Key Terms

alternation of generations a life-cycle type in which the diploid and haploid stages alternate
aneuploid an individual with an error in chromosome number; includes deletions and duplications of chromosome segments
autosome any of the non-sex chromosomes
chiasmata (singular = *chiasma*) the structure that forms at the crossover points after genetic material is exchanged
chromosome inversion the detachment, 180° rotation, and reinsertion of a chromosome arm
crossing over (also, recombination) the exchange of genetic material between homologous chromosomes resulting in chromosomes that incorporate genes from both parents of the organism forming reproductive cells
diploid-dominant a life-cycle type in which the multicellular diploid stage is prevalent
euploid an individual with the appropriate number of chromosomes for their species
fertilization the union of two haploid cells typically from two individual organisms
gametophyte a multicellular haploid life-cycle stage that produces gametes
germ cell a specialized cell that produces gametes, such as eggs or sperm
haploid-dominant a life-cycle type in which the multicellular haploid stage is prevalent
interkinesis a period of rest that may occur between meiosis I and meiosis II; there is no replication of DNA during interkinesis
karyogram the photographic image of a karyotype
karyotype the number and appearance of an individuals chromosomes, including the size, banding patterns, and centromere position
life cycle the sequence of events in the development of an organism and the production of cells that produce offspring
meiosis a nuclear division process that results in four haploid cells
meiosis I the first round of meiotic cell division; referred to as reduction division because the resulting cells are haploid
meiosis II the second round of meiotic cell division following meiosis I; sister chromatids are separated from each other, and the result is four unique haploid cells
monosomy an otherwise diploid genotype in which one chromosome is missing
nondisjunction the failure of synapsed homologs to completely separate and migrate to separate poles during the first cell division of meiosis
polyploid an individual with an incorrect number of chromosome sets
recombinant describing something composed of genetic material from two sources, such as a chromosome with both maternal and paternal segments of DNA
reduction division a nuclear division that produces daughter nuclei each having one-half as many chromosome sets as the parental nucleus; meiosis I is a reduction division
somatic cell all the cells of a multicellular organism except the gamete-forming cells
sporophyte a multicellular diploid life-cycle stage that produces spores
synapsis the formation of a close association between homologous chromosomes during prophase I
tetrad two duplicated homologous chromosomes (four chromatids) bound together by chiasmata during prophase I
translocation the process by which one segment of a chromosome dissociates and reattaches to a different, nonhomologous chromosome
trisomy an otherwise diploid genotype in which one entire chromosome is duplicated
X inactivation the condensation of X chromosomes into Barr bodies during embryonic development in females to compensate for the double genetic dose

Chapter Summary

7.1 Sexual Reproduction

Nearly all eukaryotes undergo sexual reproduction. The variation introduced into the reproductive cells by meiosis appears to be one of the advantages of sexual reproduction that has made it so successful. Meiosis and fertilization alternate in sexual life cycles. The process of meiosis produces genetically unique reproductive cells called gametes, which have half the number of chromosomes as the parent cell.

Fertilization, the fusion of haploid gametes from two individuals, restores the diploid condition. Thus, sexually reproducing organisms alternate between haploid and diploid stages. However, the ways in which reproductive cells are produced and the timing between meiosis and fertilization vary greatly. There are three main categories of life cycles: diploid-dominant, demonstrated by most animals; haploid-dominant, demonstrated by all fungi and some algae; and alternation of generations, demonstrated by plants

and some algae.

7.2 Meiosis

Sexual reproduction requires that diploid organisms produce haploid cells that can fuse during fertilization to form diploid offspring. The process that results in haploid cells is called meiosis. Meiosis is a series of events that arrange and separate chromosomes into daughter cells. During the interphase of meiosis, each chromosome is duplicated. In meiosis, there are two rounds of nuclear division resulting in four nuclei and usually four haploid daughter cells, each with half the number of chromosomes as the parent cell. During meiosis, variation in the daughter nuclei is introduced because of crossover in prophase I and random alignment at metaphase I. The cells that are produced by meiosis are genetically unique.

Meiosis and mitosis share similarities, but have distinct outcomes. Mitotic divisions are single nuclear divisions that produce daughter nuclei that are genetically identical and have the same number of chromosome sets as the original cell. Meiotic divisions are two nuclear divisions that produce four daughter nuclei that are genetically different and have one chromosome set rather than the two sets the parent cell had. The main differences between the processes occur in the first division of meiosis. The homologous chromosomes separate into different nuclei during meiosis I causing a reduction of ploidy level. The second division of meiosis is much more similar to a mitotic division.

7.3 Variations in Meiosis

The number, size, shape, and banding pattern of chromosomes make them easily identifiable in a karyogram and allow for the assessment of many chromosomal abnormalities. Disorders in chromosome number, or aneuploidies, are typically lethal to the embryo, although a few trisomic genotypes are viable. Because of X inactivation, aberrations in sex chromosomes typically have milder effects on an individual. Aneuploidies also include instances in which segments of a chromosome are duplicated or deleted. Chromosome structures also may be rearranged, for example by inversion or translocation. Both of these aberrations can result in negative effects on development, or death. Because they force chromosomes to assume contorted pairings during meiosis I, inversions and translocations are often associated with reduced fertility because of the likelihood of nondisjunction.

Visual Connection Questions

1. Figure 7.2 If a mutation occurs so that a fungus is no longer able to produce a minus mating type, will it still be able to reproduce?

Review Questions

2. What is a likely evolutionary advantage of sexual reproduction over asexual reproduction?
   a. sexual reproduction involves fewer steps
   b. less chance of using up the resources in a given environment
   c. sexual reproduction results in greater variation in the offspring
   d. sexual reproduction is more cost-effective

3. Which type of life cycle has both a haploid and diploid multicellular stage?
   a. an asexual life cycle
   b. diploid-dominant
   c. haploid-dominant
   d. alternation of generations

4. Which event leads to a diploid cell in a life cycle?
   a. meiosis
   b. fertilization
   c. alternation of generations
   d. mutation

5. Meiosis produces _____ daughter cells.
   a. two haploid
   b. two diploid
   c. four haploid
   d. four diploid

6. At which stage of meiosis are sister chromatids separated from each other?
   a. prophase I
   b. prophase II
   c. anaphase I
   d. anaphase II

7. The part of meiosis that is similar to mitosis is _____.
   a. meiosis I
   b. anaphase I
   c. meiosis II
   d. interkinesis

8. If a muscle cell of a typical organism has 32 chromosomes, how many chromosomes will be in a gamete of that same organism?
   a. 8
   b. 16
   c. 32
   d. 64

9. The genotype XXY corresponds to:
   a. Klinefelter syndrome
   b. Turner syndrome
   c. Triplo-X
   d. Jacob syndrome

10. Abnormalities in the number of X chromosomes tend to be milder than the same abnormalities in autosomes because of _____.
   a. deletions
   b. nonhomologous recombination
   c. synapsis
   d. X inactivation

11. Aneuploidies are deleterious for the individual because of what phenomenon?
   a. nondisjunction
   b. gene dosage
   c. meiotic errors
   d. X inactivation

Critical Thinking Questions

12. Explain the advantage that populations of sexually reproducing organisms have over asexually reproducing organisms?

13. Describe the two events that are common to all sexually reproducing organisms and how they fit into the different life cycles of those organisms.

14. Explain how the random alignment of homologous chromosomes during metaphase I contributes to variation in gametes produced by meiosis.

15. In what ways is meiosis II similar to and different from mitosis of a diploid cell?

16. Individuals with trisomy 21 are more likely to survive to adulthood than individuals with trisomy 18. Based on what you know about aneuploidies from this module, what can you hypothesize about chromosomes 21 and 18?

CHAPTER 8
Patterns of Inheritance

FIGURE 8.1 Experimenting with thousands of garden peas, Mendel uncovered the fundamentals of genetics. (credit: modification of work by Jerry Kirkhart)

CHAPTER OUTLINE

INTRODUCTION Genetics is the study of heredity. Johann Gregor Mendel set the framework for genetics long before chromosomes or genes had been identified, at a time when meiosis was not well understood. Mendel selected a simple biological system and conducted methodical, quantitative analyses using large sample sizes. Because of Mendel's work, the fundamental principles of heredity were revealed. We now know that genes, carried on chromosomes, are the basic functional units of heredity with the ability to be replicated, expressed, or mutated. Today, the postulates put forth by Mendel form the basis of classical, or Mendelian, genetics. Not all traits are transmitted from parents to offspring according to Mendelian genetics, but Mendel's experiments serve as an excellent starting point for thinking about inheritance.

8.1 Mendel's Experiments

LEARNING OBJECTIVES

By the end of this section, you will be able to:

- Explain the scientific reasons for the success of Mendel's experimental work
- Describe the expected outcomes of monohybrid crosses involving dominant and recessive alleles

FIGURE 8.2 Johann Gregor Mendel set the framework for the study of genetics.

Johann Gregor Mendel (1822–1884) (Figure 8.2) was a lifelong learner, teacher, scientist, and man of faith. As a young adult, he joined the Augustinian Abbey of St. Thomas in Brno in what is now the Czech Republic. Supported by the monastery, he taught physics, botany, and natural science courses at the secondary and university levels. In 1856, he began a decade-long research pursuit involving inheritance patterns in honeybees and plants, ultimately settling on pea plants as his primary **model system** (a system with convenient characteristics that is used to study a specific biological phenomenon to gain understanding to be applied to other systems). In 1865, Mendel presented the results of his experiments with nearly 30,000 pea plants to the local natural history society. He demonstrated that traits are transmitted faithfully from parents to offspring in specific patterns. In 1866, he published his work, *Experiments in Plant Hybridization,*[1] in the proceedings of the Natural History Society of Brünn. As stated earlier, in genetics, "parent" is often used to describe the individual organism(s) that contribute genetic material to an offspring, usually in the form of gamete cells.

Mendel's work went virtually unnoticed by the scientific community, which incorrectly believed that the process of inheritance involved a blending of parental traits that produced an intermediate physical appearance in offspring. This hypothetical process appeared to be correct because of what we know now as continuous variation. **Continuous variation** is the range of small differences we see among individuals in a characteristic like human height. It does appear that offspring are a "blend" of their parents' traits when we look at characteristics that exhibit continuous variation. Mendel worked instead with traits that show **discontinuous variation**. Discontinuous variation is the variation seen among individuals when each individual shows one of two—or a very few—easily distinguishable traits, such as violet or white flowers. Mendel's choice of these kinds of traits allowed him to see experimentally that the traits were not blended in the offspring as would have been expected at the time, but that they were inherited as distinct traits. In 1868, Mendel became abbot of the monastery and exchanged his scientific pursuits for his

1 Johann Gregor Mendel, "Versuche über Pflanzenhybriden." *Verhandlungen des naturforschenden Vereines in Brünn,* Bd. IV für das Jahr, 1865 Abhandlungen (1866):3–47. [for English translation, see http://www.mendelweb.org/Mendel.plain.html]

pastoral duties. He was not recognized for his extraordinary scientific contributions during his lifetime; in fact, it was not until 1900 that his work was rediscovered, reproduced, and revitalized by scientists on the brink of discovering the chromosomal basis of heredity.

Mendel's Crosses

Mendel's seminal work was accomplished using the garden pea, *Pisum sativum*, to study inheritance. This species naturally self-fertilizes, meaning that pollen encounters ova within the same flower. Because every pea plant has both male reproductive organs and female reproductive organs, each plant produces both types of gametes required for reproduction—both pollen and ova. In plants, just as in animals, reproductive organs are classified by the size of the gametes produced. The organs producing the smaller pollen are called male reproductive organs, while the organs producing the larger ova are called female reproductive organs.

In garden peas, the flower petals remain sealed tightly until pollination is completed to prevent the pollination of other plants. The result is highly inbred, or "true-breeding," pea plants. These are plants that always produce offspring that look like the parent. By experimenting with true-breeding pea plants, Mendel avoided the appearance of unexpected traits in offspring that might occur if the plants were not true-breeding. The garden pea also grows to maturity within one season, meaning that several generations could be evaluated over a relatively short time. Finally, large quantities of garden peas could be cultivated simultaneously, allowing Mendel to conclude that his results did not come about simply by chance.

Mendel performed **hybridizations**, which involve mating two true-breeding individuals that have different traits. In the pea, which is naturally self-pollinating, this is done by manually transferring pollen from the anther of a mature pea plant of one variety to the stigma of a separate mature pea plant of the second variety.

Plants used in first-generation crosses were called P, or parental generation, plants (Figure 8.3). Mendel collected the seeds produced by the P plants that resulted from each cross and grew them the following season. These offspring were called the $\mathbf{F_1}$, or the first filial (filial = daughter or son), generation. Once Mendel examined the characteristics in the F_1 generation of plants, he allowed them to self-fertilize naturally. He then collected and grew the seeds from the F_1 plants to produce the $\mathbf{F_2}$, or second filial, generation. Mendel's experiments extended beyond the F_2 generation to the F_3 generation, F_4 generation, and so on, but it was the ratio of characteristics in the P, F_1, and F_2 generations that were the most intriguing and became the basis of Mendel's postulates.

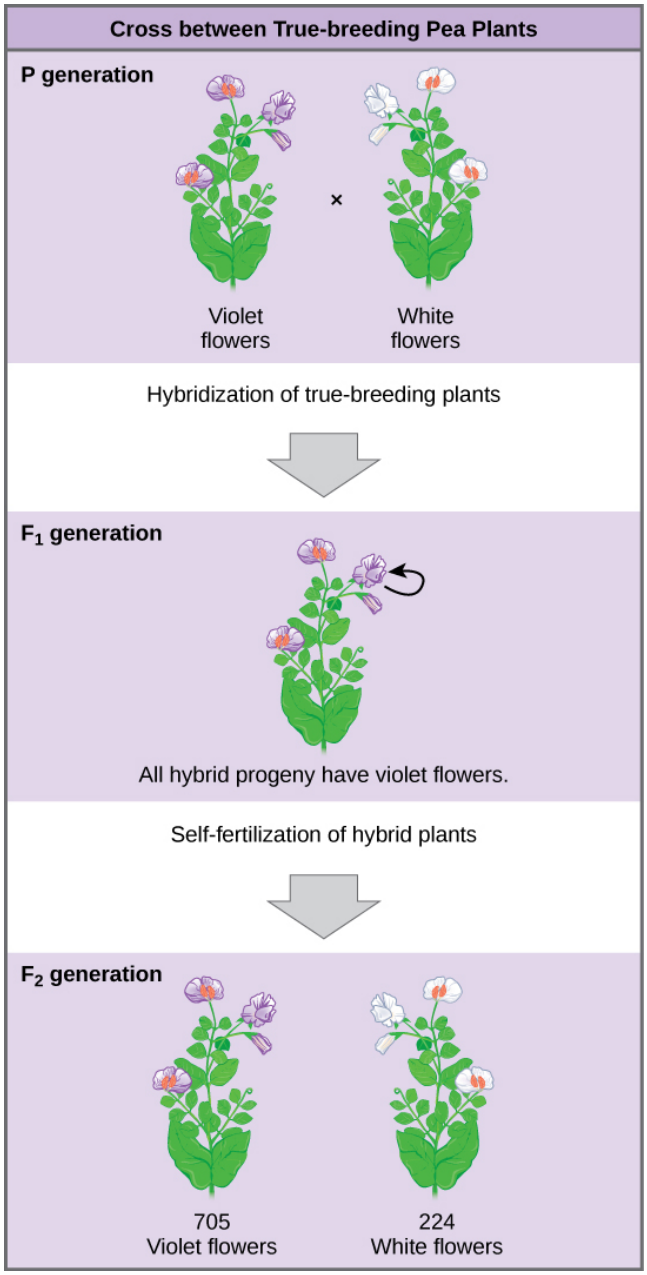


FIGURE 8.3 Mendel's process for performing crosses included examining flower color.

Garden Pea Characteristics Revealed the Basics of Heredity

In his 1865 publication, Mendel reported the results of his crosses involving seven different characteristics, each with two contrasting traits. A **trait** is defined as a variation in the physical appearance of a heritable characteristic. The characteristics included plant height, seed texture, seed color, flower color, pea-pod size, pea-pod color, and flower position. For the characteristic of flower color, for example, the two contrasting traits were white versus violet. To fully examine each characteristic, Mendel generated large numbers of F_1 and F_2 plants and reported results from thousands of F_2 plants.

What results did Mendel find in his crosses for flower color? First, Mendel confirmed that he was using plants that bred true for white or violet flower color. Irrespective of the number of generations that Mendel examined, all self-crossed offspring of parents with white flowers had white flowers, and all self-crossed offspring of parents with violet flowers had violet flowers. In addition, Mendel confirmed that, other than flower color, the pea plants were physically identical. This was an important check to make sure that the two varieties of pea plants only differed with respect to one trait, flower color.

Once these validations were complete, Mendel applied the pollen from a plant with violet flowers to the stigma of a plant with white flowers. After gathering and sowing the seeds that resulted from this cross, Mendel found that 100 percent of the F_1 hybrid generation had violet flowers. Conventional wisdom at that time would have predicted the hybrid flowers to be pale violet or for hybrid plants to have equal numbers of white and violet flowers. In other words, the contrasting parental traits were expected to blend in the offspring. Instead, Mendel's results demonstrated that the white flower trait had completely disappeared in the F_1 generation.

Importantly, Mendel did not stop his experimentation there. He allowed the F_1 plants to self-fertilize and found that 705 plants in the F_2 generation had violet flowers and 224 had white flowers. This was a ratio of 3.15 violet flowers to one white flower, or approximately 3:1. Mendel performed an additional experiment to ascertain differences in inheritance of traits carried in the pollen versus the ovum. When Mendel transferred pollen from a plant with violet flowers to fertilize the ova of a plant with white flowers and vice versa, he obtained approximately the same ratio irrespective of which gamete contributed which trait. This is called a **reciprocal cross**—a paired cross in which the respective traits of the male and female in one cross become the respective traits of the female and male in the other cross. For the other six characteristics that Mendel examined, the F_1 and F_2 generations behaved in the same way that they behaved for flower color. One of the two traits would disappear completely from the F_1 generation, only to reappear in the F_2 generation at a ratio of roughly 3:1 (Figure 8.4).

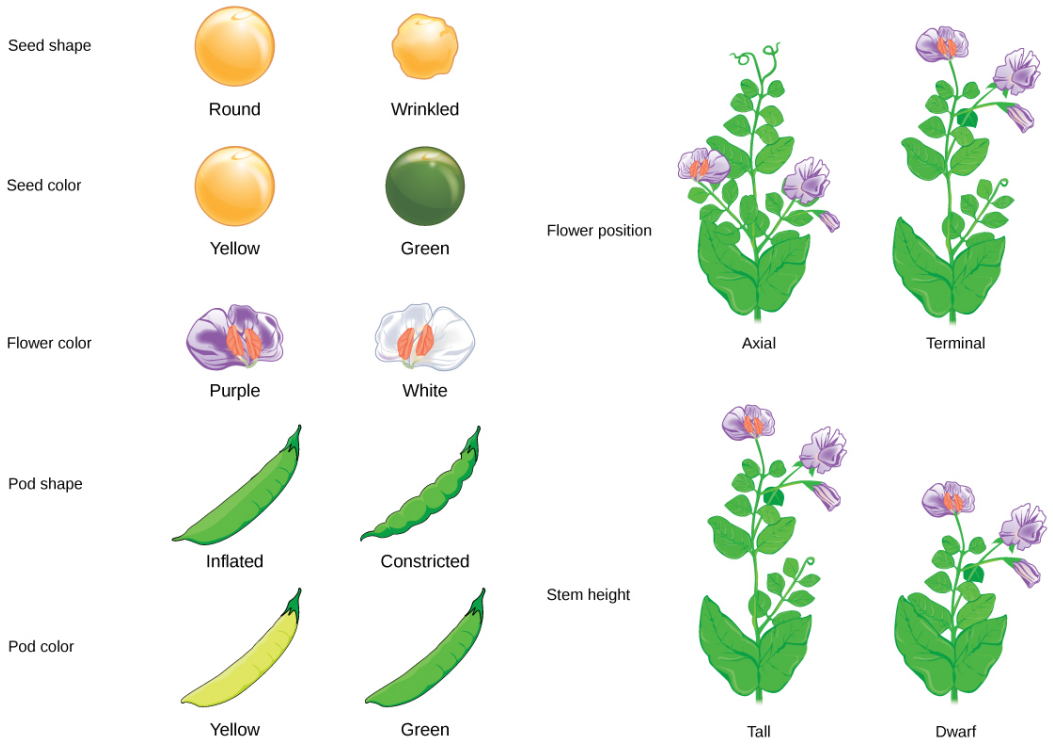


FIGURE 8.4 Mendel identified seven pea plant characteristics.

Upon compiling his results for many thousands of plants, Mendel concluded that the characteristics could be divided into expressed and latent traits. He called these dominant and recessive traits, respectively. **Dominant** traits are those that are inherited unchanged in a hybridization. **Recessive** traits become latent, or disappear in the offspring of a hybridization. The recessive trait does, however, reappear in the progeny of the hybrid offspring. An example of a dominant trait is the violet-colored flower trait. For this same characteristic (flower color), white-colored flowers are a recessive trait. The fact that the recessive trait reappeared in the F_2 generation meant that the traits remained separate (and were not blended) in the plants of the F_1 generation. Mendel proposed that this was because the plants possessed two copies of the trait for the flower-color characteristic, and that each parent transmitted one of their two copies to their offspring, where they came together. Moreover, the physical observation of a dominant trait could mean that the genetic composition of the organism included two dominant versions of the characteristic, or that it included one dominant and one recessive version. Conversely, the observation of a recessive trait meant that the organism lacked any dominant versions of this characteristic.

8.2 Laws of Inheritance

LEARNING OBJECTIVES

By the end of this section, you will be able to:

- Explain the relationship between genotypes and phenotypes in dominant and recessive gene systems
- Use a Punnett square to calculate the expected proportions of genotypes and phenotypes in a monohybrid cross
- Explain Mendel's law of segregation and independent assortment in terms of genetics and the events of meiosis
- Explain the purpose and methods of a test cross

The seven characteristics that Mendel evaluated in his pea plants were each expressed as one of two versions, or traits. Mendel deduced from his results that each individual had two discrete copies of the characteristic that are passed individually to offspring. We now call those two copies genes, which are carried on chromosomes. The reason we have two copies of each gene is that we inherit one from each parent. In fact, it is the chromosomes we inherit and the two copies of each gene are located on paired chromosomes. Recall that in meiosis these chromosomes are separated out into haploid gametes. This separation, or segregation, of the homologous chromosomes means also that only one of the copies of the gene gets moved into a gamete. The offspring are formed when that gamete unites with one from another parent and the two copies of each gene (and chromosome) are restored.

For cases in which a single gene controls a single characteristic, a diploid organism has two genetic copies that may or may not encode the same version of that characteristic. For example, one individual may carry a gene that determines white flower color and a gene that determines violet flower color. Gene variants that arise by mutation and exist at the same relative locations on homologous chromosomes are called **alleles**. Mendel examined the inheritance of genes with just two allele forms, but it is common to encounter more than two alleles for any given gene in a natural population.

Phenotypes and Genotypes

Two alleles for a given gene in a diploid organism are expressed and interact to produce physical characteristics. The observable traits expressed by an organism are referred to as its **phenotype**. An organism's underlying genetic makeup, consisting of both the physically visible and the non-expressed alleles, is called its **genotype**. Mendel's hybridization experiments demonstrate the difference between phenotype and genotype. For example, the phenotypes that Mendel observed in his crosses between pea plants with differing traits are connected to the diploid genotypes of the plants in the P, F_1, and F_2 generations. We will use a second trait that Mendel investigated, seed color, as an example. Seed color is governed by a single gene with two alleles. The yellow-seed allele is dominant and the green-seed allele is recessive. When true-breeding plants were cross-fertilized, in which one parent had yellow seeds and one had green seeds, all of the F_1 hybrid offspring had yellow seeds. That is, the hybrid offspring were phenotypically identical to the true-breeding parent with yellow seeds. However, we know that the allele donated by the parent with green seeds was not simply lost because it reappeared in some of the F_2 offspring (Figure 8.5). Therefore, the F_1 plants must have been genotypically different from the parent with yellow seeds.

The P plants that Mendel used in his experiments were each homozygous for the trait he was studying. Diploid organisms that are **homozygous** for a gene have two identical alleles, one on each of their homologous chromosomes. The genotype is often written as *YY* or *yy*, for which each letter represents one of the two alleles in the genotype. The dominant allele is capitalized and the recessive allele is lower case. The letter used for the gene (seed color in this case) is usually related to the dominant trait (yellow allele, in this case, or "*Y*"). Mendel's parental pea plants always bred true because both produced gametes carried the same allele. When P plants with contrasting traits were cross-fertilized, all of the offspring were **heterozygous** for the contrasting trait, meaning their genotype had different alleles for the gene being examined. For example, the F_1 yellow plants that received a *Y* allele from their yellow parent and a *y* allele from their green parent had the genotype *Yy*.

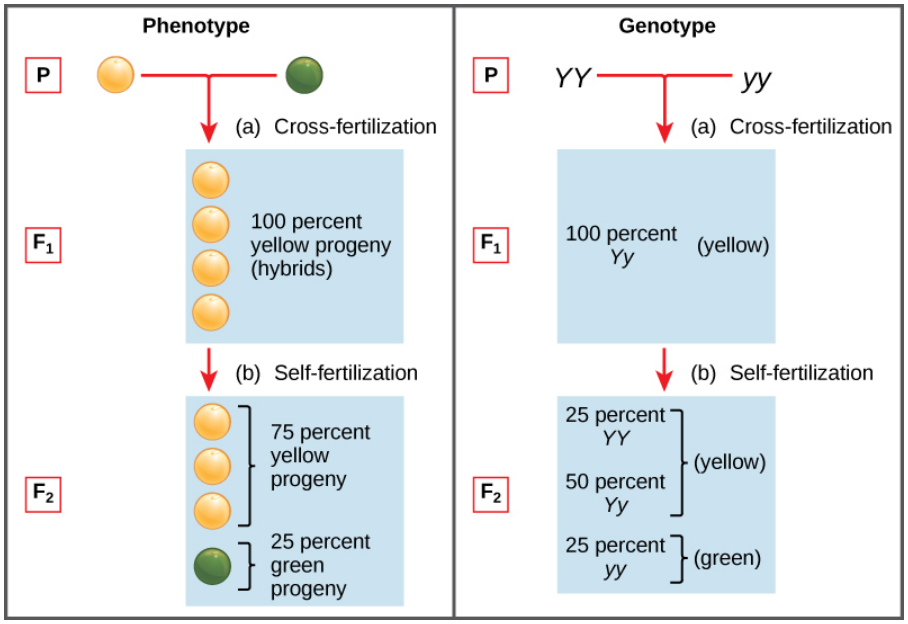


FIGURE 8.5 Phenotypes are physical expressions of traits that are transmitted by alleles. Capital letters represent dominant alleles and lowercase letters represent recessive alleles. The phenotypic ratios are the ratios of visible characteristics. The genotypic ratios are the ratios of gene combinations in the offspring, and these are not always distinguishable in the phenotypes.

Law of Dominance

Our discussion of homozygous and heterozygous organisms brings us to why the F_1 heterozygous offspring were identical to one of the parents, rather than expressing both alleles. In all seven pea-plant characteristics, one of the two contrasting alleles was dominant, and the other was recessive. Mendel called the dominant allele the expressed unit factor; the recessive allele was referred to as the latent unit factor. We now know that these so-called unit factors are actually genes on homologous chromosomes. For a gene that is expressed in a dominant and recessive pattern, homozygous dominant and heterozygous organisms will look identical (that is, they will have different genotypes but the same phenotype), and the traits of the recessive allele will only be observed in homozygous recessive individuals (Table 8.1).

Correspondence between Genotype and Phenotype for a Dominant-Recessive Characteristic.

	Homozygous	Heterozygous	Homozygous
Genotype	*YY*	*Yy*	*yy*
Phenotype	yellow	yellow	green

TABLE 8.1

Mendel's **law of dominance** states that in a heterozygote, one trait will conceal the presence of another trait for the same characteristic. For example, when crossing true-breeding violet-flowered plants with true-breeding white-flowered plants, all of the offspring were violet-flowered, even though they all had one allele for violet and one allele for white. Rather than both alleles contributing to a phenotype, the dominant allele will be expressed exclusively. The recessive allele will remain latent, but will be transmitted to offspring in the same manner as that by which the dominant allele is transmitted. The recessive trait will only be expressed by offspring that have two copies of this allele (Figure 8.6), and these offspring will breed true when self-crossed.

FIGURE 8.6 The allele for albinism, expressed here in humans, is recessive. Both of this child's parents carried the recessive allele.

Monohybrid Cross and the Punnett Square

When fertilization occurs between two true-breeding parents that differ by only the characteristic being studied, the process is called a **monohybrid** cross, and the resulting offspring are called monohybrids. Mendel performed seven types of monohybrid crosses, each involving contrasting traits for different characteristics. Out of these crosses, all of the F_1 offspring had the phenotype of one parent, and the F_2 offspring had a 3:1 phenotypic ratio. On the basis of these results, Mendel postulated that each parent in the monohybrid cross contributed one of two paired unit factors to each offspring, and every possible combination of unit factors was equally likely.

The results of Mendel's research can be explained in terms of probabilities, which are mathematical measures of likelihood. The probability of an event is calculated by the number of times the event occurs divided by the total number of opportunities for the event to occur. A probability of one (100 percent) for some event indicates that it is guaranteed to occur, whereas a probability of zero (0 percent) indicates that it is guaranteed to not occur, and a probability of 0.5 (50 percent) means it has an equal chance of occurring or not occurring.

To demonstrate this with a monohybrid cross, consider the case of true-breeding pea plants with yellow versus green seeds. The dominant seed color is yellow; therefore, the parental genotypes were *YY* for the plants with yellow seeds and *yy* for the plants with green seeds. A **Punnett square**, devised by the British geneticist Reginald Punnett, is useful for determining probabilities because it is drawn to predict all possible outcomes of all possible random fertilization events and their expected frequencies. Figure 8.9 shows a Punnett square for a cross between a plant with yellow peas and one with green peas. To prepare a Punnett square, all possible combinations of the parental alleles (the genotypes of the gametes) are listed along the top (for one parent) and side (for the other parent) of a grid. The combinations of egg and sperm gametes are then made in the boxes in the table on the basis of which alleles are combining. Each box then represents the diploid genotype of a zygote, or fertilized egg. Because each possibility is equally likely, genotypic ratios can be determined from a Punnett square. If the pattern of inheritance (dominant and recessive) is known, the phenotypic ratios can be inferred as well. For a monohybrid cross of two true-breeding parents, each parent contributes one type of allele. In this case, only one genotype is possible in the F_1 offspring. All offspring are *Yy* and have yellow seeds.

When the F_1 offspring are crossed with each other, each has an equal probability of contributing either a *Y* or a *y* to the F_2 offspring. The result is a 1 in 4 (25 percent) probability of both parents contributing a *Y*, resulting in an offspring with a yellow phenotype; a 25 percent probability of parent A contributing a *Y* and parent B a *y*, resulting in offspring with a yellow phenotype; a 25 percent probability of parent A contributing a *y* and parent B a *Y*, also resulting in a yellow phenotype; and a (25 percent) probability of both parents contributing a *y*, resulting in a green

phenotype. When counting all four possible outcomes, there is a 3 in 4 probability of offspring having the yellow phenotype and a 1 in 4 probability of offspring having the green phenotype. This explains why the results of Mendel's F_2 generation occurred in a 3:1 phenotypic ratio. Using large numbers of crosses, Mendel was able to calculate probabilities, found that they fit the model of inheritance, and use these to predict the outcomes of other crosses.

Law of Segregation

Observing that true-breeding pea plants with contrasting traits gave rise to F_1 generations that all expressed the dominant trait and F_2 generations that expressed the dominant and recessive traits in a 3:1 ratio, Mendel proposed the **law of segregation**. This law states that paired unit factors (genes) must segregate equally into gametes such that offspring have an equal likelihood of inheriting either factor. For the F_2 generation of a monohybrid cross, the following three possible combinations of genotypes result: homozygous dominant, heterozygous, or homozygous recessive. Because heterozygotes could arise from two different pathways (receiving one dominant and one recessive allele from either parent), and because heterozygotes and homozygous dominant individuals are phenotypically identical, the law supports Mendel's observed 3:1 phenotypic ratio. The equal segregation of alleles is the reason we can apply the Punnett square to accurately predict the offspring of parents with known genotypes. The physical basis of Mendel's law of segregation is the first division of meiosis in which the homologous chromosomes with their different versions of each gene are segregated into daughter nuclei. This process was not understood by the scientific community during Mendel's lifetime (Figure 8.7).

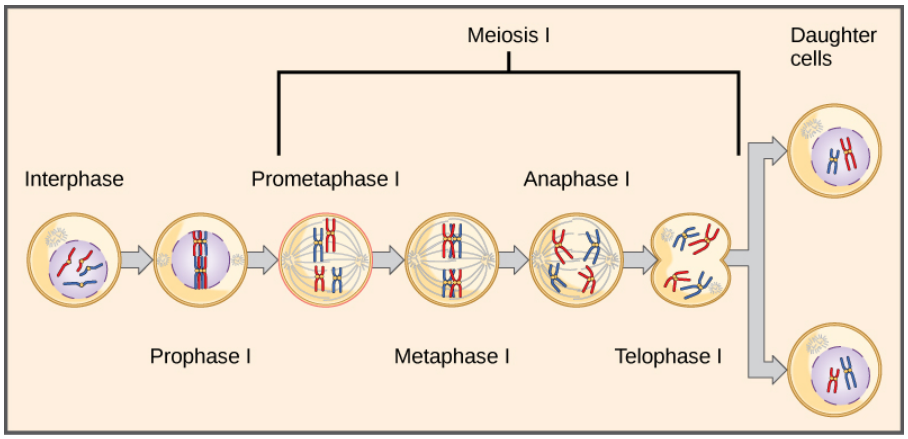


FIGURE 8.7 The first division in meiosis is shown.

Test Cross

Beyond predicting the offspring of a cross between known homozygous or heterozygous parents, Mendel also developed a way to determine whether an organism that expressed a dominant trait was a heterozygote or a homozygote. Called the **test cross**, this technique is still used by plant and animal breeders. In a test cross, the dominant-expressing organism is crossed with an organism that is homozygous recessive for the same characteristic. If the dominant-expressing organism is a homozygote, then all F_1 offspring will be heterozygotes expressing the dominant trait (Figure 8.8). Alternatively, if the dominant-expressing organism is a heterozygote, the F_1 offspring will exhibit a 1:1 ratio of heterozygotes and recessive homozygotes (Figure 8.8). The test cross further validates Mendel's postulate that pairs of unit factors segregate equally.

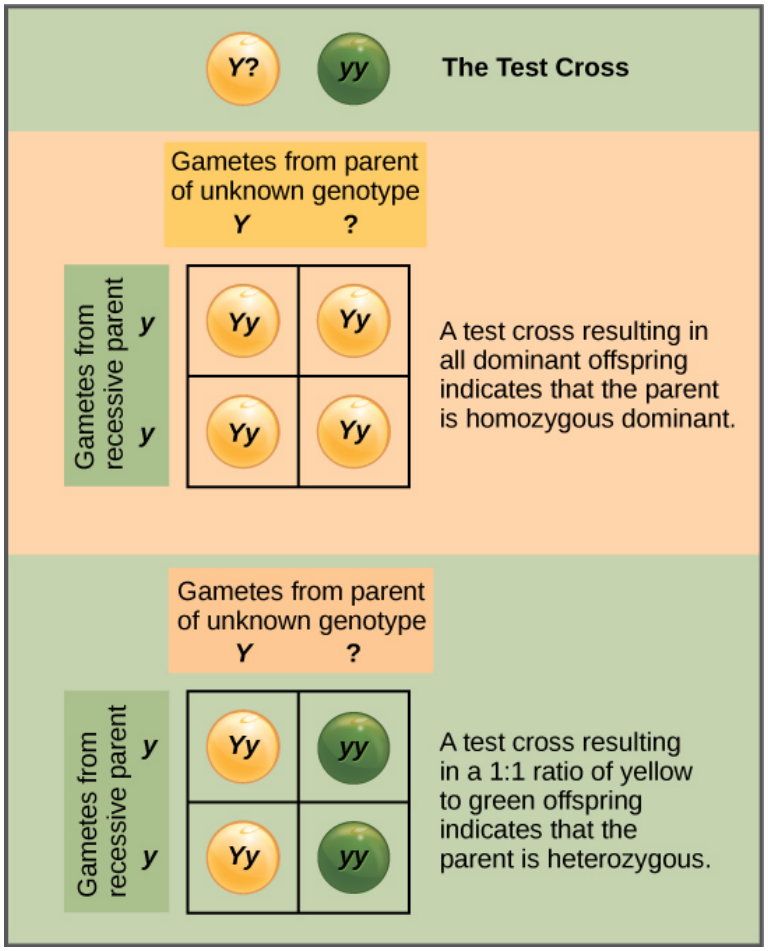


FIGURE 8.8 A test cross can be performed to determine whether an organism expressing a dominant trait is a homozygote or a heterozygote.

VISUAL CONNECTION

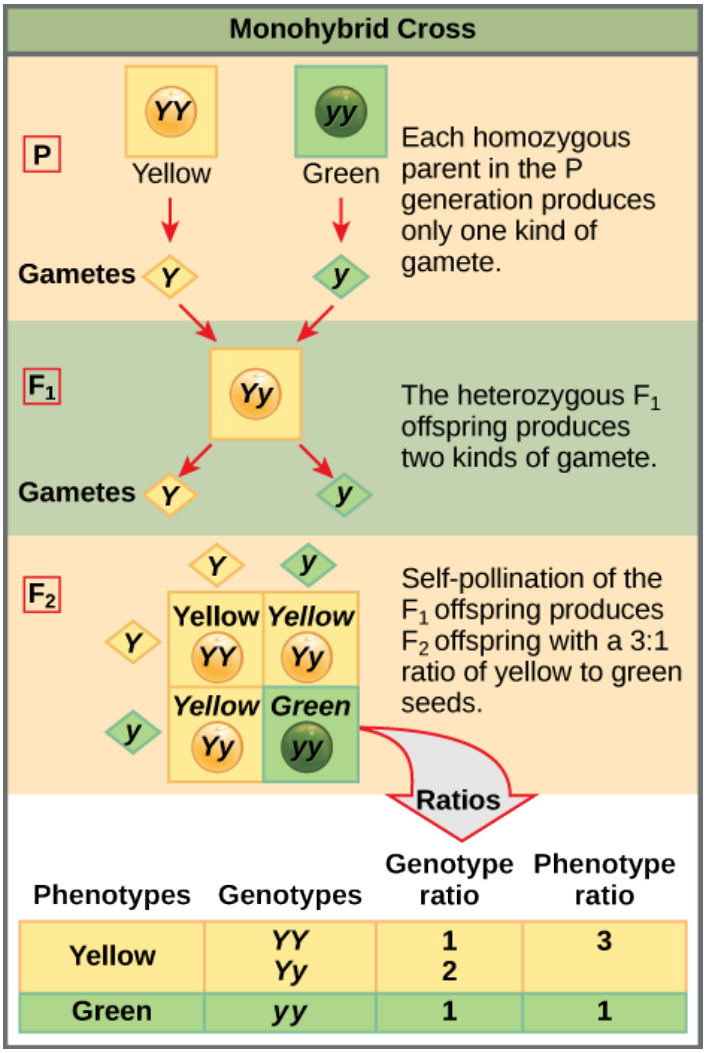


FIGURE 8.9 This Punnett square shows the cross between plants with yellow seeds and green seeds. The cross between the true-breeding P plants produces F_1 heterozygotes that can be self-fertilized. The self-cross of the F_1 generation can be analyzed with a Punnett square to predict the genotypes of the F_2 generation. Given an inheritance pattern of dominant–recessive, the genotypic and phenotypic ratios can then be determined.

In pea plants, round peas (*R*) are dominant to wrinkled peas (*r*). You do a test cross between a pea plant with wrinkled peas (genotype *rr*) and a plant of unknown genotype that has round peas. You end up with three plants, all which have round peas. From this data, can you tell if the parent plant is homozygous dominant or heterozygous?

Law of Independent Assortment

Mendel's **law of independent assortment** states that genes do not influence each other with regard to the sorting of alleles into gametes, and every possible combination of alleles for every gene is equally likely to occur. Independent assortment of genes can be illustrated by the **dihybrid** cross, a cross between two true-breeding parents that express different traits for two characteristics. Consider the characteristics of seed color and seed texture for two pea plants, one that has wrinkled, green seeds (*rryy*) and another that has round, yellow seeds (*RRYY*). Because each parent is homozygous, the law of segregation indicates that the gametes for the wrinkled–green plant all are *ry*, and the gametes for the round–yellow plant are all *RY*. Therefore, the F_1 generation of offspring all are *RrYy* (Figure 8.10).

VISUAL CONNECTION

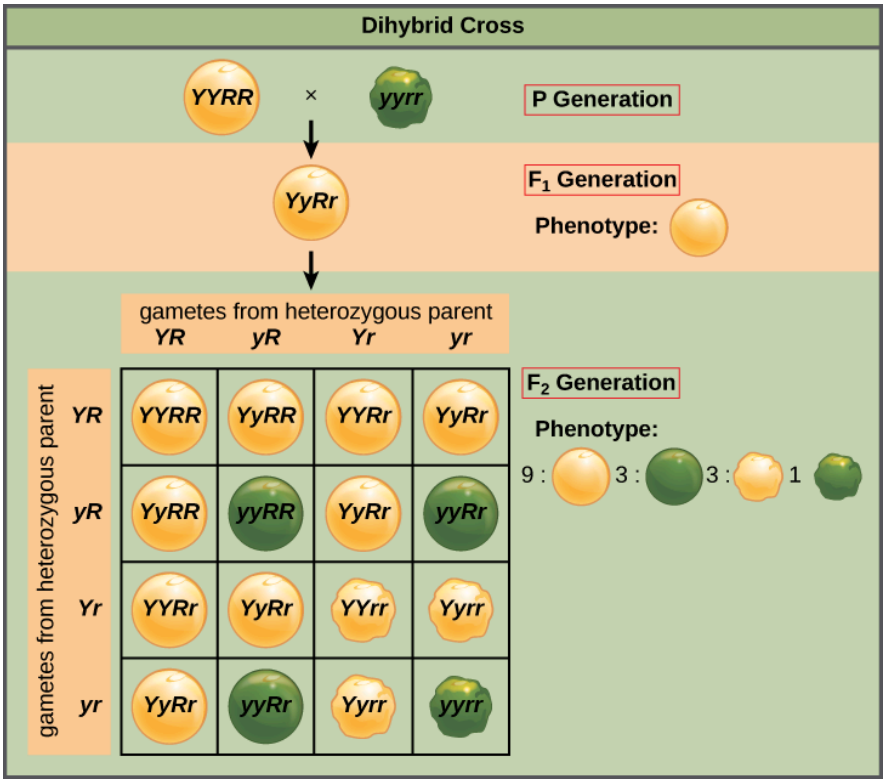


FIGURE 8.10 A dihybrid cross in pea plants involves the genes for seed color and texture. The P cross produces F_1 offspring that are all heterozygous for both characteristics. The resulting 9:3:3:1 F_2 phenotypic ratio is obtained using a Punnett square.

In pea plants, round seed shape (*R*) is dominant to wrinkled seed shape (*r*) and yellow peas (*Y*) are dominant to green peas (*y*). What are the possible genotypes and phenotypes for a cross between *RrYY* and *rrYy* pea plants? How many squares do you need to do a Punnett square analysis of this cross?

The gametes produced by the F_1 individuals must have one allele from each of the two genes. For example, a gamete could get an *R* allele for the seed shape gene and either a *Y* or a *y* allele for the seed color gene. It cannot get both an *R* and an *r* allele; each gamete can have only one allele per gene. The law of independent assortment states that a gamete into which an *r* allele is sorted would be equally likely to contain either a *Y* or a *y* allele. Thus, there are four equally likely gametes that can be formed when the *RrYy* heterozygote is self-crossed, as follows: *RY*, *rY*, *Ry*, and *ry*. Arranging these gametes along the top and left of a 4 × 4 Punnett square (Figure 8.10) gives us 16 equally likely genotypic combinations. From these genotypes, we find a phenotypic ratio of 9 round–yellow:3 round–green:3 wrinkled–yellow:1 wrinkled–green (Figure 8.10). These are the offspring ratios we would expect, assuming we performed the crosses with a large enough sample size.

The physical basis for the law of independent assortment also lies in meiosis I, in which the different homologous pairs line up in random orientations. Each gamete can contain any combination of paternal and maternal chromosomes (and therefore the genes on them) because the orientation of tetrads on the metaphase plane is random (Figure 8.11).

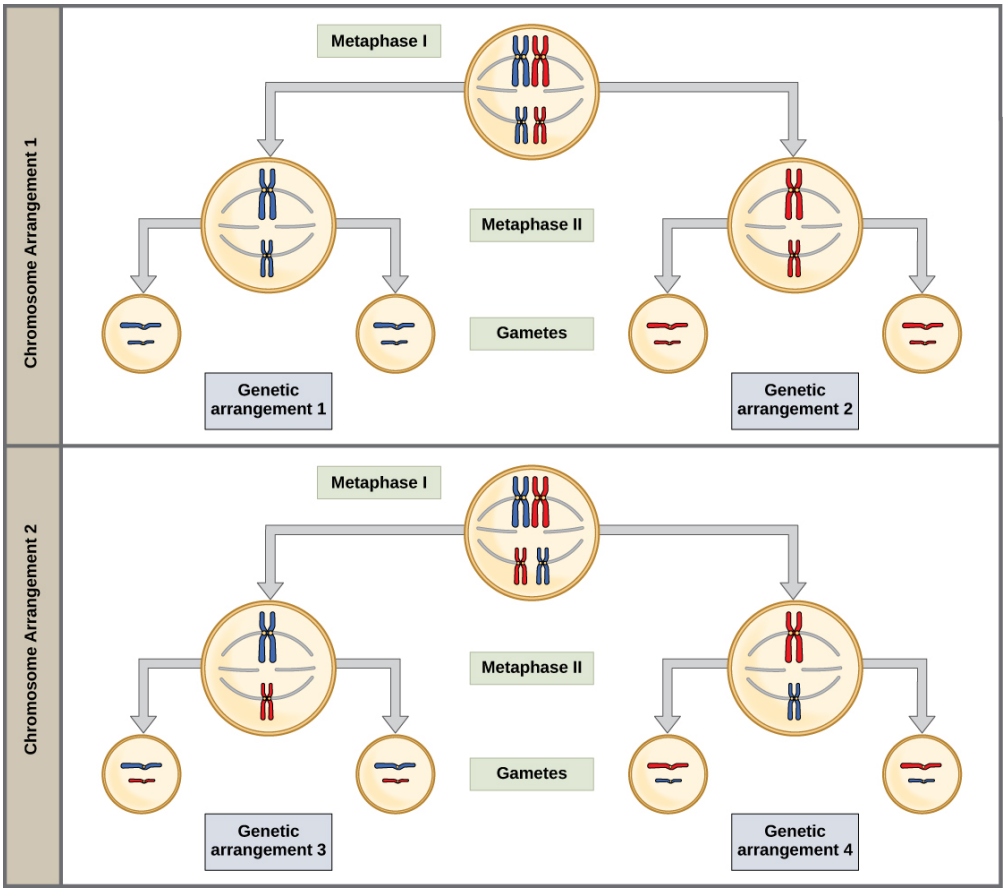


FIGURE 8.11 The random segregation into daughter nuclei that happens during the first division in meiosis can lead to a variety of possible genetic arrangements.

8.3 Extensions of the Laws of Inheritance

LEARNING OBJECTIVES

By the end of this section, you will be able to:

- Identify non-Mendelian inheritance patterns such as incomplete dominance, codominance, multiple alleles, and sex linkage from the results of crosses
- Explain the effect of linkage and recombination on gamete genotypes
- Explain the phenotypic outcomes of epistatic effects among genes

Mendel studied traits with only one mode of inheritance in pea plants. The inheritance of the traits he studied all followed the relatively simple pattern of dominant and recessive alleles for a single characteristic. There are several important modes of inheritance, discovered after Mendel's work, that do not follow the dominant and recessive, single-gene model.

Alternatives to Dominance and Recessiveness

Mendel's experiments with pea plants suggested that: 1) two types of "units" or alleles exist for every gene; 2) alleles maintain their integrity in each generation (no blending); and 3) in the presence of the dominant allele, the recessive allele is hidden, with no contribution to the phenotype. Therefore, recessive alleles can be "carried" and not expressed by individuals. Such heterozygous individuals are sometimes referred to as "carriers." Since then, genetic studies in other organisms have shown that much more complexity exists, but that the fundamental principles of Mendelian genetics still hold true. In the sections to follow, we consider some of the extensions of Mendelism.

Incomplete Dominance

Mendel's results, demonstrating that traits are inherited as dominant and recessive pairs, contradicted the view at that time that offspring exhibited a blend of their parents' traits. However, the heterozygote phenotype occasionally

does appear to be intermediate between the two parents. For example, in the snapdragon, *Antirrhinum majus* (Figure 8.12), a cross between a homozygous parent with white flowers (C^WC^W) and a homozygous parent with red flowers (C^RC^R) will produce offspring with pink flowers (C^RC^W). (Note that different genotypic abbreviations are used for Mendelian extensions to distinguish these patterns from simple dominance and recessiveness.) This pattern of inheritance is described as **incomplete dominance**, meaning that one of the alleles appears in the phenotype in the heterozygote, but not to the exclusion of the other, which can also be seen. The allele for red flowers is incompletely dominant over the allele for white flowers. However, the results of a heterozygote self-cross can still be predicted, just as with Mendelian dominant and recessive crosses. In this case, the genotypic ratio would be 1 C^RC^R:2 C^RC^W:1 C^WC^W, and the phenotypic ratio would be 1:2:1 for red:pink:white. The basis for the intermediate color in the heterozygote is simply that the pigment produced by the red allele (anthocyanin) is diluted in the heterozygote and therefore appears pink because of the white background of the flower petals.

FIGURE 8.12 These pink flowers of a heterozygote snapdragon result from incomplete dominance. (credit: "storebukkebruse"/Flickr)

Codominance

A variation on incomplete dominance is **codominance**, in which both alleles for the same characteristic are simultaneously expressed in the heterozygote. An example of codominance occurs in the ABO blood groups of humans. The A and B alleles are expressed in the form of A or B molecules present on the surface of red blood cells. Homozygotes (I^AI^A and I^BI^B) express either the A or the B phenotype, and heterozygotes (I^AI^B) express both phenotypes equally. The I^AI^B individual has blood type AB. In a self-cross between heterozygotes expressing a codominant trait, the three possible offspring genotypes are phenotypically distinct. However, the 1:2:1 genotypic ratio characteristic of a Mendelian monohybrid cross still applies (Figure 8.13).

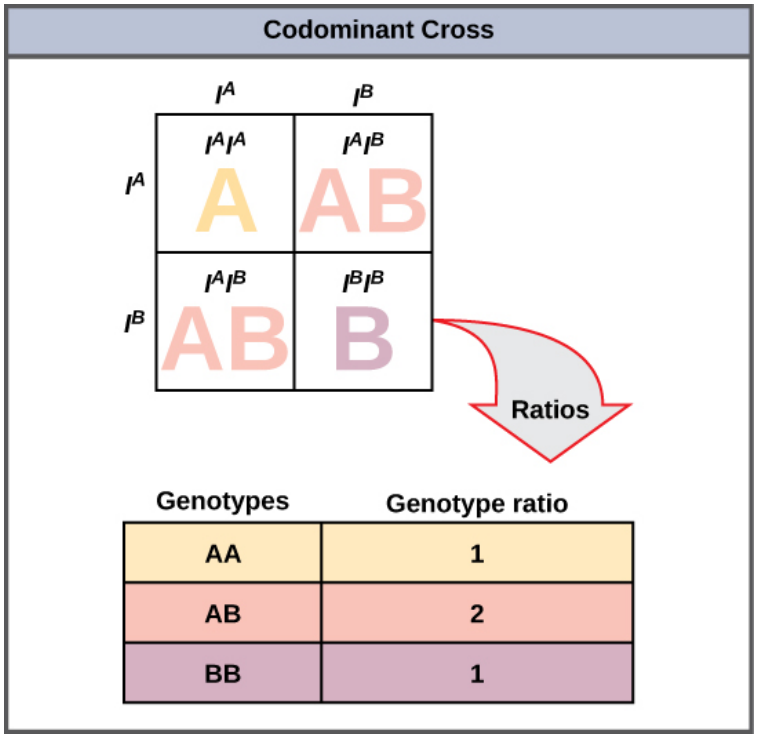


FIGURE 8.13 This Punnet square shows an AB/AB blood type cross

Multiple Alleles

Mendel implied that only two alleles, one dominant and one recessive, could exist for a given gene. We now know that this is an oversimplification. Although individual humans (and all diploid organisms) can only have two alleles for a given gene, multiple alleles may exist at the population level, such that many combinations of two alleles are observed. Note that when many alleles exist for the same gene, the convention is to denote the most common phenotype or genotype in the natural population as the **wild type** (often abbreviated "+"). All other phenotypes or genotypes are considered variants (mutants) of this typical form, meaning they deviate from the wild type. The variant may be recessive or dominant to the wild-type allele.

An example of multiple alleles is the ABO blood-type system in humans. In this case, there are three alleles circulating in the population. The I^A allele codes for A molecules on the red blood cells, the I^B allele codes for B molecules on the surface of red blood cells, and the *i* allele codes for no molecules on the red blood cells. In this case, the I^A and I^B alleles are codominant with each other and are both dominant over the *i* allele. Although there are three alleles present in a population, each individual only gets two of the alleles from their parents. This produces the genotypes and phenotypes shown in Figure 8.14. Notice that instead of three genotypes, there are six different genotypes when there are three alleles. The number of possible phenotypes depends on the dominance relationships between the three alleles.

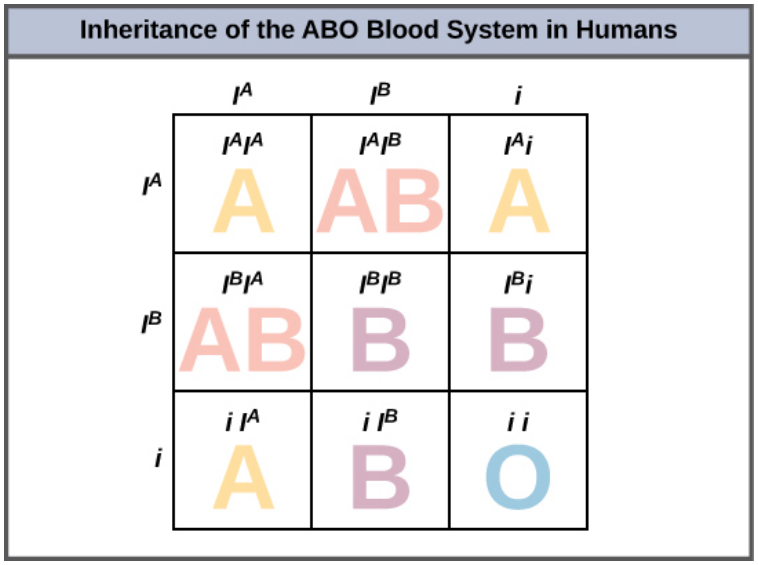


FIGURE 8.14 Inheritance of the ABO blood system in humans is shown.

EVOLUTION CONNECTION

Multiple Alleles Confer Drug Resistance in the Malaria Parasite

Malaria is a parasitic disease in humans that is transmitted by infected female mosquitoes, including *Anopheles gambiae*, and is characterized by cyclic high fevers, chills, flu-like symptoms, and severe anemia. *Plasmodium falciparum* and *P. vivax* are the most common causative agents of malaria, and *P. falciparum* is the most deadly. When promptly and correctly treated, *P. falciparum* malaria has a mortality rate of 0.1 percent. However, in some parts of the world, the parasite has evolved resistance to commonly used malaria treatments, so the most effective malarial treatments can vary by geographic region. Ninety percent of malaria victims live in Africa, most of them children under the age of five.

In Southeast Asia, Africa, and South America, *P. falciparum* has developed resistance to the anti-malarial drugs chloroquine, mefloquine, and sulfadoxine-pyrimethamine. *P. falciparum*, which is haploid during the life stage in which it is infective to humans, has evolved multiple drug-resistant mutant alleles of the *dhps* gene. Varying degrees of sulfadoxine resistance are associated with each of these alleles. Being haploid, *P. falciparum* needs only one drug-resistant allele to express this trait.

In Southeast Asia, different sulfadoxine-resistant alleles of the *dhps* gene are localized to different geographic regions. This is a common evolutionary phenomenon that comes about because drug-resistant mutants arise in a population and interbreed with other *P. falciparum* isolates in close proximity. Sulfadoxine-resistant parasites cause considerable human hardship in regions in which this drug is widely used as an over-the-counter malaria remedy. As is common with pathogens that multiply to large numbers within an infection cycle, *P. falciparum* evolves relatively rapidly (over a decade or so) in response to the selective pressure of commonly used anti-malarial drugs. For this reason, scientists must constantly work to develop new drugs or drug combinations to combat the worldwide malaria burden.[2]

n late 2021, R21/Matrix-M became the first vaccine to be recommended for widespread use by the World Health Organization. At least ten other candidate vaccines are in development. The effort is an multinational one involving governments, universities, nonprofits, philanthropists, and pharmaceutical companies. Much of the recent progress can be credited to organizations within the most affected countries, such as the Malaria Research and Training Center in Mali. Founded by Ogobara Duombo and Yeya Touré in the 1990s, the center has emerged as a primary front-line research driver, including running many of the critical clinical trials that are so important to vaccine development and approval.

2 Sumiti Vinayak et al., "Origin and Evolution of Sulfadoxine Resistant *Plasmodium falciparum*," *PLoS Pathogens* 6 (2010): e1000830.

Sex-Linked Traits

In humans, as well as in many other animals and some plants, the sex of the individual is determined by sex chromosomes—one pair of non-homologous chromosomes. Humans may identify as being male, female, neither of these, both, or other gender(s) independently of these chromosomes, but the sex chromosomes can be associated with certain traits. Until now, we have only considered inheritance patterns among non-sex chromosomes, or autosomes. In addition to 22 homologous pairs of autosomes, human females have a homologous pair of X chromosomes, whereas human males have an XY chromosome pair. Although the Y chromosome contains a small region of similarity to the X chromosome so that they can pair during meiosis, the Y chromosome is much shorter and contains fewer genes. When a gene being examined is present on the X, but not the Y, chromosome, it is **X-linked**.

Eye color in *Drosophila*, the common fruit fly, was the first X-linked trait to be identified. Thomas Hunt Morgan mapped this trait to the X chromosome in 1910. Like humans, *Drosophila* males have an XY chromosome pair, and females are XX. In flies the wild-type eye color is red (X^W) and is dominant to white eye color (X^w) (Figure 8.15). Because of the location of the eye-color gene, reciprocal crosses do not produce the same offspring ratios. Males are said to be **hemizygous**, in that they have only one allele for any X-linked characteristic. Hemizygosity makes descriptions of dominance and recessiveness irrelevant for XY males. *Drosophila* males lack the white gene on the Y chromosome; that is, their genotype can only be X^WY or X^wY. In contrast, females have two allele copies of this gene and can be X^WX^W, X^WX^w, or X^wX^w.

FIGURE 8.15 In *Drosophila*, the gene for eye color is located on the X chromosome. Red eye color is wild-type and is dominant to white eye color.

In an X-linked cross, the genotypes of F_1 and F_2 offspring depend on whether the recessive trait was expressed by the male or the female in the P generation. With respect to *Drosophila* eye color, when the P male expresses the white-eye phenotype and the female is homozygously red-eyed, all members of the F_1 generation exhibit red eyes (Figure 8.16). The F_1 females are heterozygous (X^WX^w), and the males are all X^WY, having received their X chromosome from the homozygous dominant P female and their Y chromosome from the P male. A subsequent cross between the X^WX^w female and the X^WY male would produce only red-eyed females (with X^WX^W or X^WX^w genotypes) and both red- and white-eyed males (with X^WY or X^wY genotypes). Now, consider a cross between a homozygous white-eyed female and a male with red eyes. The F_1 generation would exhibit only heterozygous red-eyed females (X^WX^w) and only white-eyed males (X^wY). Half of the F_2 females would be red-eyed (X^WX^w) and half would be white-eyed (X^wX^w). Similarly, half of the F_2 males would be red-eyed (X^WY) and half would be white-eyed (X^wY).

VISUAL CONNECTION

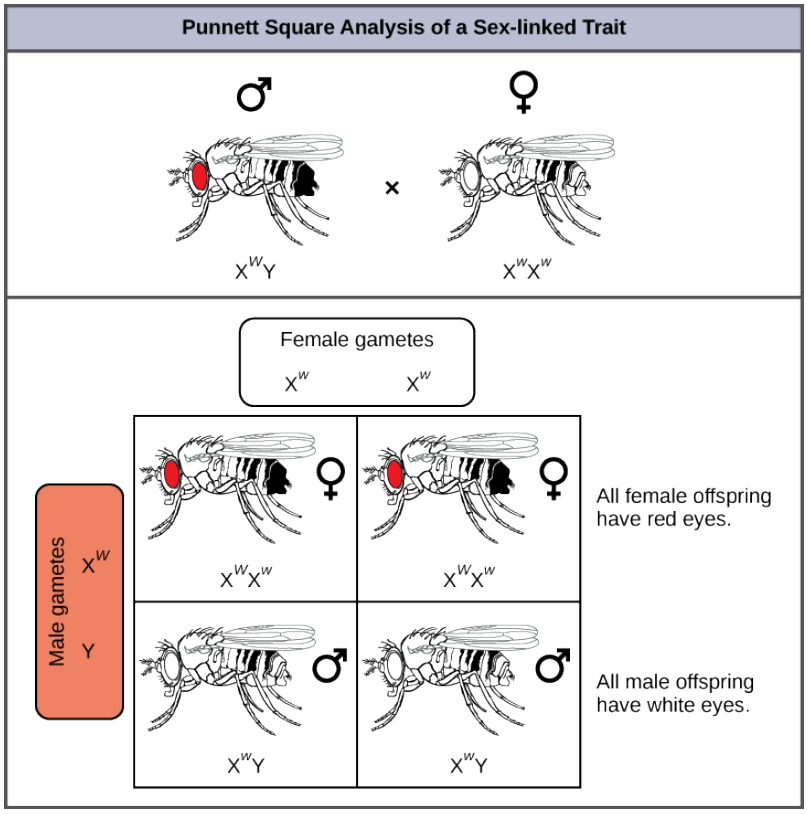


FIGURE 8.16 Crosses involving sex-linked traits often give rise to different phenotypes for the different sexes of offspring, as is the case for this cross involving red and white eye color in *Drosophila*. In the diagram, *w* is the white-eye mutant allele and *W* is the wild-type, red-eye allele.

What ratio of offspring would result from a cross between a white-eyed male and a female that is heterozygous for red eye color?

Discoveries in fruit fly genetics can be applied to human genetics. When a female parent is homozygous for a recessive X-linked trait, the parent will pass the trait on to 100 percent of the male offspring, because the males will receive the Y chromosome from the male parent. In humans, the alleles for certain conditions (some color-blindness, hemophilia, and muscular dystrophy) are X-linked. Females who are heterozygous for these diseases are said to be carriers and may not exhibit any phenotypic effects. These females will pass the disease to half of their male offspring and will pass carrier status to half of their female offspring; therefore, X-linked traits appear more frequently in males than females.

In some groups of organisms with sex chromosomes, the sex with the non-homologous sex chromosomes is the female rather than the male. This is the case for all birds. In this case, sex-linked traits will be more likely to appear in the female, in whom they are hemizygous.

LINK TO LEARNING

Watch this video (http://openstax.org/l/sex-linked_trts) to learn more about sex-linked traits.

Linked Genes Violate the Law of Independent Assortment

Although all of Mendel's pea plant characteristics behaved according to the law of independent assortment, we now know that some allele combinations are not inherited independently of each other. Genes that are located on separate, non-homologous chromosomes will always sort independently. However, each chromosome contains

hundreds or thousands of genes, organized linearly on chromosomes like beads on a string. The segregation of alleles into gametes can be influenced by **linkage**, in which genes that are located physically close to each other on the same chromosome are more likely to be inherited as a pair. However, because of the process of recombination, or "crossover," it is possible for two genes on the same chromosome to behave independently, or as if they are not linked. To understand this, let us consider the biological basis of gene linkage and recombination.

Homologous chromosomes possess the same genes in the same order, though the specific alleles of the gene can be different on each of the two chromosomes. Recall that during interphase and prophase I of meiosis, homologous chromosomes first replicate and then synapse, with like genes on the homologs aligning with each other. At this stage, segments of homologous chromosomes exchange linear segments of genetic material (Figure 8.17). This process is called **recombination**, or crossover, and it is a common genetic process. Because the genes are aligned during recombination, the gene order is not altered. Instead, the result of recombination is that maternal and paternal alleles are combined onto the same chromosome. Across a given chromosome, several recombination events may occur, causing extensive shuffling of alleles.

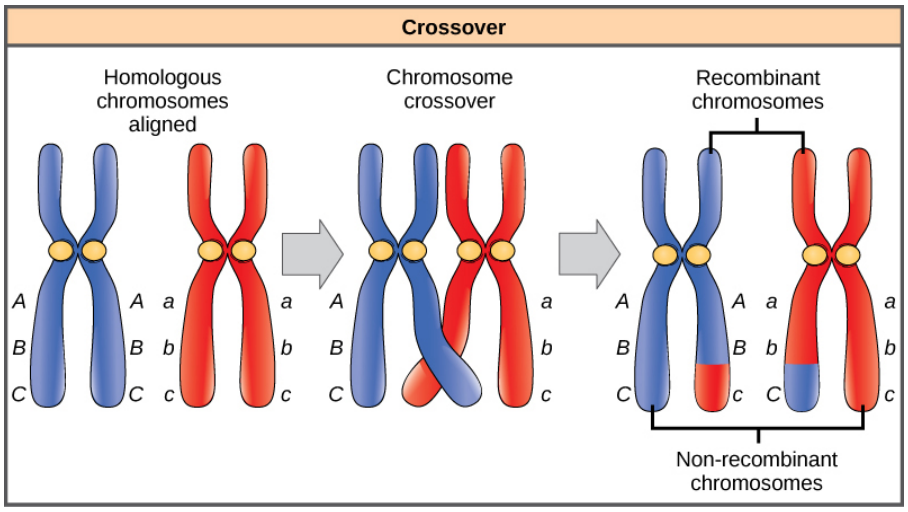


FIGURE 8.17 The process of crossover, or recombination, occurs when two homologous chromosomes align and exchange a segment of genetic material.

When two genes are located on the same chromosome, they are considered linked, and their alleles tend to be transmitted through meiosis together. To exemplify this, imagine a dihybrid cross involving flower color and plant height in which the genes are next to each other on the chromosome. If one homologous chromosome has alleles for tall plants and red flowers, and the other chromosome has genes for short plants and yellow flowers, then when the gametes are formed, the tall and red alleles will tend to go together into a gamete and the short and yellow alleles will go into other gametes. These are called the parental genotypes because they have been inherited intact from the parents of the individual producing gametes. But unlike if the genes were on different chromosomes, there will be no gametes with tall and yellow alleles and no gametes with short and red alleles. If you create a Punnett square with these gametes, you will see that the classical Mendelian prediction of a 9:3:3:1 outcome of a dihybrid cross would not apply. As the distance between two genes increases, the probability of one or more crossovers between them increases and the genes behave more like they are on separate chromosomes. Geneticists have used the proportion of recombinant gametes (the ones not like the parents) as a measure of how far apart genes are on a chromosome. Using this information, they have constructed linkage maps of genes on chromosomes for well-studied organisms, including humans.

Mendel's seminal publication makes no mention of linkage, and many researchers have questioned whether he encountered linkage but chose not to publish those crosses out of concern that they would invalidate his independent assortment postulate. The garden pea has seven pairs of chromosomes, and some have suggested that his choice of seven characteristics was not a coincidence. However, even if the genes he examined were not located on separate chromosomes, it is possible that he simply did not observe linkage because of the extensive shuffling effects of recombination.

Epistasis

Mendel's studies in pea plants implied that the sum of an individual's phenotype was controlled by genes (or as he called them, unit factors), such that every characteristic was distinctly and completely controlled by a single gene. In fact, single observable characteristics are almost always under the influence of multiple genes (each with two or more alleles) acting in unison. For example, at least eight genes contribute to eye color in humans.

LINK TO LEARNING

Eye color in humans is determined by multiple alleles. Use the Eye Color Calculator (http://openstax.org/l/eye_color_calc) to predict the eye color of children from parental eye color.

In some cases, several genes can contribute to aspects of a common phenotype without their gene products ever directly interacting. In the case of organ development, for instance, genes may be expressed sequentially, with each gene adding to the complexity and specificity of the organ. Genes may function in complementary or synergistic fashions, such that two or more genes expressed simultaneously affect a phenotype. An apparent example of this occurs with human skin color, which appears to involve the action of at least three (and probably more) genes. Cases in which inheritance for a characteristic like skin color or human height depend on the combined effects of numerous genes are called polygenic inheritance.

Genes may also oppose each other, with one gene suppressing the expression of another. In **epistasis**, the interaction between genes is antagonistic, such that one gene masks or interferes with the expression of another. "Epistasis" is a word composed of Greek roots meaning "standing upon." The alleles that are being masked or silenced are said to be hypostatic to the epistatic alleles that are doing the masking. Often the biochemical basis of epistasis is a gene pathway in which expression of one gene is dependent on the function of a gene that precedes or follows it in the pathway.

An example of epistasis is pigmentation in mice. The wild-type coat color, agouti (AA) is dominant to solid-colored fur (aa). However, a separate gene C, when present as the recessive homozygote (cc), negates any expression of pigment from the A gene and results in an albino mouse (Figure 8.18). Therefore, the genotypes *AAcc*, *Aacc*, and *aacc* all produce the same albino phenotype. A cross between heterozygotes for both genes (*AaCc* x *AaCc*) would generate offspring with a phenotypic ratio of 9 agouti:3 black:4 albino (Figure 8.18). In this case, the *C* gene is epistatic to the *A* gene.

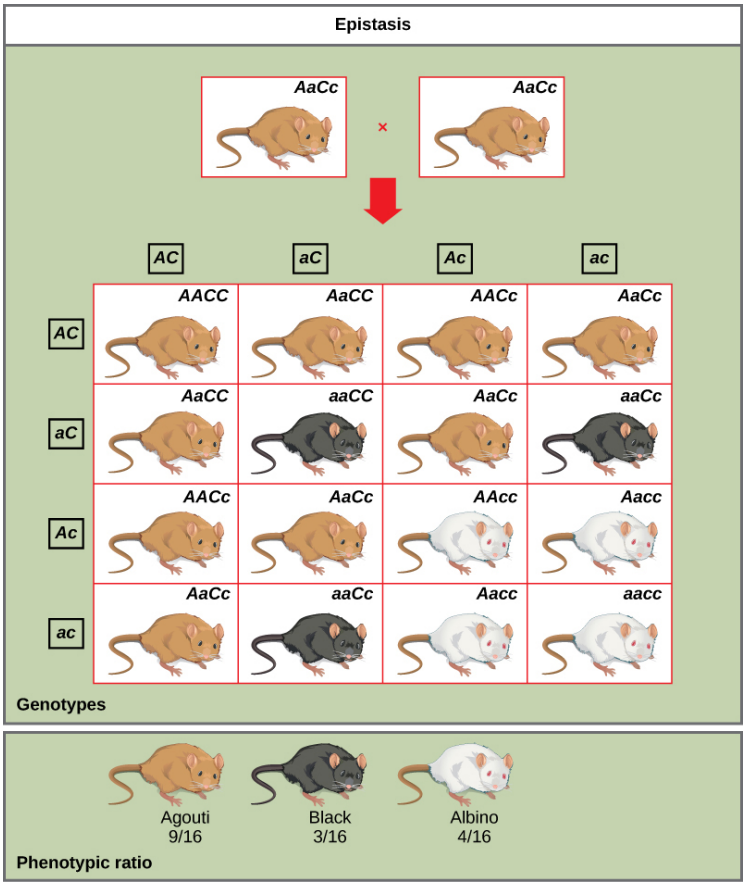


FIGURE 8.18 In this example of epistasis, one gene (*C*) masks the expression of another (*A*) for coat color. When the *C* allele is present, coat color is expressed; when it is absent (*cc*), no coat color is expressed. Coat color depends on the *A* gene, which shows dominance, with the recessive homozygote showing a different phenotype than the heterozygote or dominant homozygote.

Key Terms

allele one of two or more variants of a gene that determines a particular trait for a characteristic

codominance in a heterozygote, complete and simultaneous expression of both alleles for the same characteristic

continuous variation a variation in a characteristic in which individuals show a range of traits with small differences between them

dihybrid the result of a cross between two true-breeding parents that express different traits for two characteristics

discontinuous variation a variation in a characteristic in which individuals show two, or a few, traits with large differences between them

dominant describes a trait that masks the expression of another trait when both versions of the gene are present in an individual

epistasis an interaction between genes such that one gene masks or interferes with the expression of another

F_1 the first filial generation in a cross; the offspring of the parental generation

F_2 the second filial generation produced when F_1 individuals are self-crossed or fertilized with each other

genotype the underlying genetic makeup, consisting of both physically visible and non-expressed alleles, of an organism

hemizygous the presence of only one allele for a characteristic, as in X-linkage; hemizygosity makes descriptions of dominance and recessiveness irrelevant

heterozygous having two different alleles for a given gene on the homologous chromosomes

homozygous having two identical alleles for a given gene on the homologous chromosomes

hybridization the process of mating two individuals that differ, with the goal of achieving a certain characteristic in their offspring

incomplete dominance in a heterozygote, expression of two contrasting alleles such that the individual displays an intermediate phenotype

law of dominance in a heterozygote, one trait will conceal the presence of another trait for the same characteristic

law of independent assortment genes do not influence each other with regard to sorting of alleles into gametes; every possible combination of alleles is equally likely to occur

law of segregation paired unit factors (i.e., genes) segregate equally into gametes such that offspring have an equal likelihood of inheriting any combination of factors

linkage a phenomenon in which alleles that are located in close proximity to each other on the same chromosome are more likely to be inherited together

model system a species or biological system used to study a specific biological phenomenon to gain understanding that will be applied to other species

monohybrid the result of a cross between two true-breeding parents that express different traits for only one characteristic

parental generation (P) the first generation in a cross

phenotype the observable traits expressed by an organism

Punnett square a visual representation of a cross between two individuals in which the gametes of each individual are denoted along the top and side of a grid, respectively, and the possible zygotic genotypes are recombined at each box in the grid

recessive describes a trait whose expression is masked by another trait when the alleles for both traits are present in an individual

reciprocal cross a paired cross in which the respective traits of the male and female in one cross become the respective traits of the female and male in the other cross

recombination the process during meiosis in which homologous chromosomes exchange linear segments of genetic material, thereby dramatically increasing genetic variation in the offspring and separating linked genes

test cross a cross between a dominant expressing individual with an unknown genotype and a homozygous recessive individual; the offspring phenotypes indicate whether the unknown parent is heterozygous or homozygous for the dominant trait

trait a variation in an inherited characteristic

wild type the most commonly occurring genotype or phenotype for a given characteristic found in a population

X-linked a gene present on the X chromosome, but not the Y chromosome

Chapter Summary

8.1 Mendel's Experiments

Working with garden pea plants, Mendel found that crosses between parents that differed for one trait produced F_1 offspring that all expressed one parent's traits. The traits that were visible in the F_1 generation

are referred to as dominant, and traits that disappear in the F_1 generation are described as recessive. When the F_1 plants in Mendel's experiment were self-crossed, the F_2 offspring exhibited the dominant trait or the recessive trait in a 3:1 ratio, confirming that the recessive trait had been transmitted faithfully from the original P parent. Reciprocal crosses generated identical F_1 and F_2 offspring ratios. By examining sample sizes, Mendel showed that traits were inherited as independent events.

8.2 Laws of Inheritance

When true-breeding, or homozygous, individuals that differ for a certain trait are crossed, all of the offspring will be heterozygous for that trait. If the traits are inherited as dominant and recessive, the F_1 offspring will all exhibit the same phenotype as the parent homozygous for the dominant trait. If these heterozygous offspring are self-crossed, the resulting F_2 offspring will be equally likely to inherit gametes carrying the dominant or recessive trait, giving rise to offspring of which one quarter are homozygous dominant, half are heterozygous, and one quarter are homozygous recessive. Because homozygous dominant and heterozygous individuals are phenotypically identical, the observed traits in the F_2 offspring will exhibit a ratio of three dominant to one recessive.

Mendel postulated that genes (characteristics) are inherited as pairs of alleles (traits) that behave in a dominant and recessive pattern. Alleles segregate into gametes such that each gamete is equally likely to receive either one of the two alleles present in a diploid individual. In addition, genes are assorted into gametes independently of one another. That is, in general, alleles are not more likely to segregate into a gamete with a particular allele of another gene.

8.3 Extensions of the Laws of Inheritance

Alleles do not always behave in dominant and recessive patterns. Incomplete dominance describes situations in which the heterozygote exhibits a phenotype that is intermediate between the homozygous phenotypes. Codominance describes the simultaneous expression of both of the alleles in the heterozygote. Although diploid organisms can only have two alleles for any given gene, it is common for more than two alleles for a gene to exist in a population. In humans, as in many animals and some plants, females have two X chromosomes and males have one X and one Y chromosome. Genes that are present on the X but not the Y chromosome are said to be X-linked, such that males only inherit one allele for the gene, and females inherit two.

According to Mendel's law of independent assortment, genes sort independently of each other into gametes during meiosis. This occurs because chromosomes, on which the genes reside, assort independently during meiosis and crossovers cause most genes on the same chromosomes to also behave independently. When genes are located in close proximity on the same chromosome, their alleles tend to be inherited together. This results in offspring ratios that violate Mendel's law of independent assortment. However, recombination serves to exchange genetic material on homologous chromosomes such that maternal and paternal alleles may be recombined on the same chromosome. This is why alleles on a given chromosome are not always inherited together. Recombination is a random event occurring anywhere on a chromosome. Therefore, genes that are far apart on the same chromosome are likely to still assort independently because of recombination events that occurred in the intervening chromosomal space.

Whether or not they are sorting independently, genes may interact at the level of gene products, such that the expression of an allele for one gene masks or modifies the expression of an allele for a different gene. This is called epistasis.

Visual Connection Questions

1. Figure 8.9 In pea plants, round peas (*R*) are dominant to wrinkled peas (*r*). You do a test cross between a pea plant with wrinkled peas (genotype *rr*) and a plant of unknown genotype that has round peas. You end up with three plants, all which have round peas. From this data, can you tell if the parent plant is homozygous dominant or heterozygous?

2. Figure 8.10 In pea plants, round seed shape (*R*) is dominant to wrinkled seed shape (*r*) and yellow peas (*Y*) are dominant to green peas (*y*). What are the possible genotypes and phenotypes for a cross between *RrYY* and *rrYy* pea plants? How many squares do you need to do a Punnett square analysis of this cross?

3. Figure 8.16 What ratio of offspring would result from a cross between a white-eyed male and a female that is heterozygous for red eye color?

Review Questions

4. Imagine that you are performing a cross involving seed color in garden pea plants. What traits would you expect to observe in the F_1 offspring if you cross true-breeding parents with green seeds and yellow seeds? Yellow seed color is dominant over green.
   a. only yellow-green seeds
   b. only yellow seeds
   c. 1:1 yellow seeds:green seeds
   d. 1:3 green seeds:yellow seeds

5. Imagine that you are performing a cross involving seed texture in garden pea plants. You cross true-breeding round and wrinkled parents to obtain F_1 offspring. Which of the following experimental results in terms of numbers of plants are closest to what you expect in the F_2 progeny?
   a. 810 round seeds
   b. 810 wrinkled seeds
   c. 405:395 round seeds:wrinkled seeds
   d. 610:190 round seeds:wrinkled seeds

6. The observable traits expressed by an organism are described as its _____.
   a. phenotype
   b. genotype
   c. alleles
   d. zygote

7. A recessive trait will be observed in individuals that are _____ for that trait.
   a. heterozygous
   b. homozygous or heterozygous
   c. homozygous
   d. diploid

8. What are the types of gametes that can be produced by an individual with the genotype *AaBb*?
   a. *Aa*, *Bb*
   b. *AA*, *aa*, *BB*, *bb*
   c. *AB*, *Ab*, *aB*, *ab*
   d. *AB*, *ab*

9. What is the reason for doing a test cross?
   a. to identify heterozygous individuals with the dominant phenotype
   b. to determine which allele is dominant and which is recessive
   c. to identify homozygous recessive individuals in the F_2
   d. to determine if two genes assort independently

10. If black and white true-breeding mice are mated and the result is all gray offspring, what inheritance pattern would this be indicative of?
   a. dominance
   b. codominance
   c. multiple alleles
   d. incomplete dominance

11. The ABO blood groups in humans are expressed as the I^A, I^B, and i alleles. The I^A allele encodes the A blood group antigen, I^B encodes B, and i encodes O. Both A and B are dominant to O. If a heterozygous blood type A parent (I^Ai) and a heterozygous blood type B parent (I^Bi) mate, one quarter of their offspring are expected to have the AB blood type (I^AI^B) in which both antigens are expressed equally. Therefore, ABO blood groups are an example of:
   a. multiple alleles and incomplete dominance
   b. codominance and incomplete dominance
   c. incomplete dominance only
   d. multiple alleles and codominance

12. In a cross between a homozygous red-eyed female fruit fly and a white-eyed male fruit fly, what is the expected outcome?
   a. all white-eyed male offspring
   b. all white-eyed female offspring
   c. all red-eyed offspring
   d. half white-eyed make offspring

13. When a population has a gene with four alleles circulating, how many possible genotypes are there?
   a. 3
   b. 6
   c. 10
   d. 16

Critical Thinking Questions

14. Describe one of the reasons that made the garden pea an excellent choice of model system for studying inheritance.

15. Use a Punnett square to predict the offspring in a cross between a dwarf pea plant (homozygous recessive) and a tall pea plant (heterozygous). What is the phenotypic ratio of the offspring?

16. Use a Punnett square to predict the offspring in a cross between a tall pea plant (heterozygous) and a tall pea plant (heterozygous). What is the genotypic ratio of the offspring?

17. Can a male be a carrier of red-green color blindness?

18. Could an individual with blood type O (genotype *ii*) be a legitimate child of parents in which one parent had blood type A and the other parent had blood type B?

CHAPTER 9
Molecular Biology

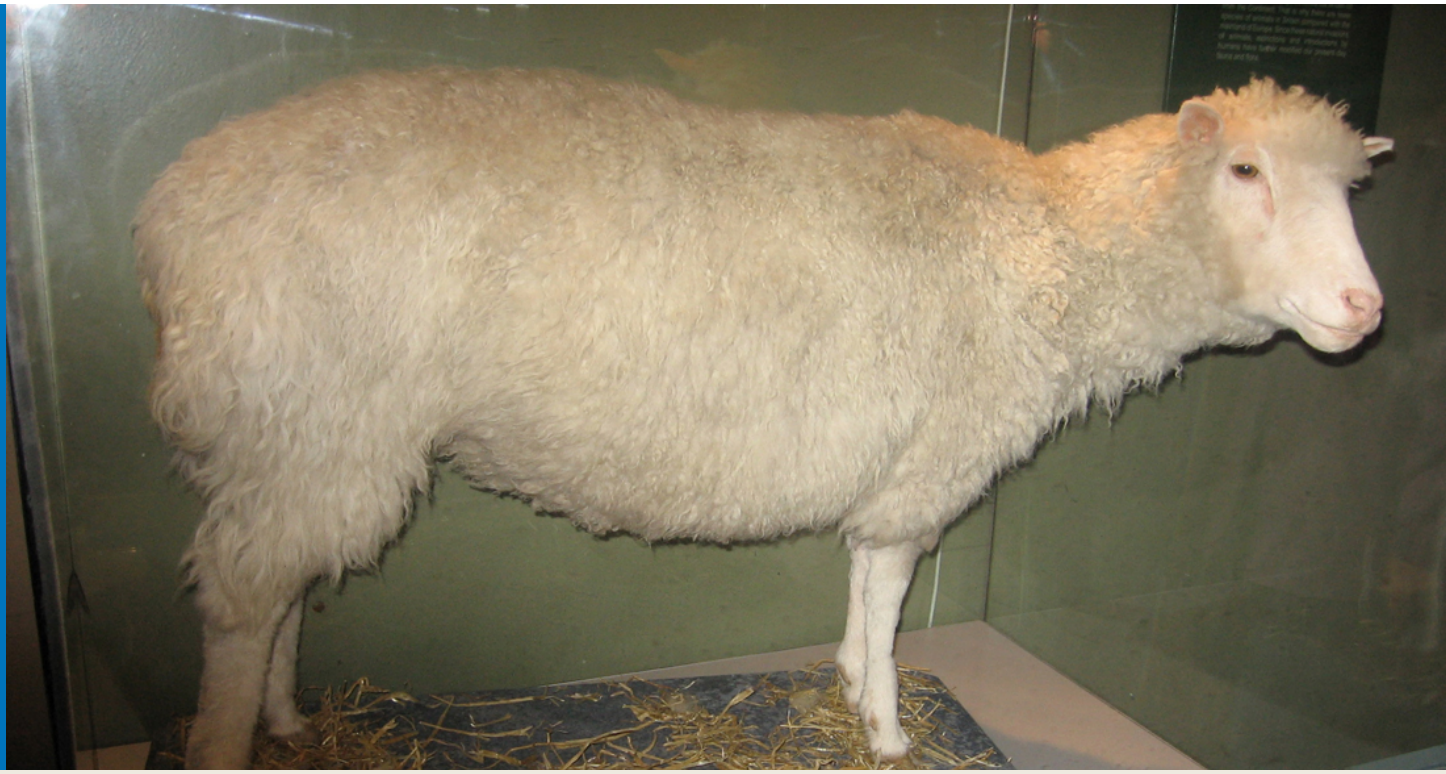


FIGURE 9.1 Dolly the sheep was the first cloned mammal.

CHAPTER OUTLINE

INTRODUCTION The three letters "DNA" have now become associated with crime solving, paternity testing, human identification, and genetic testing. DNA can be retrieved from hair, blood, or saliva. With the exception of identical twins, each person's DNA is unique and it is possible to detect differences between human beings on the basis of their unique DNA sequence.

DNA analysis has many practical applications beyond forensics and paternity testing. DNA testing is used for tracing genealogy and identifying pathogens. In the medical field, DNA is used in diagnostics, new vaccine development, and cancer therapy. It is now possible to determine predisposition to many diseases by analyzing genes.

DNA is the genetic material passed from parent to offspring for all life on Earth. The technology of molecular genetics developed in the last half century has enabled us to see deep into the history of life to deduce the relationships between living things in ways never thought possible. It also allows us to understand the workings of evolution in populations of organisms. Over a thousand species have had their entire genome sequenced, and there have been thousands of individual human genome sequences completed. These sequences will allow us to understand human disease and the relationship of humans to the rest of the tree of life. Finally, molecular genetics

techniques have revolutionized plant and animal breeding for human agricultural needs. All of these advances in biotechnology depended on basic research leading to the discovery of the structure of DNA in 1953, and the research since then that has uncovered the details of DNA replication and the complex process leading to the expression of DNA in the form of proteins in the cell.

9.1 The Structure of DNA

LEARNING OBJECTIVES

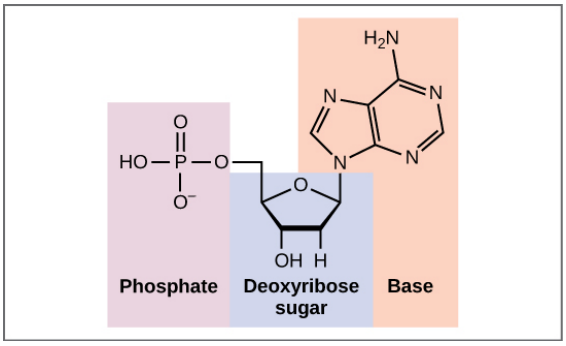
By the end of this section, you will be able to:

- Describe the structure of DNA
- Describe how eukaryotic and prokaryotic DNA is arranged in the cell

In the 1950s, Francis Crick and James Watson worked together at the University of Cambridge, England, to determine the structure of DNA. Other scientists, such as Linus Pauling and Maurice Wilkins, were also actively exploring this field. Pauling had discovered the secondary structure of proteins using X-ray crystallography. X-ray crystallography is a method for investigating molecular structure by observing the patterns formed by X-rays shot through a crystal of the substance. The patterns give important information about the structure of the molecule of interest. In Wilkins' lab, researcher Rosalind Franklin was using X-ray crystallography to understand the structure of DNA. Watson and Crick were able to piece together the puzzle of the DNA molecule using Franklin's data (Figure 9.2). Watson and Crick also had key pieces of information available from other researchers such as Chargaff's rules. Chargaff had shown that of the four kinds of monomers (nucleotides) present in a DNA molecule, two types were always present in equal amounts and the remaining two types were also always present in equal amounts. This meant they were always paired in some way. In 1962, James Watson, Francis Crick, and Maurice Wilkins were awarded the Nobel Prize in Medicine for their work in determining the structure of DNA.

FIGURE 9.2 Scientist Rosalind Franklin discovered the X-ray diffraction pattern of DNA, which helped to elucidate its double helix structure. (credit: modification of work by NIH)

Now let's consider the structure of the two types of nucleic acids, deoxyribonucleic acid (DNA) and ribonucleic acid (RNA). The building blocks of DNA are nucleotides, which are made up of three parts: a **deoxyribose** (5-carbon sugar), a **phosphate group**, and a **nitrogenous base** (Figure 9.3). There are four types of nitrogenous bases in DNA. Adenine (A) and guanine (G) are double-ringed purines, and cytosine (C) and thymine (T) are smaller, single-ringed pyrimidines. The nucleotide is named according to the nitrogenous base it contains.



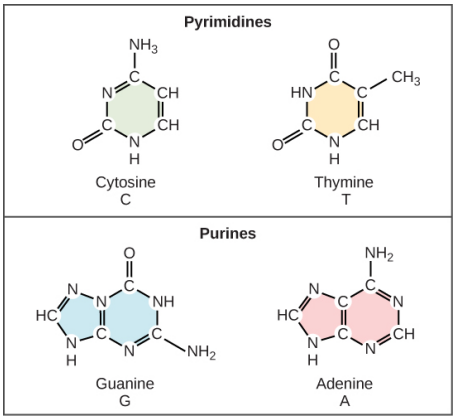


FIGURE 9.3 (a) Each DNA nucleotide is made up of a sugar, a phosphate group, and a base. (b) Cytosine and thymine are pyrimidines. Guanine and adenine are purines.

The phosphate group of one nucleotide bonds covalently with the sugar molecule of the next nucleotide, and so on, forming a long polymer of nucleotide monomers. The sugar–phosphate groups line up in a "backbone" for each single strand of DNA, and the nucleotide bases stick out from this backbone. The carbon atoms of the five-carbon sugar are numbered clockwise from the oxygen as 1', 2', 3', 4', and 5' (1' is read as "one prime"). The phosphate group is attached to the 5' carbon of one nucleotide and the 3' carbon of the next nucleotide. In its natural state, each DNA molecule is actually composed of two single strands held together along their length with hydrogen bonds between the bases.

Watson and Crick proposed that the DNA is made up of two strands that are twisted around each other to form a right-handed helix, called a **double helix**. Base-pairing takes place between a purine and pyrimidine: namely, A pairs with T, and G pairs with C. In other words, adenine and thymine are complementary base pairs, and cytosine and guanine are also complementary base pairs. This is the basis for Chargaff's rule; because of their complementarity, there is as much adenine as thymine in a DNA molecule and as much guanine as cytosine. Adenine and thymine are connected by two hydrogen bonds, and cytosine and guanine are connected by three hydrogen bonds. The two strands are anti-parallel in nature; that is, one strand will have the 3' carbon of the sugar in the "upward" position, whereas the other strand will have the 5' carbon in the upward position. The diameter of the DNA double helix is uniform throughout because a purine (two rings) always pairs with a pyrimidine (one ring) and their combined lengths are always equal. (Figure 9.4).

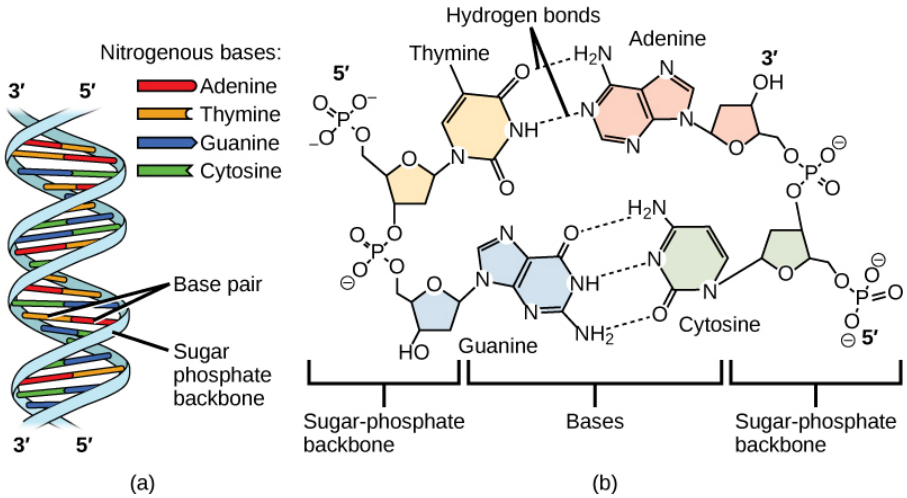


FIGURE 9.4 DNA (a) forms a double stranded helix, and (b) adenine pairs with thymine and cytosine pairs with guanine. (credit a: modification of work by Jerome Walker, Dennis Myts)

The Structure of RNA

There is a second nucleic acid in all cells called ribonucleic acid, or RNA. Like DNA, RNA is a polymer of nucleotides. Each of the nucleotides in RNA is made up of a nitrogenous base, a five-carbon sugar, and a phosphate group. In the case of RNA, the five-carbon sugar is ribose, not deoxyribose. Ribose has a hydroxyl group at the 2' carbon, unlike

deoxyribose, which has only a hydrogen atom (Figure 9.5).

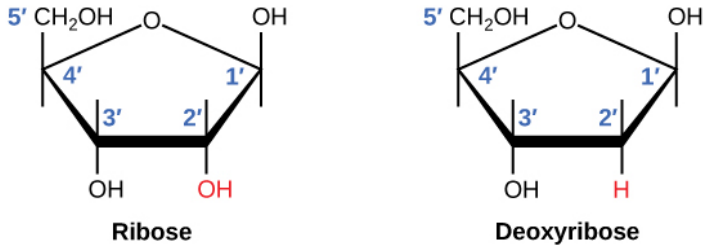


FIGURE 9.5 The difference between the ribose found in RNA and the deoxyribose found in DNA is that ribose has a hydroxyl group at the 2' carbon.

RNA nucleotides contain the nitrogenous bases adenine, cytosine, and guanine. However, they do not contain thymine, which is instead replaced by uracil, symbolized by a "U." RNA exists as a single-stranded molecule rather than a double-stranded helix. Molecular biologists have named several kinds of RNA on the basis of their function. These include messenger RNA (mRNA), transfer RNA (tRNA), and ribosomal RNA (rRNA)—molecules that are involved in the production of proteins from the DNA code.

How DNA Is Arranged in the Cell

DNA is a working molecule; it must be replicated when a cell is ready to divide, and it must be "read" to produce the molecules, such as proteins, to carry out the functions of the cell. For this reason, the DNA is protected and packaged in very specific ways. In addition, DNA molecules can be very long. Stretched end-to-end, the DNA molecules in a single human cell would come to a length of about 2 meters. Thus, the DNA for a cell must be packaged in a very ordered way to fit and function within a structure (the cell) that is not visible to the naked eye. The chromosomes of prokaryotes are much simpler than those of eukaryotes in many of their features (Figure 9.6). Most prokaryotes contain a single, circular chromosome that is found in an area in the cytoplasm called the nucleoid.

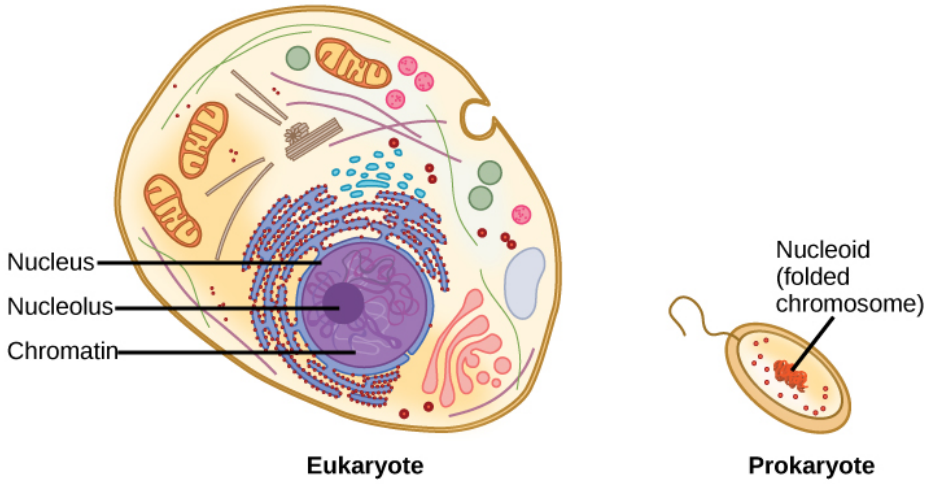


FIGURE 9.6 A eukaryote contains a well-defined nucleus, whereas in prokaryotes, the chromosome lies in the cytoplasm in an area called the nucleoid.

The size of the genome in one of the most well-studied prokaryotes, *Escherichia coli,* is 4.6 million base pairs, which would extend a distance of about 1.6 mm if stretched out. So how does this fit inside a small bacterial cell? The DNA is twisted beyond the double helix in what is known as supercoiling. Some proteins are known to be involved in the supercoiling; other proteins and enzymes help in maintaining the supercoiled structure.

Eukaryotes, whose chromosomes each consist of a linear DNA molecule, employ a different type of packing strategy to fit their DNA inside the nucleus (Figure 9.7). Before the structure of DNA was even uncovered, Marie Maynard Daly and Arthur E. Mirsky conducted extensive research in the 1940s and 1950s to understand the molecules and structures in involved. At the most basic level, DNA is wrapped around proteins known as histones to form structures called nucleosomes. The DNA is wrapped tightly around the histone core. This nucleosome is linked to the next one by a short strand of DNA that is free of histones. This is also known as the "beads on a string"

structure; the nucleosomes are the “beads” and the short lengths of DNA between them are the “string.” The nucleosomes, with their DNA coiled around them, stack compactly onto each other to form a 30-nm–wide fiber. This fiber is further coiled into a thicker and more compact structure. At the metaphase stage of mitosis, when the chromosomes are lined up in the center of the cell, the chromosomes are at their most compacted. They are approximately 700 nm in width, and are found in association with scaffold proteins.

In interphase, the phase of the cell cycle between mitoses at which the chromosomes are decondensed, eukaryotic chromosomes have two distinct regions that can be distinguished by staining. There is a tightly packaged region that stains darkly, and a less dense region. The darkly staining regions usually contain genes that are not active, and are found in the regions of the centromere and telomeres. The lightly staining regions usually contain genes that are active, with DNA packaged around nucleosomes but not further compacted.

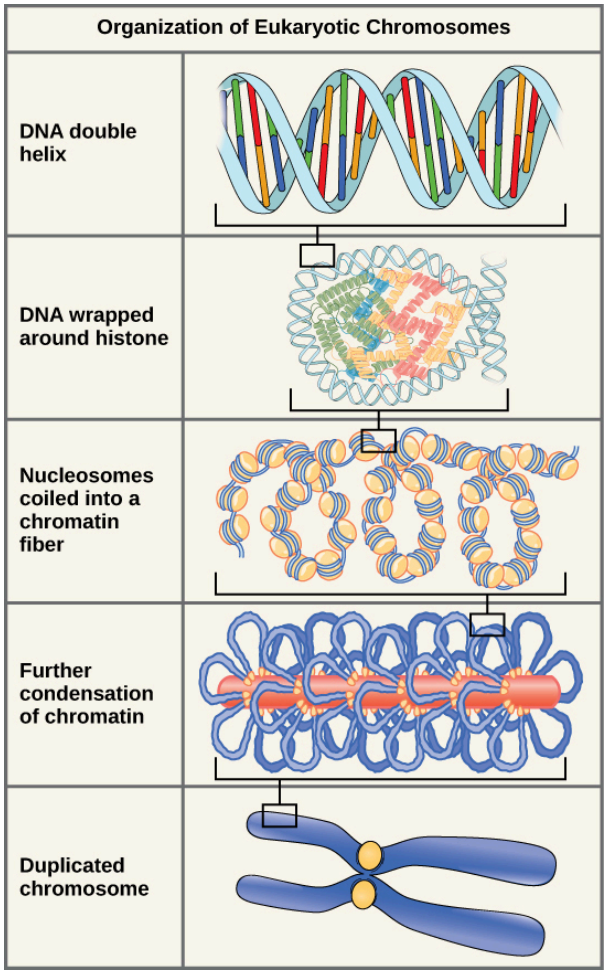


FIGURE 9.7 These figures illustrate the compaction of the eukaryotic chromosome.

LINK TO LEARNING

Watch this animation (http://openstax.org/l/DNA_packaging) of DNA packaging.

9.2 DNA Replication

LEARNING OBJECTIVES

By the end of this section, you will be able to:

- Explain the process of DNA replication
- Explain the importance of telomerase to DNA replication
- Describe mechanisms of DNA repair

When a cell divides, it is important that each daughter cell receives an identical copy of the DNA. This is accomplished by the process of DNA replication. The replication of DNA occurs during the synthesis phase, or S phase, of the cell cycle, before the cell enters mitosis or meiosis.

The elucidation of the structure of the double helix provided a hint as to how DNA is copied. Recall that adenine nucleotides pair with thymine nucleotides, and cytosine with guanine. This means that the two strands are complementary to each other. For example, a strand of DNA with a nucleotide sequence of AGTCATGA will have a complementary strand with the sequence TCAGTACT (Figure 9.8).

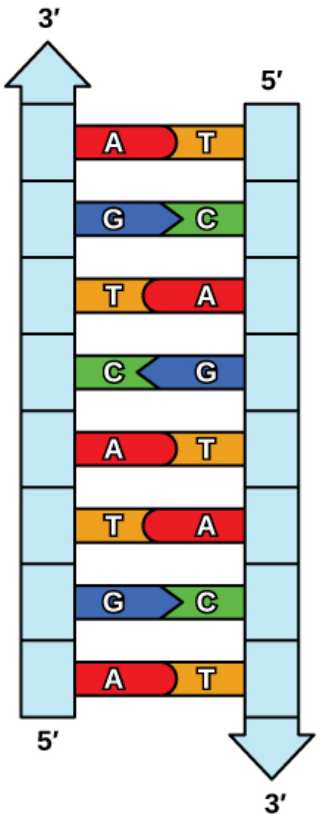


FIGURE 9.8 The two strands of DNA are complementary, meaning the sequence of bases in one strand can be used to create the correct sequence of bases in the other strand.

Because of the complementarity of the two strands, having one strand means that it is possible to recreate the other strand. This model for replication suggests that the two strands of the double helix separate during replication, and each strand serves as a template from which the new complementary strand is copied (Figure 9.9).

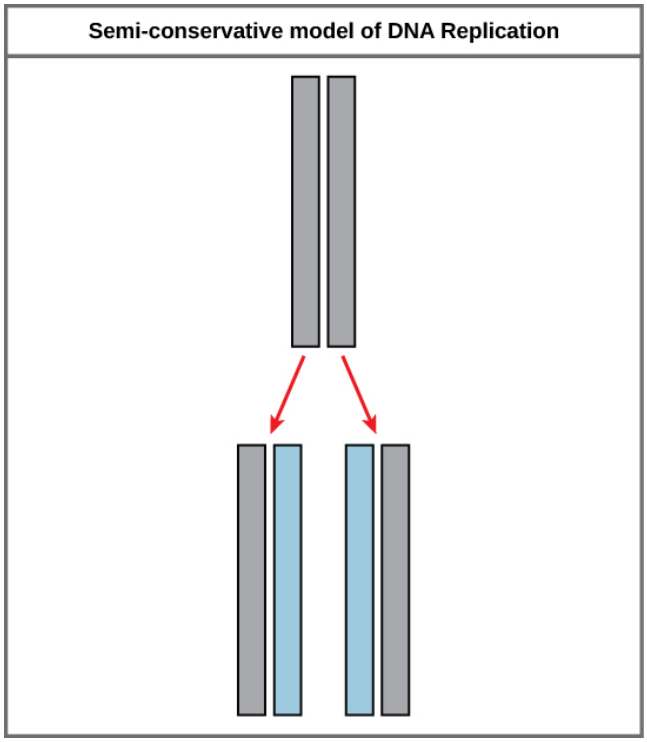


FIGURE 9.9 The semiconservative model of DNA replication is shown. Gray indicates the original DNA strands, and blue indicates newly synthesized DNA.

During DNA replication, each of the two strands that make up the double helix serves as a template from which new strands are copied. The new strand will be complementary to the parental or "old" strand. Each new double strand consists of one parental strand and one new daughter strand. This is known as **semiconservative replication**. When two DNA copies are formed, they have an identical sequence of nucleotide bases and are divided equally into two daughter cells.

DNA Replication in Eukaryotes

Because eukaryotic genomes are very complex, DNA replication is a very complicated process that involves several enzymes and other proteins. It occurs in three main stages: initiation, elongation, and termination.

Recall that eukaryotic DNA is bound to proteins known as histones to form structures called nucleosomes. During initiation, the DNA is made accessible to the proteins and enzymes involved in the replication process. How does the replication machinery know where on the DNA double helix to begin? It turns out that there are specific nucleotide sequences called origins of replication at which replication begins. Certain proteins bind to the origin of replication while an enzyme called **helicase** unwinds and opens up the DNA helix. As the DNA opens up, Y-shaped structures called **replication forks** are formed (Figure 9.10). Two replication forks are formed at the origin of replication, and these get extended in both directions as replication proceeds. There are multiple origins of replication on the eukaryotic chromosome, such that replication can occur simultaneously from several places in the genome.

During elongation, an enzyme called **DNA polymerase** adds DNA nucleotides to the 3' end of the template. Because DNA polymerase can only add new nucleotides at the end of a backbone, a **primer** sequence, which provides this starting point, is added with complementary RNA nucleotides. This primer is removed later, and the nucleotides are replaced with DNA nucleotides. One strand, which is complementary to the parental DNA strand, is synthesized continuously toward the replication fork so the polymerase can add nucleotides in this direction. This continuously synthesized strand is known as the **leading strand**. Because DNA polymerase can only synthesize DNA in a 5' to 3' direction, the other new strand is put together in short pieces called **Okazaki fragments**. The Okazaki fragments each require a primer made of RNA to start the synthesis. The strand with the Okazaki fragments is known as the **lagging strand**. As synthesis proceeds, an enzyme removes the RNA primer, which is then replaced with DNA nucleotides, and the gaps between fragments are sealed by an enzyme called **DNA ligase**.

The process of DNA replication can be summarized as follows:

1. DNA unwinds at the origin of replication.
2. New bases are added to the complementary parental strands. One new strand is made continuously, while the other strand is made in pieces.
3. Primers are removed, new DNA nucleotides are put in place of the primers and the backbone is sealed by DNA ligase.

VISUAL CONNECTION

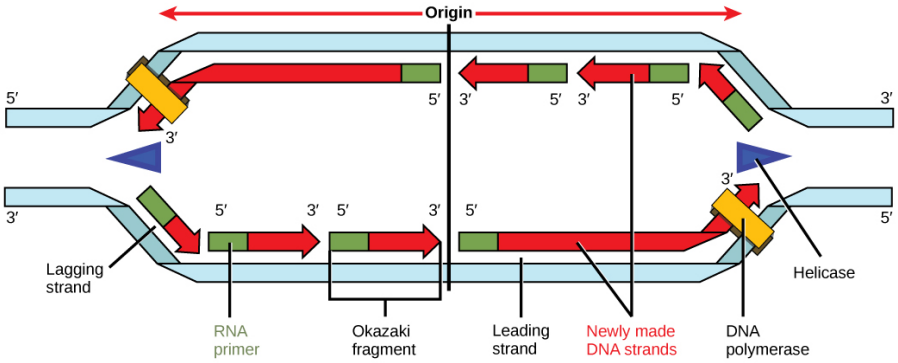


FIGURE 9.10 A replication fork is formed by the opening of the origin of replication, and helicase separates the DNA strands. An RNA primer is synthesized, and is elongated by the DNA polymerase. On the leading strand, DNA is synthesized continuously, whereas on the lagging strand, DNA is synthesized in short stretches. The DNA fragments are joined by DNA ligase (not shown).

You isolate a cell strain in which the joining together of Okazaki fragments is impaired and suspect that a mutation has occurred in an enzyme found at the replication fork. Which enzyme is most likely to be mutated?

Telomere Replication

Because eukaryotic chromosomes are linear, DNA replication comes to the end of a line in eukaryotic chromosomes. As you have learned, the DNA polymerase enzyme can add nucleotides in only one direction. In the leading strand, synthesis continues until the end of the chromosome is reached; however, on the lagging strand there is no place for a primer to be made for the DNA fragment to be copied at the end of the chromosome. This presents a problem for the cell because the ends remain unpaired, and over time these ends get progressively shorter as cells continue to divide. The ends of the linear chromosomes are known as **telomeres**, which have repetitive sequences that do not code for a particular gene. As a consequence, it is telomeres that are shortened with each round of DNA replication instead of genes. For example, in humans, a six base-pair sequence, TTAGGG, is repeated 100 to 1000 times. The discovery of the enzyme **telomerase** (Figure 9.11) helped in the understanding of how chromosome ends are maintained. The telomerase attaches to the end of the chromosome, and complementary bases to the RNA template are added on the end of the DNA strand. Once the lagging strand template is sufficiently elongated, DNA polymerase can now add nucleotides that are complementary to the ends of the chromosomes. Thus, the ends of the chromosomes are replicated.

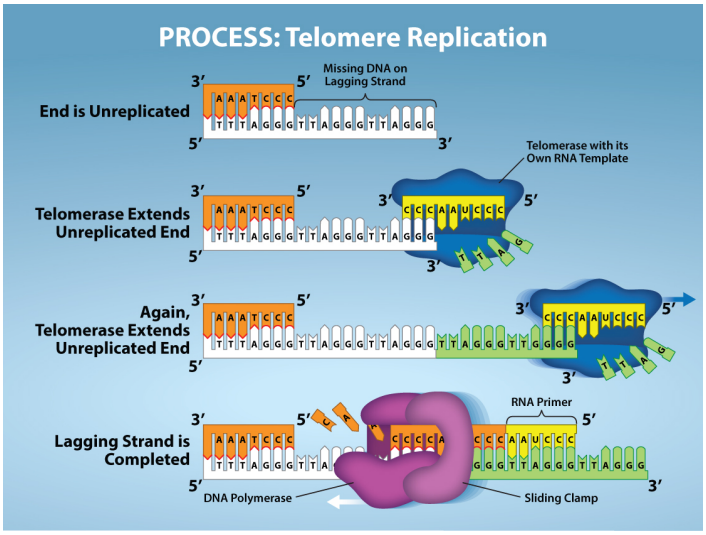


FIGURE 9.11 The ends of linear chromosomes are maintained by the action of the telomerase enzyme.

Telomerase is typically found to be active in germ cells, adult stem cells, and some cancer cells. For her discovery of telomerase and its action, Elizabeth Blackburn (Figure 9.12) received the Nobel Prize for Medicine and Physiology in 2009. Later research using HeLa cells (obtained from Henrietta Lacks) confirmed that telomerase is present in human cells. And in 2001, researchers including Diane L. Wright found that telomerase is necessary for cells in human embryos to rapidly proliferate.

FIGURE 9.12 Elizabeth Blackburn, 2009 Nobel Laureate, was the scientist who discovered how telomerase works. (credit: U.S. Embassy, Stockholm, Sweden)

Telomerase is not active in adult somatic cells. Adult somatic cells that undergo cell division continue to have their telomeres shortened. This essentially means that telomere shortening is associated with aging. In 2010, scientists found that telomerase can reverse some age-related conditions in mice, and this may have potential in regenerative medicine.[1] Telomerase-deficient mice were used in these studies; these mice have tissue atrophy, stem-cell depletion, organ system failure, and impaired tissue injury responses. Telomerase reactivation in these mice caused extension of telomeres, reduced DNA damage, reversed neurodegeneration, and improved functioning of the testes, spleen, and intestines. Thus, telomere reactivation may have potential for treating age-related diseases in humans.

DNA Replication in Prokaryotes

Recall that the prokaryotic chromosome is a circular molecule with a less extensive coiling structure than eukaryotic chromosomes. The eukaryotic chromosome is linear and highly coiled around proteins. While there are many

1 Mariella Jaskelioff, et al., "Telomerase reactivation reverses tissue degeneration in aged telomerase-deficient mice," *Nature*, 469 (2011):102–7.

similarities in the DNA replication process, these structural differences necessitate some differences in the DNA replication process in these two life forms.

DNA replication has been extremely well-studied in prokaryotes, primarily because of the small size of the genome and large number of variants available. *Escherichia coli* has 4.6 million base pairs in a single circular chromosome, and all of it gets replicated in approximately 42 minutes, starting from a single origin of replication and proceeding around the chromosome in both directions. This means that approximately 1000 nucleotides are added per second. The process is much more rapid than in eukaryotes. Table 9.1 summarizes the differences between prokaryotic and eukaryotic replications.

Differences between Prokaryotic and Eukaryotic Replications

Property	Prokaryotes	Eukaryotes
Origin of replication	Single	Multiple
Rate of replication	1000 nucleotides/s	50 to 100 nucleotides/s
Chromosome structure	circular	linear
Telomerase	Not present	Present

TABLE 9.1

LINK TO LEARNING

Click through a tutorial (http://openstax.org/l/DNA_replicatio2) on DNA replication.

DNA Repair

DNA polymerase can make mistakes while adding nucleotides. It edits the DNA by proofreading every newly added base. Incorrect bases are removed and replaced by the correct base, and then polymerization continues (Figure 9.13**a**). Most mistakes are corrected during replication, although when this does not happen, the **mismatch repair** mechanism is employed. Mismatch repair enzymes recognize the wrongly incorporated base and excise it from the DNA, replacing it with the correct base (Figure 9.13**b**). In yet another type of repair, **nucleotide excision repair**, the DNA double strand is unwound and separated, the incorrect bases are removed along with a few bases on the 5' and 3' end, and these are replaced by copying the template with the help of DNA polymerase (Figure 9.13**c**). Nucleotide excision repair is particularly important in correcting thymine dimers, which are primarily caused by ultraviolet light. In a thymine dimer, two thymine nucleotides adjacent to each other on one strand are covalently bonded to each other rather than their complementary bases. If the dimer is not removed and repaired it will lead to a mutation. Individuals with flaws in their nucleotide excision repair genes show extreme sensitivity to sunlight and develop skin cancers early in life.

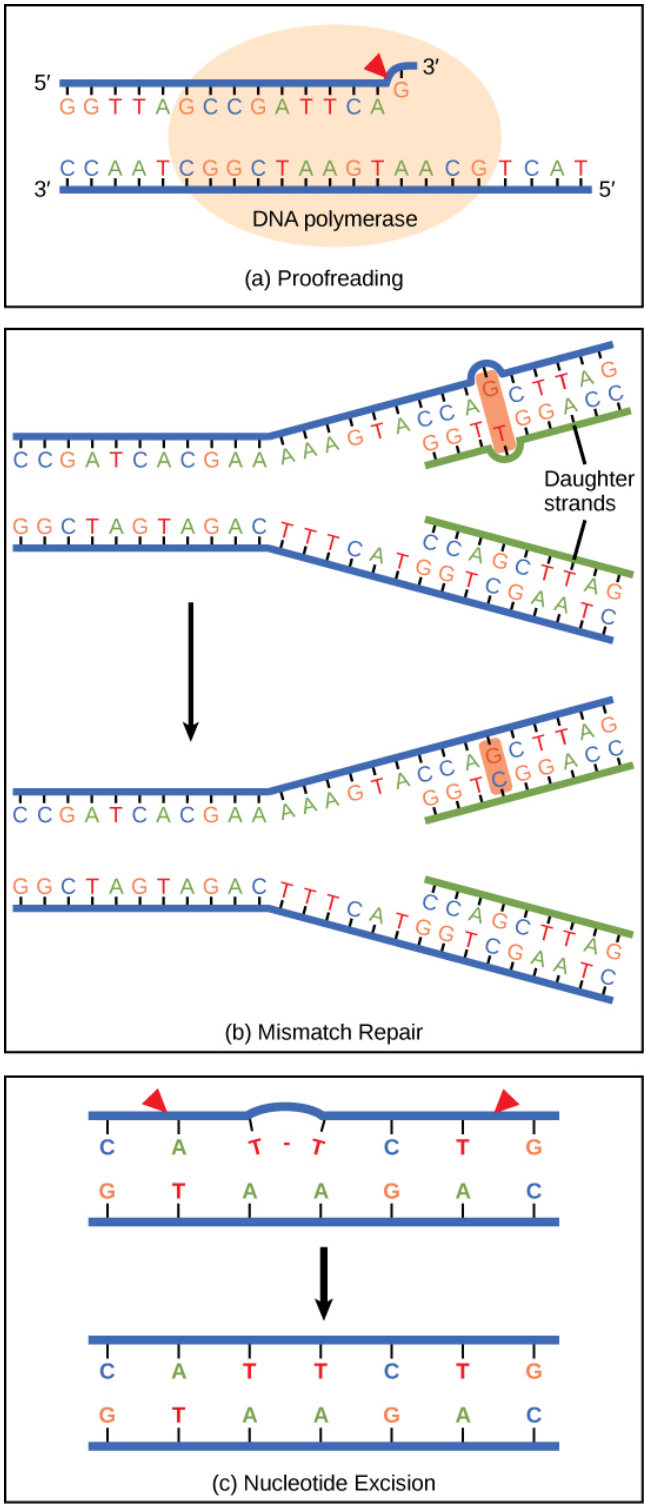


FIGURE 9.13 Proofreading by DNA polymerase (a) corrects errors during replication. In mismatch repair (b), the incorrectly added base is detected after replication. The mismatch repair proteins detect this base and remove it from the newly synthesized strand by nuclease action. The gap is now filled with the correctly paired base. Nucleotide excision (c) repairs thymine dimers. When exposed to UV, thymines lying adjacent to each other can form thymine dimers. In normal cells, they are excised and replaced.

Most mistakes are corrected; if they are not, they may result in a **mutation**—defined as a permanent change in the DNA sequence. Mutations in repair genes may lead to serious consequences like cancer.

9.3 Transcription

LEARNING OBJECTIVES

By the end of this section, you will be able to:

- Explain the central dogma
- Explain the main steps of transcription
- Describe how eukaryotic mRNA is processed

In both prokaryotes and eukaryotes, the second function of DNA (the first was replication) is to provide the information needed to construct the proteins necessary so that the cell can perform all of its functions. To do this, the DNA is "read" or transcribed into an **mRNA** molecule. The mRNA then provides the code to form a protein by a process called translation. Through the processes of transcription and translation, a protein is built with a specific sequence of amino acids that was originally encoded in the DNA. This module discusses the details of transcription.

The Central Dogma: DNA Encodes RNA; RNA Encodes Protein

The flow of genetic information in cells from DNA to mRNA to protein is described by the central dogma (Figure 9.14), which states that genes specify the sequences of mRNAs, which in turn specify the sequences of proteins.

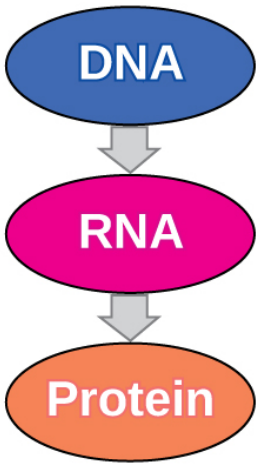


FIGURE 9.14 The central dogma states that DNA encodes RNA, which in turn encodes protein.

The copying of DNA to mRNA is relatively straightforward, with one nucleotide being added to the mRNA strand for every complementary nucleotide read in the DNA strand. The translation to protein is more complex because groups of three mRNA nucleotides correspond to one amino acid of the protein sequence. However, as we shall see in the next module, the translation to protein is still systematic, such that nucleotides 1 to 3 correspond to amino acid 1, nucleotides 4 to 6 correspond to amino acid 2, and so on.

Transcription: from DNA to mRNA

Both prokaryotes and eukaryotes perform fundamentally the same process of transcription, with the important difference of the membrane-bound nucleus in eukaryotes. With the genes bound in the nucleus, transcription occurs in the nucleus of the cell and the mRNA transcript must be transported to the cytoplasm. The prokaryotes, which include bacteria and archaea, lack membrane-bound nuclei and other organelles, and transcription occurs in the cytoplasm of the cell. In both prokaryotes and eukaryotes, transcription occurs in three main stages: initiation, elongation, and termination.

Initiation

Transcription requires the DNA double helix to partially unwind in the region of mRNA synthesis. The region of unwinding is called a **transcription bubble**. The DNA sequence onto which the proteins and enzymes involved in transcription bind to initiate the process is called a **promoter**. In most cases, promoters exist upstream of the genes they regulate. The specific sequence of a promoter is very important because it determines whether the corresponding gene is transcribed all of the time, some of the time, or hardly at all (Figure 9.15).

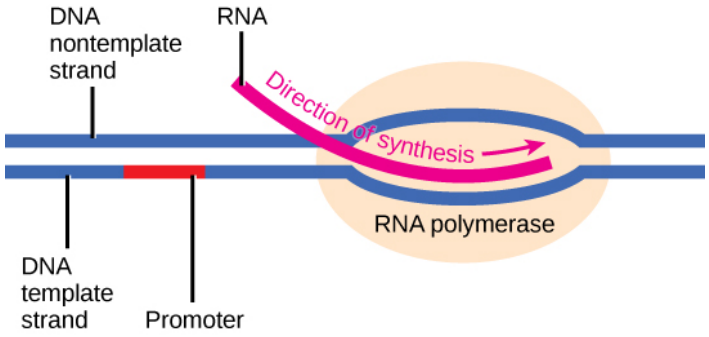


FIGURE 9.15 The initiation of transcription begins when DNA is unwound, forming a transcription bubble. Enzymes and other proteins involved in transcription bind at the promoter.

Elongation

Transcription always proceeds from one of the two DNA strands, which is called the **template strand**. The mRNA product is complementary to the template strand and is almost identical to the other DNA strand, called the **nontemplate strand**, with the exception that RNA contains a uracil (U) in place of the thymine (T) found in DNA. During elongation, an enzyme called **RNA polymerase** proceeds along the DNA template adding nucleotides by base pairing with the DNA template in a manner similar to DNA replication, with the difference that an RNA strand is being synthesized that does not remain bound to the DNA template. As elongation proceeds, the DNA is continuously unwound ahead of the core enzyme and rewound behind it (Figure 9.16).

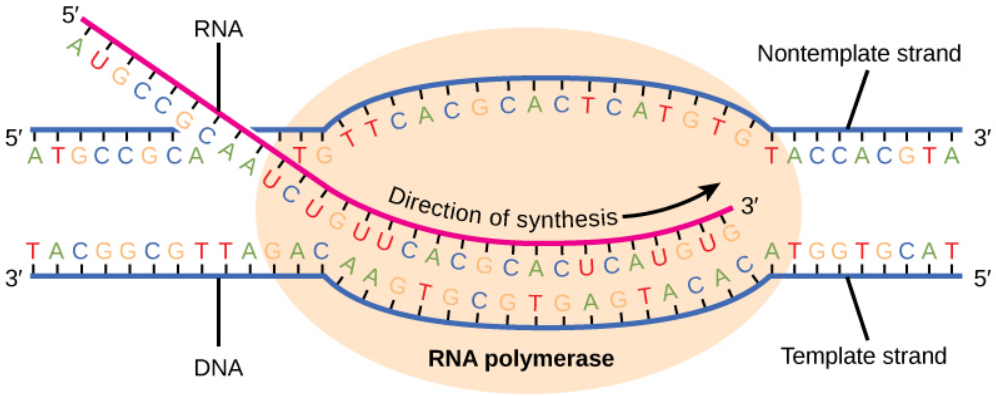


FIGURE 9.16 During elongation, RNA polymerase tracks along the DNA template, synthesizes mRNA in the 5' to 3' direction, and unwinds then rewinds the DNA as it is read.

Termination

Once a gene is transcribed, the prokaryotic polymerase needs to be instructed to dissociate from the DNA template and liberate the newly made mRNA. Depending on the gene being transcribed, there are two kinds of termination signals, but both involve repeated nucleotide sequences in the DNA template that result in RNA polymerase stalling, leaving the DNA template, and freeing the mRNA transcript.

On termination, the process of transcription is complete. In a prokaryotic cell, by the time termination occurs, the transcript would already have been used to partially synthesize numerous copies of the encoded protein because these processes can occur concurrently using multiple ribosomes (polyribosomes) (Figure 9.17). In contrast, the presence of a nucleus in eukaryotic cells precludes simultaneous transcription and translation.

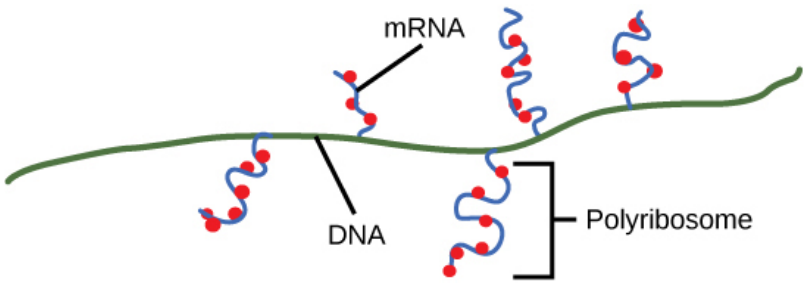


FIGURE 9.17 Multiple polymerases can transcribe a single bacterial gene while numerous ribosomes concurrently translate the mRNA transcripts into polypeptides. In this way, a specific protein can rapidly reach a high concentration in the bacterial cell.

Eukaryotic RNA Processing

The newly transcribed eukaryotic mRNAs must undergo several processing steps before they can be transferred from the nucleus to the cytoplasm and translated into a protein. The additional steps involved in eukaryotic mRNA maturation create a molecule that is much more stable than a prokaryotic mRNA. For example, eukaryotic mRNAs last for several hours, whereas the typical prokaryotic mRNA lasts no more than five seconds.

The mRNA transcript is first coated in RNA-stabilizing proteins to prevent it from degrading while it is processed and exported out of the nucleus. This occurs while the pre-mRNA still is being synthesized by adding a special nucleotide "cap" to the 5' end of the growing transcript. In addition to preventing degradation, factors involved in protein synthesis recognize the cap to help initiate translation by ribosomes.

Once elongation is complete, an enzyme then adds a string of approximately 200 adenine residues to the 3' end, called the poly-A tail. This modification further protects the pre-mRNA from degradation and signals to cellular factors that the transcript needs to be exported to the cytoplasm.

Eukaryotic genes are composed of protein-coding sequences called **exons** (*ex*-on signifies that they are *ex*pressed) and *int*ervening sequences called **introns** (*int*-ron denotes their *int*ervening role). Introns are removed from the pre-mRNA during processing. Intron sequences in mRNA do not encode functional proteins. It is essential that all of a pre-mRNA's introns be completely and precisely removed before protein synthesis so that the exons join together to code for the correct amino acids. If the process errs by even a single nucleotide, the sequence of the rejoined exons would be shifted, and the resulting protein would be nonfunctional. The process of removing introns and reconnecting exons is called **splicing** (Figure 9.18). Introns are removed and degraded while the pre-mRNA is still in the nucleus.

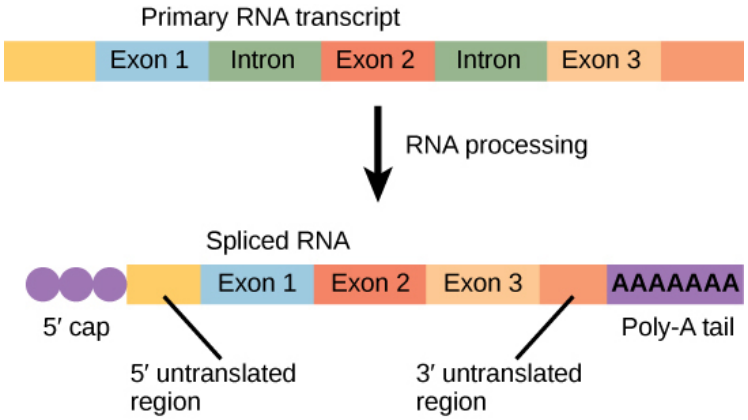


FIGURE 9.18 Eukaryotic mRNA contains introns that must be spliced out. A 5' cap and 3' tail are also added.

9.4 Translation

LEARNING OBJECTIVES

By the end of this section, you will be able to:

- Describe the different steps in protein synthesis
- Discuss the role of ribosomes in protein synthesis
- Describe the genetic code and how the nucleotide sequence determines the amino acid and the protein sequence

The synthesis of proteins is one of a cell's most energy-consuming metabolic processes. In turn, proteins account for more mass than any other component of living organisms (with the exception of water), and proteins perform a wide variety of the functions of a cell. The process of translation, or protein synthesis, involves decoding an mRNA message into a polypeptide product. Amino acids are covalently strung together in lengths ranging from approximately 50 amino acids to more than 1,000.

The Protein Synthesis Machinery

In addition to the mRNA template, many other molecules contribute to the process of translation. The composition of each component may vary across species; for instance, ribosomes may consist of different numbers of ribosomal RNAs (**rRNA**) and polypeptides depending on the organism. However, the general structures and functions of the protein synthesis machinery are comparable from bacteria to human cells. Translation requires the input of an mRNA template, ribosomes, tRNAs, and various enzymatic factors (Figure 9.19).

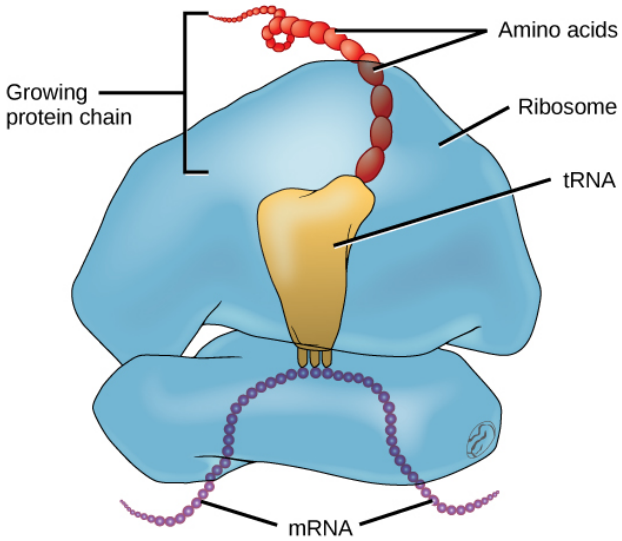


FIGURE 9.19 The protein synthesis machinery includes the large and small subunits of the ribosome, mRNA, and tRNA. (credit: modification of work by NIGMS, NIH)

In *E. coli*, there are 200,000 ribosomes present in every cell at any given time. A ribosome is a complex macromolecule composed of structural and catalytic rRNAs, and many distinct polypeptides. In eukaryotes, the nucleolus is completely specialized for the synthesis and assembly of rRNAs.

Ribosomes are located in the cytoplasm in prokaryotes and in the cytoplasm and endoplasmic reticulum of eukaryotes. Ribosomes are made up of a large and a small subunit that come together for translation. The small subunit is responsible for binding the mRNA template, whereas the large subunit sequentially binds **tRNAs**, a type of RNA molecule that brings amino acids to the growing chain of the polypeptide. Each mRNA molecule is simultaneously translated by many ribosomes, all synthesizing protein in the same direction.

Depending on the species, 40 to 60 types of tRNA exist in the cytoplasm. Serving as adaptors, specific tRNAs bind to sequences on the mRNA template and add the corresponding amino acid to the polypeptide chain. Therefore, tRNAs are the molecules that actually "translate" the language of RNA into the language of proteins. For each tRNA to function, it must have its specific amino acid bonded to it. In the process of tRNA "charging," each tRNA molecule is bonded to its correct amino acid.

The Genetic Code

To summarize what we know to this point, the cellular process of transcription generates messenger RNA (mRNA), a mobile molecular copy of one or more genes with an alphabet of A, C, G, and uracil (U). Translation of the mRNA template converts nucleotide-based genetic information into a protein product. Protein sequences consist of 20 commonly occurring amino acids; therefore, it can be said that the protein alphabet consists of 20 letters. Each amino acid is defined by a three-nucleotide sequence called the triplet **codon**. The relationship between a nucleotide codon and its corresponding amino acid is called the **genetic code**.

Given the different numbers of "letters" in the mRNA and protein "alphabets," combinations of nucleotides corresponded to single amino acids. Using a three-nucleotide code means that there are a total of 64 (4 × 4 × 4) possible combinations; therefore, a given amino acid is encoded by more than one nucleotide triplet (Figure 9.20).

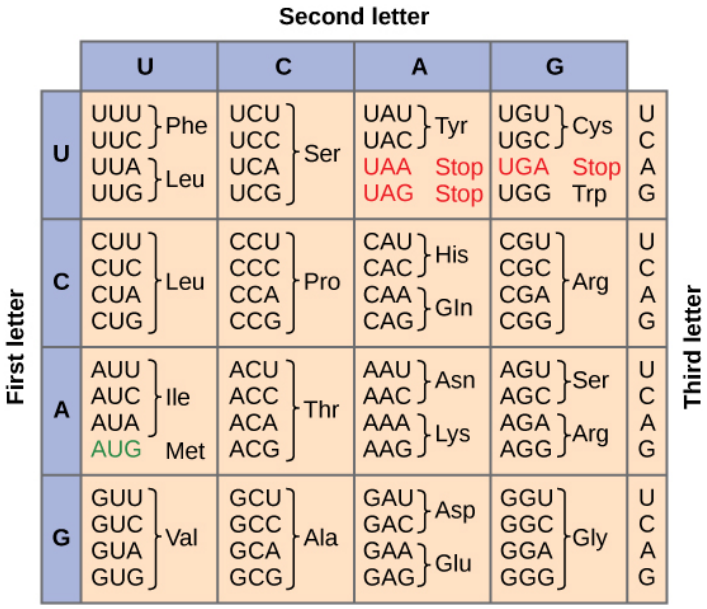


FIGURE 9.20 This figure shows the genetic code for translating each nucleotide triplet, or codon, in mRNA into an amino acid or a termination signal in a nascent protein. (credit: modification of work by NIH)

Three of the 64 codons terminate protein synthesis and release the polypeptide from the translation machinery. These triplets are called **stop codons**. Another codon, AUG, also has a special function. In addition to specifying the amino acid methionine, it also serves as the **start codon** to initiate translation. The reading frame for translation is set by the AUG start codon near the 5' end of the mRNA. The genetic code is universal. With a few exceptions, virtually all species use the same genetic code for protein synthesis, which is powerful evidence that all life on Earth shares a common origin.

The Mechanism of Protein Synthesis

Just as with mRNA synthesis, protein synthesis can be divided into three phases: initiation, elongation, and termination. The process of translation is similar in prokaryotes and eukaryotes. Here we will explore how translation occurs in *E. coli*, a representative prokaryote, and specify any differences between prokaryotic and eukaryotic translation.

Protein synthesis begins with the formation of an initiation complex. In *E. coli*, this complex involves the small ribosome subunit, the mRNA template, three initiation factors, and a special initiator tRNA. The initiator tRNA interacts with the AUG start codon, and links to a special form of the amino acid methionine that is typically removed from the polypeptide after translation is complete.

In prokaryotes and eukaryotes, the basics of polypeptide elongation are the same, so we will review elongation from the perspective of *E. coli*. The large ribosomal subunit of *E. coli* consists of three compartments: the A site binds incoming charged tRNAs (tRNAs with their attached specific amino acids). The P site binds charged tRNAs carrying amino acids that have formed bonds with the growing polypeptide chain but have not yet dissociated from their corresponding tRNA. The E site releases dissociated tRNAs so they can be recharged with free amino acids. The

ribosome shifts one codon at a time, catalyzing each process that occurs in the three sites. With each step, a charged tRNA enters the complex, the polypeptide becomes one amino acid longer, and an uncharged tRNA departs. The energy for each bond between amino acids is derived from GTP, a molecule similar to ATP (Figure 9.21). Amazingly, the *E. coli* translation apparatus takes only 0.05 seconds to add each amino acid, meaning that a 200-amino acid polypeptide could be translated in just 10 seconds.

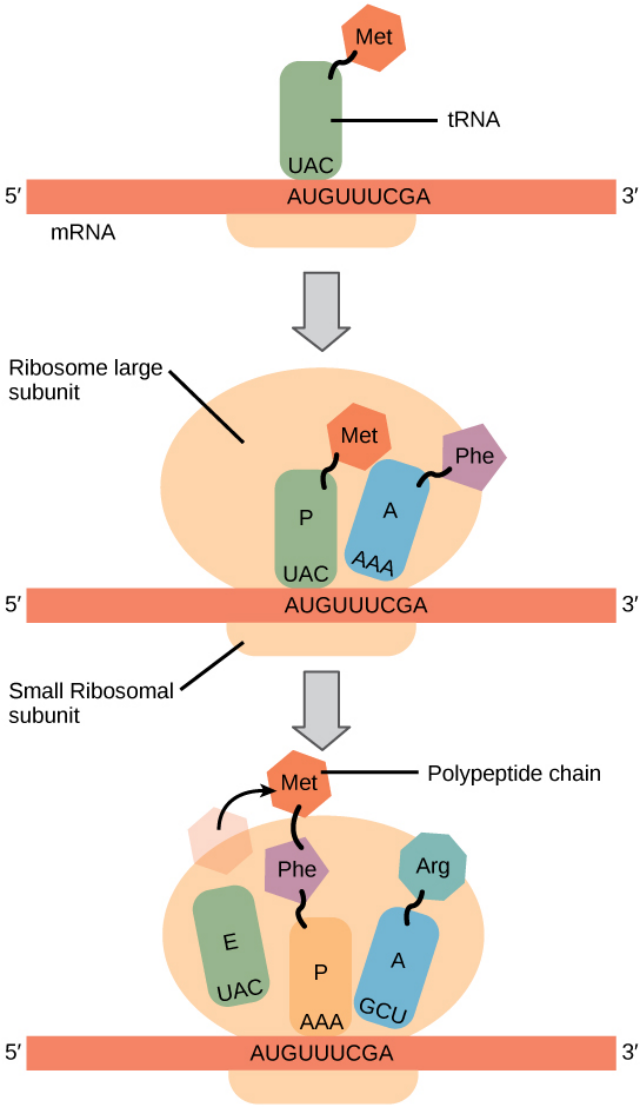


FIGURE 9.21 Translation begins when a tRNA anticodon recognizes a codon on the mRNA. The large ribosomal subunit joins the small subunit, and a second tRNA is recruited. As the mRNA moves relative to the ribosome, the polypeptide chain is formed. Entry of a release factor into the A site terminates translation and the components dissociate.

Termination of translation occurs when a stop codon (UAA, UAG, or UGA) is encountered. When the ribosome encounters the stop codon, the growing polypeptide is released and the ribosome subunits dissociate and leave the mRNA. After many ribosomes have completed translation, the mRNA is degraded so the nucleotides can be reused in another transcription reaction.

LINK TO LEARNING

Transcribe a gene and translate it to protein using complementary pairing and the genetic code at this site (http://openstax.org/l/create_protein2).

9.5 How Genes Are Regulated

LEARNING OBJECTIVES

By the end of this section, you will be able to:

- Discuss why every cell does not express all of its genes
- Describe how prokaryotic gene expression occurs at the transcriptional level
- Understand that eukaryotic gene expression occurs at the epigenetic, transcriptional, post-transcriptional, translational, and post-translational levels

For a cell to function properly, necessary proteins must be synthesized at the proper time. All organisms and cells control or regulate the transcription and translation of their DNA into protein. The process of turning on a gene to produce RNA and protein is called **gene expression**. Whether in a simple unicellular organism or in a complex multicellular organism, each cell controls when and how its genes are expressed. For this to occur, there must be a mechanism to control when a gene is expressed to make RNA and protein, how much of the protein is made, and when it is time to stop making that protein because it is no longer needed.

Cells in multicellular organisms are specialized; cells in different tissues look very different and perform different functions. For example, a muscle cell is very different from a liver cell, which is very different from a skin cell. These differences are a consequence of the expression of different sets of genes in each of these cells. All cells have certain basic functions they must perform for themselves, such as converting the energy in sugar molecules into energy in ATP. Each cell also has many genes that are not expressed, and expresses many that are not expressed by other cells, such that it can carry out its specialized functions. In addition, cells will turn on or off certain genes at different times in response to changes in the environment or at different times during the development of the organism. Unicellular organisms, both eukaryotic and prokaryotic, also turn on and off genes in response to the demands of their environment so that they can respond to special conditions.

The control of gene expression is extremely complex. Malfunctions in this process are detrimental to the cell and can lead to the development of many diseases, including cancer.

Prokaryotic versus Eukaryotic Gene Expression

To understand how gene expression is regulated, we must first understand how a gene becomes a functional protein in a cell. The process occurs in both prokaryotic and eukaryotic cells, just in slightly different fashions.

Because prokaryotic organisms lack a cell nucleus, the processes of transcription and translation occur almost simultaneously. When the protein is no longer needed, transcription stops. As a result, the primary method to control what type and how much protein is expressed in a prokaryotic cell is through the regulation of DNA transcription into RNA. All the subsequent steps happen automatically. When more protein is required, more transcription occurs. Therefore, in prokaryotic cells, the control of gene expression is almost entirely at the transcriptional level.

The first example of such control was discovered using *E. coli* in the 1950s and 1960s by French researchers and is called the *lac* operon. The *lac* operon is a stretch of DNA with three adjacent genes that code for proteins that participate in the absorption and metabolism of lactose, a food source for *E. coli*. When lactose is not present in the bacterium's environment, the *lac* genes are transcribed in small amounts. When lactose is present, the genes are transcribed and the bacterium is able to use the lactose as a food source. The operon also contains a promoter sequence to which the RNA polymerase binds to begin transcription; between the promoter and the three genes is a region called the operator. When there is no lactose present, a protein known as a repressor binds to the operator and prevents RNA polymerase from binding to the promoter, except in rare cases. Thus very little of the protein products of the three genes is made. When lactose is present, an end product of lactose metabolism binds to the repressor protein and prevents it from binding to the operator. This allows RNA polymerase to bind to the promoter and freely transcribe the three genes, allowing the organism to metabolize the lactose.

Eukaryotic cells, in contrast, have intracellular organelles and are much more complex. Recall that in eukaryotic cells, the DNA is contained inside the cell's nucleus and it is transcribed into mRNA there. The newly synthesized mRNA is then transported out of the nucleus into the cytoplasm, where ribosomes translate the mRNA into protein. The processes of transcription and translation are physically separated by the nuclear membrane; transcription occurs only within the nucleus, and translation only occurs outside the nucleus in the cytoplasm. The regulation of

gene expression can occur at all stages of the process (Figure 9.22). Regulation may occur when the DNA is uncoiled and loosened from nucleosomes to bind transcription factors (**epigenetic** level), when the RNA is transcribed (transcriptional level), when RNA is processed and exported to the cytoplasm after it is transcribed (**post-transcriptional** level), when the RNA is translated into protein (translational level), or after the protein has been made (**post-translational** level).

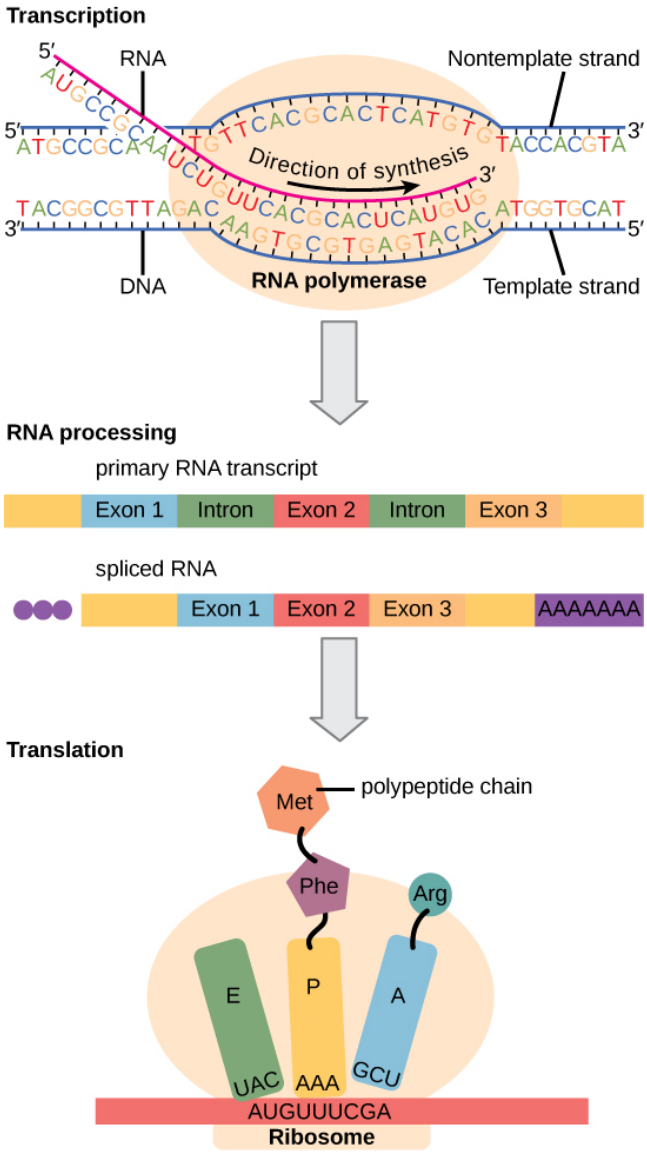


FIGURE 9.22 Eukaryotic gene expression is regulated during transcription and RNA processing, which take place in the nucleus, as well as during protein translation, which takes place in the cytoplasm. Further regulation may occur through post-translational modifications of proteins.

The differences in the regulation of gene expression between prokaryotes and eukaryotes are summarized in Table 9.2.

Differences in the Regulation of Gene Expression of Prokaryotic and Eukaryotic Organisms

Prokaryotic organisms	Eukaryotic organisms
Lack nucleus	Contain nucleus
RNA transcription and protein translation occur almost simultaneously	• RNA transcription occurs prior to protein translation, and it takes place in the nucleus. RNA translation to protein occurs in the cytoplasm. • RNA post-processing includes addition of a 5' cap, poly-A tail, and excision of introns and splicing of exons.
Gene expression is regulated primarily at the transcriptional level	Gene expression is regulated at many levels (epigenetic, transcriptional, post-transcriptional, translational, and post-translational)

TABLE 9.2

EVOLUTION CONNECTION

Alternative RNA Splicing

In the 1970s, genes were first observed that exhibited **alternative RNA splicing**. Alternative RNA splicing is a mechanism that allows different protein products to be produced from one gene when different combinations of introns (and sometimes exons) are removed from the transcript (Figure 9.23). This alternative splicing can be haphazard, but more often it is controlled and acts as a mechanism of gene regulation, with the frequency of different splicing alternatives controlled by the cell as a way to control the production of different protein products in different cells, or at different stages of development. Alternative splicing is now understood to be a common mechanism of gene regulation in eukaryotes; according to one estimate, 70% of genes in humans are expressed as multiple proteins through alternative splicing.

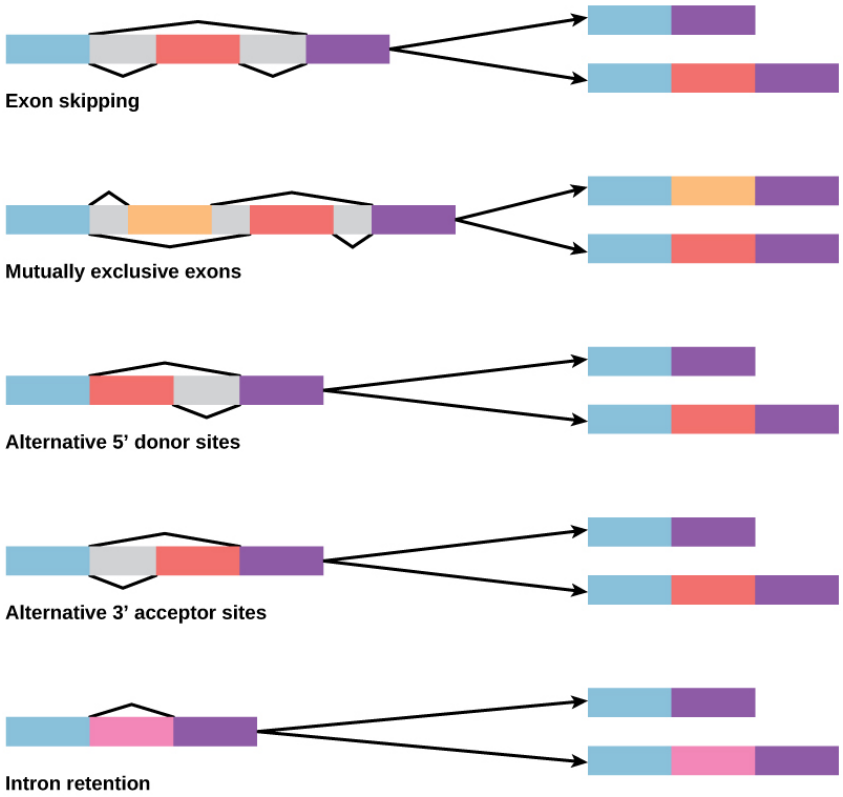


FIGURE 9.23 There are five basic modes of alternative splicing. Segments of pre-mRNA with exons shown in blue, red, orange, and pink can be spliced to produce a variety of new mature mRNA segments.

How could alternative splicing evolve? Introns have a beginning and ending recognition sequence, and it is easy to imagine the failure of the splicing mechanism to identify the end of an intron and find the end of the next intron, thus removing two introns and the intervening exon. In fact, there are mechanisms in place to prevent such exon skipping, but mutations are likely to lead to their failure. Such "mistakes" would more than likely produce a nonfunctional protein. Indeed, the cause of many genetic diseases is alternative splicing rather than mutations in a sequence. However, alternative splicing would create a protein variant without the loss of the original protein, opening up possibilities for adaptation of the new variant to new functions. Gene duplication has played an important role in the evolution of new functions in a similar way—by providing genes that may evolve without eliminating the original functional protein.

Key Terms

alternative RNA splicing a post-transcriptional gene regulation mechanism in eukaryotes in which multiple protein products are produced by a single gene through alternative splicing combinations of the RNA transcript

codon three consecutive nucleotides in mRNA that specify the addition of a specific amino acid or the release of a polypeptide chain during translation

deoxyribose a five-carbon sugar molecule with a hydrogen atom rather than a hydroxyl group in the 2' position; the sugar component of DNA nucleotides

DNA ligase the enzyme that catalyzes the joining of DNA fragments together

DNA polymerase an enzyme that synthesizes a new strand of DNA complementary to a template strand

double helix the molecular shape of DNA in which two strands of nucleotides wind around each other in a spiral shape

epigenetic describing non-genetic regulatory factors, such as changes in modifications to histone proteins and DNA that control accessibility to genes in chromosomes

exon a sequence present in protein-coding mRNA after completion of pre-mRNA splicing

gene expression processes that control whether a gene is expressed

genetic code the amino acids that correspond to three-nucleotide codons of mRNA

helicase an enzyme that helps to open up the DNA helix during DNA replication by breaking the hydrogen bonds

intron non–protein-coding intervening sequences that are spliced from mRNA during processing

lagging strand during replication of the 3' to 5' strand, the strand that is replicated in short fragments and away from the replication fork

leading strand the strand that is synthesized continuously in the 5' to 3' direction that is synthesized in the direction of the replication fork

mismatch repair a form of DNA repair in which non-complementary nucleotides are recognized, excised, and replaced with correct nucleotides

mRNA messenger RNA; a form of RNA that carries the nucleotide sequence code for a protein sequence that is translated into a polypeptide sequence

mutation a permanent variation in the nucleotide sequence of a genome

nitrogenous base a nitrogen-containing molecule that acts as a base; often referring to one of the purine or pyrimidine components of nucleic acids

nontemplate strand the strand of DNA that is not used to transcribe mRNA; this strand is identical to the mRNA except that T nucleotides in the DNA are replaced by U nucleotides in the mRNA

nucleotide excision repair a form of DNA repair in which the DNA molecule is unwound and separated in the region of the nucleotide damage, the damaged nucleotides are removed and replaced with new nucleotides using the complementary strand, and the DNA strand is resealed and allowed to rejoin its complement

Okazaki fragments the DNA fragments that are synthesized in short stretches on the lagging strand

phosphate group a molecular group consisting of a central phosphorus atom bound to four oxygen atoms

post-transcriptional control of gene expression after the RNA molecule has been created but before it is translated into protein

post-translational control of gene expression after a protein has been created

primer a short stretch of RNA nucleotides that is required to initiate replication and allow DNA polymerase to bind and begin replication

promoter a sequence on DNA to which RNA polymerase and associated factors bind and initiate transcription

replication fork the Y-shaped structure formed during the initiation of replication

RNA polymerase an enzyme that synthesizes an RNA strand from a DNA template strand

rRNA ribosomal RNA; molecules of RNA that combine to form part of the ribosome

semiconservative replication the method used to replicate DNA in which the double-stranded molecule is separated and each strand acts as a template for a new strand to be synthesized, so the resulting DNA molecules are composed of one new strand of nucleotides and one old strand of nucleotides

splicing the process of removing introns and reconnecting exons in a pre-mRNA

start codon the AUG (or, rarely GUG) on an mRNA from which translation begins; always specifies methionine

stop codon one of the three mRNA codons that specifies termination of translation

telomerase an enzyme that contains a catalytic part and an inbuilt RNA template; it functions to maintain telomeres at chromosome ends

telomere the DNA at the end of linear chromosomes

template strand the strand of DNA that specifies the

complementary mRNA molecule

transcription bubble the region of locally unwound DNA that allows for transcription of mRNA

tRNA transfer RNA; an RNA molecule that contains a specific three-nucleotide anticodon sequence to pair with the mRNA codon and also binds to a specific amino acid

Chapter Summary

9.1 The Structure of DNA

The model of the double-helix structure of DNA was proposed by Watson and Crick. The DNA molecule is a polymer of nucleotides. Each nucleotide is composed of a nitrogenous base, a five-carbon sugar (deoxyribose), and a phosphate group. There are four nitrogenous bases in DNA, two purines (adenine and guanine) and two pyrimidines (cytosine and thymine). A DNA molecule is composed of two strands. Each strand is composed of nucleotides bonded together covalently between the phosphate group of one and the deoxyribose sugar of the next. From this backbone extend the bases. The bases of one strand bond to the bases of the second strand with hydrogen bonds. Adenine always bonds with thymine, and cytosine always bonds with guanine. The bonding causes the two strands to spiral around each other in a shape called a double helix. Ribonucleic acid (RNA) is a second nucleic acid found in cells. RNA is a single-stranded polymer of nucleotides. It also differs from DNA in that it contains the sugar ribose, rather than deoxyribose, and the nucleotide uracil rather than thymine. Various RNA molecules function in the process of forming proteins from the genetic code in DNA.

Prokaryotes contain a single, double-stranded circular chromosome. Eukaryotes contain double-stranded linear DNA molecules packaged into chromosomes. The DNA helix is wrapped around proteins to form nucleosomes. The protein coils are further coiled, and during mitosis and meiosis, the chromosomes become even more greatly coiled to facilitate their movement. Chromosomes have two distinct regions which can be distinguished by staining, reflecting different degrees of packaging and determined by whether the DNA in a region is being expressed (euchromatin) or not (heterochromatin).

9.2 DNA Replication

DNA replicates by a semi-conservative method in which each of the two parental DNA strands act as a template for new DNA to be synthesized. After replication, each DNA has one parental or “old” strand, and one daughter or “new” strand.

Replication in eukaryotes starts at multiple origins of replication, while replication in prokaryotes starts from a single origin of replication. The DNA is opened with enzymes, resulting in the formation of the replication fork. Primase synthesizes an RNA primer to initiate synthesis by DNA polymerase, which can add nucleotides in only one direction. One strand is synthesized continuously in the direction of the replication fork; this is called the leading strand. The other strand is synthesized in a direction away from the replication fork, in short stretches of DNA known as Okazaki fragments. This strand is known as the lagging strand. Once replication is completed, the RNA primers are replaced by DNA nucleotides and the DNA is sealed with DNA ligase.

The ends of eukaryotic chromosomes pose a problem, as polymerase is unable to extend them without a primer. Telomerase, an enzyme with an inbuilt RNA template, extends the ends by copying the RNA template and extending one end of the chromosome. DNA polymerase can then extend the DNA using the primer. In this way, the ends of the chromosomes are protected. Cells have mechanisms for repairing DNA when it becomes damaged or errors are made in replication. These mechanisms include mismatch repair to replace nucleotides that are paired with a non-complementary base and nucleotide excision repair, which removes bases that are damaged such as thymine dimers.

9.3 Transcription

In prokaryotes, mRNA synthesis is initiated at a promoter sequence on the DNA template. Elongation synthesizes new mRNA. Termination liberates the mRNA and occurs by mechanisms that stall the RNA polymerase and cause it to fall off the DNA template. Newly transcribed eukaryotic mRNAs are modified with a cap and a poly-A tail. These structures protect the mature mRNA from degradation and help export it from the nucleus. Eukaryotic mRNAs also undergo splicing, in which introns are removed and exons are reconnected with single-nucleotide accuracy. Only finished mRNAs are exported from the nucleus to the cytoplasm.

9.4 Translation

The central dogma describes the flow of genetic information in the cell from genes to mRNA to proteins. Genes are used to make mRNA by the process of

transcription; mRNA is used to synthesize proteins by the process of translation. The genetic code is the correspondence between the three-nucleotide mRNA codon and an amino acid. The genetic code is "translated" by the tRNA molecules, which associate a specific codon with a specific amino acid. The genetic code is degenerate because 64 triplet codons in mRNA specify only 20 amino acids and three stop codons. This means that more than one codon corresponds to an amino acid. Almost every species on the planet uses the same genetic code.

The players in translation include the mRNA template, ribosomes, tRNAs, and various enzymatic factors. The small ribosomal subunit binds to the mRNA template. Translation begins at the initiating AUG on the mRNA. The formation of bonds occurs between sequential amino acids specified by the mRNA template according to the genetic code. The ribosome accepts charged tRNAs, and as it steps along the mRNA, it catalyzes bonding between the new amino acid and the end of the growing polypeptide. The entire mRNA is translated in three-nucleotide "steps" of the ribosome. When a stop codon is encountered, a release factor binds and dissociates the components and frees the new protein.

9.5 How Genes Are Regulated

While all somatic cells within an organism contain the same DNA, not all cells within that organism express the same proteins. Prokaryotic organisms express the entire DNA they encode in every cell, but not necessarily all at the same time. Proteins are expressed only when they are needed. Eukaryotic organisms express a subset of the DNA that is encoded in any given cell. In each cell type, the type and amount of protein is regulated by controlling gene expression. To express a protein, the DNA is first transcribed into RNA, which is then translated into proteins. In prokaryotic cells, these processes occur almost simultaneously. In eukaryotic cells, transcription occurs in the nucleus and is separate from the translation that occurs in the cytoplasm. Gene expression in prokaryotes is regulated only at the transcriptional level, whereas in eukaryotic cells, gene expression is regulated at the epigenetic, transcriptional, post-transcriptional, translational, and post-translational levels.

Visual Connection Questions

1. Figure 9.10 You isolate a cell strain in which the joining together of Okazaki fragments is impaired and suspect that a mutation has occurred in an enzyme found at the replication fork. Which enzyme is most likely to be mutated?

Review Questions

2. Which of the following does cytosine pair with?
- a. guanine
- b. thymine
- c. adenine
- d. a pyrimidine

3. Prokaryotes contain a _____chromosome, and eukaryotes contain _____ chromosomes.
- a. single-stranded circular; single-stranded linear
- b. single-stranded linear; single-stranded circular
- c. double-stranded circular; double-stranded linear
- d. double-stranded linear; double-stranded circular

4. DNA replicates by which of the following models?
- a. conservative
- b. semiconservative
- c. dispersive
- d. none of the above

5. The initial mechanism for repairing nucleotide errors in DNA is _____.
- a. mismatch repair
- b. DNA polymerase proofreading
- c. nucleotide excision repair
- d. thymine dimers

6. A promoter is _____.
- a. a specific sequence of DNA nucleotides
- b. a specific sequence of RNA nucleotides
- c. a protein that binds to DNA
- d. an enzyme that synthesizes RNA

7. Portions of eukaryotic mRNA sequence that are removed during RNA processing are _____.
- a. exons
- b. caps
- c. poly-A tails
- d. introns

8. The RNA components of ribosomes are synthesized in the _____.
   a. cytoplasm
   b. nucleus
   c. nucleolus
   d. endoplasmic reticulum

9. How long would the peptide be that is translated from this MRNA sequence: 5'-AUGGGCUACCGA-3'?
   a. 0
   b. 2
   c. 3
   d. 4

10. Control of gene expression in eukaryotic cells occurs at which level(s)?
   a. only the transcriptional level
   b. epigenetic and transcriptional levels
   c. epigenetic, transcriptional, and translational levels
   d. epigenetic, transcriptional, post-transcriptional, translational, and post-translational levels

11. Post-translational control refers to:
   a. regulation of gene expression after transcription
   b. regulation of gene expression after translation
   c. control of epigenetic activation
   d. period between transcription and translation

Critical Thinking Questions

12. Describe the organization of the eukaryotic chromosome.

13. Describe the structure and complementary base pairing of DNA.

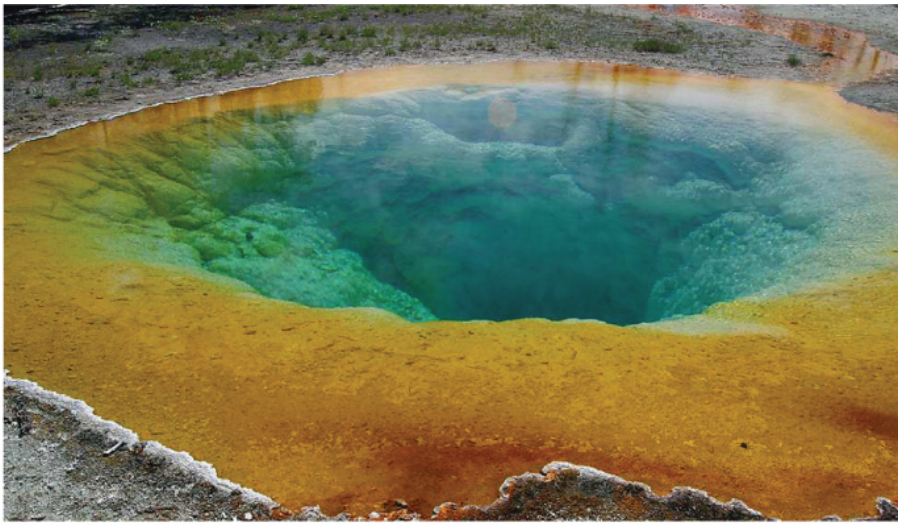
14. How do the linear chromosomes in eukaryotes ensure that its ends are replicated completely?

15. Transcribe and translate the following DNA sequence (nontemplate strand): 5'-ATGGCCGGTTATTAAGCA-3'

16. Describe how controlling gene expression will alter the overall protein levels in the cell.

CHAPTER 10
Biotechnology

(a)



(b)

FIGURE 10.1 (a) A thermal cycler, such as the one shown here, is a basic tool used to study DNA in a process called the polymerase chain reaction (PCR). The polymerase enzyme most often used with PCR comes from a strain of bacteria that lives in (b) the hot springs of Yellowstone National Park. (credit a: modification of work by Magnus Manske; credit b: modification of work by Jon Sullivan)

CHAPTER OUTLINE

INTRODUCTION The latter half of the twentieth century began with the discovery of the structure of DNA, then progressed to the development of the basic tools used to study and manipulate DNA. These advances, as well as advances in our understanding of and ability to manipulate cells, have led some to refer to the twenty-first century as the biotechnology century. The rate of discovery and of the development of new applications in medicine, agriculture, and energy is expected to accelerate, bringing huge benefits to humankind and perhaps also significant risks. Many of these developments are expected to raise significant ethical and social questions that human societies have not yet had to consider.

10.1 Cloning and Genetic Engineering

LEARNING OBJECTIVES

By the end of this section, you will be able to:

- Explain the basic techniques used to manipulate genetic material
- Explain molecular and reproductive cloning

Biotechnology is the use of artificial methods to modify the genetic material of living organisms or cells to produce novel compounds or to perform new functions. Biotechnology has been used for improving livestock and crops since the beginning of agriculture through selective breeding. Since the discovery of the structure of DNA in 1953, and particularly since the development of tools and methods to manipulate DNA in the 1970s, biotechnology has become synonymous with the manipulation of organisms' DNA at the molecular level. The primary applications of this technology are in medicine (for the production of vaccines and antibiotics) and in agriculture (for

the genetic modification of crops). Biotechnology also has many industrial applications, such as fermentation, the treatment of oil spills, and the production of biofuels, as well as many household applications such as the use of enzymes in laundry detergent.

Manipulating Genetic Material

To accomplish the applications described above, biotechnologists must be able to extract, manipulate, and analyze nucleic acids.

Review of Nucleic Acid Structure

To understand the basic techniques used to work with nucleic acids, remember that nucleic acids are macromolecules made of nucleotides (a sugar, a phosphate, and a nitrogenous base). The phosphate groups on these molecules each have a net negative charge. An entire set of DNA molecules in the nucleus of eukaryotic organisms is called the genome. DNA has two complementary strands linked by hydrogen bonds between the paired bases.

Unlike DNA in eukaryotic cells, RNA molecules leave the nucleus. Messenger RNA (mRNA) is analyzed most frequently because it represents the protein-coding genes that are being expressed in the cell.

Isolation of Nucleic Acids

To study or manipulate nucleic acids, the DNA must first be extracted from cells. Various techniques are used to extract different types of DNA (Figure 10.2). Most nucleic acid extraction techniques involve steps to break open the cell, and then the use of enzymatic reactions to destroy all undesired macromolecules. Cells are broken open using a detergent solution containing buffering compounds. To prevent degradation and contamination, macromolecules such as proteins and RNA are inactivated using enzymes. The DNA is then brought out of solution using alcohol. The resulting DNA, because it is made up of long polymers, forms a gelatinous mass.

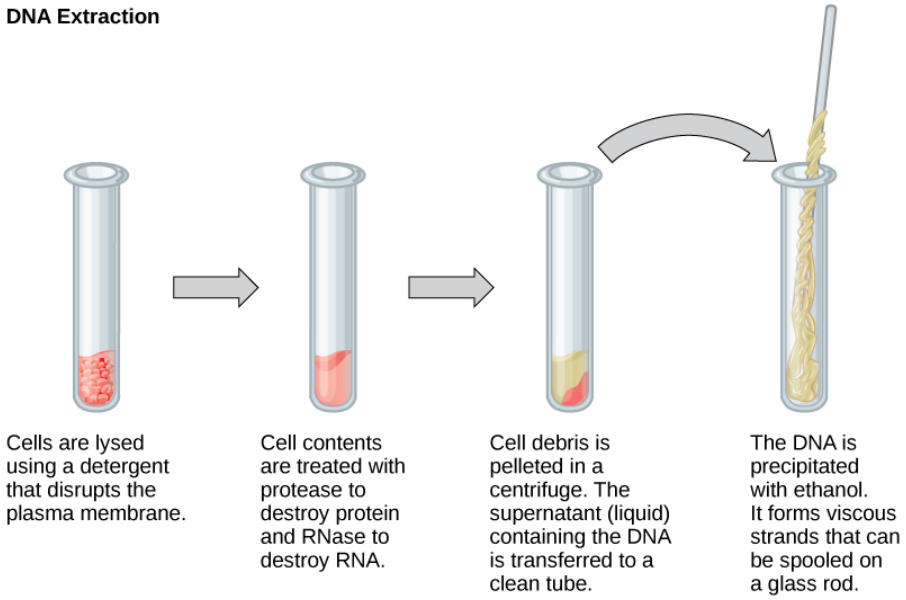


FIGURE 10.2 This diagram shows the basic method used for the extraction of DNA.

RNA is studied to understand gene expression patterns in cells. RNA is naturally very unstable because enzymes that break down RNA are commonly present in nature. Some are even secreted by our own skin and are very difficult to inactivate. Similar to DNA extraction, RNA extraction involves the use of various buffers and enzymes to inactivate other macromolecules and preserve only the RNA.

Gel Electrophoresis

Because nucleic acids are negatively charged ions at neutral or alkaline pH in an aqueous environment, they can be moved by an electric field. **Gel electrophoresis** is a technique used to separate charged molecules on the basis of size and charge. The nucleic acids can be separated as whole chromosomes or as fragments. The nucleic acids are loaded into a slot at one end of a gel matrix, an electric current is applied, and negatively charged molecules are pulled toward the opposite end of the gel (the end with the positive electrode). Smaller molecules move through the pores in the gel faster than larger molecules; this difference in the rate of migration separates the fragments on the basis of size. The nucleic acids in a gel matrix are invisible until they are stained with a compound that allows them to be seen, such as a dye. Distinct fragments of nucleic acids appear as bands at specific distances from the top of the gel (the negative electrode end) that are based on their size (Figure 10.3). A mixture of many fragments of varying sizes appear as a long smear, whereas uncut genomic DNA is usually too large to run through the gel and forms a single large band at the top of the gel.

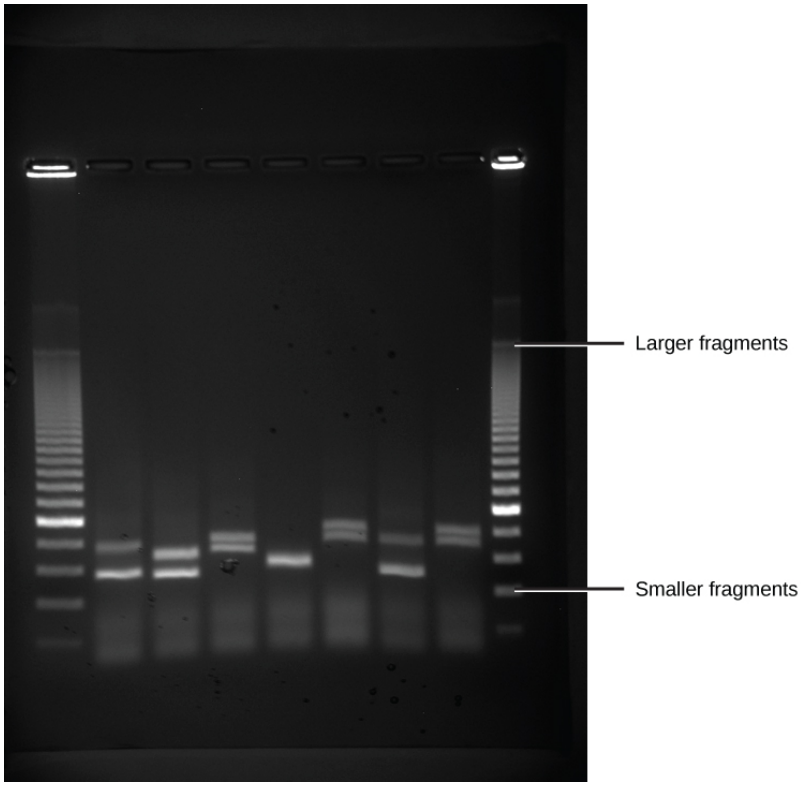


FIGURE 10.3 Shown are DNA fragments from six samples run on a gel, stained with a fluorescent dye and viewed under UV light. (credit: modification of work by James Jacob, Tompkins Cortland Community College)

Polymerase Chain Reaction

DNA analysis often requires focusing on one or more specific regions of the genome. It also frequently involves situations in which only one or a few copies of a DNA molecule are available for further analysis. These amounts are insufficient for most procedures, such as gel electrophoresis. **Polymerase chain reaction (PCR)** is a technique used to rapidly increase the number of copies of specific regions of DNA for further analyses (Figure 10.4). PCR uses a special form of DNA polymerase, the enzyme that replicates DNA, and other short nucleotide sequences called primers that base pair to a specific portion of the DNA being replicated. PCR is used for many purposes in laboratories. These include: 1) the identification of the owner of a DNA sample left at a crime scene; 2) paternity analysis; 3) the comparison of small amounts of ancient DNA with modern organisms; and 4) determining the sequence of nucleotides in a specific region.

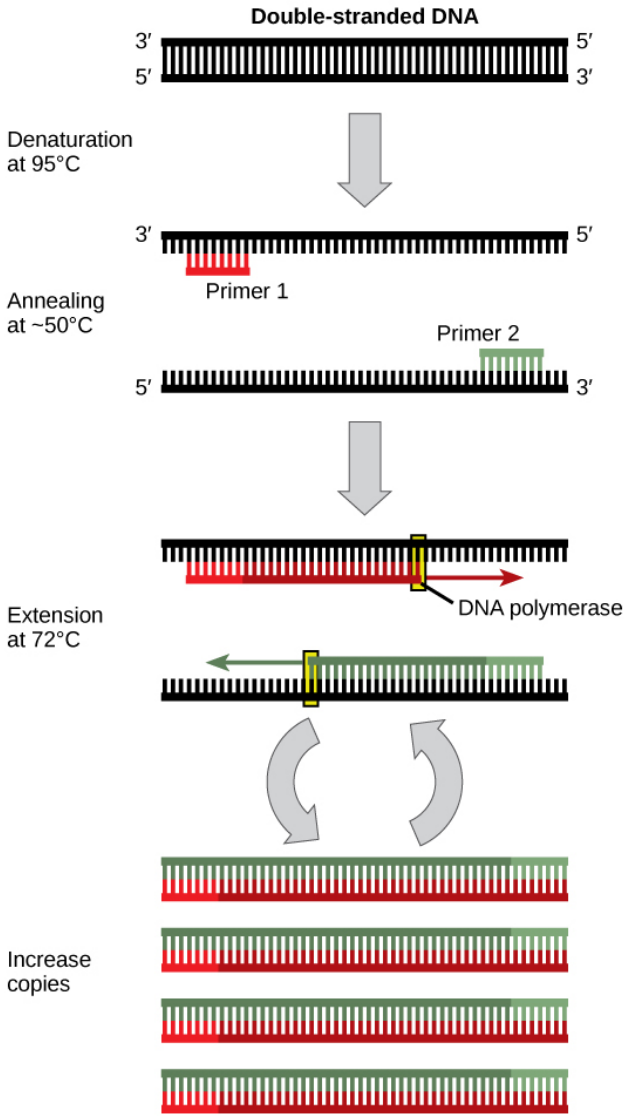


FIGURE 10.4 Polymerase chain reaction, or PCR, is used to produce many copies of a specific sequence of DNA using a special form of DNA polymerase.

Cloning

In general, **cloning** means the creation of a perfect replica. Typically, the word is used to describe the creation of a genetically identical copy. In biology, the re-creation of a whole organism is referred to as “reproductive cloning.” Long before attempts were made to clone an entire organism, researchers learned how to copy short stretches of DNA—a process that is referred to as molecular cloning. The technique offered methods to create new medicines and to overcome difficulties with existing ones. When Lydia Villa-Komaroff, working in the Gilbert Lab at Harvard, published the first paper outlining the technique for producing synthetic insulin, diabetes researchers and patients received new hope in fighting the disease. Insulin at that time was only produced using pig and cow pancreases, and the life-saving substance was often in short supply. Synthetic insulin, once mass produced, would solve that problem for many patients. These early discoveries led to the "BioTech Boom," and spurred continued research and funding for newer and better ways to improve health.

Molecular Cloning

Cloning allows for the creation of multiple copies of genes, expression of genes, and study of specific genes. To get the DNA fragment into a bacterial cell in a form that will be copied or expressed, the fragment is first inserted into a plasmid. A **plasmid** (also called a vector in this context) is a small circular DNA molecule that replicates independently of the chromosomal DNA in bacteria. In cloning, the plasmid molecules can be used to provide a