In recent years, nuclei taken from cells of adult animals have been used to produce new animals. In this procedure, the nucleus is removed from a body cell (e.g., skin or blood cell) of a donor animal and introduced into an unfertilized mammalian egg that has been deprived of its own nucleus. In a step that has now been done with mice, cows, sheep, mules, and some other animals, the egg with its donor nucleus is implanted into a foster mother. The ability of such a donor nucleus to direct the development of an entire animal shows that all the information required for life is retained in the nuclei of some adult cells. Since all the cells in an animal produced in this way have the genes of the single original donor cell, the new animal is a genetic clone of the donor (Figure 1-8), though the animals may differ anyway due to their distinct environments and experiences. Repeating the process can give rise to many clones. Nuclei taken from ES cells work especially well, whereas nuclei from other parts of the body at later times in life work far less well. The majority of embryos produced by this technique do not survive due to birth defects, so the donor nuclei may not have all the needed information or the nuclei may be damaged by the cloning process. Even those animals that are born alive have abnormalities, including accelerated aging. The "rooting" of plants, in contrast, is a type of cloning that is readily accomplished by gardeners, farmers, and laboratory technicians.

Scientific interest in the cloning of humans is very limited. Virtually all scientists oppose it because of its high risk to the embryo (also, most people don't believe there is a critical shortage of twins and triplets). Of much greater scientific and medical interest is the ability to generate specific cell types starting from embryonic or adult stem cells. This procedure, **somatic cell nuclear transfer** (SCNT), produces cells that are grown in culture and never turned into an embryo. The scientific interest in such cells comes from learning the signals that can unleash the potential of the genes to form a certain cell type. The medical interest comes from the possibility of treating the numerous diseases in which particular cell types are damaged or missing and of repairing wounds more completely. The cells may also be useful in culture to test the effects of drugs or other treatments. If the cells are produced using a donor nucleus from a patient, the properties of the cells may allow them to escape rejection by the patient's immune system, opening new possibilities for cell-transplant therapies.

FIGURE 1-8 Five genetically identical cloned sheep. An early embryo was divided into five groups of cells and each was immediately implanted into a surrogate mother, much like the natural process of twinning. At an early stage the cells are able to adjust and form an entire animal; later in development the cells become progressively restricted and can no longer do so. An alternative way to make animals is to replace the nuclei of multiple single-celled embryos with donor nuclei from cells of an adult sheep. Each embryo is genetically identical to the adult from which the nucleus was derived. Low percentages of embryos survive these procedures to healthy animals, and the full impact of the techniques on the animals is not yet known. [Geoff Tompkinson/Science Photo Library/Photo Researchers, Inc.]

Reading #1

1.2 The Molecules of a Cell

Molecular cell biologists explore how all the remarkable properties of the cell arise from underlying molecular events: the assembly of large molecules, binding of large molecules to each other, catalytic effects that promote particular chemical reactions, and the deployment of information carried by giant molecules. Here we review the most important kinds of molecules that form the chemical foundations of cell structure and function.

Small Molecules Carry Energy, Transmit Signals, and Are Linked into Macromolecules

Much of the cell's content is a watery soup flavored with small molecules (e.g., simple sugars, amino acids, vitamins) and ions (e.g., sodium, chloride, calcium ions). The locations and concentrations of small molecules and ions within the cell are controlled by numerous proteins inserted in cellular membranes. These pumps, transporters, and ion channels move nearly all small molecules and ions into or out of the cell and its organelles (Chapter 11).

One of the best-known small molecules is **adenosine triphosphate** (ATP), which stores readily available chemical energy in two of its chemical bonds (see Figure 2-31). When cells split apart these energy-rich bonds in ATP, the released energy can be harnessed to power an energy-requiring process such as muscle contraction or protein biosynthesis. To obtain energy for making ATP, cells break down food molecules. For instance, when sugar is degraded to carbon dioxide and water, the energy stored in the original chemical bonds is released and much of it can be "captured" in ATP (Chapter 12). Bacterial, plant, and animal cells can all make ATP by this process. In addition, plants and a few other organisms can harvest energy from sunlight to form ATP in **photosynthesis.**

Other small molecules act as signals both within and between cells; such signals direct numerous cellular activities (Chapters 15 and 16). The powerful effect on our bodies of a frightening event comes from the instantaneous flooding of

the body with epinephrine, a small-molecule **hormone** that mobilizes the "fight-or-flight" response. The movements needed to fight or flee are triggered by nerve impulses that flow from the brain to our muscles with the aid of **neurotransmitters**, another type of small-molecule signal that we discuss in Chapter 23.

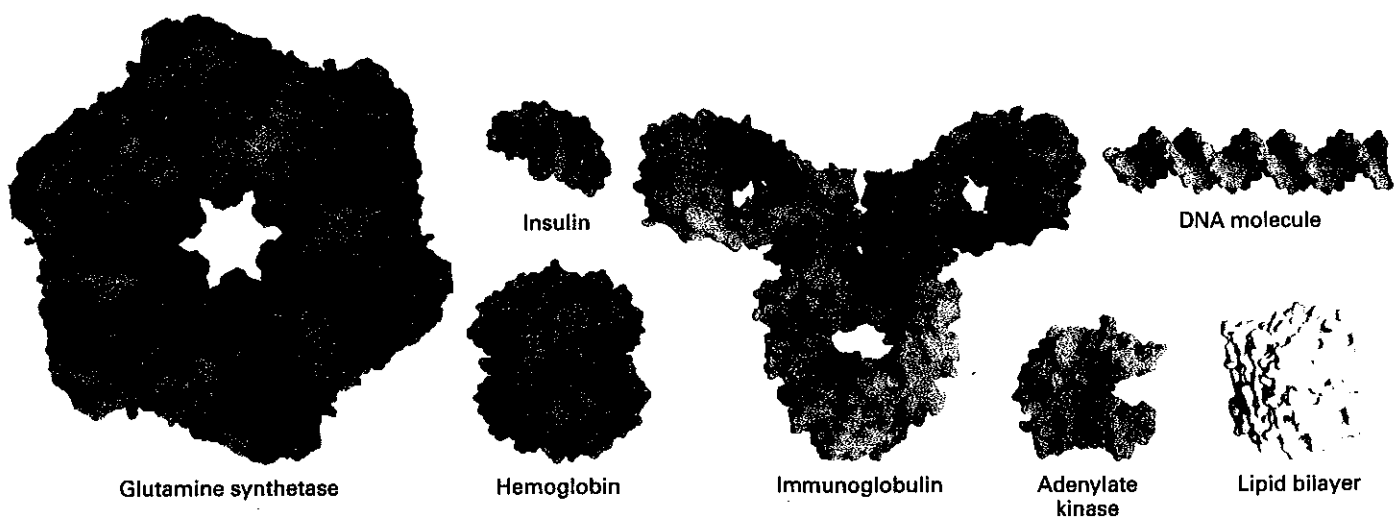
Certain small molecules (**monomers**) in the cellular soup can be joined to form **polymers** through repetition of a single type of chemical-linkage reaction (see Figure 2-1). Cells produce three types of large polymers, commonly called **macromolecules**: polysaccharides, proteins, and nucleic acids. Sugars, for example, are the monomers used to form **polysaccharides**. These macromolecules are critical structural components of plant cell walls and insect skeletons. A typical polysaccharide is a linear or branched chain of repeating identical sugar units. Such a chain carries information: the number of units. However, if the units are *not* identical, then the order and type of units carry additional information. As we will see in Chapter 6, some polysaccharides exhibit the greater informational complexity associated with a linear code made up of different units assembled in a particular order. But this property is most typical of the two other types of biological macromolecules—**proteins** and **nucleic acids**.

Proteins Give Cells Structure and Perform Most Cellular Tasks

The varied, intricate structures of proteins enable them to carry out numerous functions. Cells string together 20 different **amino acids** in a linear chain to form a protein (see Figure 2-14). Proteins commonly range in length from 100 to 1000 amino acids, but some are much shorter and others longer. We obtain amino acids either by synthesizing them from other molecules or by breaking down proteins that we eat. The "essential" amino acids, from a dietary standpoint, are the eight that we cannot synthesize and must obtain from food. Beans and corn together have all eight, making their combination particularly nutritious. Once a chain of amino acids is formed, it folds into a complex shape, conferring a distinctive three-dimensional structure and function on each protein (Figure 1-9).

Some proteins are similar to one another and therefore can be considered members of a **protein family**. A few hundred such families have been identified. Most proteins are designed to work in particular places within a cell or to be released into the extracellular (*extra*, "outside") space. Elaborate cellular pathways ensure that proteins are transported to their proper intracellular (*intra*, "within") locations or secreted (Chapters 13 and 14).

Proteins can serve as structural components of a cell, for example, by forming an internal skeleton (Chapters 10, 17, and 18). They can be sensors that change shape as temperature, ion concentrations, or other properties of the cell change. They can import and export substances across the plasma membrane (Chapter 11). They can be **enzymes**, causing chemical reactions to occur much more rapidly than they would without the aid of these protein **catalysts** (Chapter 3). They can bind to a specific gene, turning it on or off (Chapter 7). They can be extracellular signals, released from one cell to communicate with other cells, or intracellular signals, carrying information within the cell (Chapters 15 and 16). They can be motors that move other molecules around, burning chemical energy (ATP) to do so (Chapters 17 and 18).



▲ FIGURE 1-9 Proteins vary greatly in size, shape, and function. These models of the water-accessible surface of some representative proteins are drawn to a common scale and reveal the numerous projections and crevices on the surface. Each protein has a defined three-dimensional shape (conformation) that is stabilized by numerous chemical interactions discussed in Chapters 2 and 3. The illustrated proteins include enzymes (glutamine synthetase and adenylate kinase), an antibody (immunoglobulin), a hormone (insulin), and the blood's oxygen carrier (hemoglobin). Models of a segment of the nucleic acid DNA and a small region of the lipid bilayer that forms cellular membranes (see Section 1.3) demonstrate the relative width of these structures compared with that of typical proteins. [Courtesy of Gareth White.]

How can 20 amino acids form all the different proteins needed to perform these varied tasks? It seems impossible at first glance. But if a "typical" protein is about 400 amino acids long, there are 20^{400} possible different protein sequences. Even assuming that many of these would be functionally equivalent, unstable, or otherwise discountable, the number of possible proteins is well along toward infinity.

Next we might ask how many protein molecules a cell needs to operate and maintain itself. To estimate this number, let's take a typical eukaryotic cell, such as a hepatocyte (liver cell). This cell, roughly a cube 15 μm (0.0015 cm) on a side, has a volume of 3.4×10^{-9} cm^3 (or milliliters). Assuming a cell density of 1.03 g/ml, the cell would weigh 3.5×10^{-9} g. Since protein accounts for approximately 20 percent of a cell's weight, the total weight of cellular protein is 7×10^{-10} g. The average yeast protein has a molecular weight of 52,700 (g/mol). Assuming this value is typical of eukaryotic proteins, we can calculate the total number of protein molecules per liver cell as about 7.9×10^9 from the total protein weight and Avogadro's number, the number of molecules per mole of any chemical compound (6.02×10^{23}). To carry this calculation one step further, consider that a liver cell contains about 10,000 different proteins; thus a cell contains close to a million molecules of each type of protein on average. In fact, the abundance of different proteins varies widely, from the quite rare insulin-binding receptor protein (20,000 molecules) to the abundant structural protein actin (5×10^8 molecules).

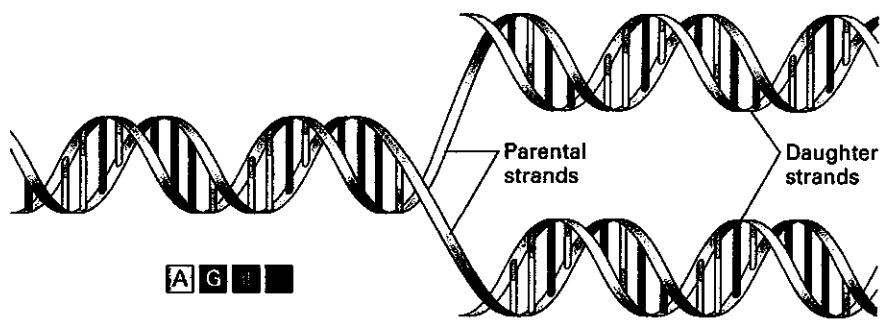
Nucleic Acids Carry Coded Information for Making Proteins at the Right Time and Place

The information about how, when, and where to produce each kind of protein is carried in the genetic material, a polymer called **deoxyribonucleic acid (DNA).** The three-dimensional structure of DNA consists of two long helical strands that are coiled around a common axis, forming a **double helix.** DNA strands are composed of monomers called **nucleotides;** these often are referred to as *bases* because their structures contain cyclic organic bases (Chapter 4).

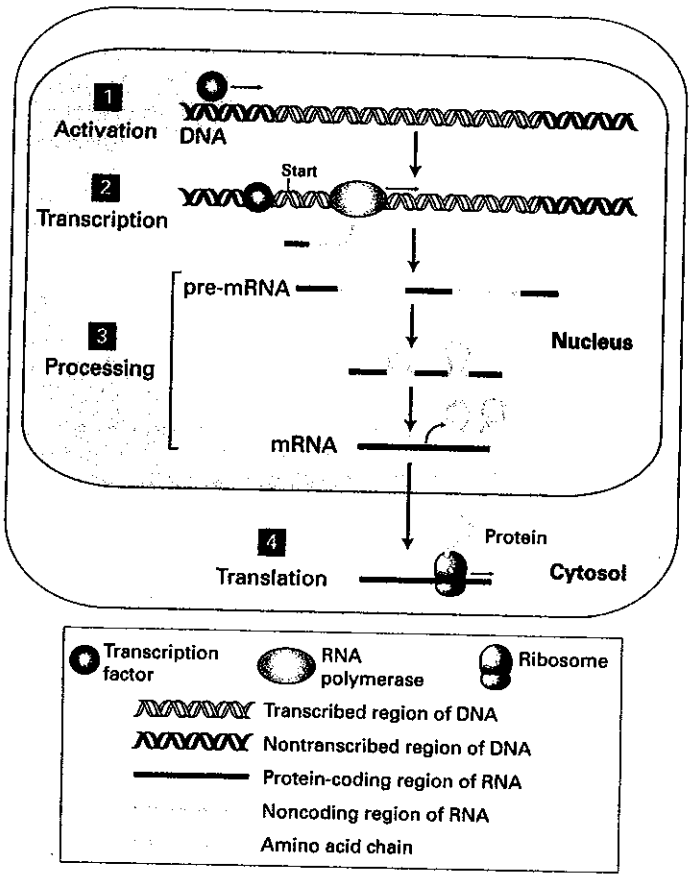
Four different nucleotides, abbreviated A, T, C, and G, are joined end to end in a DNA strand, with the base parts projecting out from the helical backbone of the strand. Each DNA double helix has a simple construction: wherever one strand has an A, the other strand has a T, and each C is matched with a G (Figure 1-10). This **complementary** matching of the two strands is so strong that if complementary strands are separated, they will spontaneously zip back together in the right salt and temperature conditions. Such **nucleic acid hybridization** is extremely useful for detecting one strand using the other. For example, if one strand is purified and attached to a piece of paper, soaking the paper in a solution containing the other complementary strand will lead to zippering, even if the solution also contains many other DNA strands that do not match.

The genetic information carried by DNA resides in its sequence, the linear order of nucleotides along a strand. The information-bearing portion of DNA is divided into discrete functional units, the **genes,** which typically are 5000 to 100,000 nucleotides long. Most bacteria have a few thousand genes; humans, about 20,000–25,000. The genes that carry instructions for making proteins commonly contain two parts: a *coding region* that specifies the amino acid sequence of a protein and a *regulatory region* that controls when and in which cells the protein is made.

Cells use two processes in series to convert the coded information in DNA into proteins (Figure 1-11). In the first, called **transcription,** the coding region of a gene is copied into a single-stranded **ribonucleic acid (RNA)** version of the double-stranded DNA. A large enzyme, **RNA polymerase,** catalyzes the linkage of nucleotides into a RNA chain using DNA as a template. In eukaryotic cells, the initial RNA product is processed into a smaller **messenger RNA (mRNA)** molecule, which moves to the cytoplasm. Here the **ribosome,** an enormously complex molecular machine composed of both RNA and protein, carries out the second process, called **translation.** During translation, the ribosome assembles and links together amino acids in the precise order dictated by the mRNA sequence according to the nearly universal **genetic code.** We examine the cell components that carry out transcription and translation in detail in Chapter 4.



▲ **FIGURE 1-10 DNA consists of two complementary strands wound around each other to form a double helix.** *(Left)* The double helix is stabilized by weak hydrogen bonds between the A and T bases and between the C and G bases. *(Right)* During replication, the two strands are unwound and used as templates to produce complementary strands. The outcome is two copies of the original double helix, each containing one of the original strands and one new daughter (complementary) strand.



▲ **FIGURE 1-11 The coded information in DNA is converted into the amino acid sequences of proteins by a multistep process.** Step 1: Transcription factors bind to the regulatory regions of the specific genes they control and activate them. Step 2: Following assembly of a multiprotein initiation complex bound to the DNA, RNA polymerase begins transcription of an activated gene at a specific location, the start site. The polymerase moves along the DNA linking nucleotides into a single-stranded pre-mRNA transcript using one of the DNA strands as a template. Step 3: The transcript is processed to remove noncoding sequences. Step 4: In a eukaryotic cell, the mature messenger RNA (mRNA) moves to the cytoplasm, where it is bound by ribosomes that read its sequence and assemble a protein by chemically linking amino acids into a linear chain.

In addition to its role in transferring information from nucleus to cytoplasm, RNA can serve as a framework for building a molecular machine. For example, the ribosome has four RNA chains that team up with more than 50 proteins to make a remarkably precise and efficient mRNA reader and protein synthesizer. Recently, RNA has also been found to play a remarkably important role in regulating many aspects of gene activity, including chromosome structure and RNA processing and stability. In many cases small RNAs, 20–200 nucleotides long, specifically regulate the structure and function of chromosomes, the stability of larger RNA molecules, and the translation of mRNA molecules into protein.

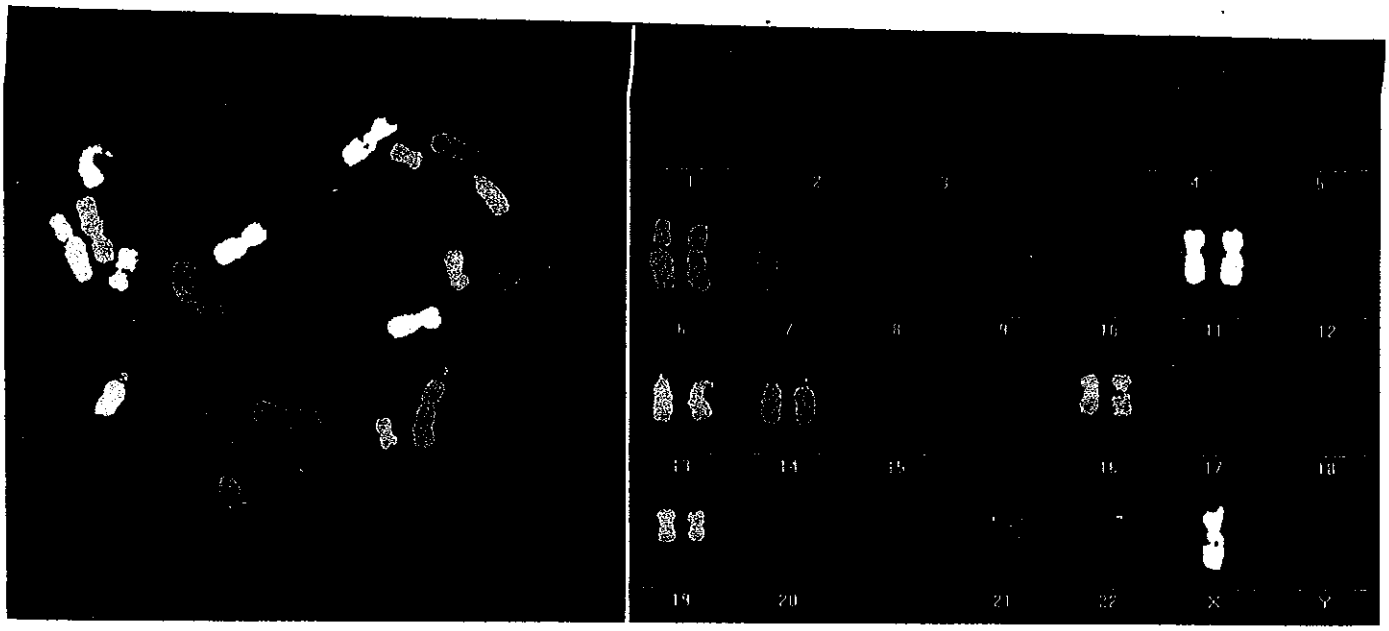
All organisms have ways to control when and where their genes can be transcribed. For instance, nearly all the cells in our bodies contain the full set of human genes, but in each cell type only some of these genes are active, or turne on, and used to make proteins. That's why liver cells pr duce some proteins that are not produced by kidney cel and vice versa. Moreover, many cells can respond to extern signals or changes in external conditions by turning specif genes on or off, thereby adapting their repertoire of protei to meet current needs. Such control of gene activity depen on DNA-binding proteins called **transcription factors**, whic bind to DNA and act as switches, either activating or re pressing transcription of particular genes (Chapter 7).

Transcription factors are shaped so precisely that the are able to bind preferentially to the regulatory regions o just a few genes out of the thousands present in a cell's DNA Typically a DNA-binding protein will recognize shor DNA sequences about 6–12 base pairs long. A segment o DNA containing 10 base pairs can have 4^{10} possible se quences (1,048,576) since each position can be any of fou nucleotides. Only a few copies of each such sequence will oc cur in the DNA of a cell, ensuring the specificity of gene ac tivation and repression. Multiple copies of one type of tran scription factor can coordinately regulate a set of genes i binding sites for that factor exist near each gene in the set Transcription factors often work as multiprotein complexes with more than one protein contributing its own DNA-binding specificity to selecting the regulated genes. In complex organisms, hundreds of different transcription factors are employed to form an exquisite control system that activates the right genes in the right cells at the right times. Small RNA molecules can have a dramatic effect on gene expression, regulating production and stability of gene transcripts. By some estimates small RNAs may regulate most or all genes, though the mechanisms and ubiquity of this type of regulation are still being explored.

The Genome Is Packaged into Chromosomes and Replicated During Cell Division

Most of the DNA in eukaryotic cells is located in the nucleus, extensively folded into the familiar structures we know as **chromosomes** (Chapter 6). Each chromosome contains a single linear DNA molecule associated with certain proteins. In prokaryotic cells, most or all of the genetic information resides in a single circular DNA molecule about a millimeter in length; this molecule lies, folded back on itself many times, in the central region of the cell (see Figure 1-2a). The **genome** of an organism comprises its entire complement of DNA. With the exception of eggs and sperm, every normal human cell has 46 chromosomes (Figure 1-12). Half of these, and thus half of the genes, can be traced back to Mom; the other half, to Dad.

Every time a cell divides, a large multiprotein replication machine, the replisome, separates the two strands of double-helical DNA in the chromosomes and uses each strand as a template to assemble nucleotides into a new complementary strand (see Figure 1-10). The outcome is a pair of double helices, each identical to the original. **DNA polymerase**, which is responsible for linking nucleotides into a DNA strand; the many other components of the replisome are described in



▲ **FIGURE 1-12 Chromosomes can be "painted" for easy identification.** A normal human has 23 pairs of morphologically distinct chromosomes; one member of each pair is inherited from the mother and the other member from the father. *(Left)* A chromosome spread from a human body cell midway through mitosis, when the chromosomes are fully condensed. This preparation was treated with fluorescent-labeled staining reagents that allow each of the 22 pairs and the X and Y chromosomes to appear in a different color when viewed in a fluorescence microscope. This technique of multiplex fluorescence in situ hybridization (M-FISH) sometimes is called chromosome painting (Chapter 6). *(Right)* Chromosomes from the preparation on the left arranged in pairs in descending order of size, an array called a karyotype. The presence of X and Y chromosomes identifies the sex of the individual as male. [Courtesy of M. R. Speicher.]

Chapter 4. The molecular design of DNA and the remarkable properties of the replisome ensure rapid, highly accurate copying. Many DNA polymerase molecules work in concert, each one copying part of a chromosome. The entire genome of fruit flies, about 1.2×10^8 nucleotides long, can be copied in three minutes! Because of the accuracy of DNA replication, nearly all the cells in our bodies carry the same genetic instructions, and we can inherit Mom's brown hair and Dad's blue eyes.

A rather dramatic example of gene control involves inactivation of an entire chromosome in human females. Women have two X chromosomes, whereas men have one X chromosome and one Y chromosome, which has different genes than the X chromosome. Yet the genes on the X chromosome must, for the most part, be equally active in female cells (XX) and male cells (XY). To achieve this balance, one of the X chromosomes in female cells is chemically modified and condensed into a very small mass called a Barr body, which is inactive and never transcribed.

Surprisingly, we inherit a small amount of genetic material entirely and uniquely from our mothers. This is the circular DNA present in mitochondria, the organelles in eukaryotic cells that synthesize ATP using the energy released by the breakdown of nutrients. Mitochondria contain multiple copies of their own DNA genomes, which code for some of the mitochondrial proteins (Chapter 6). Because each human inherits mitochondrial DNA only from his or her mother (it comes with the egg but not the sperm), the distinctive features of a particular mitochondrial DNA can be used to trace maternal history. Chloroplasts, the organelles that carry out photosynthesis in plants, also have their own circular genomes. Both mitochondria and chloroplasts are believed to be derived from **endosymbionts,** bacteria that took up residence inside eukaryotic cells in a mutually beneficial partnership. The mitochondrial and chloroplast circular DNAs appear to have originated as bacterial genomes, which also are usually circular, though the organelle genomes have lost most of the bacterial genes.

Mutations May Be Good, Bad, or Indifferent

Mistakes occasionally do occur spontaneously during DNA replication, causing changes in the sequence of nucleotides. Such changes, or **mutations,** also can arise from radiation that causes damage to the nucleotide chain or from chemical poisons, such as those in cigarette smoke, that lead to errors during the DNA-copying process (Chapter 25). Mutations come in various forms: a simple swap of one nucleotide for another; the deletion, insertion, or inversion of one to millions of nucleotides in the DNA of one chromosome; and translocation of a stretch of DNA from one chromosome to another.

In sexually reproducing animals such as ourselves, mutations can be inherited only if they are present in cells that potentially contribute to the formation of offspring. Such **germ-line cells** include eggs, sperm, and their precursor cells.