recognizable traits, physical and biochemical, associated with the genotype. Expressivity refers to the expression of a gene in the phenotype, and penetrance is the ability of a gene to express its function. The point on the DNA molecule that controls the inheritance of a particular trait is called a *gene locus.* Alternate forms of a gene at one gene locus are called alleles. The alleles at a gene locus may carry recessive or dominant traits. A recessive trait is one expressed only when there are two copies (homozygous) of the recessive alleles. Dominant traits are expressed with either homozygous or heterozygous pairing of the alleles. A pedigree is a graphic method for portraying a family history of an inherited trait.

GENE MAPPING AND TECHNOLOGY

The genome is the gene complement of an organism. Genomic mapping is the assignment of genes to specific chromosomes or parts of the chromosome. The Human Genome Project, which started in 1990, is an international project to identify and localize the estimated 35,000 to 140,000 genes in the human genome. This is a phenomenal undertaking because there are approximately 3.2 billion base pairs in the human genome. The national effort toward genomic mapping is jointly coordinated by the National Institutes of Health and the Department of Energy. Organizers of the U.S. Human Genome Project hope to have the completed map by the year 2005. It is also anticipated that the project will reveal the chemical basis for as many as 4000 genetic diseases. It is expected to provide tests for screening and diagnosing genetic disorders and the basis for new treatments. To date, chromosomes 14, 20, 21 and 22 have been completely mapped.

Although the human genome project has received most of the publicity, genome sequencing also is underway in a variety of animals, such as the mouse and fruit fly, and other organisms, from bacteria to plants. These and other endeavors will be of benefit in animal husbandry and agriculture and the study of genetic similarities among various species. The human genome project has created an anticipation of future benefits from its undertakings as well as ethical concerns that must be dealt with by both scientists and the public.

■ **FIGURE 3-13** ■ The Punnett square showing all possible combinations for transmission of a single gene trait (dimpled cheeks). The example shown is when both parents are heterozygous (dD) for the trait. The alleles carried by the mother are on the left and those carried by the father are on the top. The D allele is dominant and the d allele is recessive. The DD and Dd offspring have dimples and the dd offspring does not.

Genomic Mapping

Two types of genomic maps exist: genetic maps and physical maps. Genetic maps are like highway maps. They use linkage studies (*e.g.,* dosage, hybridization) to estimate the distances between chromosomal landmarks (*i.e.,* gene markers). Physical maps are similar to a surveyor's map. They measure the actual physical distance between chromosomal elements in biochemical units, the smallest being the nucleotide base.

Genetic maps and physical maps have been refined over the decades. The earliest mapping efforts localized genes on the X chromosome. The initial assignment of a gene to a particular chromosome was made in 1911 for the color blindness gene inherited from the mother (*i.e.,* following the X-linked pattern of inheritance). In 1968, the specific location of the Duffy blood group on the long arm of chromosome 1 was determined. The locations of more than 2300 expressed human genes have been mapped to a specific chromosome and most of them to a specific region on the chromosome. However, genetic mapping is continuing so rapidly that these numbers are constantly being updated. Documentation of gene assignments to specific human chromosomes is updated almost daily in the *Online Mendelian Inheritance in Man* (http://www.ncbi.nlm.nih.gov/Omim), an encyclopedia of expressed gene loci. Another source is the Human Genome Project, which is the central database for mapped genes and an international repository for most mapping information. Many methods have been used for developing genetic maps. The most important ones are family linkage studies, gene dosage methods, and hybridization studies. Often, the specific assignment of a gene is made using information from several mapping techniques.

Most of the genome mapping has been accomplished by a method called *transcript mapping.* The two-part process begins with isolating mRNAs immediately after they are transcribed. Next, the complementary DNA molecule is prepared. Although transcript mapping may provide knowledge of the gene sequence, it does not automatically mean that the function of the genetic material has been determined.

Linkage Studies

Mendel's laws were often insufficient to explain the transmission of several well-known traits, such as color blindness and hemophilia A. In some families it was noted that these two conditions were transmitted together.

Linkage studies assume that genes occur in a linear array along the chromosomes. During meiosis, the paired chromosomes of the diploid germ cell exchange genetic material because of the crossing-over phenomenon (see Fig. 3-7). This exchange usually involves more than one gene; large blocks of genes (representing large portions of the chromosome) usually are exchanged. Although the point at which the block separates from another occurs randomly, the closer together two genes are on the same chromosome, the greater the chance is that they will be passed on together to the offspring. When two inherited traits occur together at a rate significantly greater than would occur by chance, they are said to be *linked.*

Several methods take advantage of the crossing over and recombination of genes to map a particular gene. In one method, any gene that is already assigned to a chromosome can be used as a marker to assign other linked genes. For example, it

was found that an extra long chromosome 1 and the Duffy blood group were inherited as a dominant trait, placing the position of the blood group gene close to the extra material on chromosome 1. Color blindness has been linked to classic hemophilia A (*i.e.,* lack of factor VIII) in some pedigrees; hemophilia A has been linked to glucose-6-phosphate dehydrogenase deficiency in others; and color blindness has been linked to glucose-6-phosphate dehydrogenase deficiency in still others. Because the gene for color blindness is found on the X chromosome, all three genes must be located in a small section of the X chromosome. Linkage analysis can be used clinically to identify affected persons in a family with a known genetic defect. Males, because they have one X and one Y chromosome, are said to be *hemizygous* for sex-linked traits. Females can be homozygous (normal or mutant) or heterozygous for sex-linked traits. Heterozygous females are known as *carriers* for X-linked defects.

One autosomal recessive disorder that has been successfully diagnosed prenatally by linkage studies using amniocentesis is congenital adrenal hyperplasia (due to 21-hydroxylase deficiency), which is linked to an immune response gene (human leukocyte antigen [HLA] type). Postnatally, linkage studies have been used in diagnosing hemochromatosis, which is closely linked to another HLA type. Persons with this disorder are unable to metabolize iron, and it accumulates in the liver and other organs. It cannot be diagnosed by conventional means until irreversible damage has been done. Given a family history of the disorder, HLA typing can determine if the gene is present, and if it is present, dietary restriction of iron intake may be used to prevent organ damage.

Dosage Studies

Dosage studies involve measuring enzyme activity. Autosomal genes are normally arranged in pairs, and normally both are expressed. If both alleles are present and both are expressed, the activity of the enzyme should be 100%. If one member of the gene pair is missing, only 50% of the enzyme activity is present, reflecting the activity of the remaining normal allele.

Hybridization Studies

A recent biologic discovery revealed that two somatic cells from different species, when grown together in the same culture, occasionally fuse to form a new hybrid cell. Two types of hybridization methods are used in genomic studies: somatic cell hybridization and in situ hybridization.

Somatic cell hybridization involves the fusion of human somatic cells with those of a different species (typically, the mouse) to yield a cell containing the chromosomes of both species. Because these hybrid cells are unstable, they begin to lose chromosomes of both species during subsequent cell divisions. This makes it possible to obtain cells with different partial combinations of human chromosomes. By studying the enzymes that these cells produce, it is possible to determine that an enzyme is produced only when a certain chromosome is present; the coding for that enzyme must be located on that chromosome.

In situ hybridization involves the use of a specific sequence of DNA or RNA to locate genes that do not express themselves in cell culture. DNA and RNA can be chemically tagged with radioactive or fluorescent markers. These chemically tagged

DNA or RNA sequences are used as probes to determine gene location. The probe is added to a chromosome spread after the DNA strands have been separated. If the probe matches the complementary DNA of a chromosome segment, it hybridizes and remains at the precise location (thus the term *in situ*) on a chromosome. Radioactive or fluorescent markers are used to determine the location of the probe.

Recombinant DNA Technology

During the past several decades, genetic engineering has provided the methods for manipulating nucleic acids and recombining genes (recombinant DNA) into hybrid molecules that can be inserted into unicellular organisms and reproduced many times over. Each hybrid molecule produces a genetically identical population, called a *clone*, that reflects its common ancestor.

The techniques of gene isolation and cloning rely on the fact that the genes of all organisms, from bacteria through mammals, are based on similar molecular organization. Gene cloning requires cutting a DNA molecule apart, modifying and reassembling its fragments, and producing copies of the modified DNA, its mRNA, and its gene product. The DNA molecule is cut apart by using a bacterial enzyme, called a *restriction enzyme*, that binds to DNA wherever a particular short sequence of base pairs is found and cleaves the molecule at a specific nucleotide site. In this way, a long DNA molecule can be broken into smaller, discrete fragments with the intent that one fragment contains the gene of interest. More than 100 restriction enzymes are commercially available that cut DNA at different recognition sites.

The selected gene fragment is replicated through insertion into a unicellular organism, such as a bacterium. To do this, a cloning vector such as a bacterial virus or a small DNA circle that is found in most bacteria, called a *plasmid,* is used. Viral and plasmid vectors replicate autonomously in the host bacterial cell. During gene cloning, a bacterial vector and the DNA fragment are mixed and joined by a special enzyme called a *DNA ligase.* The recombinant vectors formed are then introduced into a suitable culture of bacteria, and the bacteria are allowed to replicate and express the recombinant vector gene. Sometimes, mRNA taken from a tissue that expresses a high level of the gene is used to produce a complementary DNA molecule that can be used in the cloning process. Because the fragments of the entire DNA molecule are used in the cloning process, additional steps are taken to identify and separate the clone that contains the gene of interest.

In terms of biologic research and technology, cloning makes it possible to identify the DNA sequence in a gene and produce the protein product encoded by a gene. The specific nucleotide sequence of a cloned DNA fragment can often be identified by analyzing the amino acid sequence and mRNA codons of its protein product. Short sequences of base pairs can be synthesized, radioactively labeled, and subsequently used to identify their complementary sequence. In this way, identifying normal and abnormal gene structures is possible. Proteins that formerly were available only in small amounts can now be made in large quantities once their respective genes have been isolated. For example, genes encoding for insulin and growth hormone have been cloned to produce these hormones for pharmacologic use.

Gene Therapy

Although quite different from inserting genetic material into a unicellular organism such as bacteria, techniques are available for inserting genes into the genome of intact multicellular plants and animals. Promising delivery vehicles for these genes are the adenoviruses. These viruses are ideal vehicles because their DNA does not become integrated into the host genome; however, repeated inoculations are often needed because the body's immune system usually targets cells expressing adenovirus proteins. Sterically stable liposomes also show promise as DNA delivery mechanisms. This type of therapy is one of the more promising methods for the treatment of genetic disorders, certain cancers, cystic fibrosis, and many infectious diseases.

Two main approaches are used in gene therapy: transferred genes either replace defective genes or selectively inhibit deleterious genes. Cloned DNA sequences or ribosomes usually are the compounds used in gene therapy. However, the introduction of the cloned gene into the multicellular organism can influence only the few cells that get the gene. An answer to this problem would be the insertion of the gene into a sperm or ovum; after fertilization, the gene would be replicated in all of the differentiating cell types. Even so, techniques for cell insertion are limited. Not only are moral and ethical issues involved, but these techniques cannot direct the inserted DNA to attach to a particular chromosome or supplant an existing gene by knocking it out of its place.

DNA Fingerprinting

The technique of DNA fingerprinting is based in part on those techniques used in recombinant DNA technology and those originally used in medical genetics to detect slight variations in the genomes of different individuals. Using restrictive endonucleases, DNA is cleaved at specific regions. The DNA fragments are separated according to size by electrophoresis (*i.e.,* Southern blot) and transferred to a nylon membrane. The fragments are then broken apart and subsequently annealed with a series of radioactive probes specific for regions in each fragment. An autoradiograph reveals the DNA fragments on the membrane. When used in forensic pathology, this procedure is undertaken on specimens from the suspect and the forensic specimen. Banding patterns are then analyzed to see if they match. With conventional methods of analysis of blood and serum enzymes, a 1 in 100 to 1000 chance exists that the two specimens match because of chance. With DNA fingerprinting, these odds are 1 in 100,000 to 1 million.

In summary, the genome is the gene complement of an organism. Genomic mapping is a method used to assign genes to particular chromosomes or parts of a chromosome. The most important ones used are family linkage studies, gene dosage methods, and hybridization studies. Often the specific assignment of a gene is determined by using information from several mapping techniques. Linkage studies assign a chromosome location to genes based on their close association with other genes of known location. Recombinant DNA studies involve the extraction of specific types of messenger RNA used in synthesis of complementary DNA strands. The

complementary DNA strands, labeled with a radioisotope, bind with the genes for which they are complementary and are used as gene probes. Now underway is an international project to identify and localize all 35,000 to 140,000 genes in the human genome. Genetic engineering has provided the methods for manipulating nucleic acids and recombining genes (recombinant DNA) into hybrid molecules that can be inserted into unicellular organisms and reproduced many times over. As a result, proteins that formerly were available only in small amounts can now be made in large quantities once their respective genes have been isolated. DNA fingerprinting, which relies on recombinant DNA technologies and those of genetic mapping, is often used in forensic investigations.

REVIEW QUESTIONS

■ Explain the mechanisms by which genes control cell function and the characteristics of the next generation.

■ Define *gene locus* and *allele.*

■ Explain the difference between a person's genotype and phenotype.

■ Describe the concepts of *induction* and *repression* as they apply to gene function.

■ Describe the pathogenesis of gene mutation.

■ Explain how gene expressivity and penetrance determine the effects of a mutant gene that codes for the production of an essential enzyme.

■ Differentiate between autosomes and sex chromosomes and between meiosis and mitosis.

■ Explain the significance of the Barr body.

■ Construct a hypothetical pedigree for a recessive and dominant trait according to Mendel's law.

connection—○

Visit the Connection site at connection.lww.com/go/porth for links to chapter-related resources on the Internet.

BIBLIOGRAPHY

Alberts B., Johnson A., Lewis J., et al. (2002). *Molecular biology of the cell* (4th ed.). New York: Garland Publishing.

Carlson B.M. (1999). *Human embryology and developmental biology* (2nd ed., pp. 2–23, 128–145). St. Louis: Mosby.

Guyton A.C., Hall J.E. (2000). *Textbook of medical physiology* (10th ed., pp. 24–37). Philadelphia: W.B. Saunders.

Hawley R.S., Mori C.A. (1999). *The human genome: A user's guide.* San Diego: Harcourt Academic Press.

Human Genome Project. (2000). *Columbia encyclopedia* (6th ed.). [On-line]. Available: http://www.bartleby.com/65/hu/_HumanG.html. Accessed October 22, 2000.

The International RH Mapping Consortium. (2000). *A new gene map of the human genome.* [On-line]. Available: http://www.ornl.gov/TechResources/Human_Genome. Accessed March 2003.

Jain H.K. (1999). *Genetics: Principles, concepts and implications.* Enfield, NH: Science Publishers.

Moore K.L., Persaud T.V.N. (1998). *The developing human: Clinically orientated embryology* (6th ed., pp. 17–46). Philadelphia: W.B. Saunders.

Nussbaum, R.L., McInnes R.R., Willard H.F. (2001). *Thompson & Thompson genetics in medicine.* Philadelphia: W.B. Saunders.

Ostrer H. (1998). *Non-Mendelian genetics in humans.* New York: Oxford University Press.

Sadler T.W. (2000). *Langman's medical embryology* (8th ed., pp. 3–30, 112–135). Philadelphia: Lippincott Williams & Wilkins.

Sapienza C. (1990). Parenteral imprinting of genes. *Scientific American* 263(4), 52–60.

Shapiro L.J. (2000). Molecular basis of genetic disorders. In Behrman R.E., Kliegman R.M., Jenson H. (Eds.), *Nelson's textbook of pediatrics* (16th ed., pp. 313–317). Philadelphia: W.B. Saunders.

Snustad D.P., Simmons J.M. (Eds.). (2000). *Principles of genetics* (2nd ed., pp. 3–21, 52–71, 91–115, 665–669). New York: John Wiley & Sons.