Inheritance is governed by information stored in discrete factors called genes.

Genes are transmitted from generation to generation on vehicles called chromosomes.

Chromosomes, which exist in pairs, provide the basis of biparental inheritance.

During gamete formation, chromosomes are distributed according to postulates first described by Gregor Mendel, based on his nineteenth-century research with the garden pea.

Mendelian postulates prescribe that homologous chromosomes segregate from one another and assort independently with other segregating homologs during gamete formation.

Genetic ratios, expressed as probabilities, are subject to chance deviation and may be evaluated statistically.

The analysis of pedigrees allows predictions involving the genetic nature of human traits.
Although inheritance of biological traits has been recognized for thousands of years, the first significant insights into the mechanisms involved occurred about 135 years ago. In 1866, Gregor Johann Mendel published the results of a series of experiments that would lay the foundation for the formal discipline of genetics. Mendel’s work went largely unnoticed until the turn of the century, but in the ensuing years the concept of the gene as a distinct hereditary unit was established. The ways in which genes, as members of chromosomes, are transmitted to offspring and control traits were clarified. Research continued unabated throughout the twentieth century—indeed, studies in genetics, most recently at the molecular level, have remained continually at the forefront of biological research since the early 1900s.

When Mendel began his studies of inheritance using *Pisum sativum*, the garden pea, chromosomes and the role and mechanism of meiosis were totally unknown. Nevertheless, he determined that discrete units of inheritance exist and predicted their behavior during the formation of gametes. Subsequent investigators, with access to cytological data, were able to relate their observations of chromosome behavior during meiosis to Mendel’s principles of inheritance. Once this correlation was made, Mendel’s postulates were accepted as the basis for the study of what is known as transmission genetics.

### How Do We Know?

In this chapter, we focus on how Mendel was able to derive the essential postulates that explain inheritance. As you study this topic, you should try to answer several fundamental questions:

1. How did Mendel know that unit factors existed as fundamental genetic components if he could not directly observe them?
2. How do we know that an organism expressing a dominant trait is homozygous or heterozygous?
3. In genetic data, how do we know that deviation from the expected ratio is due to chance rather than another independent factor?
4. How do we know how a trait is inherited in humans?

### 3.1 Mendel Used a Model Experimental Approach to Study Patterns of Inheritance

Johann Mendel was born in 1822 to a peasant family in the central European village of Heinzendorf. An excellent student in high school, he studied philosophy for several years afterward, and in 1843 he was admitted to the Augustinian Monastery of St. Thomas in Brno, now part of the Czech Republic, taking the name Gregor. In 1849, he was relieved of his monastic duties and accepted a teaching appointment that lasted several years. From 1851 to 1853, he attended the University of Vienna, where he studied physics and botany. He returned to Brno in 1854, where he taught physics and natural science for the next 16 years. Mendel received support from the monastery for his studies and research throughout his life.

In 1856, Mendel performed his first set of hybridization experiments with the garden pea. The research phase of his career lasted until 1868, when he was elected abbot of the monastery. Although he retained his interest in genetics, his new responsibilities demanded most of his time. In 1884, Mendel died of a kidney disorder. The local newspaper paid him the following tribute:

“His death deprives the poor of a benefactor, and mankind at large of a man of the noblest character, one who was a warm friend, a promoter of the natural sciences, and an exemplary priest.”

Mendel first reported the results of some simple genetic crosses between certain strains of the garden pea in 1865. Although his was not the first attempt to provide experimental evidence pertaining to inheritance, Mendel’s success where others failed can be attributed, at least in part, to his elegant model of experimental design and analysis.

Mendel showed remarkable insight into the methodology necessary for good experimental biology. He chose an organism that is easy to grow and hybridize artificially. The pea plant is self-fertilizing in nature but is easy to crossbreed experimentally. It reproduces well and grows to maturity in a single season. Mendel followed seven visible features (unit characters), each represented by two contrasting forms, or traits (Figure 3–1). For the character stem height, for example, he experimented with the traits tall and dwarf. He selected six other visibly contrasting pairs of traits involving seed shape and color, pod shape and color, and pod and flower arrangement. From local seed merchants, Mendel obtained true-breeding strains—those in which each trait appeared unchanged generation after generation in self-fertilizing plants.

There were several reasons for Mendel’s success. In addition to his choice of a suitable organism, he restricted his examination to one or very few pairs of contrasting traits in each experiment. He also kept accurate quantitative records, a necessity in genetic experiments. From the analysis of his data, Mendel derived certain postulates that became principles of transmission genetics.

The results of Mendel’s experiments were unappreciated until the turn of the century, well after his death. However, once Mendel’s publications were rediscovered by geneticists investigating the function and behavior of chromosomes, the implications of his postulates were immediately apparent. He had discovered the basis for the transmission of hereditary traits!

### 3.2 The Monohybrid Cross Reveals How One Trait Is Transmitted from Generation to Generation

Mendel’s simplest crosses involved only one pair of contrasting traits. Each such experiment is a monohybrid cross, which is made by mating true-breeding individuals from two parent strains, each exhibiting one of the two contrasting forms of the character under study. Initially, we examine the first
Seeds
- round/wrinkled
- yellow/green

F1 results: all round
F2 results: 5474 round, 1850 wrinkled
F2 ratio: 2.96:1

Pods
- full/constricted
- green/yellow

F1 results: all full
F2 results: 882 full, 299 constricted
F2 ratio: 2.95:1

Flower color
- violet/white

F1 results: all violet
F2 results: 705 violet, 224 white
F2 ratio: 3.15:1

Flower position
- axial/terminal

F1 results: all axial
F2 results: 651 axial, 207 terminal
F2 ratio: 3.14:1

Stem length
- tall/dwarf

F1 results: all tall
F2 results: 787 tall, 277 dwarf
F2 ratio: 2.84:1

FIGURE 3-1 Seven pairs of contrasting traits and the results of Mendel’s seven monohybrid crosses of the garden pea (Pisum sativum).
In each case, pollen derived from plants exhibiting one trait was used to fertilize the ova of plants exhibiting the other trait. In the F1 generation, one of the two traits was exhibited by all plants. The contrasting trait reappeared in approximately 1/4 of the F2 plants.

generation of offspring of such a cross, and then we consider the results of selfing, the offspring of self-fertilizing individuals from this first generation. The original parents constitute the P1, or parental generation, their offspring are the F1, or first filial generation, and the individuals resulting from the selfed F1 generation are the F2, or second filial generation. We can continue to follow subsequent generations.

The cross between true-breeding pea plants with tall stems and dwarf stems is representative of Mendel’s monohybrid crosses. Tall and dwarf are contrasting traits of the character of stem height. Unless tall or dwarf plants are crossed together or with another strain, they will undergo self-fertilization and breed true, producing their respective traits generation after generation. However, when Mendel crossed tall plants with dwarf plants, the resulting F1 generation consisted only of tall plants. When members of the F1 generation were selfed, Mendel observed that 787 of 1064 F2 plants were tall, while the remaining 277 were dwarf. Note that in this cross (Figure 3–1) the dwarf trait disappears in the F1, only to reappear in the F2 generation.

Genetic data are usually expressed and analyzed as ratios. In this particular example, many identical P1 crosses were made, and many F1 plants—all tall—were produced. Of the 1064 F2 offspring, 787 were tall and 277 were dwarf—a ratio of 2.84:1.0, or about 3:1.

Mendel made similar crosses between pea plants exhibiting other pairs of contrasting traits; the results of these crosses are shown in Figure 3–1. In every case, the outcome was similar to the tall/dwarf cross just described. All F1 offspring were identical to one of the parents, but in the F2 offspring, an approximate ratio of 3:1 was obtained. That is, three-fourths looked like the F1 plants, while one-fourth exhibited the contrasting trait, which had disappeared in the F1 generation.

We will point out one further aspect of Mendel’s monohybrid crosses. In each cross, the F1 and F2 patterns of inheritance were similar regardless of which P1 plant served as the source of pollen (sperm) and which served as the source of the ovum (egg). The crosses could be made either way—pollination of dwarf plants by tall plants or vice versa. These are called reciprocal crosses. Therefore, the results of Mendel’s monohybrid crosses were not sex-dependent.

To explain these results, Mendel proposed the existence of particular unit factors for each trait. He suggested that these factors serve as the basic units of heredity and are passed unchanged from generation to generation, determining the various traits expressed by each individual plant. Using these general ideas, Mendel proceeded to hypothesize precisely how unit factors could account for the results of the monohybrid crosses.

Mendel’s First Three Postulates

Using the consistent pattern of results in the monohybrid crosses, Mendel derived the following three postulates or principles of inheritance.
1. **Unit Factors in Pairs**

*Genetic characters are controlled by unit factors that exist in pairs in individual organisms.*

In the monohybrid cross involving tall and dwarf stems, a specific unit factor exists for each trait. Because the factors occur in pairs, three combinations are possible: two factors for tallness, two factors for dwarfness, or one factor for each trait. Every individual contains one of these three combinations, which determines stem height.

2. **Dominance/Recessiveness**

*When two unlike unit factors responsible for a single character are present in a single individual, one unit factor is dominant to the other, which is said to be recessive.*

In each monohybrid cross, the trait expressed in the F1 generation is controlled by the dominant unit factor. The trait not expressed is controlled by the recessive unit factor. Note that this dominance/recessiveness relationship pertains only when unlike unit factors are present in pairs. The terms dominant and recessive are also used to designate traits. In this case, tall stems are said to be dominant over the recessive dwarf stems.

3. **Segregation**

*During the formation of gametes, the paired unit factors separate or segregate randomly so that each gamete receives one or the other with equal likelihood.*

If an individual contains a pair of like unit factors (e.g., both specific for tall), then all gametes receive one tall unit factor. If an individual contains unlike unit factors (e.g., one for tall and one for dwarf), then each gamete has a 50 percent probability of receiving either the tall or the dwarf unit factor.

These postulates provide a suitable explanation for the results of the monohybrid crosses. Let’s use the tall/dwarf cross to illustrate. Mendel reasoned that P1 tall plants contain identical paired unit factors, as do the P1 dwarf plants. The gametes of tall plants all receive one tall unit factor as a result of segregation. Similarly, the gametes of dwarf plants all receive one dwarf unit factor. Following fertilization, all F1 plants receive one unit factor from each parent: a tall factor from one and a dwarf factor from the other, reestablishing the paired relationship—but because tall is dominant to dwarf, all F1 plants are tall.

When F1 plants form gametes, the postulate of segregation demands that each gamete randomly receives either the tall or the dwarf unit factor. Following random fertilization events during F1 selfing, four F2 combinations result in equal frequency:

1. tall/tall
2. tall/dwarf
3. dwarf/tall
4. dwarf/dwarf

Combinations (1) and (4) result in tall and dwarf plants, respectively. According to the postulate of dominance/recessiveness, combinations (2) and (3) both yield tall plants. Therefore, the F2 is predicted to consist of 3/4 tall and 1/4 dwarf, or a ratio of 3:1. This is approximately what Mendel observed in the cross between tall and dwarf plants. A similar pattern was observed in each of the other monohybrid crosses (see Figure 3–1).

**Modern Genetic Terminology**

To illustrate the monohybrid cross and Mendel’s first three postulates, we must first introduce several new terms as well as a symbol convention for the unit factors.

Traits such as tall or dwarf are visible expressions of the information contained in unit factors. The physical appearance of a trait is the phenotype of the individual. Mendel’s unit factors represent units of inheritance called genes by modern geneticists. For any given character, such as plant height, the phenotype is determined by alternative forms of a single gene called alleles. For example, the unit factors representing tall and dwarf are alleles determining the height of the pea plant.

The convention we will use is to choose the first letter of the recessive trait to symbolize the character in question—the lowercase italic letter designates the allele for the recessive trait, and the uppercase italic letter designates the allele for the dominant trait. Thus, for Mendel’s pea plants, we use d for the dwarf allele and D for the tall allele. When alleles are written in pairs to represent the two unit factors present in any individual (DD, Dd, or dd), these symbols are called the genotype. This term reflects the genetic makeup of an individual, whether it is haploid or diploid. By reading the genotype, we know the phenotype of the individual: DD and Dd are tall, and dd is dwarf. When both alleles are the same (DD or dd), the individual is homozygous or a homozygote; when the alleles are different (Dd), we use the term heterozygous or a heterozygote. These symbols and terms are used in Figure 3–2 to illustrate the monohybrid cross.

Because he operated without the hindsight that modern geneticists enjoy, Mendel’s analytical reasoning must be considered a truly outstanding scientific achievement. On the basis of rather simple but precisely executed breeding experiments, he not only proposed that discrete particular units of heredity exist, he also explained how they are transmitted from one generation to the next.

**Now Solve This**

Problem 5 on page 58 involves a Mendelian cross where you must determine the mode of inheritance and the genotypes of the parents in a number of instances.

**Hint:** The first step is to determine how many genes are involved. To do so, convert the data to ratios that are characteristic of Mendelian crosses. In the case of this problem, ask first whether any of the F2 ratios match Mendel’s 3:1 monohybrid ratio.
**Punnett Squares**

The genotypes and phenotypes resulting from the recombination of gametes during fertilization can be easily visualized by constructing a **Punnett square**, named after the person who first devised this approach, Reginald C. Punnett. **Figure 3-3** demonstrates this method of analysis for our $F_1 \times F_1$ monohybrid cross. Each of the possible gametes is assigned to a column or a row; the vertical column represents those of the female parent, and the horizontal row represents those of the male parent. After putting the gametes into the rows and columns, the new generation is predicted by combining the male and female gametic information for each combination by entering the resulting genotypes in the boxes. This process thus lists all possible

**FIGURE 3-2** The monohybrid cross between tall ($D$) and dwarf ($d$) pea plants. Individuals are shown in rectangles, and gametes in circles.

**FIGURE 3-3** A Punnett square generating the $F_2$ ratio of the $F_1 \times F_1$ cross shown in Figure 3–2.
random fertilization events. The genotypes and phenotypes of all potential offspring are ascertained by reading the entries in the boxes.

The Punnett square method is particularly useful when you are first learning about genetics and how to solve problems. Note the ease with which the 3:1 phenotypic ratio and the 1:2:1 genotypic ratio is derived in the F2 generation in Figure 3–3.

The Testcross: One Character

Tall plants produced in the F2 generation are predicted to be either the DD or Dd genotype. You might ask if there is a way to distinguish the genotype. Mendel devised a rather simple method that is still used today in breeding plants and animals: the testcross. The organism expressing the dominant phenotype, but of unknown genotype, is crossed to a known homozygous recessive individual. For example, as shown in Figure 3–4(a), if a tall plant of genotype DD is testcrossed to a dwarf plant, which must have the dd genotype, all offspring will be tall phenotypically and Dd genotypically. However, as shown in Figure 3–4(b), if a tall plant is Dd and it is crossed to a dwarf plant (dd), then one-half of the offspring will be tall (Dd) and the other half will be dwarf (dd). Therefore, a 1:1 tall/dwarf ratio demonstrates the heterozygous nature of the tall plant of unknown genotype. The test cross reinforced Mendel’s conclusion that separate unit factors control traits.

### How Mendel’s Peas Become Wrinkled: A Molecular Explanation

Only recently, well over a hundred years after Mendel used wrinkled peas in his groundbreaking hybridization experiments, have we come to find out how the wrinkled gene makes peas wrinkled. The wild-type allele of the gene encodes a protein called **starch-branching enzyme (SBEI)**. This enzyme catalyzes the formation of highly branched starch molecules as the seed matures.

Wrinkled peas, which result from the homozygous presence of the mutant form of the gene, lack the activity of this enzyme. The production of branch points is inhibited during the synthesis of starch within the seed, which in turn leads to the accumulation of more sucrose and a higher water content while the seed develops. Osmotic pressure inside rises, causing the seed to lose water internally, and ultimately results in the wrinkled appearance of the seed at maturation. In contrast, developing seeds that bear at least one copy of the normal gene (being either homozygous or heterozygous for the dominant allele) synthesize starch and reach an osmotic balance that minimizes the loss of water. The end result is a smooth-textured outer coat.

The SBEI gene has been cloned and analyzed, providing greater insight into the relationship between genotypes and phenotypes. Interestingly, the mutant gene contains a foreign sequence of some 800 base pairs that disrupts the normal coding sequence. This foreign segment closely resembles other such sequences, called **transposable elements**. These sequences have the ability to move from place to place in the genome of organisms. Transposable elements have been found in maize (corn), parsley, and snapdragons, fruit flies, and humans, among many other organisms.
Web Tutorial 3.2
Independent Assortment

Figure 3–5 will occur: the F1 offspring will all be yellow and round. It is therefore apparent that yellow is dominant to green, and that round is dominant to wrinkled. When the F1 individuals are selfed, approximately 9/16 of the F2 plants express yellow and round, 3/16 express yellow and wrinkled, 3/16 express green and round, and 1/16 express green and wrinkled.

A variation of this cross is also shown in Figure 3–5. Instead of crossing one P1 parent with both dominant traits (yellow, round) and one with both recessive traits (green, wrinkled), plants with yellow, wrinkled seeds are crossed with those with green, round seeds. In spite of the change in the P1 phenotypes, both the F1 and F2 results remain unchanged. It will become clear in the next section why this is so.

Mendel’s Fourth Postulate:
Independent Assortment

We can most easily understand the results of a dihybrid cross if we consider it theoretically as consisting of two monohybrid crosses conducted separately. Think of the two sets of traits as inherited independently of each other; that is, the chance of any plant having yellow or green seeds is not at all influenced by the chance that this plant will have round or wrinkled seeds. Thus, because yellow is dominant to green, all F1 plants in the first theoretical cross would have yellow seeds. In the second theoretical cross, all F1 plants would have round seeds because round is dominant to wrinkled. When Mendel examined the F1 plants of the dihybrid cross, all were yellow and round, as we just predicted.

The predicted F2 results of the first cross are 3/4 yellow and 1/4 green. Similarly, the second cross would yield 3/4 round and 1/4 wrinkled. Figure 3–5 shows that in the dihybrid cross, 12/16 F2 plants are yellow while 4/16 are green, exhibiting the expected 3:1 (3/4:1/4) ratio. Similarly, 12/16 F2 plants have round seeds while 4/16 have wrinkled seeds, again revealing the 3:1 ratio.

It is evident that the two pairs of contrasting traits are inherited independently, so we can predict the frequencies of all possible F2 phenotypes by applying the product law of probabilities: When two independent events occur simultaneously, the combined probability of the two outcomes is equal to the product of their individual probabilities of occurrence. For example, the probability of an F2 plant having yellow and round seeds is (3/4)(3/4), or 9/16, because 3/4 of all F2 plants should be yellow and (3/4) of all F2 plants should be round. In a like manner, the probabilities of the other three F2 phenotypes can be calculated: yellow (3/4) and wrinkled (1/4) are predicted to be present together 3/16 of the time; green (1/4) and round (3/4) are predicted 3/16 of the time; and green (1/4) and wrinkled (1/4) are predicted 1/16 of the time. These calculations are shown in Figure 3–6.

It is now apparent why the F1 and F2 results are identical whether the initial cross is yellow, round plants bred with green, wrinkled plants, or if yellow, wrinkled plants are bred with green, round plants. In both crosses, the F1 genotype of all plants is identical. Each plant is heterozygous for both gene pairs. As a result, the F2 generation is also identical in both crosses.

<table>
<thead>
<tr>
<th>F1</th>
<th>yellow, round × yellow, round</th>
</tr>
</thead>
<tbody>
<tr>
<td>F2</td>
<td>Of all offspring</td>
</tr>
<tr>
<td></td>
<td>3/4 are round and 1/4 are wrinkled</td>
</tr>
<tr>
<td></td>
<td>3/4 are round and 1/4 are wrinkled</td>
</tr>
<tr>
<td></td>
<td>1/4 are green</td>
</tr>
<tr>
<td></td>
<td>1/4 are green</td>
</tr>
</tbody>
</table>

Figure 3–6 Computation of the combined probabilities of each F2 phenotype for two independently inherited characters. The probability of each plant’s being yellow or green is independent of the probability of its bearing round or wrinkled seeds.
On the basis of similar results in numerous dihybrid crosses, Mendel proposed a fourth postulate called **independent assortment**: During gamete formation, segregating pairs of unit factors assort independently of each other.

This postulate stipulates that segregation of any pair of unit factors occurs independently of all others. As a result of random segregation, each gamete receives one member of every pair of unit factors. For one pair, whichever unit factor is received does not influence the outcome of segregation of any other pair. Thus, according to the postulate of independent assortment, all possible combinations of gametes are formed in equal frequency.

The Punnett square in Figure 3–7 shows how independent assortment works in the formation of the F₂ generation.

**FIGURE 3-7** Analysis of the dihybrid crosses shown in Figure 3–5. The F₁ heterozygous plants are self-fertilized to produce an F₂ generation, which is computed using a Punnett square. Both the phenotypic and genotypic F₂ ratios are shown.
Examine the formation of gametes by the F1 plants; segregation preserves that every gamete receives either a G or g allele and a W or w allele. Independent assortment stipulates that all four combinations (GW, Gw, gW, and gw) will be formed with equal probabilities.

In every F1 × F1 fertilization event, each zygote has an equal probability of receiving one of the four combinations from each parent. If many offspring are produced, 9/16 have yellow, round seeds, 3/16 have yellow, wrinkled seeds, 3/16 have green, round seeds, and 1/16 have green, wrinkled seeds, yielding what is designated as Mendel’s 9:3:3:1 dihybrid ratio. This is an ideal ratio based on probability events involving segregation, independent assortment, and random fertilization. Because of deviation due to chance, particularly if small numbers of offspring are produced, actual results are highly unlikely to match the ideal ratio.

### The Testcross: Two Characters

The testcross can also be applied to individuals that express two dominant traits but whose genotypes are unknown. For example, the expression of the yellow, round seed phenotype in the F2 generation just described may result from the expression of one or more of these four genotypes. To determine the genotypes of the parents, we can perform a testcross.

For example, consider the cross shown in Figure 3-8, where the gene pairs of theoretical contrasting traits are represented by the symbols A, a, B, b, C, and c. In the cross between AAbbCc and aabbcc individuals, all F1 individuals are heterozygous for all three gene pairs. When F1 individuals serve as parents, each produces eight different gametes in equal frequencies. At this point, we could construct a Punnett square with 64 separate boxes and read out the phenotypes—but such a method is cumbersome in a cross involving so many factors. Therefore another method has been devised to calculate the predicted ratio.

### The Forked-Line Method

It is much less difficult to consider each contrasting pair of traits separately and then to combine these results by using the *forked-line method*, first shown in Figure 3-6. This method, also called a branch diagram, relies on the simple application of the laws of probability established for the dihybrid cross. Each gene pair is assumed to behave independently during gamete formation.

When the monohybrid cross AA × aa is made, we know that:

1. All F1 individuals have the genotype Aa and express the phenotype represented by the A allele, which is called the A phenotype in the following discussion.
2. The F2 generation consists of individuals with either the A phenotype or the a phenotype in the ratio of 3:1.

The same generalizations can be made for the BB × bb and CC × cc crosses. Thus, in the F2 generation, 3/4 of all organisms express phenotype A, 3/4 express B, and 3/4 express C. Similarly, 1/4 of all organisms express phenotype a, 1/4 express b, and 1/4 express c. The proportions of organisms that express each phenotypic combination can be predicted by assuming that fertilization, following the independent assortment of these three gene pairs during...
The source of natural variation intrigued students of evolutionary theory. These individuals, stimulated by the proposal developed by Charles Darwin and Alfred Russel Wallace, believed in **continuous variation**, whereby offspring were a *blend* of their parents’ phenotypes. As we mentioned earlier, Mendel theorized that variation was due to discrete or particulate units, resulting in **discontinuous variation**. For example, Mendel proposed that the F₂ offspring of a dihybrid cross are expressing traits produced by new combinations of previously existing unit factors. As a result, Mendel’s theories did not fit well with the evolutionists’ preconceptions about causes of variation.

In the latter part of the nineteenth century, a remarkable observation set the scene for the rebirth of Mendel’s work: Walter Flemming’s discovery of chromosomes in the nuclei of salamander cells. In 1879, Flemming described the behavior of these threadlike structures during cell division. As a result of his findings and the work of many other cytologists, the presence of discrete units within the nucleus soon became an integral part of ideas about inheritance. It was this mind-set that prompted scientists to reexamine Mendel’s findings.

In the early twentieth century, research led to renewed interest in Mendel’s work. Hybridization experiments similar to Mendel’s were performed independently by three botanists: Hugo de Vries, Karl Correns, and Erich Tschermak. De Vries’s work demonstrated the principle of segregation in his experiments with several plant species. Apparently, he searched the existing literature and found that Mendel’s work was not appreciated for about 35 years. Vries’s proposal developed by Charles Darwin intrigued students of evolutionary theory. These individuals, stimulated by the chromosome theory of inheritance, which was developed during the next two decades.

### Unit Factors, Genes, and Homologous Chromosomes

Because the correlation between Sutton’s and Boveri’s observations and Mendelian principles serves as the foundation for the modern interpretation of transmission genetics, we will examine this correlation in some detail before moving on to other topics.

As we know, each species possesses a specific number of chromosomes in each somatic cell nucleus (except in
gametes). For diploid organisms, this number is called the **diploid number** \((2n)\) and is characteristic of that species. During the formation of gametes, this number is precisely halved \((n)\), and when two gametes combine during fertilization, the diploid number is reestablished. During meiosis, however, the chromosome number is not reduced in a random manner. It was apparent to early cytologists that the diploid number of chromosomes is composed of homologous pairs identifiable by their morphological appearance and behavior. The gametes contain one member of each pair—thus the chromosome complement of a gamete is quite specific, and the number of chromosomes in each gamete is equal to the haploid number.

With this basic information, we can see the correlation between the behavior of unit factors and chromosomes and genes. Figure 3–10 shows three of Mendel’s postulates and the chromosomal explanation of each. Unit factors are really genes located on homologous pairs of chromosomes [Figure 3–10(a)]. Members of each pair of homologs separate, or segregate, during gamete formation [Figure 3–10(b)]. Two different alignments are possible, both of which are shown.

To illustrate the principle of independent assortment, we must distinguish between members of any given homologous pair of chromosomes. One member of each pair comes from

**FIGURE 3-10** Illustrated correlation between the Mendelian postulates of (a) unit factors in pairs, (b) segregation, and (c) independent assortment, showing the presence of genes located on homologous chromosomes and their behavior during meiosis.
the maternal parent, while the other member comes from the paternal parent. (We represent the different parental origins with different colors.) As shown in Figure 3–10(c), following independent assortment of each pair of homologs, each gamete receives one member from each pair of chromosomes. All possible combinations are formed with equal probability. If we add the symbols used in Mendel’s dihybrid cross \((G, g, W, w)\) to the diagram, we can see why equal numbers of the four types of gametes are formed. The independent behavior of Mendel’s pairs of unit factors \((G\) and \(W\) in this example) is due to their presence on separate pairs of homologous chromosomes.

Observations of the phenotypic diversity of living organisms make it logical to assume that there are many more genes than chromosomes. Therefore, each homolog must carry genetic information for more than one trait. The currently accepted concept is that a chromosome is composed of a large number of linearly ordered, information-containing genes. Mendel’s unit factors (which determine tall or dwarf stems, for example) actually constitute a pair of genes located on one pair of homologous chromosomes. The location on a given chromosome where any particular gene occurs is called its locus (pl., loci). The different forms taken by a given gene, called alleles \((G\) or \(g\)), contain slightly different genetic information that determines the same character (seed color in this case). Although we have examined only genes with two alternative alleles, most genes have more than two allelic forms. We conclude this section by reviewing the criteria necessary to classify two chromosomes as a homologous pair:

1. During mitosis and meiosis, when chromosomes are visible as distinct figures, both members of a homologous pair are the same size and exhibit identical centromere locations. The sex chromosomes are an exception.
2. During early stages of meiosis, homologous chromosomes form pairs, or synapse.
3. Although not generally microscopically visible, homologs contain identical, linearly ordered gene loci.

### 3.6 Independent Assortment Leads to Extensive Genetic Variation

One major consequence of independent assortment is the production by an individual of genetically dissimilar gametes. Genetic variation results because the two members of any homologous pair of chromosomes are rarely, if ever, genetically identical. Therefore, because independent assortment leads to the production of all possible chromosome combinations, extensive genetic diversity results.

We have seen that the number of possible gametes, each with different chromosome compositions, is \(2^n\), where \(n\) equals the haploid number. Thus, if a species has haploid number of \(n = 4\), then \(2^4 = 16\) different gamete combinations can be formed as a result of independent assortment. Although this number is not high, consider the human species, where \(n = 23\). If \(2^{23}\) is calculated, we find that in excess of \(8 \times 10^6\), or over 8 million, different types of gametes are represented. Because fertilization represents an event involving only one of approximately \(8 \times 10^6\) possible gametes from each of two parents, each offspring represents only one of \((8 \times 10^6)^2\), or one of only \(64 \times 10^{12}\) potential genetic combinations! No wonder that, except for identical twins, each member of the human species demonstrates a distinctive appearance and individuality—this number of combinations is far greater than the number of humans who have ever lived on Earth! Genetic variation resulting from independent assortment has been extremely important to the process of evolution in all sexually reproducing organisms.

### 3.7 Laws of Probability Help to Explain Genetic Events

Recall that genetic ratios are expressed as probabilities—for example, \(3/4\) tall:1/4 dwarf. These values predict the outcome of each fertilization event, such that the probability of each zygote having the genetic potential for becoming tall is \(3/4\), while the potential for becoming dwarf is 1/4. Probabilities range from 0.0, when an event is certain not to occur, to 1.0, when an event is certain to occur. When two or more events occur independently but at the same time, we can calculate the probability of possible outcomes when they occur together. This is accomplished by applying the product law—the probability of two or more events occurring simultaneously is equal to the product of their individual probabilities. Two or more events are independent of one another if the outcome of one does not affect the outcome of any of the others under consideration.

To illustrate the product law, consider the possible results if you toss a penny \((P)\) and a nickel \((N)\) at the same time and examine all combinations of heads \((H)\) and tails \((T)\) that can occur. There are four possible outcomes:

\[
(P_H, N_H) = (1/2)(1/2) = 1/4 \\
(P_T, N_H) = (1/2)(1/2) = 1/4 \\
(P_H, N_T) = (1/2)(1/2) = 1/4 \\
(P_T, N_T) = (1/2)(1/2) = 1/4
\]

The probability of obtaining a head or a tail in the toss of either coin is 1/2 and is unrelated to the outcome of the toss of the other coin. Thus, all four possible combinations are predicted to occur with equal probability.

If we want to calculate the probability where the possible outcomes of two events are independent of one another but can be accomplished in more than one way, we apply the sum law. For example, what is the probability of tossing our penny and nickel and obtaining one head and one tail? In such a case, we do not care whether it is the penny or the nickel that comes up heads, provided the other coin has the alternative outcome. As we saw above, there are two ways in which the desired outcome can be accomplished, each with a probability of 1/4.
Thus, according to the sum law, the overall probability is equal to
\[ (1/4) + (1/4) = 1/2 \]
One-half of all coin tosses are predicted to yield the desired outcome.

These simple probability laws will be useful throughout our discussions of transmission genetics and for solving genetics problems. In fact, we already applied the product law when we used the forked-line method to calculate the phenotypic results of Mendel’s dihybrid and trihybrid crosses. When we wish to know the results of a cross, we need only calculate the probability of each possible outcome. The results of this calculation then allow us to predict the proportion of offspring expressing each phenotype or each genotype.

An important point to remember when you deal with probability is that predictions of possible outcomes are based on large sample sizes. If we predict that 9/16 of the offspring of a dihybrid cross will express both dominant traits, it is very unlikely that, in a small sample, exactly 9 of every 16 offspring will express this phenotype. Instead, our prediction is that, of a large number of offspring, approximately 9/16 of them will do so. The deviation from the predicted ratio in smaller sample sizes is attributed to chance, a subject we examine in our discussion of statistics in the next section. As you shall see, the impact of deviation due strictly to chance diminishes as the sample size increases.

### 3.8 Chi-Square Analysis Evaluates the Influence of Chance on Genetic Data

Mendel’s 3:1 monohybrid and 9:3:3:1 dihybrid ratios are hypothetical predictions based on the following assumptions: (1) Each allele is dominant or recessive; (2) segregation is operative; (3) independent assortment occurs; and (4) fertilization is random. The final two assumptions are influenced operative; (3) independent assortment occurs; and (4) fertilization is random. The final two assumptions are influenced

The concept of chance deviation is most easily illustrated by tossing a single coin numerous times and recording the number of heads and tails observed. In each toss, there is a probability of 1/2 that a head will occur and a probability of 1/2 that a tail will occur. Therefore, the expected ratio of many tosses is 1:1. If a coin is tossed 1000 times, usually about 500 heads and 500 tails will be observed. Any reasonable fluctuation from this hypothetical ratio (e.g., 486 heads and 514 tails) is attributed to chance.

As the total number of tosses is reduced, the impact of chance deviation increases. For example, if a coin is tossed only four times, you would not be too surprised if all four tosses result in only heads or only tails. For 1000 tosses, however, 1000 heads or 1000 tails would be most unexpected. In fact, you might believe that such a result would be impossible. Actually, all heads or all tails in 1000 tosses can be predicted to occur with a probability of \((1/2)^{1000}\). Since \((1/2)^{20}\) is equivalent to less than one in a million times, an event occurring with a probability of \((1/2)^{1000}\) is virtually impossible. Two major points are significant before we consider chi-square analysis:

1. The outcomes of independent assortment and fertilization, like coin tossing, are subject to random fluctuations from their predicted occurrences as a result of chance deviation.

2. As the sample size increases, the average deviation from the expected results decreases. Therefore, a larger sample size diminishes the impact of chance deviation on the final outcome.

### Chi-Square Calculations and the Null Hypothesis

In genetics, the ability to evaluate observed deviation is a crucial skill. When we assume that data will fit a given ratio such as 1:1, 3:1, or 9:3:3:1, we establish what is called the null hypothesis \((H_0)\). It is so named because the hypothesis assumes that no real difference exists between measured values (or ratio) and predicted values (or ratio). The apparent difference can be attributed purely to chance. The null hypothesis is evaluated using statistical analysis. On this basis, the null hypothesis may either (1) be rejected or (2) fail to be rejected. If it is rejected, the observed deviation from the expected result is not attributed to chance alone. The null hypothesis and the underlying assumptions leading to it must be reexamined. If the null hypothesis fails to be rejected, any observed deviations are attributed to chance.

One of the simplest statistical tests devised to assess the null hypothesis is chi-square (\(\chi^2\)) analysis. This test takes into account the observed deviation in each component of an expected ratio as well as the sample size and reduces them to a single numerical value. The value for \(\chi^2\) is then used to estimate how frequently the observed deviation can be expected to occur strictly as a result of chance. The formula for chi-square analysis is

\[
\chi^2 = \sum \frac{(o - e)^2}{e}
\]

where \(o\) is the observed value for a given category, \(e\) is the expected value for that category, and \(\sum\) (the Greek letter sigma) represents the sum of the calculated values for each category of the ratio. Because \((o - e)\) is the deviation \((d)\) in each case, the equation reduces to

\[
\chi^2 = \sum \frac{d^2}{e}
\]

Table 3.1(a) shows a \(\chi^2\) calculation for the \(F_2\) results of a hypothetical monohybrid cross. To analyze these data, you work from left to right, calculating and entering the appropriate numbers in each column. Regardless of whether the deviation \(d\) is positive or negative, \(d^2\) always becomes positive after the number is squared. In Table 3.1(b) the \(F_2\) results of a hypothetical dihybrid cross are analyzed. Be sure that you understand how each number was calculated in the dihybrid example.
The final step in chi-square analysis is to interpret the $\chi^2$ value. To do so, you must initially determine the value of the **degrees of freedom (df)**, which is equal to $n - 1$, where $n$ is the number of different categories into which each datum point may fall. For the 3:1 ratio, $n = 2$, so $df = 2 - 1 = 1$. For the 9:3:3:1 ratio, $n = 4$ and $df = 3$. Degrees of freedom must be taken into account because the greater the number of categories, the more deviation is expected as a result of chance.

Once you have determined the degrees of freedom, we can interpret the $\chi^2$ value in terms of a corresponding **probability value (p)**. Since this calculation is complex, we usually take the $p$ value from a standard table or graph. Figure 3–11 shows a wide range of $\chi^2$ and $p$ values for various degrees of freedom in both a graph and a table. Let’s use the graph to determine the $p$ value. The caption for Figure 3–11(b) explains how to use the table.

To determine $p$, execute the following steps:

1. Locate the $\chi^2$ value on the abscissa (the horizontal or $x$-axis).
2. Draw a vertical line from this point up to the angled line on the graph representing the appropriate $df$.
3. Extend a horizontal line from this point to the left until it intersects the ordinate (the vertical or $y$-axis).
4. Estimate, by interpolation, the corresponding $p$ value.

We used these steps for the monohybrid cross in Table 3.1(a) to estimate the $p$ value of 0.48 shown in Figure 3–11(a). For the dihybrid cross, try this method to see if you can determine the $p$ value. Since the $\chi^2$ value is 4.16 and $df = 3$, an approximate $p$ value is 0.26. Checking this result in the table confirms that $p$ values for both the monohybrid and dihybrid crosses are between 0.20 and 0.50.

### Interpreting Probability Values

So far, we have been concerned with calculating $\chi^2$ values and determining the corresponding $p$ values. The most important aspect of chi-square analysis is understanding the meaning of the $p$ value. Let’s use the example of the dihybrid cross in Table 3.1(b) ($p = 0.26$). In these discussions, it is simplest to think of the $p$ value as a percentage (e.g., $0.26 = 26\%$). In our example, the $p$ value indicates that if we repeat the same experiment many times, 26 percent of the trials would be expected to exhibit chance deviation as great or greater than that seen in the initial trial. Conversely, 74 percent of the trials would show less deviation than initially observed as a result of chance. Thus, the $p$ value reveals that a hypothesis (the 9:3:3:1 ratio in this case) is never proved or disproved absolutely. Instead, a relative standard is set that enables us to either reject or fail to reject the null hypothesis—this standard is most often a $p$ value of 0.05. When applied to chi-square analysis, a $p$ value less than 0.05 means that the observed deviation in the set of results will be obtained by chance alone less than 5 percent of the time. Such a $p$ value indicates that the difference between the observed and predicted results is substantial and thus enables us to reject the null hypothesis. On the other hand, $p$ values of 0.05 or greater (0.05 to 1.0) indicate that the observed deviation will be obtained by chance alone 5 percent or more of the time. The conclusion is not to reject the null hypothesis. Thus, for the $p$ value of 0.26, assessing the hypothesis that independent assortment accounts for the results fails to be rejected. Therefore, the observed deviation can be reasonably attributed to chance.
Thus, we fail to reject the null hypothesis.

freedom, the corresponding (lighter blue areas) justify rejecting the null hypothesis. For example, using the table in part (b), where for 1 degree of

than wingless flies. Rejection of the null hypothesis

wings. As a result, when the data are gathered, there are fewer young adults, compared to flies whose genotype gives rise to

mutant embryos may survive their preadult development or as

mutant wingless fly zygotes. However, not as many of the

mutations of segregation and independent assortment, because

other factors are operative.

A final note is relevant here for the case where the null hypothesis is rejected, that is, where \( p \leq 0.05 \). Suppose we are testing the null hypothesis that the data represented a 9:3:3:1 ratio, indicative of independent assortment. If the null hypothesis is rejected, what are alternative interpretations of the data? Researchers will reassess the assumptions that underlie the null hypothesis. In our example, we assumed that segregation operates faithfully for both gene types—that is, that all gametes are equally likely to participate in fertilization. Finally, following fertilization, we assumed that all preadult stages and adult offspring are equally viable regardless of their genotype. If any of these assumptions is incorrect, the original hypothesis is not necessarily invalid.

An example will clarify this. Suppose our null hypothesis is that a dihybrid cross between fruit flies will result in 3/16 mutant wingless fly zygotes. However, not as many of the mutant embryos may survive their preadult development or as young adults, compared to flies whose genotype gives rise to wings. As a result, when the data are gathered, there are fewer than 3/16 wingless flies. Rejection of the null hypothesis alone is not cause for us to disregard the validity of the postulates of segregation and independent assortment, because other factors are operative.

A final note is relevant here for the case where the null hypothesis is rejected, that is, where \( p \leq 0.05 \). Suppose we are testing the null hypothesis that the data represented a 9:3:3:1 ratio, indicative of independent assortment. If the null hypothesis is rejected, that is, where Suppose we

Problem 18 on page 59 asks you to apply \( \chi^2 \) analysis to a set of data and determine whether the data fit several ratios.

Hint: In calculating \( \chi^2 \), first determine the expected outcomes using the predicted ratios. Then follow a stepwise approach, determining the deviation in each case, and calculating \( d^2/e \) for each category.

We now explore how to determine the mode of inheritance of phenotypes in humans, where designed crosses are not possible and where relatively few offspring are available for study. The traditional way to study inheritance has been to construct a family tree, indicating the presence or absence of the trait in question for each member of each generation. Such a family tree is called a pedigree. By analyzing a pedigree, we may be able to predict how the trait under study is inherited—for example, is it due to a dominant or recessive allele? When many pedigrees for the same trait are studied, we can often ascertain the mode of inheritance.
3.9 Pedigrees Reveal Patterns of Inheritance of Human Traits

Pedigree Conventions

A variety of conventions are commonly used in pedigree construction. Figure 3–12 illustrates a number of such conventions: circles represent females and squares designate males. Parents are connected by a single horizontal line, and vertical lines lead to their offspring. If the parents are related (consanguineous), such as first cousins, they are connected by a double line. Offspring are called sibs (short for siblings) and are connected by a horizontal sibship line. Sibs are placed from left to right according to birth order, and are labeled with Arabic numerals. Each generation is indicated by a Roman numeral. If the sex of an individual is unknown, a diamond is used. When a pedigree traces only a single trait, the circles, squares, and diamonds are shaded if the phenotype being considered is expressed and unshaded if not. In some pedigrees, those individuals that fail to express a recessive trait, but are known with certainty to be a heterozygous carrier, have a shaded dot within their unshaded circle or square. If an individual is deceased and the phenotype is unknown, a diagonal line is placed over the circle or square.

Twins are indicated by diagonal lines stemming from a vertical line connected to the sibship line. For identical (or monozygotic) twins, the diagonal lines are linked by a horizontal line. Fraternal (or dizygotic) twins lack this connecting line. A number within one of the symbols represents numerous sibs of the same or unknown phenotypes. The individual whose phenotype first brought attention to the investigation and construction of the pedigree is called the proband and is indicated by an arrow connected to the designation p. This term applies to either a male or a female.

Pedigree Analysis

In Figure 3–13, two pedigrees are shown. The first illustrates a representative pedigree for a trait that demonstrates autosomal recessive inheritance, such as albinism. The male parent of the first generation (I-1) is affected. Characteristic of a rare recessive trait with an affected parent, the trait “appears” in the offspring of the next generation. Assuming recessiveness, we might predict that the unaffected female parent (I-2) is a homozygous normal individual because none of the offspring show the disorder. Had she been heterozygous, one half of the offspring would be expected to exhibit albinism, but none do. However, such a small sample (three offspring) prevents us from knowing for certain.

Further evidence supports the prediction of a recessive trait. If albinism were inherited as a dominant trait, individual II-3 would have to express the disorder in order to pass it to his offspring (III-3 and III-4), but he does not. Inspection of the offspring constituting the third generation (row III) provides further support for the hypothesis that albinism is a recessive trait. If it is, parents II-3 and II-4 are both heterozygous, and approximately one-fourth of their offspring should be affected. Two of the six offspring do show albinism. This deviation from the expected ratio is not unexpected in crosses with few offspring. Once we are confident that albinism is inherited as an autosomal recessive trait, we could portray the II-3 and II-4 individuals with a shaded dot within their larger square and circle.

Finally, we can note that, characteristic of pedigrees for autosomal traits, both males and females are affected with equal probability. In Chapter 4, we will examine a pedigree representing a gene located on the sex-determining X chromosome. We will see certain limitations imposed on the transmission of X-linked traits, such as that these traits are more prevalent in male offspring and are never passed from affected fathers to their sons.

The second pedigree illustrates the pattern of inheritance for a trait such as Huntington disease, which is caused by an autosomal dominant allele. The key to identifying such a pedigree that reflects a dominant trait is that all affected offspring will have a parent that also expresses the trait. It is also possible, by chance, that none of the offspring will inherit the dominant allele. If so, the trait will cease to exist in future generations. Like recessive traits, provided that the gene is autosomal, both males and females are equally affected.

When autosomal dominant diseases are rare within the population, and most are, then it is highly unlikely that affected individuals will inherit a copy of the mutant gene from both parents. Therefore, in most cases, affected individuals are heterozygous for the dominant allele. As a result, approximately one-half of the offspring inherit it. This is borne out in the second pedigree in Figure 3–13. Furthermore, if a mutation is dominant, and a single copy is sufficient to produce a mutant phenotype, homozygotes are
likely to be even more severely affected, perhaps even failing to survive. An illustration of this is the dominant gene for familial hypercholesterolemia. Heterozygotes display a defect in their receptors for low density lipoproteins, the so-called LDLs. As a result, too little cholesterol is taken up by cells from the blood, and elevated plasma levels of LDLs result. Such heterozygous individuals have heart attacks during the fourth decade of their life, or before. While heterozygotes have LDL levels about double that of a normal individual, rare homozygotes have been detected. They lack LDL receptors altogether, and have LDL levels nearly 10 times above the normal range. They are likely to have a heart attack very early in life, even before age five, and almost inevitably before they reach the age of 20.

Pedigree analysis of many traits has historically been an extremely valuable research technique in human genetic studies. However, the approach does not usually provide the certainty in drawing conclusions afforded by designed crosses yielding large numbers of offspring. Nevertheless, when many independent pedigrees of the same trait or disorder are analyzed, consistent conclusions can often be drawn. Table 3.2 lists numerous human traits and classifies them according to their recessive or dominant expression. The genes controlling some of these traits are located on the sex-determining chromosomes. We will discuss pedigrees for X-linked traits in Chapter 4.

### TABLE 3.2 Representative Recessive and Dominant Human Traits

<table>
<thead>
<tr>
<th>Recessive Traits</th>
<th>Dominant Traits</th>
</tr>
</thead>
<tbody>
<tr>
<td>Albinism</td>
<td>Achondroplasia</td>
</tr>
<tr>
<td>Alkaptonuria</td>
<td>Brachydactyly</td>
</tr>
<tr>
<td>Ataxia telangiectasia</td>
<td>Congenital stationary night blindness</td>
</tr>
<tr>
<td>Color blindness</td>
<td>Ehler-Danlos syndrome</td>
</tr>
<tr>
<td>Cystic fibrosis</td>
<td>Hypotrichosis</td>
</tr>
<tr>
<td>Duchenne muscular dystrophy</td>
<td>Huntington disease</td>
</tr>
<tr>
<td>Galactosemia</td>
<td>Hypercholesterolemia</td>
</tr>
<tr>
<td>Hemophilia</td>
<td>Marfan syndrome</td>
</tr>
<tr>
<td>Lesch–Nyhan syndrome</td>
<td>Neurofibromatosis</td>
</tr>
<tr>
<td>Phenylketonuria</td>
<td>Phenylthiocarbamide (PTC) tasting</td>
</tr>
<tr>
<td>Sickle-cell anemia</td>
<td>Porphyria</td>
</tr>
<tr>
<td>Tay–Sachs disease</td>
<td>Widow’s peak</td>
</tr>
</tbody>
</table>

**FIGURE 3-13** Representative pedigrees for two characteristics, each followed through three generations.

- **(a) Autosomal Recessive Trait**
  - Either I-3 or I-4 must be heterozygous
  - Recessive traits typically skip generations
  - Recessive autosomal traits appear equally in both sexes

- **(b) Autosomal Dominant Trait**
  - I-1 is heterozygous for a dominant allele
  - Dominant traits almost always appear in each generation.
  - Affected individuals all have an affected parent.
  - Dominant autosomal traits appear equally in both sexes.

**Web Tutorial 3.4**

Analysis of Human Pedigrees

Now Solve This

Problem 26 on page 59 asks you to examine a pedigree for myopia and predict whether the trait is dominant or recessive.

**Hint:** One of the first things to look for are individuals who express the trait, but neither of whose parents also expresses the trait. Such an observation makes it highly unlikely that the trait is dominant.
5. The discovery of chromosomes in the late 1800s, along with Mendel's postulate of independent assortment states that each chromosome in the human genome consists of two homologous chromosomes. This discovery led to the understanding that traits are inherited in pairs rather than individually. The study of genetics has since expanded to include the molecular basis of these traits, including the identification of genes and the mechanisms by which they are passed from one generation to the next.

6. The Punnett square and the forked-line methods are used to predict the probabilities of phenotypes (and genotypes) from crosses involving two or more gene pairs. Genetic ratios are expressed as probabilities. Thus, deriving outcomes of genetic crosses requires an understanding of the laws of probability.

7. Statistical analysis is used to test the validity of experimental outcomes. In genetics, variations from the expected ratios due to chance deviations can be anticipated. Chi-square analysis allows us to assess the null hypothesis, which states that there is no real difference between the expected and observed values. As such, it tests the probability of whether observed variations can be attributed to chance deviation.

8. Pedigree analysis is a method for studying the inheritance pattern of human traits over several generations. It frequently provides the basis for determining the mode of inheritance of human characteristics and disorders.
KEY TERMS

albinism, 53
allele, 41
branch diagram, 46
chance deviation, 50
chi-square (χ²) analysis, 50
chromosomal theory of inheritance, 47
consanguineous, 53
continuous variation, 47
degrees of freedom (df), 51
dihybrid cross, 43
dinozygotic twins, 53
diploid number (2n), 48
discontinuous variation, 47
dizygotic twins, 53
dominant, 41
familial hypercholesterolemia, 54
F₁ generation (first filial), 40
forked-line method, 46
fraternal twins, 53
F₂ generation (second filial), 40
gene, 41
genotype, 41
heterozygote, 41
hexosaminidase A (Hex-A), 55
homozygote, 41
identical twins, 53
independent assortment, 45
locus, 49
maternal parent, 49
Mendel’s 9:3:3:1 dihybrid ratio, 46
monohybrid cross, 39
monozygotic twins, 53
null hypothesis, 50
paternal parent, 49
pedigree, 52
phenotype, 41
P₁ generation (parental), 40
probability value (p), 51
proband, 53
product law, 44
Punnett square, 42
recessive, 41
reciprocal cross, 40
segregation, 41
selfing, 40
sib (sibling), 53
sibship line, 53
starch-branching enzyme (SBEI), 43
sum law, 49
Tay-Sachs disease, 55
testcross, 43
traits, 39
transmission genetics, 39
transposable elements, 43
trihybrid cross, 46
unit factor, 41
unit of inheritance, 39

INSIGHTS AND SOLUTIONS

As a student, you will be asked to demonstrate your knowledge of transmission genetics by solving genetics problems. Success at this task represents not only comprehension of theory but its application to more practical genetic situations. Most students find problem solving in genetics to be challenging and rewarding. This section will provide you with basic insights into the reasoning essential to this process.

Genetics problems are in many ways similar to word problems in algebra. The approach to solving them is identical: (1) Analyze the problem carefully; (2) translate words into symbols, first defining each one; and (3) choose and apply a specific technique to solve the problem. The first two steps are critical. The third step is largely mechanical.

The simplest problems state all necessary information about the P₁ generation and ask you to find the expected ratios of the F₁ and F₂ genotypes and/or phenotypes. Always follow these steps when you encounter this type of problem:

1. Determine insofar as possible the genotypes of the individuals in the P₁ generation.
2. Determine what gametes may be formed by the P₁ parents.
3. Recombine gametes by the Punnett square or the forked-line methods, or if the situation is very simple, by inspection. Read the F₁ phenotypes.
4. Repeat the process to obtain information about the F₂ generation.

Determining the genotypes from the given information requires that you understand the basic theory of transmission genetics.

Consider this problem: A recessive mutant allele, black, causes a very dark body in Drosophila (a fruit fly) when homozygous. The wild-type (normal) color is gray. What F₁ phenotypic ratio is predicted when a black female is crossed with a gray male whose father was black?

To work out this problem, you must understand dominance and recessiveness, as well as the principle of segregation. Furthermore, you must use the information about the male parent’s father. Here is one way to solve this problem:

1. The female parent is black, so she must be homozygous for the mutant allele (bb).
2. The male parent is gray; therefore, he must have at least one dominant allele (B). His father was black (bb), and he received one of the chromosomes bearing these alleles, so the male parent must be heterozygous (Bb).
3. With this information, the problem is simple:
Apply this approach to the following problems.

1. Mendel found that full pods are dominant over constricted pods while round seeds are dominant over wrinkled seeds. One of his crosses was between full, round plants and constricted, wrinkled plants. From this cross, he obtained an F1 generation that was all full and round. In the F2 generation, Mendel obtained his classic 9:3:3:1 ratio. Using this information, determine the expected F1 and F2 results of a cross between homozygous constricted, round plants and full, wrinkled plants.

**Solution:** Define gene symbols for each of contrasting traits. Use the lowercase first letter of the recessive traits to designate those phenotypes and the uppercase first letter to designate the dominant traits. Thus, C and c indicate full and constricted, and W and w indicate round and wrinkled phenotypes, respectively.

Determine the genotypes of the F1 generation, form the gametes, reconstitute the F1 generation, and read off the phenotype(s):

\[
\begin{align*}
P1: & \quad ccWW \quad \times \quad CCww \\
\text{Gametes:} & \quad cW \quad \text{CcWw} \\
F1: & \quad CcWw \quad \text{full, round}
\end{align*}
\]

You can immediately see that the F1 generation expresses both dominant phenotypes and is heterozygous for both gene pairs. Thus, you expect that the F2 generation will yield the classic Mendelian ratio of 9:3:3:1. Let’s work it out anyway to confirm this, using the forked-line method. Both gene pairs are heterozygous and can be expected to assort independently, so we can predict the F2 outcomes from each gene pair separately and then proceed with the forked-line method.

Every F2 offspring is subject to the following probabilities:

\[
\begin{align*}
Cc \times Cc & \quad Ww \times Ww \\
\downarrow & \quad \downarrow \\
CC & \quad WW \\
Cc & \quad WW \quad \text{full, round} \\
cC & \quad Ww \quad \text{round} \\
cc & \quad ccww \quad \text{constricted, wrinkled}
\end{align*}
\]

The forked-line method then confirms the 9:3:3:1 phenotypic ratio. Remember that this represents proportions of 9:16/3:16: 3/16:1/16. Note that we are applying the product law as we compute the final probabilities:

\[
\begin{align*}
3/4 \text{ round} & \quad \frac{3/4}{3/4} \quad 9/16 \text{ full, round} \\
1/4 \text{ wrinkled} & \quad \frac{3/4}{1/4} \quad 3/16 \text{ full, wrinkled} \\
3/4 \text{ round} & \quad \frac{1/4}{3/4} \quad 3/16 \text{ constricted, round} \\
1/4 \text{ constricted} & \quad \frac{1/4}{1/4} \quad 1/16 \text{ constricted, wrinkled}
\end{align*}
\]

2. In another cross involving parent plants of unknown genotype and phenotype, the following offspring were obtained.

<table>
<thead>
<tr>
<th>F1</th>
<th>3/8 full, round</th>
<th>3/8 full, wrinkled</th>
<th>1/8 constricted, round</th>
<th>1/8 constricted, wrinkled</th>
</tr>
</thead>
</table>

Determine the genotypes and phenotypes of the parents.

**Solution:** This problem is more difficult and requires keener insight because you must work backward. The best approach is to consider the outcomes of pod shape separately from those of seed texture. Of all the plants, 3/8 + 3/8 = 3/4 are full and 1/8 + 1/8 = 1/4 are constricted. Of the various genotypic combinations that can serve as parents, which combination will give rise to a ratio of 3/4:1/4? This ratio is identical to Mendel’s monohybrid F2 results, and we can propose that both unknown parents share the same genetic characteristic as the monohybrid F1 parents; they must both be heterozygous for the genes controlling pod shape and thus areCc.

Before we accept this hypothesis, let’s consider the possible genotypic combinations that control seed texture. If we consider this characteristic alone, we see that the traits are expressed in a ratio of 3/8 + 1/8 = 1/2 round: 3/8 + 1/8 = 1/2 wrinkled. To generate such a ratio, the parents cannot both be heterozygous, or their offspring would yield a 3/4 : 1/4 phenotypic ratio. They cannot both be homozygous, or all of their offspring would express a single phenotype. Thus, we are left with testing the hypothesis that one parent is homozygous and one is heterozygous for the alleles controlling texture. The potential case of Ww × Ww does not work, since it yields only a single phenotype. This leaves us with the potential case of CcWw × Ww. Offspring in such a mating will yield 1/2 Ww (round): 1/2 ww (winkled), exactly the outcome we are seeking.

Now, let’s combine the hypotheses and predict the outcome of the cross. In our solution, we use a dash (—) to indicate that the second allele may be either dominant or recessive, since we are only predicting phenotypes.

\[
\begin{align*}
3/4 \ Cc & \rightarrow 3/8 \ C–Ww \quad \text{full, round} \\
1/2 \ ww & \rightarrow 3/8 \ C–ww \quad \text{full, wrinkled} \\
1/2 \ Ww & \rightarrow 1/8 \ ccWw \quad \text{constricted, round} \\
1/4 \ cc & \rightarrow 1/8 \ ccww \quad \text{constricted, wrinkled}
\end{align*}
\]

As you can see, this cross produces offspring according to our initial information, and we have solved the problem. Note that in this solution, we used genotypes in the forked-line method, in contrast to the use of phenotypes in the earlier solution.

3. Determine the probability that a plant of genotype CcWw will be produced from parental plants with the genotypes CcWw and CcWw.

**Solution:** The two gene pairs demonstrate straightforward dominance and recessiveness and assort independently during gamete formation. We need only calculate the individual probabilities of obtaining the two separate outcomes (Cc and Ww) and apply the product law to calculate the final probability:

\[
\begin{align*}
Cc \times Cc & \rightarrow 1/4 \ Cc : 1/2 \ Cc : 1/4 \ cc \\
Ww \times Ww & \rightarrow 1/2 \ Ww : 1/2 \ Ww \\
p & = (1/2 \ Cc)(1/2 \ Ww) = 1/4 \ CcWw
\end{align*}
\]

4. In the laboratory, a genetics student crossed flies that had normal, long wings with flies expressing the dumpy mutation (truncated wings), which she believed was a recessive trait. In the F1 generation, all flies had long wings. The following results were obtained in the F2 generation:

- 792 long-winged flies
- 208 dumpy-winged flies
When working out genetics problems in this and succeeding chapters, always assume that members of the P1 generation are homozygous, unless the information given, or the data, indicates otherwise.

1. In a cross between a black and a white guinea pig, all members of the F1 generation are black. The F2 generation is made up of approximately black and white guinea pigs. Diagram this cross, and show the genotypes and phenotypes.
2. Albinism in humans is inherited as a simple recessive trait. Determine the genotypes of the parents and offspring for the following families. When two alternative genotypes are possible, list both. (a) Two nonalbino (normal) parents have five children, four normal and one albino. (b) A normal male and an albino female have six children, all normal.
3. In a problem involving albinism (see Problem 2), which of Mendel’s postulates are demonstrated?
4. Why was the garden pea a good choice as an experimental organism in Mendel’s work?
5. Pigeons exhibit a checkered or plain feather pattern. In a series of controlled matings, the following data were obtained:

<table>
<thead>
<tr>
<th>F1 Progeny</th>
<th>Checkered</th>
<th>Plain</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1 Cross</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(a) checkered × checkered</td>
<td>36</td>
<td>0</td>
</tr>
<tr>
<td>(b) checkered × plain</td>
<td>38</td>
<td>0</td>
</tr>
<tr>
<td>(c) plain × plain</td>
<td>0</td>
<td>35</td>
</tr>
</tbody>
</table>

Then F1 offspring were selectively mated with the following results. (The P1 cross giving rise to each F1 pigeon is indicated in parentheses.)

<table>
<thead>
<tr>
<th>F2 Progeny</th>
<th>Checkered</th>
<th>Plain</th>
</tr>
</thead>
<tbody>
<tr>
<td>F1 × F1 Crosses</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(d) checkered (a) × plain (c)</td>
<td>34</td>
<td>0</td>
</tr>
<tr>
<td>(e) checkered (b) × plain (c)</td>
<td>17</td>
<td>14</td>
</tr>
<tr>
<td>(f) checkered (b) × checkered (b)</td>
<td>28</td>
<td>9</td>
</tr>
<tr>
<td>(g) checkered (a) × checkered (b)</td>
<td>39</td>
<td>0</td>
</tr>
</tbody>
</table>

How are the checkered and plain patterns inherited? Predict the results of the F1 × F1 mating from cross (b).

6. Mendel crossed peas having round seeds and yellow cotyledons with peas having wrinkled seeds and green cotyledons. All the F1 plants had round seeds with yellow cotyledons. Diagram this cross through the F2 generation, using both the Punnett square and forked-line methods.
7. Determine the genotypes of the parental plants by analyzing the phenotypes of the offspring from these crosses:

<table>
<thead>
<tr>
<th>Parental Plants</th>
<th>Offspring</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a) round, yellow × round, yellow</td>
<td>3/4 round, yellow</td>
</tr>
<tr>
<td></td>
<td>1/4 wrinkled, yellow</td>
</tr>
<tr>
<td>(b) round, yellow × wrinkled, yellow</td>
<td>6/16 wrinkled, yellow</td>
</tr>
<tr>
<td></td>
<td>2/16 wrinkled, green</td>
</tr>
<tr>
<td></td>
<td>6/16 round, yellow</td>
</tr>
<tr>
<td></td>
<td>2/16 round, green</td>
</tr>
<tr>
<td>(c) round, yellow × wrinkled, green</td>
<td>1/4 round, yellow</td>
</tr>
<tr>
<td></td>
<td>1/4 round, green</td>
</tr>
<tr>
<td></td>
<td>1/4 wrinkled, yellow</td>
</tr>
<tr>
<td></td>
<td>1/4 wrinkled, green</td>
</tr>
</tbody>
</table>

8. Are any of the crosses in Problem 7 testcrosses? If so, which one(s)?

9. Which of Mendel’s postulates can be demonstrated in the crosses of Problem 7 but not in those in Problems 1 and 5? State this postulate.

10. Correlate Mendel’s four postulates with what is now known about homologous chromosomes, genes, alleles, and the process of meiosis.

11. What is the basis for homology among chromosomes?

12. Distinguish between homozygosity and heterozygosity.

13. In *Drosophila*, gray body color is dominant over ebony body color, while long wings are dominant over vestigial wings. Work the following crosses through the F2 generation, and determine the genotypic and phenotypic ratios for each generation. Assume that the P1 individuals are homozygous:

   (a) gray, long × ebony, vestigial
   (b) gray, vestigial × ebony, long
   (c) gray, long × gray, vestigial

14. How many different types of gametes can be formed by individuals of the following genotypes? What are they in each case? (a) AaBb, (b) AaBB, (c) AaBbCc, (d) AaBBCc, (e) AaBbcc, and (f) AaBbCcDdEe?

15. Using the forked-line method, determine the genotypic and phenotypic ratios of these trihybrid crosses:

   (a) AaBbCc × AaBBCc, (b) AaBBCc × aBBCc, and (c) AaBbCc × AaBbCc.

16. Mendel crossed peas with round, green seeds with peas having wrinkled, yellow seeds. All F1 plants had seeds that were round and yellow. Predict the results of testcrossing these F1 plants.

17. Shown are F2 results of two of Mendel’s monohybrid crosses. State a null hypothesis that you will test using chi-square analysis. Calculate the χ² value and determine the p value for both crosses, then interpret the p values. Which cross shows a greater amount of deviation?

   | (a) Full pods | 882 |
   | Constricted pods | 299 |
   | (b) Violet flowers | 705 |
   | White flowers | 224 |

18. In one of Mendel’s dihybrid crosses, he observed 315 round, yellow; 108 round, green; 101 wrinkled, yellow; and 32 wrinkled, green F2 plants. Analyze these data using chi-square analysis to see whether (a) they fit a 9:3:3:1 ratio; (b) the round, wrinkled traits fit a 3:1 ratio; or (c) the yellow, green traits fit a 3:1 ratio.

19. A geneticist, in assessing data that fell into two phenotypic classes, observed values of 250:150. He decided to perform chi-square analysis using two different null hypotheses: (a) The data fit a 3:1 ratio; and (b) the data fit a 1:1 ratio. Calculate the χ² values for each hypothesis. What can you conclude about each hypothesis?

20. The basis for rejecting any null hypothesis is arbitrary. The researcher can set more or less stringent standards by deciding to raise or lower the critical p value. Would the use of a standard of p = 0.10 be more or less stringent in failing to reject the null hypothesis? Explain.

21. Consider three independently assorting gene pairs, *Aa*, *Bb*, and *Cc*, where each demonstrates typical dominance (*A–, B–, C–*), and recessiveness (*aa, bb, cc*). What is the probability of obtaining an offspring that is *AaBbCc* from parents that are *AaBbCc* and *AaBbCc*?

22. What is the probability of obtaining a triply recessive individual from the parents shown in Problem 21?

23. Of all offspring of the parents in Problem 21, what proportion will express all three dominant traits?

24. For the following pedigree, predict the mode of inheritance and the resulting genotypes of each individual. Assume that the alleles *A* and *a* control the expression of the trait.

25. Which of Mendel’s postulates are demonstrated by the pedigree in Problem 24? List and define these postulates.

26. The following pedigree follows the inheritance of myopia (near-sightedness) in humans. Predict whether the disorder is inherited as a dominant or a recessive trait. Based on your prediction, indicate the most probable genotype for each individual.

27. Draw all possible conclusions concerning the mode of inheritance of the trait expressed in each of the following limited pedigrees. (Each case is based on a different trait.)
28. Two true-breeding pea plants are crossed. One parent is round, terminal, violet, constricted, while the other expresses the contrasting phenotypes of wrinkled, axial, white, full. The four pairs of contrasting traits are controlled by four genes, each located on a separate chromosome. In the F1 generation, only round, axial, violet, and full are expressed. In the F2 generation, all possible combinations of these traits are expressed in ratios consistent with Mendelian inheritance.

(a) What conclusion can you draw about the inheritance of these traits based on the F1 results?
(b) Which phenotype appears most frequently in the F2 results? Write a mathematical expression that predicts the frequency of occurrence of this phenotype.
(c) Which F2 phenotype is expected to occur least frequently? Write a mathematical expression that predicts this frequency.
(d) How often is either P1 phenotype likely to occur in the F2 generation?
(e) If the F1 plant is testcrossed, how many different phenotypes will be produced, and how does this number compare to the number of different phenotypes in the F2 generation discussed in part (b)?

29. Tay-Sachs disease (TSD) is an inborn error of metabolism that results in death, usually before the age of five. You are a genetic counselor, and you interview a phenotypically normal couple who consult you because the man had a female first cousin (on his father’s side) who died from TSD, and the woman had a maternal uncle with TSD. There are no other known cases in either family, and none of the matings were between related individuals. Assume that this trait is rare in this population.

(a) Using standard pedigree symbols, draw a pedigree of these individuals’ families, showing the relevant individuals.
(b) The couple asks you to calculate the probability that they both are heterozygous for the TSD allele.
(c) They also want to know the probability that neither of them is heterozygous.
(d) They also ask you for the probability that one of them is heterozygous but the other is not. [Hint: The answers to (b), (c), and (d) should add up to 1.0.]

30. The wild-type (normal) fruit fly, *Drosophila melanogaster*, has straight wings and long bristles. Mutant strains have been isolated with either curled wings or short bristles. The genes representing these two mutant traits are located on separate chromosomes. Carefully examine the data from the five crosses below. (a) For each mutation, determine whether it is dominant or recessive. In each case, identify which crosses support your answer; and (b) define gene symbols, and determine the genotypes of the parents for each cross.

<table>
<thead>
<tr>
<th>Cross</th>
<th>Number of Progeny</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>straight wings, long bristles</td>
</tr>
<tr>
<td>1 straight, short × straight, short</td>
<td>30</td>
</tr>
<tr>
<td>2 straight, long × straight, long</td>
<td>120</td>
</tr>
<tr>
<td>3 curled, long × straight, short</td>
<td>40</td>
</tr>
<tr>
<td>4 straight, short × straight, short</td>
<td>40</td>
</tr>
<tr>
<td>5 curled, short × straight, short</td>
<td>20</td>
</tr>
</tbody>
</table>

31. To assess Mendel’s law of segregation using tomatoes, a true-breeding tall variety (SS) is crossed with a true-breeding short variety (ss). The heterozygous tall plants (Ss) were crossed to produce the two sets of F2 data shown below:

<table>
<thead>
<tr>
<th>Set I</th>
<th>Set II</th>
</tr>
</thead>
<tbody>
<tr>
<td>30 tall</td>
<td>300 tall</td>
</tr>
<tr>
<td>5 short</td>
<td>50 short</td>
</tr>
</tbody>
</table>

(a) Using chi-square analysis, analyze the results for both data sets. Calculate \( \chi^2 \) values, and estimate the \( p \) values in both cases.
(b) From the analysis in part (a), what can you conclude about the importance of generating large data sets in experimental settings?
Chromosomes in eukaryotes contain many genes whose locations are fixed along the length of the chromosomes. Unless separated by crossing over, alleles present on a chromosome segregate as a unit during gamete formation.

Crossing over between homologs during meiosis creates recombinant gametes with different combinations of alleles that enhance genetic variation.

Crossing over between homologs serves as the basis for the construction of chromosome maps.

Genetic maps depict relative locations of genes on chromosomes in a species.
Walter Sutton, along with Theodor Boveri, was instrumental in uniting the fields of cytology and genetics. As early as 1903, Sutton pointed out the likelihood that there must be many more “unit factors” than chromosomes in most organisms. Soon thereafter, genetics investigations revealed that certain genes segregate as if they were somehow joined or linked together. Further investigations showed that such genes are part of the same chromosome and may indeed be transmitted as a single unit. We now know that most chromosomes contain a very large number of genes. Those that are part of the same chromosome are said to be linked and to demonstrate linkage in genetic crosses.

Because the chromosome, not the gene, is the unit of transmission during meiosis, linked genes are not free to undergo independent assortment. Instead, the alleles at all loci of one chromosome should, in theory, be transmitted as a unit during gamete formation. However, in many instances this does not occur. During the first meiotic prophase, when homologs are paired or synapsed, a reciprocal exchange of chromosome segments can take place. This crossing over results in the reshuffling, or recombination, of the alleles between homologs and always occurs during the tetrad stage.

The degree of crossing over between any two loci on a single chromosome is proportional to the distance between them. Therefore, depending on which loci are being studied, the percentage of recombinant gametes varies. This correlation allows us to construct chromosome maps, which give the relative locations of genes on chromosomes.

In this chapter, we will discuss linkage, crossing over, and chromosome mapping in more detail. We will conclude by demonstrating linkage in genetic crosses.

### How Do We Know?

In this chapter, we will focus on genes located on the same eukaryotic chromosome, a concept called linkage. We are particularly interested in how such linked genes are transmitted to offspring and how the results of such transmission enable us to map them. As you study this topic, you should try to answer several fundamental questions:

1. How do we know that specific genes are linked on a single chromosome, in contrast to being located on separate chromosomes?
2. How was it determined that linked genes on homologous chromosomes are recombined during meiosis?
3. How do we know the order of, and intergenic distance between, genes found on the same chromosome?
4. In organisms where designed matings do not occur (such as humans), how do we know that genes are linked, and how do we map them?
5. How do we know that crossing over results from a physical exchange between chromatids?

### Now Solve This

Problem 9 on page 160 asks you to contrast the results of a testcross when two genes are unlinked versus linked, and when they are linked, if they are very far apart or relatively close together.

**Hint:** The results are indistinguishable when two genes are unlinked compared to the case where they are linked but so far apart that crossing over always intervenes between them during meiosis.
Chapter 7 Linkage and Chromosome Mapping in Eukaryotes

The four types (two parental and two recombinant gametes). In this case, transmission of two linked genes is indistinguishable from that of two unlinked, independently assorting genes. That is, the proportion of the four possible genotypes is identical, as shown in Figures 7–1(a) and (c).

The Linkage Ratio

If complete linkage exists between two genes because of their close proximity, and organisms heterozygous at both loci are mated, a unique F2 phenotypic ratio results, which we designate the **linkage ratio**. To illustrate this ratio, let’s consider a cross involving the closely linked, recessive, mutant genes brown eye (bw) and heavy wing vein (hv) in Drosophila melanogaster (Figure 7–2). The normal, wild-type alleles bw+ and hv+ are both dominant and result in red eyes and thin wing veins, respectively.

In this cross, flies with mutant brown eyes and normal thin wing veins are mated to flies with normal red eyes and mutant heavy wing veins. In more concise terms, brown-eyed flies are crossed with heavy-veined flies. If we extend the system of genetic symbols established in Chapter 4, linked genes are represented by placing their allele designations above and below a single or double horizontal line. Those above the line are located at loci on one homolog, and those below are located at the homologous loci on the other homolog. Thus, we represent the generation as follows:

\[
F_1: \frac{bw^+ hv}{bw hv} \times \frac{bw^+ hv}{bw hv}
\]

brown, thin red, heavy

These genes are located on an autosome, so no distinction between males and females is necessary.

In the F1 generation, each fly receives one chromosome of each pair from each parent. All flies are heterozygous for both gene pairs and exhibit the dominant traits of red eyes and thin wing veins:

\[
F_2: \frac{bw hv}{bw hv}
\]

As shown in Figure 7–2(a), when the F1 generation is interbred, each F1 individual forms only parental gametes because of complete linkage. After fertilization, the F2 generation is produced in a 1:2:1 phenotypic and genotypic ratio. One-fourth of this generation shows brown eyes and thin wing veins; one-half shows both wild-type traits, namely, red eyes and thin wing veins; and one-fourth shows red eyes and heavy wing veins. In more concise terms, the ratio is 1 brown: 2 wild:1 heavy. Such a 1:2:1 ratio is characteristic of complete linkage. Complete linkage is usually observed only when genes are very close together and the number of progeny is relatively small.

Figure 7–2(b) also gives the results of a testcross with the F1 flies. Such a cross produces a 1:1 ratio of brown, thin and red, heavy flies. Had the genes controlling these traits been incompletely linked or located on separate autosomes, the testcross would have produced four phenotypes rather than two.
7.1 Genes Linked on the Same Chromosome Segregate Together

FIGURE 7–2 Results of a cross involving two genes located on the same chromosome where complete linkage is demonstrated.

(a) The F₂ results of the cross. (b) The results of a testcross involving the F₁ progeny.

Because of complete linkage, F₁ individuals form only parental gametes.
When large numbers of mutant genes present in any given species are investigated, genes located on the same chromosome show evidence of linkage to one another. As a result, linkage groups can be established, one for each chromosome. In theory, the number of linkage groups should correspond to the haploid number of chromosomes. In diploid organisms in which large numbers of mutant genes are available for genetic study, this correlation has been confirmed.

### 7.2 Crossing Over Serves as the Basis of Determining the Distance Between Genes During Mapping

It is highly improbable that two randomly selected genes linked on the same chromosome will be so close to one another along the chromosome that they demonstrate complete linkage. Instead, crosses involving two such genes almost always produce a percentage of offspring resulting from recombinant gametes. This percentage is variable and depends on the distance between the two genes along the chromosome. This phenomenon was first explained around 1910 by two *Drosophila* geneticists, Thomas H. Morgan and his undergraduate student, Alfred H. Sturtevant.

**Morgan, Sturtevant, and Crossing Over**

In his studies, Morgan investigated numerous *Drosophila* mutations located on the X chromosome. When he analyzed crosses involving only one trait, he deduced the mode of X-linked inheritance. However, when he made crosses involving two X-linked genes, his results were initially puzzling. For example, female flies expressing the mutant yellow body (*y*) and white eyes (*w*) alleles were crossed with wild-type males (gray bodies and red eyes). The F1 females were wild type, while the F1 males expressed both mutant traits. In the F2, the vast majority of the offspring showed the expected parental phenotypes—either yellow-bodied, white-eyed flies or wild-type flies (gray-bodied, red-eyed). However, the remaining flies, less than 1.0 percent, were either yellow-bodied with red eyes or gray-bodied with white eyes. It was as if the two mutant alleles had somehow separated from each other on the homolog during gamete formation in the F1 female flies. This cross is illustrated in cross A of Figure 7–3, using data later compiled by Sturtevant.

When Morgan studied other X-linked genes, the same basic pattern was observed, but the proportion of the unexpected F2 phenotypes differed. For example, in a cross involving the mutant white eye, miniature wing (*m*) alleles, the majority of the F2 again showed the parental phenotypes, but a much higher proportion of the offspring appeared as if the mutant genes had separated during gamete formation. This is illustrated in cross B of Figure 7–3, again using data subsequently compiled by Sturtevant.

Morgan was faced with two questions: (1) What was the source of gene separation, and (2) why did the frequency of the apparent separation vary depending on the genes being studied? The answer he proposed for the first question was based on his knowledge of earlier cytological observations made by F. A. Janssens and others. Janssens had observed that synapsed homologous chromosomes in meiosis wrap around each other, creating chiasmata (sing., chiasma) where points of overlap are evident (see the photo on p. 136). Morgan proposed that chiasmata could represent points of genetic exchange.

In the crosses shown in Figure 7–3, Morgan postulated that if an exchange occurs during gamete formation between the mutant genes on the two X chromosomes of the F1 females, the unique phenotypes will occur. He suggested that such exchanges led to recombinant gametes in both the yellow–white cross and the white–miniature cross, in contrast to the parental gametes that have undergone no exchange. On the basis of this and other experiments, Morgan concluded that linked genes exist in a linear order along the chromosome and that a variable amount of exchange occurs between any two genes during gamete formation.

In answer to the second question, Morgan proposed that two genes located relatively close to each other along a chromosome are less likely to have a chiasma form between them than if the two genes are farther apart on the chromosome. Therefore, the closer two genes are, the less likely a genetic exchange will occur between them. Morgan was the first to propose the term crossing over to describe the physical exchange leading to recombination.

**Sturtevant and Mapping**

Morgan’s student, Alfred H. Sturtevant, was the first to realize that his mentor’s proposal could be used to map the sequence of linked genes. According to Sturtevant,

> In a conversation with Morgan ... I suddenly realized that the variations in strength of linkage, already attributed by Morgan to differences in the spatial separation of the genes, offered the possibility of determining sequences in the linear dimension of a chromosome. I went home and spent most of the night (to the neglect of my undergraduate homework) in producing the first chromosomal map.

Sturtevant compiled data from numerous crosses made by Morgan and other geneticists involving recombination between the genes represented by the yellow, white, and miniature mutants. These data are shown in Figure 7–3. The following recombination between each pair of these three genes, published in Sturtevant’s paper in 1913, is as follows:

1. **yellow–white** 0.5%
2. **white–miniature** 34.5%
3. **yellow–miniature** 35.4%

Because the sum of (1) and (2) approximately equals (3), Sturtevant suggested that the recombination frequencies between linked genes are additive. On this basis, he predicted that the order of the genes on the X chromosome is yellow–white–miniature. In arriving at this conclusion, he reasoned as follows: The yellow and white genes are apparently close to each other because the recombination frequency is
7.2 Crossing Over Serves as the Basis of Determining the Distance Between Genes During Mapping

Crossing Over Serves as the Basis of Determining the Distance Between Genes During Mapping

The data in Figure 7–3 depict the results of crosses involving the yellow body (y), white eye (w) mutations (cross A), and the white eye, miniature wing (m) mutations (cross B), as compiled by Sturtevant. In cross A, 0.5 percent of the F2 flies (males and females) demonstrate recombinant phenotypes, which express either white or yellow. In cross B, 34.5 percent of the F2 flies (males and females) demonstrate recombinant phenotypes, which express either miniature or white.

In honor of Morgan’s work, 1 map unit (mu) is now referred to as a centimorgan (cM).
Chapter 7  Linkage and Chromosome Mapping in Eukaryotes

The fact that these do not add up perfectly is due to the imprecision of independently-conducted mapping experiments.

In addition to these three genes, Sturtevant considered two other genes on the X chromosome and produced a more extensive map that included all five genes. He and a colleague, Calvin Bridges, soon began a search for autosomal linkage in *Drosophila*. By 1923, they had clearly shown that linkage and crossing over are not restricted to X-linked genes but can also be demonstrated with autosomes. During this work, they made another interesting observation. Crossing over in *Drosophila* was shown to occur only in females. The fact that no crossing over occurs in males made genetic mapping much less complex to analyze in *Drosophila*. However, crossing over does occur in both sexes in most other organisms.

Although many refinements in chromosome mapping have been developed since Sturtevant’s initial work, his basic principles are considered to be correct. These principles are used to produce detailed chromosome maps of organisms for which large numbers of linked mutant genes are known. Sturtevant’s findings are also historically significant to the broader field of genetics. In 1910, the chromosomal theory of inheritance was still widely disputed—even Morgan was skeptical of this theory before he conducted his experiments. Research has now firmly established that chromosomes contain genes in a linear order and that these genes are the equivalent of Mendel’s unit factors.

**Single Crossovers**

Why should the relative distance between two loci influence the amount of recombination and crossing over observed between them? During meiosis, a limited number of crossover events occur in each tetrad. These recombinant events occur randomly along the length of the tetrad. Therefore, the closer two loci reside along the axis of the chromosome, the less likely any single crossover event will occur between them. The same reasoning suggests that the farther apart two linked loci, the more likely a random crossover event will occur between them.

In Figure 7–5(a), a single crossover occurs between two nonsister chromatids but not between the two loci; therefore, the crossover is not detected because no recombinant gametes are produced. In Figure 7–5(b), where two loci are quite far apart, crossover does occur between them, yielding recombinant gametes.

When a single crossover occurs between two nonsister chromatids, the other two chromatids of the tetrad are not involved in this exchange and enter the gamete unchanged. Even if a single crossover occurs 100 percent of the time between two linked genes, recombination is subsequently observed in only 50 percent of the potential gametes formed. This concept is diagrammed in Figure 7–6. Theoretically, if we consider only single exchanges and observe 20 percent recombinant gametes, crossing over actually occurred in 40 percent of the
Determining the Gene Sequence During Mapping Relies on the Analysis of Multiple Crossovers

The study of single crossovers between two linked genes provides the basis of determining the distance between them. However, when many linked genes are studied, their sequence along the chromosome is more difficult to determine. Fortunately, the discovery that multiple exchanges occur between the chromatids of a tetrad has facilitated the process of producing more extensive chromosome maps. As we shall see next, when three or more linked genes are investigated simultaneously, it is possible to determine first the sequence of the genes and then the distances between them.

Multiple Crossovers

It is possible that in a single tetrad, two, three, or more exchanges will occur between nonsister chromatids as a result of several crossover events. Double exchanges of genetic material result from double crossovers (DCOs), as shown in Figure 7–7. For a double exchange to be studied, three gene pairs must be investigated, each heterozygous for two alleles. Before we determine the frequency of recombination among all three loci, let's review some simple probability calculations.

As we have seen, the probability of a single exchange occurring between the A and B or the B and C genes relates directly to the distance between the respective loci. The closer A is to B and B is to C, the less likely a single exchange will occur between either of the two sets of loci. In the case of a double crossover, two separate and independent events or exchanges must occur simultaneously. The mathematical probability of
two independent events occurring simultaneously is equal to the product of the individual probabilities (the product law).

Suppose that crossover gametes resulting from single exchanges are recovered 20 percent of the time \((p = 0.20)\) between \(A\) and \(B\), and 30 percent of the time \((p = 0.30)\) between \(B\) and \(C\). The probability of recovering a double-crossover gamete arising from two exchanges (between \(A\) and \(B\), and between \(B\) and \(C\)) is predicted to be \((0.20)(0.30) = 0.06\), or 6 percent. It is apparent from this calculation that the frequency of double-crossover gametes is always expected to be much lower than that of either single-crossover class of gametes.

If three genes are relatively close together along one chromosome, the expected frequency of double-crossover gametes is extremely low. For example, suppose the \(A-B\) distance in Figure 7–7 is 3 mu and the \(B-C\) distance is 2 mu. The expected double-crossover frequency is \((0.03)(0.02) = 0.0006\), or 0.06 percent. This translates to only 6 events in 10,000. Thus, in a mapping experiment where closely linked genes are involved, very large numbers of offspring are required to detect double-crossover events. In this example, it is unlikely that a double crossover will be observed even if 1000 offspring are examined. Thus, it is evident that if four or five genes are being mapped, even fewer triple and quadruple crossovers can be expected to occur.

### Three-Point Mapping in Drosophila

The information in the preceding section enables us to map three or more linked genes in a single cross. To illustrate the mapping process in its entirety, we examine two situations involving three linked genes in two quite different organisms.

To execute a successful mapping cross, three criteria must be met:

1. The genotype of the organism producing the crossover gametes must be heterozygous at all loci under consideration.

2. The cross must be constructed so that genotypes of all gametes can be determined accurately by observing the phenotypes of the resulting offspring. This is necessary because the gametes and their genotypes can never be observed directly. To overcome this problem, each phenotypic class must reflect the genotype of the gametes of the parents producing it.

3. A sufficient number of offspring must be produced in the mapping experiment to recover a representative sample of all crossover classes.

These criteria are met in the three-point mapping cross from *D. melanogaster* shown in Figure 7–8. In this cross, three X-linked recessive mutant genes—yellow body color (\(y\)), white eye color (\(w\)), and echinus eye shape (\(ec\))—are considered. To diagram the cross, we must assume some theoretical sequence, even though we do not yet know if it is correct. In Figure 7–8, we initially assume the sequence of the three genes to be \(y-w-ec\). If this is incorrect, our analysis will demonstrate this and reveal the correct sequence.

In the \(P_1\) generation, males hemizygous for all three wild-type alleles are crossed to females that are homozygous for all three recessive mutant alleles. Therefore, the \(P_1\) males are wild type with respect to body color, eye color, and eye shape. They are said to have a *wild-type phenotype*. The females, on the other hand, exhibit the three mutant traits—yellow body color, white eyes, and echinus eye shape.

This cross produces an \(F_1\) generation consisting of females that are heterozygous at all three loci and males that, because of the \(Y\) chromosome, are hemizygous for the three mutant alleles. Phenotypically, all \(F_1\) females are wild type, while all \(F_1\) males are yellow, white, and echinus. The genotype of the \(F_1\) females fulfills the first criterion for mapping the three linked genes; that is, it is heterozygous at the three loci and can provide the source of recombinant gametes generated by crossing over. Note that because of the genotypes of the \(P_1\) parents, all three mutant alleles are on one homolog and all three wild-type alleles are on the other homolog. Other arrangements are possible. For example, the heterozygous \(F_1\) female might have the \(y\) and \(ec\) mutant alleles on one homolog and the \(w\) allele on the other. This would occur if, in the \(P_1\) cross, one parent was yellow, echinus and the other parent was white.

In our cross, the second criterion is met by virtue of the gametes formed by the \(F_1\) males. Every gamete contains either an \(X\) chromosome bearing the three mutant alleles or a \(Y\) chromosome, which is genetically inert for the three loci being considered. Whichever type participates in fertilization, the genotype of the gamete produced by the \(F_1\) female will be expressed phenotypically in the \(F_2\) male and female offspring derived from it. Thus, all \(F_1\) noncrossover and crossover gametes can be detected by observing the \(F_2\) phenotypes.

With these two criteria met, we can now construct a chromosome map from the crosses shown in Figure 7–8. First, we determine which \(F_2\) phenotypes correspond to the various noncrossover and crossover categories. To determine the noncrossover \(F_2\) phenotypes, we must identify individuals derived from the parental gametes formed by the \(F_1\) female. Each such gamete contains an \(X\) chromosome unaffected by crossing over. As a result of segregation, approximately equal proportions of the two types of gametes and, subsequently, the \(F_2\) phenotypes, are produced. Because they derive from a heterozygote, the genotypes of the two parental gametes and the resultant \(F_2\) phenotypes complement one another. For
7.3 Determining the Gene Sequence During Mapping Relies on the Analysis of Multiple Crossovers

Determining the Gene Sequence During Mapping Relies on the Analysis of Multiple Crossovers. The analysis of multiple crossovers helps in determining the gene sequence during mapping. The figure illustrates a three-point mapping cross involving the yellow (y or y+), white (w or w+), and echinus (ec or ec+) genes in *D. melanogaster*. NCO, SCO, and DCO refer to noncrossover, single-crossover, and double-crossover groups, respectively. Centromeres are not included on the chromosomes, and only two nonsister chromatids are shown initially.

### Table

<table>
<thead>
<tr>
<th>Origin of female gametes</th>
<th>Gametes</th>
<th>F2 phenotype</th>
<th>Observed Number</th>
<th>Category, total, and percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>NCO</td>
<td>y w ec</td>
<td>y w ec</td>
<td>4685</td>
<td>Non-crossover</td>
</tr>
<tr>
<td></td>
<td>y w ec</td>
<td>y w ec</td>
<td>4759</td>
<td></td>
</tr>
<tr>
<td></td>
<td>y w ec</td>
<td>y w ec</td>
<td>9444</td>
<td>94.44%</td>
</tr>
<tr>
<td></td>
<td>y w ec</td>
<td>y w ec</td>
<td>150</td>
<td>1.50%</td>
</tr>
<tr>
<td></td>
<td>y w ec</td>
<td>y w ec</td>
<td>70</td>
<td>1.50%</td>
</tr>
<tr>
<td></td>
<td>y w ec</td>
<td>y w ec</td>
<td>193</td>
<td>Single crossover between y and w</td>
</tr>
<tr>
<td></td>
<td>y w ec</td>
<td>y w ec</td>
<td>207</td>
<td></td>
</tr>
<tr>
<td></td>
<td>y w ec</td>
<td>y w ec</td>
<td>400</td>
<td>4.00%</td>
</tr>
<tr>
<td></td>
<td>y w ec</td>
<td>y w ec</td>
<td>3</td>
<td>Double crossover between y and w</td>
</tr>
<tr>
<td></td>
<td>y w ec</td>
<td>y w ec</td>
<td>6</td>
<td>0.06%</td>
</tr>
</tbody>
</table>

**Map of y, w, and ec loci**

**FIGURE 7-8** A three-point mapping cross involving the yellow (y or y+), white (w or w+), and echinus (ec or ec+) genes in *D. melanogaster*. NCO, SCO, and DCO refer to noncrossover, single-crossover, and double-crossover groups, respectively. Centromeres are not included on the chromosomes, and only two nonsister chromatids are shown initially.
example, if one is wild type, the other is completely mutant. This is the case in the cross being considered. In other situations, if one chromosome shows one mutant allele, the second chromosome shows the other two mutant alleles, and so on. They are therefore called reciprocal classes of gametes and phenotypes.

The two noncrossover phenotypes are most easily recognized because they exist in the greatest proportion. Figure 7–8 shows that gametes 1 and 2 are present in the greatest numbers. Therefore, flies that express yellow, white, and echinus phenotypes and flies that are normal (or wild type) for all three characters constitute the noncrossover category and represent 94.44 percent of the F2 offspring.

The second category that can be easily detected is represented by the double-crossover phenotypes. Because of their low probability of occurrence, they must be present in the least numbers. Remember that this group represents two independent but simultaneous single-crossover events. Two reciprocal phenotypes can be identified: gamete 7, which shows the mutant traits yellow, echinus but normal eye color; and gamete 8, which shows the mutant trait white but normal body color and eye shape. Together these double-crossover phenotypes constitute only 0.06 percent of the F2 offspring.

The remaining four phenotypic classes represent two categories resulting from single crossovers. Gametes 3 and 4, reciprocal phenotypes produced by single-crossover events occurring between the yellow and white loci, are equal to 1.50 percent of the F2 offspring; gametes 5 and 6, constituting 4.00 percent of the F2 offspring, represent the reciprocal phenotypes resulting from single-crossover events occurring between the white and echinus loci.

The map distances separating the three loci can now be calculated. The distance between y and w or between w and ec is equal to the percentage of all detectable exchanges occurring between them. For any two genes under consideration, this includes all appropriate single crossovers as well as all double crossovers. The latter are included because they represent two simultaneous single crossovers. For the y and w genes, this includes gametes 3, 4, 7, and 8, totaling 1.50% + 0.06%, or 1.56 mu. Similarly, the distance between w and ec is equal to the percentage of offspring resulting from an exchange between these two loci: gametes 5, 6, 7, and 8, totaling 4.00% + 0.06%, or 4.06 mu. The map of these three loci on the X chromosome is shown at the bottom of Figure 7–8.

**Determining the Gene Sequence**

In the preceding example, the sequence (or order) of the three genes along the chromosome was assumed to be y–w–ec. Our analysis shows this sequence to be consistent with the data. However, in most mapping experiments the gene sequence is not known, and this constitutes another variable in the analysis. In our example, had the gene sequence been unknown, it could have been determined using a straightforward method.

This method is based on the fact that there are only three possible arrangements, each containing one of the three genes between the other two:

1. y–w–ec (y in the middle)
2. y–ec–w (ec in the middle)
3. y–w–ec (w in the middle)

The steps are shown in Figure 7–9, using our y–w–ec cross. The three possible arrangements are labeled I, II, and III, as shown above.

1. Assuming that y is between w and ec, arrangement I of alleles along the homologs of the F1 heterozygote is

   \[
   w^+ y^+ ec^+
   \]

   We know this because of the way in which the P1 generation was crossed: The P1 female contributes an X chromosome bearing the w, y, and ec alleles, while the P1 male contributes an X chromosome bearing the w+, y+, and ec+ alleles.

2. A double crossover within that arrangement yields the following gametes:

   \[
   w y^+ ec \text{ and } w^+ y ec^+
   \]

   Following fertilization, if y is in the middle, the F2 double-crossover phenotypes will correspond to these gametic genotypes, yielding offspring that express the white, echinus phenotype and offspring that express the yellow phenotype. Instead, however, determination of the actual double-crossover phenotypes reveals them to be yellow, echinus flies and white flies. Therefore, our assumed order is incorrect.

3. If we consider arrangement II with the ec+ec+ alleles in the middle or arrangement III with the w+w+ alleles in the middle:

   \[
   (\text{II}) \quad y^+ ec^+ w^+ \quad \text{or} \quad (\text{III}) \quad y w^+ ec^+
   \]

   we see that arrangement II again provides predicted double-crossover phenotypes that do not correspond to the actual (observed) double-crossover phenotypes. The pre-
The resulting phenotypes produced in a testcross. For simplicity, the two noncrossover chromatids of each tetrad are omitted.

virescent seedling. A sufficient sample size must be available for complete analysis. To determine the arrangement of alleles, one must be heterozygous for all traits under consideration; (2) the map distances between the genes. What genotype must the male plant have to allow successful mapping? To meet the second criterion, the male must be homozygous for all three recessive mutant alleles. Otherwise, offspring of this cross showing a given phenotype might represent more than one genotype, making accurate mapping impossible.

To summarize, this method is rather straightforward: First determine the arrangement of alleles on the homologs of the heterozygote yielding the crossover gametes by locating the reciprocal noncrossover phenotypes. Then, test each of three possible orders to determine which yields the observed double-crossover phenotypes—the one that does so represents the correct order.

A Mapping Problem in Maize

Having established the basic principles of chromosome mapping, we now consider a related problem in maize (corn), in which the gene sequence and interlocus distances are unknown. This analysis differs from the preceding example in two ways. First, the previous mapping cross involved X-linked genes. Second, in the discussion of this cross we have changed our use of symbols, as first suggested in Chapter 4. Instead of using the gene symbols and superscripts (e.g., $bm^+$, $v^+$, and $pr^+$), we simply use + to denote each wild-type allele. This system is easier to manipulate but requires a better understanding of mapping procedures.

When we look at three autosomal linked genes in maize, the experimental cross must still meet the same three criteria we established for the X-linked genes in Drosophila: (1) One parent must be heterozygous for all traits under consideration; (2) the gametic genotypes produced by the heterozygote must be apparent from observing the phenotypes of the offspring; and (3) a sufficient sample size must be available for complete analysis.

In maize, the recessive mutant genes brown midrib ($bm$), virescent seedling ($v$), and purple aleurone ($pr$) are linked on chromosome 5. Assume that a female plant is known to be heterozygous for all three traits, but we do not know (1) the arrangement of the mutant alleles on the maternal and paternal homologs of this heterozygote, (2) the sequence of genes, or (3) the map distances between the genes. What genotype must the male plant have to allow successful mapping? To meet the second criterion, the male must be homozygous for all three recessive mutant alleles. Otherwise, offspring of this cross showing a given genotype might represent more than one genotype, making accurate mapping impossible.

Figure 7–10 diagrams this cross. As shown, we know neither the arrangement of alleles nor the sequence of loci in the heterozygous female. Several possibilities are shown, but we have yet to determine which is correct. We don’t know the sequence in the testcross male parent either, and so we must designate it randomly. Note that we have initially placed $v$ in the middle. This may or may not be correct.

The offspring are arranged in groups of two for each pair of reciprocal phenotypic classes. The two members of each reciprocal class are derived from no crossing over (NCO), one of two possible single-crossover events (SCO), or a double crossover (DCO).

To solve this problem, refer to Figures 7–10 and 7–11 as you consider the following questions.

1. What is the correct heterozygous arrangement of alleles in the female parent? Determine the two noncrossover classes, those that occur with the highest frequency. In this case, they are $v bm$ and $v bm$. Therefore, the alleles on the homologs of the female parent must be arranged as shown in Figure 7–11(a). These homologs segregate into gametes, unaffected by any recombination event. Any other arrangement of alleles will not yield the observed noncrossover classes. (Remember that $v bm$ is equivalent to $v^+ bm^+$, and that $v^+ bm^+$ is equivalent to $v^+ bm^+$).
2. **What is the correct sequence of genes?** We know that the arrangement of alleles is

\[
\text{+ pr +} \quad \text{v} \quad \text{bm} \quad \text{pr} \\
\text{+ + +} \quad + \quad +
\]

But is the gene sequence correct? That is, will a double-crossover event yield the observed double-crossover phenotypes after fertilization? **Observation shows it will not** [Figure 7–11(b)]. Now try the other two orders [Figures 7–11(c) and (d)] maintaining the same arrangement:

\[
\begin{align*}
\text{+ + bm} & \quad \text{v} \quad \text{pr} \\
\text{pr + +} & \quad + \quad +
\end{align*}
\]

Only the order on the right yields the observed double-crossover *gametes* [Figure 7–11(d)]. Therefore, the *pr* gene is in the middle. From this point on, work the problem using this arrangement and sequence, with the *pr* locus in the middle.

3. **What is the distance between each pair of genes?** Having established the sequence of loci as *v–pr–bm*, we can determine the distance between *v* and *pr* and between *pr* and *bm*.
and \( bm \). Remember that the map distance between two genes is calculated on the basis of all detectable recombination events occurring between them. This includes both single- and double-crossover events.

Figure 7–11(e) shows that the phenotypes \( v\ pr\ + + \) and \( + + bm\) result from single crossovers between the \( v \) and \( pr \) loci, accounting for 14.5 percent of the offspring [according to data in Figure 10(b)]. By adding the percentage of double crossovers (7.8%) to the number obtained for single crossovers, the total distance between the \( v \) and \( pr \) loci is calculated to be 22.3 mu.

Figure 7–11(f) shows that the phenotypes \( v\ + + \) and \( + pr\ bm\) result from single crossovers between the \( pr \) and \( bm \) loci, totaling 35.6 percent. Added to the double crossovers (7.8%), the distance between \( pr \) and \( bm \) is calculated to be 43.4 mu. The final map for all three genes in this example is shown in Figure 7–11(g).

### FIGURE 7–11 Producing a map of the three genes in the testcross of Figure 7–10, where neither the arrangement of alleles nor the sequence of genes in the heterozygous female parent is known.
As the Distance Between Two Genes Increases, Mapping Estimates Become More Inaccurate

So far, we have assumed that crossover frequencies are directly proportional to the distance between any two loci along the chromosome. However, it is not always possible to detect all crossover events. A case in point is a double exchange that occurs between the two loci in question. As shown in Figure 7–12(a), if a double exchange occurs, the original arrangement of alleles on each non-sister homolog is recovered. Therefore, even though crossing over has occurred, it is impossible to detect. This phenomenon is true for all even-numbered exchanges between two loci.

Furthermore, as a result of complications posed by multiple-strand exchanges, mapping determinations usually underestimate the actual distance between two genes. The farther apart two genes are, the greater the probability that undetected crossovers will occur. While the discrepancy is minimal for two genes relatively close together, the degree of inaccuracy increases as the distance increases, as shown in the graph of recombination frequency versus map distance in Figure 7–12(b). There, the theoretical frequency where a direct correlation between recombination and map distance exists is contrasted with the actual frequency observed as the distance between two genes increases. The most accurate maps are constructed from experiments where genes are relatively close together.

Interference and the Coefficient of Coincidence

As shown in our maize example, we can predict the expected frequency of multiple exchanges, such as double crossovers, once the distance between genes is established. For example, in the maize cross, the distance between v and pr is 22.3 μ, and the distance between pr and bm is 43.4 μ. If the two single crossovers that make up a double crossover occur independently of one another, we can calculate the expected frequency of double crossovers (DCOexp):

\[
DCO_{\text{exp}} = (0.223) \times (0.434) = 0.097 = 9.7\%
\]

Often in mapping experiments, the observed DCO frequency is less than the expected number of DCOs. In the maize cross, for example, only 7.8 percent DCOs are observed when 9.7 percent are expected. Interference (I), the phenomenon where a crossover event in one region of the chromosome inhibits a second event in nearby regions, causes this reduction.

To quantify the disparities that result from interference, we calculate the coefficient of coincidence (C):

\[
C = \frac{\text{Observed DCO}}{\text{Expected DCO}}
\]

In the maize cross, we have

\[
C = \frac{0.078}{0.097} = 0.804
\]

Once we have found C, we can quantify interference using the simple equation

\[
I = 1 - C
\]

In the maize cross, we have

\[
I = 1.000 - 0.804 = 0.196
\]

If interference is complete and no double crossovers occur, then \(I = 1.0\). If fewer DCOs than expected occur, \(I\) is a posi-
tive number and positive interference has occurred. If more DCOs than expected occur, $I$ is a negative number and negative interference has occurred. In the maize example, $I$ is a positive number (0.196), indicating that 19.6 percent fewer double crossovers occurred than expected.

Positive interference is most often observed in eukaryotic systems. In general, the closer genes are to one another along the chromosome, the more positive interference occurs. In organisms such as *Drosophila*, maize, and the mouse, where large numbers of mutations have been discovered and where mapping crosses are possible, extensive chromosome maps have been constructed. Figure 7–13 shows partial maps for the four chromosomes (I, II, III, and IV) of *Drosophila*.

**FIGURE 7–13** A partial genetic map of the four chromosomes of *D. melanogaster*. The circle on each chromosome represents the position of the centromere. Chromosome I is the X chromosome. Chromosome IV is not drawn to scale.

### 7.5 *Drosophila* Genes Have Been Extensively Mapped

In organisms such as *Drosophila*, maize, and the mouse, where large numbers of mutations have been discovered and where mapping crosses are possible, extensive chromosome maps have been constructed. Figure 7–13 shows partial maps for the four chromosomes (I, II, III, and IV) of *Drosophila*. 

- **I (X)**
  - 0.0 yellow body, y, scute bristles, sc
  - 1.5 white eyes, w
  - 3.0 facet eyes, fa
  - 5.5 echinus eyes, ec
  - 7.5 ruby eyes, rb
  - 13.7 crossveinless wings, cv
  - 20.0 cut wings, ct
  - 21.0 singed bristles, sb
  - 27.5 tan body, t
  - 27.7 lozenge eyes, lz
  - 33.0 vermilion eyes, v
  - 36.1 miniature wings, m
  - 43.0 sable body, s
  - 44.0 garnet eyes, g
  - 51.5 scalloped wings, sdl
  - 56.7 forked bristles, ft
  - 57.0 Bar eyes, B
  - 59.5 fused veins, fu
  - 62.5 carnation eyes, car
  - 66.0 bobbed hairs, bb
  - 68.1 little fly, lf

- **II**
  - 0.0 aristale antenna, al
  - 1.3 Star eyes, S
  - 6.1 Curly wing, Cy
  - 13.0 dumpy wings, dp
  - 16.5 clot eyes, cl
  - 22.0 Sternopleural bristles, Sp
  - 31.0 dachs tarsus, d
  - 36.0 corrugated, corr
  - 39.3 daughterless, da
  - 41.0 jammed wings, j
  - 48.5 black body, b
  - 51.0 reduced bristles, rd
  - 54.5 purple eyes, pr
  - 57.5 cinnabar eyes, cn
  - 61.0 withered wing, whd
  - 67.0 vestigial wings, vg
  - 72.0 Lobe eyes, L
  - 75.5 curved wings, c
  - 83.1 adipose, adp
  - 90.0 disrupted wing, dsr
  - 91.5 smooth abdomen, sm
  - 100.5 plexus wings, px
  - 104.0 heavy wing veins, hv
  - 104.5 brown eyes, bw
  - 107.0 speck body, sp

- **III**
  - 0.0 roughoid eyes, ru
  - 0.2 veinlet veins, ve
  - 1.4 Roughened eye, R
  - 11.0 female sterile, fs(3)G2
  - 17.0 raisin eye, rai
  - 19.2 javelin bristles, jv
  - 26.0 sepias eyes, se
  - 26.5 hairy body, h
  - 35.5 eyes gone, eyg
  - 40.5 Lyra wings, Ly
  - 41.0 Dichaete bristles, D
  - 43.2 thread arista, th
  - 44.0 scarlet eyes, st
  - 58.2 Stubble bristles, Sb
  - 58.5 spineless bristles, ss
  - 62.0 stripe body, sr
  - 66.2 Delta veins, Dl
  - 69.5 Hairless bristles, H
  - 70.7 ebony body, e
  - 74.7 cardinal eyes, cd
  - 77.5 obtuse wing, obt
  - 88.0 mahogany eyes, mah
  - 91.1 rough eyes, ro
  - 95.5 suppression of purple, su-pr
  - 100.7 claret eyes, ca
  - 106.2 Minute bristles, M(3)g
Virtually every morphological feature of the fruit fly has been subjected to mutations. The locus of each mutant gene is first localized to one of the four chromosomes (or linkage groups) and then mapped in relation to all other genes present on that chromosome. Based on cytological evidence, the relative lengths of these genetic maps correlate roughly with the relative physical lengths of these chromosomes.

7.6 **Lod Score Analysis and Somatic Cell Hybridization Were Historically Important in Creating Human Chromosome Maps**

In humans, where neither designed matings nor large numbers of offspring are available, the earliest linkage studies were based on pedigree analysis. Attempts were made to establish whether a trait was X-linked or autosomal. For autosomal traits, geneticists tried to distinguish whether pairs of traits demonstrate linkage or independent assortment. In this way, it was hoped that a human gene map could be created.

The difficulty arises, however, when two genes of interest are separated on a chromosome such that recombinant gametes are formed, obscuring linkage in a pedigree. In such cases, the demonstration of linkage is enhanced by an approach that relies on probability, called the **lod score method**. First devised by J. B. S. Haldane and C. A. Smith in 1947, and refined by Newton Morton in 1955, the lod score (log of the odds favoring linkage) assesses the probability that a particular pedigree involving two traits reflects linkage or not. First, the probability is calculated that family data (pedigrees) concerning two or more traits demonstrate linkage or independent assortment. In this way, it was hoped that a human gene map could be created.

The difficulty arises, however, when two genes of interest are separated on a chromosome such that recombinant gametes are formed, obscuring linkage in a pedigree. In such cases, the demonstration of linkage is enhanced by an approach that relies on probability, called the **lod score method**. First devised by J. B. S. Haldane and C. A. Smith in 1947, and refined by Newton Morton in 1955, the lod score (log of the odds favoring linkage) assesses the probability that a particular pedigree involving two traits reflects linkage or not. First, the probability is calculated that family data (pedigrees) concerning two or more traits demonstrate linkage or independent assortment. In this way, it was hoped that a human gene map could be created.

Accuracy using the lod score method is limited by the extent of the family data, but nevertheless represents an important advance in assigning human genes to specific chromosomes and constructing preliminary human chromosome maps. The initial results were discouraging because of the method’s inherent limitations and the relatively high haploid number of human chromosomes (23), and by 1960, almost no linkage or mapping information had become available.

In the 1960s, a new technique, **somatic cell hybridization**, proved to be an immense aid in assigning human genes to their respective chromosomes. This technique, first discovered by Georges Barsky, relies on the fact that two cells in culture can be induced to fuse into a single hybrid cell. Barsky used two mouse-cell lines, but it soon became evident that cells from different organisms could also be fused. When fusion occurs, an initial cell type called a **heterokaryon** is produced. The hybrid cell contains two nuclei in a common cytoplasm. By using the proper techniques, it is possible to fuse human and mouse cells, for example, and isolate the hybrids from the parental cells.

As the heterokaryons are cultured in vitro, two interesting changes occur. The nuclei eventually fuse, creating a **syndaryon**. Then, as culturing is continued for many generations, chromosomes from one of the two parental species are gradually lost. In the case of the human–mouse hybrid, human chromosomes are lost randomly until eventually the syndaryon has a full complement of mouse chromosomes and only a few human chromosomes. As we shall see, it is the preferential loss of human chromosomes (rather than mouse chromosomes) that makes possible the assignment of human genes to the chromosomes on which they reside.

The experimental rationale is straightforward. If a specific human gene product is synthesized in a syndaryon containing one to three human chromosomes, then the gene responsible for that product must reside on one of the human chromosomes remaining in the hybrid cell. On the other hand, if the human gene product is not synthesized in a syndaryon, the gene responsible is not present on those human chromosomes that remain in this hybrid cell. Ideally, a panel of 23 hybrid-cell lines, each with just one unique human chromosome, would allow the immediate assignment of any human gene for which the product could be characterized.

In practice, a panel of cell lines, each with several remaining human chromosomes, is mostly used. The correlation of the presence or absence of each chromosome with the presence or absence of each gene product is called **synteny testing**. Consider, for example, the hypothetical data provided in Figure 7–14, where four gene products (A, B, C, and D) are tested in relationship to eight human chromosomes. Let’s analyze the gene that produces product A.

<table>
<thead>
<tr>
<th>Hybrid cell lines</th>
<th>Human chromosomes present</th>
<th>Gene products expressed</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>23</td>
<td></td>
<td>+</td>
</tr>
<tr>
<td>34</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>41</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

**FIGURE 7–14** A hypothetical grid of data used in synteny testing to assign genes to their appropriate human chromosomes. Three somatic hybrid-cell lines, designated 23, 34, and 41, have each been scored for the presence or absence of human chromosomes 1–8, as well as for their ability to produce the hypothetical human gene products A, B, C, and D.
7.6 Lod Score Analysis and Somatic Cell Hybridization Were Historically Important in Creating Human Chromosome Maps

1. Product A is not produced by cell line 23, but chromosomes 1, 2, 3, and 4 are present in cell line 23. Therefore, we rule out the presence of gene A on those four chromosomes and conclude that it must be on chromosome 5, 6, 7, or 8.

2. Product A is produced by cell line 34, which contains chromosomes 5 and 6 but not 7 and 8. Therefore, gene A is on chromosome 5 or 6, but cannot be on 7 or 8 because they are absent even though product A is produced.

3. Product A is also produced by cell line 41, which contains chromosome 5 but not chromosome 6. Using similar reasoning we see that gene A must be on chromosome 5.

Using a similar approach, we can assign gene B to chromosome 3. Perform this analysis for yourself to demonstrate that this is correct.

Gene C presents a unique situation. The data indicate that it is not present on chromosomes 1–7. While it might be on chromosome 8, no direct evidence supports this conclusion, and other panels are needed. We leave gene D for you to analyze—on what chromosome does it reside?

Using this technique, researchers have assigned literally hundreds of human genes to one chromosome or another. Some of the assignments shown in Figure 7–15 were either derived or confirmed with the use of somatic cell hybridization techniques. To map genes for which the products have yet

**FIGURE 7–15** Representative regional gene assignments for human chromosome 1 and the X chromosome. Many assignments were initially derived using somatic cell hybridization techniques.

<table>
<thead>
<tr>
<th>Key</th>
<th>Gene Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AMY</td>
<td>Amylase (salivary and pancreatic)</td>
</tr>
<tr>
<td>AT3</td>
<td>Antithrombin (clotting factor IV)</td>
</tr>
<tr>
<td>CB</td>
<td>Color Blindness</td>
</tr>
<tr>
<td>DMD</td>
<td>Duchenne Muscular Dystrophy</td>
</tr>
<tr>
<td>FHM</td>
<td>Fumarate Hydratase (mitochondrial)</td>
</tr>
<tr>
<td>GDH</td>
<td>Glucose Dehydrogenase</td>
</tr>
<tr>
<td>G6PD</td>
<td>Glucose-6-Phosphate Dehydrogenase</td>
</tr>
<tr>
<td>HCPRT</td>
<td>Hypoxanthine-Guanine-Phosphoribosyl Transferase</td>
</tr>
<tr>
<td>HEMA</td>
<td>Hemophilia A (classic)</td>
</tr>
<tr>
<td>HGPRT</td>
<td>Hypoxanthine-Guanine-Phosphoribosyl Transferase (Lesch–Nyhan syndrome)</td>
</tr>
<tr>
<td>PEPC</td>
<td>Peptidase C</td>
</tr>
<tr>
<td>PGK</td>
<td>Phosphoglycerate Kinase</td>
</tr>
<tr>
<td>PGM</td>
<td>Phosphoglucomutase</td>
</tr>
<tr>
<td>Rh</td>
<td>Rhesus Blood Group (erythroblastosis fetalis)</td>
</tr>
</tbody>
</table>
to be discovered, researchers have had to rely on other approaches. For example, by using recombinant DNA technology in conjunction with pedigree analysis, it has been possible to assign the genes responsible for Huntington disease, cystic fibrosis, and neurofibromatosis to their respective chromosomes, 4, 7, and 17.

### 7.7 Linkage and Mapping Studies Can Be Performed in Haploid Organisms

Many of the single-celled eukaryotes are haploid during the vegetative stages of their life cycle. The alga Chlamydomonas and the mold Neurospora demonstrate this genetic condition. But these organisms do form reproductive cells that fuse during fertilization, producing a diploid zygote. However, this structure soon undergoes meiosis, resulting in haploid vegetative cells that then propagate by mitotic divisions. In genetic studies, small haploid organisms have several important advantages compared with diploid eukaryotes. They can be cultured and manipulated in genetic crosses much more easily. In addition, a haploid organism contains only a single allele of each gene, which is expressed directly in the phenotype. This greatly simplifies genetic analysis. As a result, organisms such as Chlamydomonas and Neurospora serve as the subjects of research investigations in many areas of genetics, including linkage and mapping studies.

In order to perform genetics experiments with such organisms, crosses are made, and following fertilization, the meiotic structures may be isolated. Because all four meiotic products give rise to spores, the structures bearing these products (asci) are called tetrads. The term tetrad has a different meaning here than earlier when it was used to describe a precise chromosome configuration during meiosis. Individual tetrads are isolated, and the resultant cells are grown and analyzed separately from those of other tetrads. In the results we are about to describe, the data reflect the proportion of tetrads that show one combination of genotypes, the proportion that show another combination, and so on. Such experimentation is called tetrad analysis.

#### Gene-to-Centromere Mapping

When a single gene in Neurospora is analyzed (Figure 7–16), the data can be used to calculate the map distance between the gene and the centromere. This process is sometimes referred to as mapping the centromere. It is accomplished by experimentally determining the frequency of recombination using tetrad data. Note that once the four meiotic products of the tetrad are formed, a mitotic division occurs, resulting in eight ordered products (ascospores). If no crossover event occurs between the gene under study and the centromere, the pattern of ascospores contained within an ascus (pl., asci) appears as shown in Figure 7–16(a), aaaa+++++.*

*The pattern (+++++aaaa) can also be formed. However, it is indistinguishable from (aaaa+++++).

This pattern represents first-division segregation because the two alleles are separated during the first meiotic division. However, a crossover event will alter this pattern, as shown in Figure 7–16(b), aa++aa+++, and Figure 7–16(c), ++aaaa++. Two other recombinant patterns occur but are not shown: ++aa++aa and aa+++++aa. These depend on the chromatid orientation during the second meiotic division. These four patterns reflect second-division segregation because the two alleles are not separated until the second meiotic division. Usually, the ordered tetrad data are condensed to reflect the genotypes of the identical ascospor pairs, and six combinations are possible.

**First-Division Segregation**

<table>
<thead>
<tr>
<th>Genotypes</th>
<th>Second-Division Segregation</th>
</tr>
</thead>
<tbody>
<tr>
<td>aa++</td>
<td>a+a+ +aa+ +aa+ +aa+</td>
</tr>
<tr>
<td>++aa</td>
<td>+a+a +a+a +a+a +a+a</td>
</tr>
</tbody>
</table>

To calculate the distance between the gene and the centromere, a large number of asci resulting from a controlled cross are counted. We then use these data to calculate the distance ($d$):

$$d = \frac{1}{2} \left( \frac{\text{second-division segregant asci}}{\text{total asci scored}} \right)$$

The distance ($d$) reflects the percentage of recombination and is only half the number of second-division segregant asci. This is because crossing over in each occurs in only two of the four chromatids during meiosis.

To illustrate, we use $a$ for albino and $+$ for wild type in Neurospora. In crosses between the two genetic types, suppose we observe 65 first-division segregants, and 70 second-division segregants. The distance between $a$ and the centromere is

$$d = \frac{(1/2)(70)}{135} = 0.259$$

or about 26 mu.

As the distance increases to 50 mu, all asci should reflect second-division segregation. However, numerous factors prevent this. As in diploid organisms, mapping accuracy based on crossover events is greatest when the gene and centromere are relatively close together.

We can also analyze haploid organisms to distinguish between linkage and independent assortment of two genes—mapping distances between gene loci are calculated once linkage is established. As a result, detailed maps of organisms such as Neurospora and Chlamydomonas are now available.

### 7.8 Other Aspects of Genetic Exchange

Careful analysis of crossing over during gamete formation allows us to construct chromosome maps in both diploid and haploid organisms. However, we should not lose sight of the real biological significance of crossing over, which is to generate genetic variation in gametes and, subsequently, in the
offspring derived from the resultant eggs and sperm. Because of the critical role of crossing over in generating variation, the study of genetic exchange is a key topic for study in genetics. But many important questions remain. For example, does crossing over involve an actual exchange of chromosome arms? Does exchange occur between paired sister chromatids during mitosis? We shall briefly consider possible answers to these questions.

Cytological Evidence for Crossing Over

Once genetic mapping was understood, it was of great interest to investigate the relationship between chiasmata observed in meiotic prophase I and crossing over. For example, are chiasmata visible manifestations of crossover events? If so, then crossing over in higher organisms appears to result from an actual physical exchange between homologous chromosomes. That this is the case was demonstrated independently in the 1930s by Harriet Creighton and Barbara McClintock in *Zea mays*, and by Curt Stern in *Drosophila*.

Since the experiments are similar, we will consider only the work with maize. Creighton and McClintock studied two linked genes on chromosome 9. At one locus, the alleles *colorless* (c) and *colored* (C) control endosperm coloration. At the other locus, the alleles *starchy* (Wx) and *waxy* (wx) control the carbohydrate characteristics of the endosperm. The maize plant studied is heterozygous at both loci. The key to this experiment is that one of the homologs contains two unique cytological markers. The markers consist of a densely stained knob at one end of the chromosome and a translocated piece of another chromosome (8) at the other end. The arrangements of alleles and cytological markers can be detected cytologically and are shown in Figure 7–17.

FIGURE 7–16 Three ways in which different ascospore patterns can be generated in *Neurospora*. Analysis of these patterns is the basis of gene-to-centromere mapping. The photograph shows a variety of ascospore arrangements within *Neurospora* asci.

### Table: Three Ways in Which Different Ascospore Patterns Can Be Generated in *Neurospora*

<table>
<thead>
<tr>
<th>Condition</th>
<th>Four-strand stage</th>
<th>Chromosomes following meiosis</th>
<th>Chromosomes following mitotic division</th>
<th>Ascospores in ascus</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a) No crossover</td>
<td><img src="image1" alt="Diagram" /></td>
<td><img src="image2" alt="Diagram" /></td>
<td><img src="image3" alt="Diagram" /></td>
<td></td>
</tr>
<tr>
<td>(b) One form of crossover in four-strand stage</td>
<td><img src="image4" alt="Diagram" /></td>
<td><img src="image5" alt="Diagram" /></td>
<td><img src="image6" alt="Diagram" /></td>
<td></td>
</tr>
<tr>
<td>(c) An alternate crossover in four-strand stage</td>
<td><img src="image7" alt="Diagram" /></td>
<td><img src="image8" alt="Diagram" /></td>
<td><img src="image9" alt="Diagram" /></td>
<td></td>
</tr>
</tbody>
</table>

First-division segregation

Second-division segregation

Second-division segregation
Chapter 7  Linkage and Chromosome Mapping in Eukaryotes

FIGURE 7–17 The phenotypes and chromosome compositions of parents and recombinant offspring in Creighton and McClintock’s experiment in maize. The knob and translocated segment are the cytological markers that established that crossing over involves an actual exchange of chromosome arms.

Creighton and McClintock crossed this plant to a plant homozygous for the colored allele (c) and heterozygous for the endosperm alleles. They obtained a variety of different phenotypes in the offspring, but they were most interested in a crossover result involving the chromosome with the unique cytological markers. They examined the chromosomes of this plant with the colorless, waxy phenotype (case I in Figure 7–17) for the presence of the cytological markers. If physical exchange between homologs accompanies genetic crossing over, the translocated chromosome will still be present, but the knob will not—this is exactly what happened. In a second plant (case II), the phenotype colored, starchy should result from either nonrecombinant gametes or from crossing over. Some of the plants then ought to contain chromosomes with the dense knob but not the translocated chromosome. This condition was also found, and the conclusion that a physical exchange takes place was again supported. Along with Stern’s findings with Drosophila, this work clearly established that crossing over has a cytological basis.

Sister Chromatid Exchanges
Knowing that crossing over occurs between synapsed homologs in meiosis, we might ask whether a similar physical exchange occurs between homologs during mitosis. While homologous chromosomes do not usually pair up or synapse in somatic cells (Drosophila is an exception), each individual chromosome in prophase and metaphase of mitosis consists of two identical sister chromatids, joined at a common centromere. Surprisingly, several experimental approaches have demonstrated that reciprocal exchanges similar to crossing over occur between sister chromatids. These sister chromatid exchanges (SCEs) do not produce new allelic combinations, but evidence is accumulating that attaches significance to these events.

Identification and study of SCEs are facilitated by several modern staining techniques. In one technique, cells replicate for two generations in the presence of the thymidine analog bromodeoxyuridine (BUdR).* Following two rounds of replication, each pair of sister chromatids has one member with one strand of DNA “labeled” with BUdR and the other member with both strands labeled with BUdR. Using a differential stain, chromatids with the analog in both strands stain less brightly than chromatids with BUdR in only one strand. As a result, SCEs are readily detectable if they occur. In Figure 7–18, numerous instances of SCE events are clearly evident. These sister chromatids are sometimes referred to as harlequin chromosomes because of their patchlike appearance.

While the significance of SCEs is still uncertain, several observations have led to great interest in this phenomenon. We know, for example, that agents that induce chromosome damage (viruses, X-rays, ultraviolet light, and certain chem-

*The abbreviation BrdU is also used to denote bromodeoxyuridine.

FIGURE 7–18 Light micrograph of sister chromatid exchanges (SCEs) in mitotic chromosomes. Sometimes called harlequin chromosomes because of their patchlike appearance, chromatids with the thymidine analog BUdR in both DNA strands fluoresce less brightly than do those with the analog in only one strand. These chromosomes were stained with 33258-Hoechst reagent and acridine orange and then viewed using fluorescence microscopy.
ical mutagens) increase the frequency of SCEs. The frequency of SCEs is also elevated in Bloom syndrome, a human disorder caused by a mutation in the BLM gene on chromosome 15. This rare, recessively inherited disease is characterized by prenatal and postnatal retardation of growth, a great sensitivity of the facial skin to the sun, immune deficiency, a predisposition to malignant and benign tumors, and abnormal behavior patterns. The chromosomes from cultured leukocytes, bone marrow cells, and fibroblasts derived from homozygotes are very fragile and unstable compared to those of homozygous and heterozygous normal individuals. Increased breaks and rearrangements between nonhomologous chromosomes are observed in addition to excessive amounts of sister chromatid exchanges. Work by James German and colleagues suggests that the BLM gene encodes an enzyme called DNA helicase, which is best known for its role in DNA replication.

7.9 Did Mendel Encounter Linkage?

We conclude this chapter by examining a modern-day interpretation of the experiments that form the cornerstone of transmission genetics—Mendel’s crosses with garden peas. Some observers believe that Mendel had extremely good fortune in his classic experiments. He did not encounter apparent linkage relationships between the seven mutant characters in any of his crosses. Had Mendel obtained highly variable data characteristic of linkage and crossing over, these unorthodox ratios might have hindered his successful analysis and interpretation.

The article by Stig Blixt, reprinted in its entirety in the following box, demonstrates the inadequacy of this hypothesis. As we shall see, some of Mendel’s genes were indeed linked. We leave it to Stig Blixt to enlighten you as to why Mendel did not detect linkage.

**Why Didn’t Gregor Mendel Find Linkage?**

It is quite often said that Mendel was very fortunate not to run into the complication of linkage during his experiments. He used seven genes, and the pea has only seven chromosomes. Some have said that had he taken just one more, he would have had problems. This, however, is a gross oversimplification. The actual situation, most probably, is shown in Table 7.1. This shows that Mendel worked with three genes in chromosome 4, two genes in chromosome 1, and one gene in each of chromosomes 5 and 7. It seems at first glance that, out of the 21 dihybrid combinations Mendel theoretically could have studied, no fewer than four (that is, a–i, v–fa, v–le, fa–le) ought to have resulted in linkages. However, as found in hundreds of crosses and shown by the genetic map of the pea, a and i in chromosome 1 are so distantly located on the chromosome that no linkage is normally detected. The same is true for v and le on the one hand, and fa on the other, in chromosome 4. This leaves v–le, which ought to have shown linkage.

Mendel, however, seems not to have published this particular combination and thus, presumably, never made the appropriate cross to obtain both genes segregating simultaneously. It is therefore not so astonishing that Mendel did not run into the complication of linkage, although he did not avoid it by choosing one gene from each chromosome.

**STIG BLIXT**

Weibullsholm Plant Breeding Institute, Landskrona, Sweden, and Centro Energia Nucleare na Agricultura, Piracicaba, SP, Brazil.


<table>
<thead>
<tr>
<th>TABLE 7.1 Relationship between modern genetic terminology and character pairs used by Mendel</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Character Pair Used by Mendel</strong></td>
</tr>
<tr>
<td>Seed color, yellow–green</td>
</tr>
<tr>
<td>Seed coat and flowers, colored–white</td>
</tr>
<tr>
<td>Mature pods, smooth expanded–wrinkled indented</td>
</tr>
<tr>
<td>Inflorescences, from leaf axis–umbellate in top of plant</td>
</tr>
<tr>
<td>Plant height, 0.5–1 m</td>
</tr>
<tr>
<td>Unripe pods, green–yellow</td>
</tr>
<tr>
<td>Mature seeds, smooth–wrinkled</td>
</tr>
</tbody>
</table>
1. Genes located on the same chromosome are said to be linked. Alleles located on the same homolog, therefore, may be transmitted together during gamete formation. However, crossing over between homologs during meiosis results in the reshuffling of alleles and thereby contributes to genetic variability within gametes.

2. Early in the twentieth century, geneticists realized that crossing over provides an experimental basis for mapping the location of linked genes relative to one another along the chromosome.

3. Somatic cell hybridization techniques have made possible linkage and mapping analysis of human genes.

4. Mapping studies may also be performed with haploid organisms such as Chlamydomonas and Neurospora.

5. Cytological investigations of both maize and Drosophila reveal that crossing over involves a physical exchange of segments between nonsister chromatids.

6. An exchange of genetic material between sister chromatids can occur during mitosis as well. These events are referred to as sister chromatid exchanges (SCEs). An elevated frequency of such events is seen in the human disorder Bloom syndrome.

7. Evidence now suggests that several of the genes studied by Mendel are, in fact, linked. However, in such cases, the genes are sufficiently far apart to prevent the detection of linkage.

**Key Terms**

- Bloom syndrome, 157
- bromodeoxyuridine (BUdR), 156
- chiasmata, 140
- chromosomal theory of inheritance, 142
- chromosome map, 141
- coefficient of coincidence (C), 150
- complete linkage, 000
- crossing over, 137
- crossover gametes, 137
- cystic fibrosis, 154
- double crossover (DCO), 143
- first-division segregation, 154
- harlequin chromosomes, 156
- heterokaryon, 152
- Huntington disease, 154
- interference (I), 150
- linkage, 137
- linkage group, 140
- linkage ratio, 138
- lod score method, 152
- mapping the centromere, 154
- multiple-strand exchange, 150
- neurofibromatosis, 154
- noncrossover gametes, 137, 140
- parental gametes, 137
- reciprocal classes, 146
- recombinant gametes, 137, 140
- second-division segregation, 154
- sister chromatid exchanges (SCEs), 156
- somatic cell hybridization, 152
- synkaryon, 152
- synteny testing, 152
- tetrad, 154
- tetrad analysis, 154

**Insights and Solutions**

1. In rabbits, black color (B) is dominant to brown (b), while full color (C) is dominant to chinchilla (c^th). The genes controlling these traits are linked. Rabbits that are heterozygous for both traits and express black, full color are crossed to rabbits that express brown, chinchilla with the following results:

   - 31 brown, chinchilla
   - 34 black, full
   - 16 brown, full
   - 19 black, chinchilla

Determine the arrangement of alleles in the heterozygous parents and the map distance between the two genes.

**Solution:** This is a two-point map problem, where the two most prevalent reciprocal phenotypes are the noncrossovers. The less frequent reciprocal phenotypes arise from a single crossover. The arrangement of alleles is derived from the noncrossover phenotypes because they enter gametes intact.

The single crossovers give rise to 35/100 offspring (35%). Therefore, the distance between the two genes is 35 mu.
2. In Drosophila, Lyra (Ly) and Stubble (Sb) are dominant mutations located at locus 40 and 58, respectively, on chromosome III. A recessive mutation with bright red eyes is discovered and shown also to be located on chromosome III. A map is obtained by crossing a female who is heterozygous for all three mutations to a male that is homozygous for the bright-red mutation (which we will call br), and the data in the table are generated. Determine the location of the br mutation on chromosome III.

<table>
<thead>
<tr>
<th>Phenotype</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) Ly Sb br</td>
<td>404</td>
</tr>
<tr>
<td>(2) + + +</td>
<td>422</td>
</tr>
<tr>
<td>(3) Ly + +</td>
<td>18</td>
</tr>
<tr>
<td>(4) + Sb br</td>
<td>16</td>
</tr>
<tr>
<td>(5) Ly + br</td>
<td>75</td>
</tr>
<tr>
<td>(6) + Sb +</td>
<td>59</td>
</tr>
<tr>
<td>(7) Ly Sb +</td>
<td>4</td>
</tr>
<tr>
<td>(8) + + br</td>
<td>2</td>
</tr>
<tr>
<td>Total = 1000</td>
<td></td>
</tr>
</tbody>
</table>

Solution: First, determine the arrangement of the alleles on the homologs of the heterozygous crossover parent (the female in this case). To do this, locate the most frequent reciprocal phenotypes, which arise from the noncrossover gametes—these are phenotypes (1) and (2). Each phenotype represents the arrangement of alleles on one of the homologs. Therefore, the arrangement is

```
Ly  Sb  br
+    +    +
```

Second, find the correct sequence of the three loci along the chromosome. This is done by determining which sequence yields the observed double-crossover phenotypes, which are the least frequent reciprocal phenotypes (7 and 8). If the sequence is correct as written, then the double crossover depicted here,

```
Ly  Sb  br
+    +    +
```

will yield Ly + br and + Sb + as phenotypes. Inspection shows that these categories (5 and 6) are actually single crossovers, not double crossovers. Therefore, the sequence is incorrect, as written. Only two other sequences are possible: The br gene is either to the left of Ly (case A), or it is between Ly and Sb (case B).

Comparison with the actual data shows that case B is correct. The double-crossover gametes yield flies that express Ly and Sb but not br, or express br but not Ly and Sb. Therefore, the correct arrangement and sequence are shown below.

```
Ly  Sb  br
+    +    +
```

Once this sequence is found, determine the location of br relative to Ly and Sb. A single crossover between Ly and br, as shown here,

```
Ly  br  Sb
+    +    +
```

yields flies that are Ly + + and + br Sb (phenotypes 3 and 4). Therefore, the distance between the Ly and br loci is equal to

\[
\frac{18 + 16 + 4 + 2}{1000} = \frac{40}{1000} = 0.04 = 4 \text{ mu}
\]

(Cont. on the next page)
Remember to add the double crossovers because they represent two single crossovers occurring simultaneously. You need to know the frequency of all crossovers between Ly and br, so they must be included.

Similarly, the distance between the br and Sb loci is derived mainly from single crossovers between them.

This event yields Ly br + and + + Sb phenotypes (phenotypes 5 and 6). Therefore, the distance equals

$$\frac{75 + 59 + 4 + 2}{1000} = 140/1000 = 0.14 = 14 \text{ mu}$$

The final map shows that br is located at locus 44, since Lyra and Stubble are known.

![Map Diagram]

3. Refer to Figure 7–13, and predict what gene (which we called br) was discovered on chromosome III in Problem 2. Suggest an experimental cross that could confirm your prediction.

**Solution:** Inspection of Figure 7–13 reveals that the mutation scarlet (st) is present at locus 44.0, so it is reasonable to hypothesize that the bright-red eye mutation is an allele at the scarlet locus.

To test this hypothesis, you could perform complementation analysis (see Chapter 4) by crossing females expressing the bright-red mutation with known scarlet males. If the two mutations are alleles, no complementation will occur and all progeny will reveal a bright-red mutant eye phenotype. If complementation occurs, all progeny will express normal brick-red (wild-type) eyes, since the bright-red mutation and scarlet are at different loci (they are probably very close together). In such a case, all progeny will be heterozygous at both the bright-red eye and the scarlet loci and will not express either mutation because they are both recessive. This type of complementation analysis is called an allelism test.

### PROBLEMS AND DISCUSSION QUESTIONS

1. What is the significance of crossing over (which leads to genetic recombination) to the process of evolution?
2. Describe the cytological observation that suggests that crossing over occurs during the first meiotic prophase.
3. Why does more crossing over occur between two distantly linked genes than between two genes that are very close together on the same chromosome?
4. Why is a 50 percent recovery of single-crossover products the upper limit, even when crossing over always occurs between two linked genes?
5. Why are double-crossover events expected less frequently than single-crossover events?
6. What is the proposed basis for positive interference?
7. What three essential criteria must be met in order to execute a successful mapping cross?
8. The genes dumpy wings (dp), clot eyes (cl), and apterous wings (ap) are linked on chromosome II of Drosophila. In a series of two-point mapping crosses, the genetic distances shown below were determined. What is the sequence of the three genes?

<table>
<thead>
<tr>
<th>Cross</th>
<th>Distance</th>
</tr>
</thead>
<tbody>
<tr>
<td>dp–ap</td>
<td>42</td>
</tr>
<tr>
<td>dp–cl</td>
<td>3</td>
</tr>
<tr>
<td>ap–cl</td>
<td>39</td>
</tr>
</tbody>
</table>

9. Consider two hypothetical recessive autosomal genes a and b, where a heterozygote is testcrossed to a double-homozygous mutant. Predict the phenotypic ratios under the following conditions:
   (a) a and b are located on separate autosomes.
   (b) a and b are linked on the same autosome but are so far apart that a crossover always occurs between them.

(c) a and b are linked on the same autosome but are so close together that a crossover almost never occurs.

(d) a and b are linked on the same autosome about 10 mu apart.

10. Colored aleurone in the kernels of corn is due to the dominant allele R. The recessive allele r, when homozygous, produces colorless aleurone. The plant color (not kernel color) is controlled by another gene with two alleles, T and t. The dominant T allele results in green color, whereas the homozygous presence of the recessive t allele causes the plant to appear yellow. In a testcross between a plant of unknown genotype and phenotype and a plant that is homozygous recessive for both traits, the following progeny were obtained:

<table>
<thead>
<tr>
<th>Phenotype</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>colored, green</td>
<td>88</td>
</tr>
<tr>
<td>colored, yellow</td>
<td>12</td>
</tr>
<tr>
<td>colorless, green</td>
<td>8</td>
</tr>
<tr>
<td>colorless, yellow</td>
<td>92</td>
</tr>
</tbody>
</table>

Explain how these results were obtained by determining the exact genotype and phenotype of the unknown plant, including the precise association of the two genes on the homologs (i.e., the arrangement).

11. In the cross shown here, involving two linked genes, ebony (e) and claret (ca), in Drosophila, where crossing over does not occur in males, offspring were produced in a (2 + :1 ca:1 e) phenotypic ratio:

<table>
<thead>
<tr>
<th>Cross</th>
<th>Genotype</th>
</tr>
</thead>
<tbody>
<tr>
<td>$e^{+}ca^{+}$</td>
<td></td>
</tr>
<tr>
<td>$e^{+}ca$</td>
<td></td>
</tr>
<tr>
<td>$e^{+}ca^{+}$</td>
<td></td>
</tr>
<tr>
<td>$e^{+}ca^{+}$</td>
<td></td>
</tr>
</tbody>
</table>

$$e^{+}ca^{+} \times e^{+}ca^{+}$$

$$e^{+}ca \times e^{+}ca$$
14. Two different female Drosophila, the following recombinant (single-crossover) frequencies were observed:

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>pr-advp</td>
<td>29</td>
</tr>
<tr>
<td>pr-vg</td>
<td>13</td>
</tr>
<tr>
<td>pr-c</td>
<td>21</td>
</tr>
<tr>
<td>pr-b</td>
<td>6</td>
</tr>
<tr>
<td>advp-b</td>
<td>35</td>
</tr>
<tr>
<td>advp-c</td>
<td>8</td>
</tr>
<tr>
<td>advp-vg</td>
<td>16</td>
</tr>
<tr>
<td>vg-b</td>
<td>19</td>
</tr>
<tr>
<td>vg-c</td>
<td>8</td>
</tr>
<tr>
<td>c-b</td>
<td>27</td>
</tr>
</tbody>
</table>

(a) If the advp gene is present near the end of chromosome II (locus 83), construct a map of these genes.
(b) In another set of experiments, a sixth gene (d) was tested against b and pr, and the results were d-b = 17% and d-pr = 23%. Predict the results of two-point maps between d and c, d and vg, and d and advp.

15. In Drosophila, a cross was made between females expressing the three X-linked recessive traits, scute bristles (sc), saddle body (s), and vermilion eyes (v) and wild-type males. All females were wild type in the F1, while all males expressed all three mutant traits. The cross was carried to the F2 generation, and 1000 offspring were counted, with the results shown in the table. No de

<table>
<thead>
<tr>
<th>Female A</th>
<th>Female B</th>
<th>Gamete formation</th>
</tr>
</thead>
<tbody>
<tr>
<td>d b c</td>
<td>d + +</td>
<td>(1) d b c (5) d +</td>
</tr>
<tr>
<td>+ + +</td>
<td>+ c b</td>
<td>(2) + + + (6) b c</td>
</tr>
<tr>
<td>+ + c</td>
<td>+ d c</td>
<td>(3) + + c (7) d +</td>
</tr>
<tr>
<td>d b +</td>
<td>+ b +</td>
<td>(4) d b + (8) b c</td>
</tr>
</tbody>
</table>

16. A cross in Drosophila involved the recessive, X-linked genes yellow body (y), white eyes (w), and cut wings (ct). A yellow-bodied, white-eyed female with normal wings was crossed to a male whose eyes and body were normal, but whose wings were cut. The F1 females were wild type for all three traits, while the F1 males expressed the yellow-body, white-eye traits. The cross was carried to F2 progeny, and only male offspring were tallied. On the basis of the data shown here, a genetic map was constructed. (a) Diagram the genotypes of the F1 parents. (b) Construct a map, assuming that w is at locus 1.5 on the X chromosome. (c) Were any double-crossover offspring expected? (d) Could the F2 male offspring be used to construct the map? Why or why not?

<table>
<thead>
<tr>
<th>Phenoype</th>
<th>Offspring</th>
</tr>
</thead>
<tbody>
<tr>
<td>sc s v</td>
<td>314</td>
</tr>
<tr>
<td>+ + +</td>
<td>280</td>
</tr>
<tr>
<td>+ s v</td>
<td>150</td>
</tr>
<tr>
<td>sc + +</td>
<td>156</td>
</tr>
<tr>
<td>sc + v</td>
<td>46</td>
</tr>
<tr>
<td>+ s +</td>
<td>30</td>
</tr>
<tr>
<td>+ + v</td>
<td>14</td>
</tr>
</tbody>
</table>

17. In Drosophila, Dichaete (D) is a mutation on chromosome III with a dominant effect on wing shape. It is lethal when homozygous. The genes ebony body (e) and pink eye (p) are recessive mutations on chromosome III. Flies from a Dichaete stock were crossed to homozygous ebony, pink flies, and the F1 progeny with a Dichaete phenotype were backcrossed to the ebony, pink homozygotes. Using the results of this backcross shown in the table, (a) diagram the cross, showing the genotypes of the parents and offspring of both crosses. (b) What is the sequence and interlocus distance between these three genes?
22. What conclusion can be drawn from the observations that in
21. Are sister chromatid exchanges effective in producing genetic
23. An organism of the genotype AaBbCc was testcrossed to a triplo-

20. In a plant heterozygous for two gene pairs (Ab/aB), where the
two loci are linked and 25 mu apart, two such individuals were
crossed together. Assuming that crossing over occurs during
the formation of both male and female gametes and that the A
and B alleles are dominant, determine the phenotypic ratio of
the offspring.

24. Based on our discussion of the potential inaccuracy of mapping
(see Figure 7–12), would you revise your answer to Problem
23(c)? If so, how?

25. In a plant, fruit color is either red or yellow, and fruit shape is
either oval or long. Red and oval are the dominant traits. Two
plants, both heterozygous for these traits, were testcrossed, with
the results shown below. Determine the location of the genes rel-
ative to one another and the genotypes of the two parental plants.

26. In a cross in Neurospora involving two alleles, B and b, the
tetrad patterns in the following table were observed. Calculate
the distance between the gene and the centromere.

27. In a cross involving two alleles, B and b, the

28. In Creighton and McClintock’s experiment demonstrating that
crossing over involves physical exchange between chromosomes
(see Section 7.8), explain the importance of the cytological
markers (the translocated segment and the chromosome knob) in
the experimental rationale.
29. A number of human–mouse somatic cell hybrid clones were examined for the expression of specific human genes and the presence of human chromosomes—the results are summarized in this table. Assign each gene to the chromosome upon which it is located.

<table>
<thead>
<tr>
<th>Hybrid-Cell Clone</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Genes (expressed or not)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ENO1 (enolase-1)</td>
<td>–</td>
<td>+</td>
<td>–</td>
<td>+</td>
<td>+</td>
<td>–</td>
</tr>
<tr>
<td>MDH1 (malate dehydrogenase-1)</td>
<td>+</td>
<td>+</td>
<td>–</td>
<td>+</td>
<td>–</td>
<td>+</td>
</tr>
<tr>
<td>PEPS (peptidase S)</td>
<td>+</td>
<td>+</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>PGM1 (phosphoglucomutase-1)</td>
<td>–</td>
<td>+</td>
<td>–</td>
<td>+</td>
<td>+</td>
<td>–</td>
</tr>
<tr>
<td><strong>Chromosomes (present or absent)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>–</td>
<td>+</td>
<td>–</td>
<td>+</td>
<td>+</td>
<td>–</td>
</tr>
<tr>
<td>2</td>
<td>+</td>
<td>+</td>
<td>–</td>
<td>+</td>
<td>–</td>
<td>+</td>
</tr>
<tr>
<td>3</td>
<td>+</td>
<td>+</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>4</td>
<td>+</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>5</td>
<td>–</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

30. A female of genotype

\[
\frac{a}{+} \quad \frac{b}{+} \quad \frac{c}{+} \\
\frac{+}{+} \quad \frac{+}{+} \quad \frac{+}{+}
\]

produces 100 meiotic tetrads. Of these, 68 show no crossover events. Of the remaining 32, 20 show a crossover between \(a\) and \(b\), 10 show a crossover between \(b\) and \(c\), and 2 show a double crossover between \(a\) and \(b\) and between \(b\) and \(c\). Of the 400 gametes produced, how many of each of the 8 different genotypes will be produced? Assuming the order \(a-b-c\) and the allele arrangement shown above, what is the map distance between these loci?

31. \(D. \ melanogaster\) has one pair of sex chromosomes (XX or XY) and three autosomes (chromosomes II, III, and IV). A genetics student discovered a male fly with very short (\(sh\)) legs. Using this male, the student was able to establish a pure-breeding stock of this mutant and found that it was recessive. She then incorporated the mutant into a stock containing the recessive gene black (\(b\), body color, located on chromosome II) and the recessive gene pink (\(p\), eye color, located on chromosome III). A female from the homozygous black, pink, short stock was then mated to a wild-type male. The \(F_1\) males of this cross were all wild type and were then backcrossed to the homozygous \(b, p, sh\) females. The \(F_2\) results appeared as shown in the table that follows, and no other phenotypes were observed. (a) Based on these results, the student was able to assign \(sh\) to a linkage group (a chromosome). Determine which chromosome, and include step-by-step reasoning. (b) The student repeated the experiment, making the reciprocal cross: \(F_1\) females backcrossed to homozygous \(b, p, sh\) males. She observed that 85 percent of the offspring fell into the given classes, but that 15 percent of the offspring were equally divided among \(b+p, b++, +shp, \) and \(+sh+\) phenotypic males and females. How can these results be explained, and what information can be derived from these data?

| Phenotype Female Male |
|-----------------------|------------------|
| wild                  | 63               |
| pink*                 | 58               |
| black, short          | 55               |
| black, pink, short    | 69               |

*Pink indicates that the other two traits are wild type (normal). Similarly, black, short offspring are wild type for eye color.

32. In \(Drosophila\), a female fly is heterozygous for three mutations, \(\text{Bar}\) eyes (\(B\)), \(\text{miniature}\) wings (\(m\)), and \(\text{ebony}\) body (\(e\)). (Note that \(\text{Bar}\) is a dominant mutation.) The fly is crossed to a male with normal eyes, miniature wings, and ebony body. The results of the cross are shown below. Interpret the results of this cross. If you conclude that linkage is involved between any of the genes, determine the map distance(s) between them.

<table>
<thead>
<tr>
<th>Phenotype</th>
<th>Female</th>
<th>Male</th>
</tr>
</thead>
<tbody>
<tr>
<td>miniature</td>
<td>111</td>
<td>Bar</td>
</tr>
<tr>
<td>wild</td>
<td>29</td>
<td>Bar</td>
</tr>
<tr>
<td>Bar, miniature</td>
<td>101</td>
<td>ebony</td>
</tr>
<tr>
<td>Bar, miniature, ebony</td>
<td>31</td>
<td>miniature, ebony</td>
</tr>
</tbody>
</table>
Cancer is a group of genetic diseases affecting fundamental aspects of cellular function, including DNA repair, the cell cycle, apoptosis, differentiation, cell migration and cell–cell contact.

Most cancer-causing mutations occur in somatic cells; only about 1 percent of cancers have a hereditary component.

Mutations in cancer-related genes lead to abnormal proliferation and loss of control over how cells spread and invade surrounding tissues.

The development of cancer is a multistep process requiring mutations in genes controlling many aspects of cell proliferation and metastasis.

Cancer cells show high levels of genomic instability, leading to the accumulation of multiple mutations in cancer-related genes.

Mutations in proto-oncogenes and tumor suppressor genes contribute to the development of cancers.

Oncogenic viruses introduce oncogenes into infected cells and stimulate cell proliferation.

Environmental agents contribute to cancer by damaging DNA.
Cancer is the second leading cause of death in Western countries, surpassed only by heart disease. It strikes people of all ages, and one out of three people will experience a cancer diagnosis sometime in his or her lifetime. Each year, more than 1 million cases of cancer are diagnosed in the United States and more than 500,000 people die from the disease (Table 16.1).

Over the last 30 years, scientists have discovered that cancer is a genetic disease, characterized by an interplay of mutant forms of oncogenes and tumor suppressor genes leading to the uncontrolled growth and spread of cancer cells. While some of these mutations may be inherited, as we will see, most all of them that lead to disease occur in somatic cells that then divide and lead to tumors. Completion of the Human Genome Project is opening the door to a wealth of new information about the mutations that trigger a cell to become cancerous. This new understanding of cancer genetics is also leading to new gene-specific treatments, some of which are now entering clinical trials. Some scientists predict that gene therapies will replace chemotherapies within the next 25 years.

The goal of this chapter is to highlight our current understanding of the nature and causes of cancer. As we will see, cancer is a genetic disease that arises from mutations in genes controlling many basic aspects of cellular function. We will examine the relationship between genes and cancer, and consider how mutations, chromosomal changes, and environmental agents play roles in the development of cancer.

**How Do We Know?**

In this chapter, we will focus on cancer as a genetic disease, with an emphasis on the relationship between cancer and DNA damage, as well as on the multiple genetic steps that lead to cancer. We also discuss how cancer cells show defects in cell-cycle regulation. We conclude with an investigation of the roles played by viruses and environmental agents in the development of cancer. As you study this topic, you should try to answer several fundamental questions:

1. How do we know that cancers arise from a single cell that contains mutations?
2. How do we know that cancer development requires more than one mutation?
3. How do we know that cancer cells contain defects in DNA repair?
4. How do we know that most cancers are not hereditary?
5. How do we know that some viruses contain genes that contribute to the development of cancer?

**16.1 Cancer Is a Genetic Disease at the Level of Somatic Cells**

Perhaps the most significant development in understanding the causes of cancer is the realization that cancer is a genetic disease. Genomic alterations that are associated with cancer range from single-nucleotide substitutions to large-scale chromosome rearrangements, amplifications, and deletions (Figure 16–1). However, unlike other genetic diseases, cancer is caused by mutations that occur predominantly in somatic cells. Only about 1 percent of cancers are associated with germ-line mutations that increase a person’s susceptibility to certain types of cancer. Another important difference between cancer and other genetic diseases is that cancers rarely arise from a single mutation, but from the accumulation of many mutations—as many as six to ten. The mutations that lead to cancer affect multiple cellular functions, including repair of DNA damage, cell division, apoptosis, cellular differentiation, migratory behavior, and cell–cell contact.

**What Is Cancer?**

Clinically, cancer is defined as a large number of complex diseases, up to a hundred, that behave differently depending on the cell types from which they originate. Cancers vary in their ages of onset, growth rates, invasiveness, prognoses, and responsive-
Cancer Probabilities in the United States

<table>
<thead>
<tr>
<th>Cancer Site</th>
<th>Gender</th>
<th>Birth to 39</th>
<th>40–59</th>
<th>60–79</th>
<th>Birth to Death</th>
</tr>
</thead>
<tbody>
<tr>
<td>All sites</td>
<td>Male</td>
<td>1 in 62</td>
<td>1 in 12</td>
<td>1 in 3</td>
<td>1 in 2</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>1 in 52</td>
<td>1 in 11</td>
<td>1 in 4</td>
<td>1 in 3</td>
</tr>
<tr>
<td>Breast</td>
<td>Female</td>
<td>1 in 235</td>
<td>1 in 25</td>
<td>1 in 15</td>
<td>1 in 8</td>
</tr>
<tr>
<td>Prostate</td>
<td>Male</td>
<td>&lt;1 in 10,000</td>
<td>1 in 53</td>
<td>1 in 7</td>
<td>1 in 6</td>
</tr>
<tr>
<td>Lung, bronchus</td>
<td>Male</td>
<td>1 in 3300</td>
<td>1 in 92</td>
<td>1 in 17</td>
<td>1 in 13</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>1 in 3180</td>
<td>1 in 120</td>
<td>1 in 25</td>
<td>1 in 17</td>
</tr>
<tr>
<td>Colon, rectum</td>
<td>Male</td>
<td>1 in 1500</td>
<td>1 in 124</td>
<td>1 in 29</td>
<td>1 in 18</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>1 in 1900</td>
<td>1 in 149</td>
<td>1 in 33</td>
<td>1 in 18</td>
</tr>
</tbody>
</table>

Source: American Cancer Society

Cancer is a genetic disease at the level of somatic cells. It is caused by mutations that lead to uncontrolled cell proliferation and metastasis, a single mutation is not sufficient to transform a normal cell into a tumor-forming (tumorigenic), malignant cell. If it were sufficient, then cancer would be far more prevalent than it is. In humans, mutations occur spontaneously at a rate of about $10^{-6}$ mutations per gene, per cell division, mainly due to the intrinsic error rates of DNA replication. As there are approximately $10^{16}$ cell divisions in a human body during a lifetime, a person might suffer up to $10^{10}$ mutations per gene anywhere in the body, during his or her lifetime. However, only about one person in three will suffer from cancer.

The phenomenon of age-related cancer is another indication that cancer develops from the accumulation of several mutagenic events in a single cell. The incidence of most cancers rises exponentially with age. If only a single mutation were sufficient to convert a normal cell to a malignant one, then cancer incidence would appear to be independent of age. The age-related incidence of cancer suggests that up to 10 independent mutations, occurring randomly and with a low probability, are necessary before a cell is transformed into a malignant cancer.
Cancer Cells Contain Genetic Defects Affecting Genomic Stability and DNA Repair

Cancer cells show higher than normal rates of mutation, chromosomal abnormalities, and genomic instability. In fact, many researchers believe that the fundamental defect in cancer cells is a derangement of the cells’ normal ability to repair DNA damage. This loss of genomic integrity leads to a general increase in the mutation rate in every gene, including specific genes that control aspects of cell proliferation, programmed cell death, and cell–cell contact. In turn, the accumulation of mutations in genes controlling these processes leads to cancer. The high level of genomic instability seen in cancer cells is known as the mutator phenotype.

Genomic instability in cancer cells manifests itself by the presence of gross defects such as translocations, aneuploidy, chromosome loss, DNA amplification, and chromosome deletions (Figures 16–1 and 16–2). Cancer cells that are grown in cultures in the lab also show a great deal of genomic instability—duplicating, losing, and translocating chromosomes or parts of chromosomes. Often cancer cells show specific chromosomal defects that are used to diagnose the type and stage of the cancer. For example, leukemic white blood cells from patients with chronic myelogenous leukemia (CML) bear a specific translocation involving the long arms of chromosomes 9 and 22 results in the formation of a characteristic chromosome, the Philadelphia chromosome (Figure 16–3). The BCR-ABL fusion gene codes for a chimeric BCR-ABL protein. The normal ABL protein is a protein kinase that acts within signal transduction pathways, transferring growth factor signals from the external environment to the nucleus. The BCR-ABL protein is an abnormal signal transduction molecule in CML cells, constantly stimulating these cells to proliferate even in the absence of external growth signals.

In keeping with the concept of the cancer mutator phenotype, a number of inherited cancers are caused by defects in genes that control DNA repair. For example, xeroderma pigmentosum (XP) is a rare hereditary disorder that is characterized by extreme sensitivity to ultraviolet light and other carcinogens. Patients with XP often develop skin cancer. Cells from patients with XP are defective in nucleotide excision repair, with mutations appearing in any one of seven genes whose products are necessary to carry out DNA repair. XP cells are impaired in their ability to repair DNA lesions such as thymine dimers induced by UV light. The...
Another hereditary cancer, hereditary nonpolyposis colorectal cancer (HNPCC), is also caused by mutations in genes controlling DNA repair. HNPCC is an autosomal dominant syndrome, affecting about one in every 200 people (Figure 16–4). Patients affected by HNPCC have an increased risk of developing colon, ovary, uterine, and kidney cancers. Cells from patients with HNPCC show higher than normal mutation rates and genomic instability. At least eight genes are associated with HNPCC, and four of these genes control aspects of DNA mismatch repair. Inactivation of any of these four genes—\( MSH2 \), \( MSH6 \), \( MLH1 \), and \( MLH3 \)—causes a rapid accumulation of genome-wide mutations and the subsequent development of colorectal and other cancers.

The observation that hereditary defects in genes controlling nucleotide excision repair and DNA mismatch repair lead to high rates of cancer lends support to the idea that the mutator phenotype is a significant contributor to the development of cancer.

**Problem 17 on page 374 asks you to consider how the BCR-ABL hybrid fusion protein, found in CML leukemic white blood cells, could be used as a target for cancer therapy.**

**Hint:** Most cancer therapies, including radiation and chemotherapies, aim to kill cells that are constantly dividing. However, many normal cells in the body also divide and are killed by cancer therapies, leading to side effects. The BCR-ABL fusion protein is found only in CML white blood cells. As you try to answer this problem, you may wish to learn more about the drug Gleevec (see http://www.nci.nih.gov/newscenter/qandagleevect).

**FIGURE 16–4** Pedigree of a family with HNPCC. Families with HNPCC are defined as those in which at least three relatives in two generations have been diagnosed with colon cancer, with one relative diagnosed at less than 50 years of age. Colon cancer: \( C \); stomach cancer: \( S \); endometrial cancer: \( E \); pancreatic cancer: \( P \); bladder/urinary cancer: \( B \). Blue symbols indicate family members with colon cancer; diagonal stripes mean that diagnosis is uncertain; orange symbols indicate other tumors. Symbols with slashes indicate deceased individuals. Reprinted with permission from Aaltonen et al. Clues to the pathogenesis of familial colorectal cancer. *Science* 260: 812-816. Copyright 1993 AAAS.

**16.3 Cancer Cells Contain Genetic Defects Affecting Cell-Cycle Regulation**

One of the fundamental aberrations in all cancer cells is a loss of control over cell proliferation. Cell proliferation is the process of cell growth and division that is essential for all development and tissue repair in multicellular organisms. Although some cells, such as epidermal cells of the skin or blood-forming cells in the bone marrow, continue to grow and divide throughout the organism’s lifetime, most cells in adult multicellular organisms are in a nondividing, quiescent, and differentiated state. Differentiated cells are those that are specialized for a specific function, such as photoreceptor cells of the retina or muscle cells of the heart. The most extreme examples of nonproliferating cells are nerve cells, which little, if at all, even to replace damaged tissue. In contrast, many differentiated cells, such as those in the liver and kidney, are able to grow and divide when stimulated by extracellular signals and growth factors. In this way, multicellular organisms are able to replace dead and damaged tissue. However, the growth and differentiation of cells must be strictly regulated; otherwise, the integrity of organs and tissues would be compromised by inappropriate types and quantities of cells. Normal regulation over cell proliferation involves a large number of gene products that control steps in the cell cycle, programmed cell death, and the response of cells to external growth signals. In cancer cells, many of the genes that control these functions are mutated or aberrantly expressed, leading to uncontrolled cell proliferation.

In this section, we will review steps in the cell cycle, some of the genes that control the cell cycle, and how these genes, when mutated, lead to cancer.
the nucleus even in the absence of external growth signals. In addition, malignant cells may not respond to external signals from surrounding cells—signals that would normally inhibit cell proliferation within a mature tissue.

Cell-Cycle Control and Checkpoints

In normal cells, progress through the cell cycle is tightly regulated, and the completion of each step is necessary prior to initiating the next step. There are at least three distinct points in the cell cycle at which the cell monitors external signals and internal equilibrium before proceeding to the next stage of the cell cycle. These are the **G1/S**, the **G2/M**, and **M checkpoints** (Figure 16–5). At the G1/S checkpoint, the cell monitors its size and determines if its DNA has been damaged. If the cell has not achieved an adequate size, or if the DNA has been damaged, further progress through the cell cycle is halted until these conditions are corrected. If cell size and DNA integrity are normal, the G1/S checkpoint is traversed, and the cell proceeds to S phase. The second important checkpoint is the G2/M checkpoint, where physiological conditions in the cell are monitored prior to entering mitosis. If DNA replication or repair of any DNA damage has not been completed, the cell cycle arrests until these processes are complete. The third major checkpoint occurs during mitosis and is called the M checkpoint. At this checkpoint, both the successful formation of the spindle-fiber system and the attachment of spindle fibers to the kinetochores associated with the centromeres are monitored. If spindle fibers are not properly formed or attachment is inadequate, mitosis is arrested.

In addition to regulating the cell cycle at checkpoints, the cell controls progress through the cell cycle through two classes of proteins: **cyclins** and **cyclin-dependent kinases (CDKs)**. The cell synthesizes and destroys cyclin proteins in a precise pattern during the cell cycle (Figure 16–6). When a cyclin is present, it binds to a specific CDK, triggering activity of the CDK/cyclin complex. The CDK/cyclin complexes then selectively phosphorylate and activate other proteins that in turn bring about the changes necessary to advance the cell through the cell cycle. For example, in G1 phase, CDK4/cyclin D complexes activate proteins that stimulate transcription of genes whose products (such as DNA polymerase δ and DNA ligase) are required for DNA replication during S phase. Another CDK/cyclin complex, CDK1/cyclin B, phosphorylates a number of proteins that bring about the events of early mitosis, such as nuclear membrane breakdown, chromosome condensation and cytoskeletal reorganization. Mitosis can only be completed, however, when cyclin B is degraded and the protein phosphorylations characteristic of M phase are reversed. Although a large number of different protein kinases exist in cells, only a few are involved in cell-cycle regulation.

Both cell-cycle checkpoints and cell-cycle control molecules are genetically regulated. In general, the cell cycle is regulated by an interplay of genes whose products either promote or suppress cell division. Mutation or mis-expression of any of the genes controlling the cell cycle contributes to the

---

**Figure 16–5** Checkpoints and proliferation decision points monitor the progress of the cell through the cell cycle.
Cancer Cells Contain Genetic Defects Affecting Cell-Cycle Regulation

Phases of the cell cycle

Relative amounts of cyclins

D1
G1
M
S
G2

FIGURE 16–6 Relative expression times and amounts of cyclins during the cell cycle. Cyclin D1 accumulates early in G1 and is expressed at a constant level through most of the cycle. Cyclin E accumulates in G1, reaches a peak, and declines by mid-S phase. Cyclin D2 begins accumulating in the last half of G1, reaches a peak just after the beginning of S, and then declines by early G2. Cyclin A appears in late G1, accumulates through S phase, peaks at the G2/M transition, and is rapidly degraded. Cyclin B peaks at the G2/M transition and declines rapidly in M phase.

The development of cancer in several ways. For example, if genes that control the G1/S or G2/M checkpoints are mutated, the cell may continue to cycle before repairing DNA damage. This may lead to the accumulation of mutations in genes whose product control cell proliferation or metastasis. Similarly, if genes that control progress through the cell cycle, such as those that encode the cyclins, are expressed inappropriately, the cell may cycle continuously and may be unable to exit the cell cycle into G0.

As already described, if DNA replication, repair, or chromosome assembly is aberrant, normal cells halt their progress through the cell cycle until the condition is corrected. This reduces the number of mutations and chromosomal abnormalities that accumulate in normal proliferating cells. However, if DNA or chromosomal damage is so severe that repair is impossible, the cell may initiate a second line of defense—a process called apoptosis, or programmed cell death. Apoptosis is a genetically controlled process whereby the cell commits suicide. Apoptosis is also initiated during normal multicellular development in order to eliminate certain cells that do not contribute to the final adult organism. The steps in apoptosis are the same for damaged cells and for cells eliminated during development: Nuclear DNA becomes fragmented, internal cellular structures are disrupted, and the cell dissolves into small spherical structures known as apoptotic bodies (Figure 16–7). In the final step, the apoptotic bodies are engulfed by the immune system’s phagocytic cells. A series of proteases called caspases are responsible for initiating apoptosis and for digesting intracellular components. Apoptosis is genetically controlled in that regulation of specific gene products such as the Bcl2 and BAX proteins can trigger or prevent apoptosis. By removing damaged cells,

FIGURE 16–7 (a) A normal white blood cell (bottom) and a white blood cell undergoing apoptosis (top). Apoptotic bodies appear as grape-like clusters on the cell surface. (b) The relative concentrations of the Bcl2 and BAX proteins regulate apoptosis. A normal cell contains a balance of Bcl2 and BAX, which form inactive heterodimers. A relative excess of Bcl2 results in Bcl2 homodimers, which prevent apoptosis. Cancer cells with Bcl2 overexpression are resistant to chemotherapies and radiation therapies. A relative excess of BAX results in BAX homodimers, which induce apoptosis. In normal cells, activated p53 protein induces transcription of BAX and inhibits transcription of Bcl2, leading to cell death. In many cancer cells, p53 is defective, preventing the apoptotic pathway from removing the cancer cells.
programmed cell death reduces the number of mutations that are propagated to the next generation, including those in cancer-causing genes. The same genes that regulate cell-cycle checkpoints can trigger apoptosis. As we will see, these genes are mutated in many cancers.

16.4 Many Cancer-Causing Genes Disrupt Control of the Cell Cycle

Two general categories of genes are mutated or mis-expressed in cancer cells—the proto-oncogenes and the tumor suppressor genes (Table 16.2). Proto-oncogenes are genes whose products are important for normal cell functions and promote cell growth and division. They do this by encoding transcription factors that stimulate expression of other genes, signal transduction molecules that stimulate cell division, or cell-cycle regulators that move the cell through the cell cycle. When normal cells become quiescent and cease division, they repress the expression or activity of most proto-oncogene products. In cancer cells, one or more proto-oncogenes are altered in such a way that their activities cannot be controlled in a normal fashion. This is sometimes due to a mutation in the proto-oncogene resulting in a protein product that acts abnormally. In other cases, proto-oncogenes are overexpressed or cannot be transcriptionally repressed at the correct time. In these cases, the proto-oncogene is continually in an “on” state, which may constantly stimulate the cell to divide. When a proto-oncogene is mutated or aberrantly expressed and contributes to the development of cancer, it is known as an oncogene (cancer-causing gene). Oncogenes are those that have experienced a gain-of-function alteration. As a result, only one allele of a proto-oncogene needs to be mutated or mis-expressed in order to trigger uncontrolled growth. Hence, oncogenes confer a dominant cancer phenotype.

Tumor suppressor genes are those whose products normally regulate cell-cycle checkpoints and initiate the process of apoptosis. In normal cells, proteins encoded by tumor suppressor genes halt progress through the cell cycle in response to DNA damage or growth-suppression signals from the extracellular environment. When tumor suppressor genes are mutated or inactivated, cells are unable to respond normally to cell-cycle checkpoints, or are unable to undergo programmed cell death if DNA damage is extensive. This leads to a further increase in mutations and to the inability of the cell to leave the cell cycle when it should become quiescent.

<table>
<thead>
<tr>
<th>Proto-oncogene</th>
<th>Normal Function</th>
<th>Alteration in Cancer</th>
<th>Associated Cancers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ha-ras</td>
<td>Signal transduction molecule, binds GTP/GDP</td>
<td>Point mutations</td>
<td>Colorectal, bladder, many types</td>
</tr>
<tr>
<td>c-erbB</td>
<td>Transmembrane growth factor receptor</td>
<td>Gene amplification, point mutations</td>
<td>Glioblastomas, breast cancer, cervix</td>
</tr>
<tr>
<td>c-myc</td>
<td>Transcription factor, regulates cell cycle, differentiation, apoptosis</td>
<td>Translocation, amplification, point mutations</td>
<td>Lymphomas, leukemias, lung cancer, many types</td>
</tr>
<tr>
<td>c-kit</td>
<td>Tyrosine kinase, signal transduction</td>
<td>Mutation</td>
<td>Sarcomas</td>
</tr>
<tr>
<td>RARα</td>
<td>Hormone-dependent transcription factor, differentiation</td>
<td>Chromosomal translocations with PML gene, fusion product</td>
<td>Acute promyelocytic leukemia</td>
</tr>
<tr>
<td>E6</td>
<td>Human papillomavirus encoded oncogene, inactivates p53</td>
<td>HPV infection</td>
<td>Cervical cancer</td>
</tr>
<tr>
<td>Cyclins</td>
<td>Bind to CDKs, regulate cell cycle</td>
<td>Gene amplification, overexpression</td>
<td>Lung, esophagus, many types</td>
</tr>
<tr>
<td>CDK2, 4</td>
<td>Cyclin-dependent kinases, regulate cell-cycle phases</td>
<td>Overexpression, mutation</td>
<td>Bladder, breast, many types</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Tumor Suppressor</th>
<th>Normal Function</th>
<th>Alteration in Cancer</th>
<th>Associated Cancers</th>
</tr>
</thead>
<tbody>
<tr>
<td>p53</td>
<td>Cell-cycle checkpoints, apoptosis</td>
<td>Mutation, inactivation by viral oncogene products</td>
<td>Brain, lung, colorectal, breast, many types</td>
</tr>
<tr>
<td>RB1</td>
<td>Cell-cycle checkpoints, binds E2F</td>
<td>Mutation, deletion, inactivation by viral oncogene products</td>
<td>Retinoblastoma, osteosarcoma, many types</td>
</tr>
<tr>
<td>APC</td>
<td>Cell–cell interaction</td>
<td>Mutation</td>
<td>Colorectal cancers, brain, thyroid</td>
</tr>
<tr>
<td>Bcl2</td>
<td>Apoptosis regulation</td>
<td>Overexpression blocks apoptosis</td>
<td>Lymphomas, leukemias</td>
</tr>
<tr>
<td>BRCA2</td>
<td>DNA repair</td>
<td>Point mutations</td>
<td>Breast, ovarian, prostate cancers</td>
</tr>
</tbody>
</table>
When both alleles of a tumor suppressor gene are inactivated, and other changes in the cell keep it growing and dividing, cells may become tumorigenic.

The following are examples of proto-oncogenes and tumor suppressor genes that contribute to cancer when mutated. There are more than 200 oncogenes and tumor suppressor genes now known, and more will likely be discovered as cancer research continues.

**The ras Proto-oncogenes**

Some of the most frequently mutated genes in human tumors are those of the ras gene family. These genes are mutated in more than 40 percent of human tumors. The ras gene family encodes signal transduction molecules that are associated with the cell membrane and regulate cell growth and division. Ras proteins normally transmit signals from the cell membrane to the nucleus, stimulating the cell to divide in response to external growth factors. Ras proteins alternate between an inactive (switched off) and an active (switched on) state by binding either guanosine diphosphate (GDP) or guanosine triphosphate (GTP). When a cell encounters a growth factor (such as platelet-derived growth factor or epidermal growth factor), growth factor receptors on the cell membrane bind to the growth factor, resulting in autophosphorylation of the cytoplasmic portion of the growth factor receptor. This causes recruitment of proteins known as nucleotide exchange factors to the plasma membrane. These nucleotide exchange factors cause Ras to release GDP and bind GTP, thereby activating Ras. The active, GTP-bound form of Ras then sends its signals through cascades of protein phosphorylations in the cytoplasm (Figure 16–8). The end-point of these cascades is activation of nuclear transcription factors that stimulate expression of genes whose products drive the cell from quiescence into the cell cycle. Once Ras has sent its signals to the nucleus, it hydrolyzes GTP to GDP and becomes inactive. Mutations that convert the proto-oncogene ras to an oncogene prevent the Ras protein from hydrolyzing GTP to GDP and hence freeze the Ras protein into its “on” conformation, constantly stimulating the cell to divide.

**The p53 Tumor Suppressor Gene**

The most frequently mutated gene in human cancers—occurring in more than 50 percent of all cancers—is the p53 gene. This gene encodes a nuclear protein that acts as a transcription factor that represses or stimulates transcription of more than 50 different genes.

Normally, the p53 protein is continuously synthesized but is rapidly degraded and therefore is present in cells at low levels. In addition, the p53 protein is normally bound to another protein called Mdm2, which prevents the phosphorylations and acetylations that convert the p53 protein from an inactive to an active form. Several types of events cause a rapid increase in the nuclear levels of activated p53 protein. These include chemical damage to DNA, double-stranded breaks in DNA induced by ionizing radiation, or the presence of DNA-repair intermediates generated by exposure of cells to ultraviolet light. Increases in the levels of activated p53 protein result from increases in protein phosphorylation, acetylation, and p53 protein stability.

The p53 protein initiates two different responses to DNA damage: arrest of the cell cycle followed by DNA repair, or apoptosis and cell death if DNA cannot be repaired. Each of these responses is accomplished by p53 acting as a transcription factor that stimulates or represses the expression of genes involved in each of these responses.

In normal cells, p53 can arrest the cell cycle at several phases. To arrest the cell cycle at the G1/S checkpoint, activated p53 protein stimulates transcription of a gene encoding the p21 protein. The p21 protein inhibits the CDK4/cyclin D1 complex, hence preventing the cell from moving from the G1 phase into S phase. Activated p53 protein also regulates expression of genes that retard the progress of DNA replication, thus allowing time for DNA damage to be repaired during S phase. By regulating expression of other genes, activated p53 can block cells at the G2/M checkpoint, if DNA damage occurs during S phase.

Activated p53 can also instruct a damaged cell to commit suicide by apoptosis. It does so by activating the transcription of the Bax gene and repressing transcription of the Bcl2 gene. In normal cells, the BAX protein is present in a heterodimer...
with the Bcl2 protein and the cell remains viable. When the levels of BAX protein increase in response to p53 stimulation of Bax gene transcription, BAX homodimers are formed and these homodimers activate the cellular changes that lead to cellular self-destruction (Figure 16–7). In cancer cells that lack functional p53, BAX protein levels do not increase in response to cellular damage and apoptosis may not occur.

Hence, cells lacking functional p53 are unable to arrest at cell-cycle checkpoints or to enter apoptosis in response to DNA damage. As a result, they move unchecked through the cell cycle, regardless of the condition of the cell’s DNA. Cells lacking p53 have high mutation rates and accumulate the types of mutations that lead to cancer. Because of the importance of the p53 gene to genomic integrity, it is often referred to as the “guardian of the genome.”

**The RB1 Tumor Suppressor Gene**

The loss or mutation of the *RB1* (retinoblastoma 1) tumor suppressor gene contributes to the development of many cancers, including those of the breast, bone, lung, and bladder. The *RB1* gene was originally identified as a result of studies on retinoblastoma, an inherited disorder in which tumors develop in the eyes of young children. Retinoblastoma occurs with a frequency of about 1 in 14,000 to 20,000 individuals. In the familial form of the disease, individuals inherit one mutated allele of the *RB1* gene and have an 85 percent chance of developing retinoblastomas as well as an increased chance of developing other cancers. All somatic cells of patients with hereditary retinoblastoma contain one mutated allele of the *RB1* gene. However, it is only when the second normal allele of the *RB1* gene is lost or mutated in certain retinal cells that retinoblastoma develops. In individuals that do not have this hereditary condition, retinoblastoma is extremely rare, as it requires at least two separate somatic mutations in a retinal cell in order to inactivate both copies of the *RB1* gene (Figure 16–9).

The **retinoblastoma protein** (pRB) is a tumor suppressor protein that controls the G1/S cell-cycle checkpoint. The pRB protein is found in the nuclei of all cell types at all stages of the cell cycle. However, its activity varies throughout the cell cycle, depending on its phosphorylation state. When cells are in the G0 phase of the cell cycle, the pRB protein is nonphosphorylated and binds to transcription factors such as E2F, inactivating them (Figure 16–10). When the cell is stimulated by growth factors, it enters G1 and approaches S phase. Throughout the G1 phase, the pRB protein becomes phosphorylated by the CDK4/cyclin D1 complex. Phosphorylated pRB is inactive and releases its bound regulatory proteins. When E2F and other regulators are released by pRB, they are free to induce the expression of over 30 genes whose products are required for the transition from G1 into S phase. After cells traverse S, G2, and M phases, pRB reverts to a nonphosphorylated state, binds to regulatory proteins such as E2F, and keeps them...
As discussed at the beginning of this chapter, uncontrolled growth alone is insufficient to create a malignant and life-threatening cancer. Cancer cells must also acquire the features of metastasis, which include the ability to disengage from the original tumor site, to enter the blood or lymphatic system, to invade surrounding tissues, and to develop into secondary tumors. In order to leave the site of the primary tumor and invade other tissues, tumor cells must dissociate from other cells and digest components of the extracellular matrix and basal lamina, which normally contain and separate tissues. The extracellular matrix and basal lamina are composed of proteins and carbohydrates. They form the scaffold for tissue growth and normally inhibit the migration of cells.

The ability to invade the extracellular matrix is also a property of some normal cell types. For example, implantation of the embryo in the uterine wall during pregnancy requires cell migration across the extracellular matrix. In addition, white blood cells reach the site of infection by penetrating capillary walls. The mechanisms of invasion are probably similar in these normal cells and in cancer cells. The difference is that, in normal cells, the invasive ability is tightly regulated, whereas in tumor cells, this regulation has been lost.

Although less is known about the genes that control metastasis than about those controlling the cell cycle, it is likely that metastasis is controlled by a large number of genes, including those that encode cell-adhesion molecules, cytoskeleton regulators, and proteolytic enzymes. For example, epithelial tumors have a lower than normal level of the E-cadherin glycoprotein, which is responsible for cell–cell adhesion in normal tissues. Also, proteolytic enzymes such as metalloproteinases are present at higher than normal levels in highly malignant tumors and are not susceptible to the normal controls conferred by regulatory molecules such as tissue inhibitors of metalloproteinases (TIMPs). It has been shown that the level of aggressiveness of a tumor correlates positively with the levels of proteolytic enzymes expressed by the tumor. Hence, inappropriately expressed cell adhesion and protease enzymes may assist malignant tumor cells by loosening the normal constraints on cell location and creating holes through which the tumor cells can pass into and out of the circulatory system.

**Now Solve This**

Problem 23 on page 374 asks you to explain how in the inherited Li–Fraumeni syndrome, mutations in one allele of the p53 gene can give rise to a wide variety of different cancers.

**Hint:** To answer this question, you might review the cellular functions regulated by the normal p53 tumor suppressor gene. Consider how each of these functions could be affected if the p53 gene product is either absent or defective. To understand how mutations in one allele of a tumor suppressor gene result in tumors that contain aberrations in both alleles, read about loss of heterozygosity in Section 16.6.

**Now Solve This**

Problem 25 on page 374 describes the isolation of a gene that is mutated in metastatic tumors. The gene appears to be a member of the serine protease family. You are asked to conjecture how this gene might contribute to the development of highly invasive cancers.

**Hint:** As you learned in this section, metastatic cancer cells have the ability to escape their adhesions to other cells, travel through the circulatory system, and invade distant tissues. You might think about what types of metastatic processes are facilitated by having a protease molecule that lacks normal regulation. Also consider what types of mutations in this gene could lead to increased metastasis.

**16.5 Cancer Is a Genetic Disorder Affecting Cell Adhesion**

16.6 Predisposition to Some Cancers Can Be Inherited

Although the vast majority of human cancers are sporadic, a small fraction (1–2 percent) have an hereditary or familial component. At present, about 50 forms of hereditary cancer are known (Table 16.3).
The development of hereditary colon cancer illustrates how inherited mutations in one allele of a gene contribute only one step in the multistep pathway leading to malignancy.

About 1 percent of colon cancer cases result from a genetic predisposition to cancer known as familial adenomatous polyposis (FAP). In FAP, individuals inherit one mutant copy of the APC (adenomatous polyposis) gene located on the long arm of chromosome 5. Mutations include deletions, frameshift, and point mutations. The normal function of the APC gene product is to act as a tumor suppressor controlling cell–cell contact and growth inhibition by interacting with the β-catenin protein. The presence of a heterozygous APC mutation causes the epithelial cells of the colon to partially escape cell-cycle control, and the cells divide to form small clusters of cells called polyps or adenomas. People who are heterozygous for this condition develop hundreds to thousands of colon and rectal polyps early in life. Although it is not necessary for the second allele of the APC gene to be mutated in polyps at this stage, in the majority of cases, the second APC allele becomes mutant in a later stage of cancer development. The relative order of mutations in the development of FAP is shown in Figure 16–11.

The second mutation in polyp cells that contain an APC gene mutation occurs in the ras proto-oncogene. The combined APC and ras gene mutations bring about the development of intermediate adenomas. Cells within these adenomas have defects in normal cell differentiation and will grow in culture, in the absence of contact with other cells and hence are described as transformed. The third step toward malignancy requires loss of function of both alleles of the DCC (deleted in colon cancer) gene. The DCC gene product is thought to be involved with cell adhesion and differentiation. Mutations in both DCC alleles result in the formation of late-stage adenomas with a number of finger-like outgrowths (villi). When late adenomas progress to cancerous adenomas, they usually suffer loss of functional p53 genes. The final steps toward malignancy involve mutations in an unknown number of genes associated with metastasis.

![Figure 16–11](image-url)
Viruses that cause cancer in animals have played a significant role in the search for knowledge about the genetics of human cancer. Most cancer-causing animal viruses are RNA viruses known as retroviruses. Because they transform cells into cancer cells, they are known as acute transforming retroviruses. The first of these acute transforming retroviruses was discovered in 1910 by Francis Peyton Rous. Rous was studying sarcomas (solid tumors of muscle, bone, or fat) in chickens, and he observed that extracts from these tumors caused the formation of new sarcomas when they were injected into tumor-free chickens. Several decades later, the agent within the extract that caused the sarcomas was identified as a retrovirus and was named the Rous sarcoma virus (RSV).

To understand how retroviruses cause cancer in animals, it is necessary to know how these viruses replicate in cells. When a retrovirus infects a cell, its RNA genome is copied into DNA by the reverse transcriptase enzyme, which is brought into the cell with the infecting virus. The DNA copy then enters the nucleus of the infected cell, where it integrates at random into the host cell’s genome. The integrated DNA copy of the retroviral RNA is called a provirus. The proviral DNA contains powerful enhancer and promoter elements in its U5 and U3 sequences at the ends of the provirus (Figure 16–12). The U5 promoter uses the host cell’s transcription proteins, directing transcription of the viral genes (gag, pol, and env). The products of these genes are the proteins and RNA genomes that make up the new retroviral particles. Because the provirus is integrated into the host genome, it is replicated along with the host’s DNA during the cell’s normal cell cycle. A retrovirus may not kill a cell, but it may continue to use the cell as a factory to replicate more viruses that will then infect surrounding cells.

A retrovirus may cause cancer in two different ways. First, the proviral DNA may integrate by chance near one of the cell’s normal proto-oncogenes. The strong promoters and enhancers in the provirus then stimulate high levels or inappropriate timing of transcription of the proto-oncogene, leading to stimulation of host cell proliferation. Second, a retrovirus may pick up a copy of a host proto-oncogene and integrate it into its genome (Figure 16–12). The new viral oncogene may be mutated during the process of transfer into the virus, or it may be expressed at abnormal levels because it is now under control of viral promoters. Retroviruses that carry these viral oncogenes can infect and transform normal cells into tumor cells. In the case of RSV, the oncogene that was captured from the chicken genome was the c-src gene. Through the study of many acute transforming viruses of animals, scientists have identified dozens of proto-oncogenes.

So far, no acute transforming retroviruses have been identified in humans. However, RNA and DNA viruses contribute to the development of human cancers in a variety of ways. It is thought that, worldwide, about 15 percent of cancers are associated with viruses, making virus infection the second greatest risk factor for cancer, next to tobacco smoking.

The most significant contributors to virus-induced cancers are the papillomaviruses (HPV 16 and 18), human T-cell leukemia virus (HTLV-1), hepatitis B virus, and Epstein-Barr virus (Table 16.4). Like other risk factors for cancer, including hereditary predisposition to certain cancers, virus infection alone is not sufficient to trigger human cancers. Other factors, including DNA damage or the accumulation of mutations in one or more of a cell’s oncogenes and tumor suppressor genes, are required to move a cell down the multistep pathway to cancer.

Because viruses are comprised solely of a nucleic acid genome surrounded by a proteinaceous coat, viruses must utilize the host cell’s biosynthetic machinery in order to reproduce themselves. As most viruses need the host cell to be in an actively growing state in order to access the host’s DNA synthetic enzymes, these viruses often contain genes encoding products that stimulate the cell cycle. If the virus does not kill the host cell, the potential exists for a loss of cell cycle control and the beginning of tumorigenesis.
Any substance or event that damages DNA has the potential to be carcinogenic. Unrepaired or inaccurately repaired DNA introduces mutations which, if they occur in proto-oncogenes or tumor suppressor genes, can lead to abnormal regulation of the cell cycle or disruption of controls over cell contact and invasion.

Our environment, both natural and man-made, contains abundant carcinogens. These include chemicals, radiation, some viruses, and chronic infections. Perhaps the most significant carcinogen in our environment is tobacco smoke. Epidemiologists estimate that about 30 percent of human cancer deaths are associated with cigarette smoking. Tobacco smoke contains a number of cancer-causing chemicals, some of which preferentially mutate proto-oncogenes such as ras and tumor suppressor genes such as p53.

Diet is often implicated in the development of cancer. Consumption of red meat and animal fat is associated with some cancers, such as colon, prostate, or breast cancer. The mechanisms by which these substances may contribute to carcinogenesis are not clear but may involve stimulation of cell division through hormones or creation of carcinogenic chemicals during cooking. Alcohol may cause inflammation of the liver and contribute to liver cancer.

Although most people perceive the man-made, industrial environment to be a highly significant contributor to cancer, it may account for only a small percentage of total cancers, and often in only specialized situations. Some of the most mutagenic agents, and hence potentially the most carcinogenic, are natural substances and natural processes. For example, aflatoxin, a component of a mold that grows on peanuts and corn, is one of the most carcinogenic chemicals known. Most chemical carcinogens, such as nitrosamines, are components of synthetic substances; however, many are naturally occurring. For example, natural pesticides and antibiotics found in plants may be carcinogenic, and the human body itself creates alkylating agents in the acidic environment of the gut. Nevertheless, these observations do not diminish the serious cancer risks to specific populations who are exposed to man-made carcinogens, such as synthetic pesticides or asbestos.

DNA lesions brought about by natural radiation (X-rays, ultraviolet light), natural dietary substances, and the external environment contribute the majority of environmentally caused mutations that lead to cancer. Normal metabolism creates oxidative end products that can damage DNA, proteins, and lipids. It is estimated that the human body suffers about 10,000 damaging DNA lesions per day due to the actions of oxygen free radicals. DNA repair enzymes deal successfully with most of this damage; however, some damage may lead to permanent mutations. The process of DNA replication itself is mutagenic. Hence, substances like growth factors or hormones that simulate cell division are ultimately mutagenic and perhaps carcinogenic. Chronic inflammation due to infection also stimulates tissue repair and cell division, resulting in DNA lesions accumulating during replication. These mutations may persist, particularly if cell-cycle checkpoints are compromised due to mutations or inactivation of tumor suppressor genes such as p53 or RB1.

Both ultraviolet (UV) light and ionizing radiation (such as X-rays and gamma rays) induce DNA damage. UV damage in sunlight is well accepted as an inducer of skin cancers. Ionizing radiation has clearly demonstrated itself as a carcinogen in studies of populations exposed to neutron and gamma radiation from atomic blasts such as those in Hiroshima and Nagasaki. Another significant environmental component, radon gas, may be responsible for up to 50 percent of the ionizing radiation exposure for the U.S. population and could contribute to lung cancers in some populations.

Problem 16 on page 374 asks you to provide your own estimate of what percentage of money spent on cancer research should be devoted to research on cancer prevention and what percentage should be devoted to research on cancer cures.

Hint: In answering this question, think about the relative rates of environmentally induced and spontaneous cancers. Second, consider the proportion of environmentally induced cancers that are due to lifestyle choices. (An interesting source of information on this topic can be found in Ames, B. N. et al. 1995. The causes and prevention of cancer. Proc. Natl. Acad. Sci. USA 92: 5258–5265.)
Breast Cancer: The Double-Edged Sword of Genetic Testing

These are exhilarating times for genetics and biotechnology. Close on the heels of the completion of the Human Genome Project has come a rush of optimism about future applications of genetics. Scientists and the media predict that gene technologies will soon diagnose and cure diseases as diverse as diabetes, asthma, heart disease, and Parkinson disease.

The prospect of using genetics to prevent and cure a whole range of diseases is exciting. However, in our enthusiasm, we often forget that these new technologies have significant limitations and profound ethical concerns. The story of genetic testing for breast cancer illustrates how we must temper our high expectations with respect for uncertainty.

Breast cancer is the most common cancer among women and the second leading cause of all cancer deaths (after lung cancer). Each year, more than 190,000 new cases are diagnosed in the United States. Breast cancer is not limited to women; about 1400 men are also diagnosed with the disease each year. A woman’s lifetime risk of developing breast cancer is about 12 percent, and the risk increases with age.

Approximately 5 to 10 percent of breast cancers are familial, defined by the appearance of several cases of breast or ovarian cancer among near blood relatives and the early onset of these diseases. In 1994, two genes were identified that show linkage to familial breast cancers. Germ-line mutations in these genes (BRCA1 and BRCA2) are associated with the majority of familial breast cancers. The molecular functions of BRCA1 and BRCA2 are still uncertain, although they appear to be tumor suppressor genes whose products are involved in repairing damaged DNA. Women who bear mutations in BRCA1 or BRCA2 have a 36 to 85 percent lifetime risk of developing breast cancer and a 16 to 60 percent risk of developing ovarian cancer. Men with germ-line mutations in BRCA2 have a 6 percent lifetime breast cancer risk—a hundredfold increase over the general male population.

BRCA1 and BRCA2 genetic tests detect any of the over 2000 different mutations that are known to occur within the coding regions of these genes, but the tests have limitations. They do not detect mutations in regulatory regions outside the coding region—mutations that could cause aberrant expression of these genes. Also, little is known about how any particular mutation manifests itself in terms of cancer risk, and the effects of each mutation may be modified by environmental factors and by interactions with other susceptibility genes.

Many patients at risk for familial breast cancer opt to undergo genetic testing. These patients feel that test results will help them to prevent breast or ovarian cancers, will guide them in childbearing decisions, and will allow them to inform family members at risk. But none of these benefits is clear-cut.

A woman whose BRCA test results are negative may be relieved and feel that she is not subject to familial breast cancer. However, her risk of developing breast cancer is still 12 percent (the population risk), and she should continue to monitor for the disease. Also, a negative BRCA genetic test does not eliminate the possibility that she bears an inherited mutation in another gene that increases breast cancer risk, or that BRCA1 or BRCA2 gene mutations exist in regions of the genes that are inaccessible to current genetic tests.

A woman whose test results are positive faces difficult choices. Her treatment options are poor, consisting of close monitoring, prophylactic mastectomy, or oophorectomy (removal of breasts and ovaries respectively) and taking drugs such as tamoxifen. Prophylactic surgery reduces her risks but does not eliminate them, as cancers can still occur in tissues that remain after surgery. Drugs such as tamoxifen reduce her risks but have serious side effects. Genetic tests affect not only the patient but also the patient’s entire family. People often experience fear, anxiety, and guilt on learning that they are carriers of a genetic disease. Studies show that people who refuse genetic test results often suffer from even more anxiety than those who opt to learn the results. Confidentiality is also a major concern. Patients fear that their genetic test results may be leaked to insurance companies or employers, jeopardizing their prospects for jobs or affordable health and life insurance. One study shows that a quarter of eligible patients refuse BRCA gene testing because of concerns about cost, confidentiality, and potential discrimination.

Genetic testing is such a new development that the health system has lagged behind the science. Because genetic testing has both psychological and medical ambiguities, genetic counseling is imperative for patients and their families. However, there are insufficient numbers of genetic counselors with experience in genetic testing, and even in the most qualified hands, issues are complex and difficult. Physicians often have limited knowledge of human clinical genetics and feel inadequate to advise their patients. The federal government and the insurance industries have yet to develop comprehensive policies concerning genetic tests and genetic information. Given the unclear interpretation of BRCA genetic tests, the relatively ineffective treatment options, and the potential for psychological and societal side effects, it is not surprising that only about 60 percent of familial breast cancer patients and their families decide to take the genetic tests.

The unanswered questions about BRCA1 and BRCA2 genetic testing are many and important. What cancer risks are associated with which mutations? Should all people have access to BRCA tests or only those at high risk? How can we ensure that the high costs of genetic tests and counseling do not limit this new technology to only a portion of the population? As we develop genetic tests for more and more diseases over the next few decades, our struggle with these issues will continue to grow.

Reference

Web Sites


**CHAPTER SUMMARY**

1. Cancer is a genetic disease, predominantly of somatic cells. About 1 percent of cancers are associated with germ-line mutations that increase the susceptibility to certain cancers.

2. Cancer cells show two basic properties: abnormal cell proliferation and a propensity to spread and invade other parts of the body (metastasis). Genes controlling these aspects of cellular function are either mutated or expressed inappropriately in cancer cells.

3. Cancers are clonal, meaning that all cells within a tumor originate from a single cell that contains a number of mutations.

4. The development of cancer is a multistep process, requiring mutations in several cancer-related genes.

5. Cancer cells show high rates of mutation, chromosomal abnormalities, and genomic instability. This leads to the accumulation of mutations in specific genes that control aspects of cell proliferation, apoptosis, differentiation, DNA repair, cell migration and cell-cell contact.

6. Cancer cells have defects in the regulation of cell-cycle progression, cell-cycle checkpoints, and signal transduction pathways.

7. Proto-oncogenes are normal genes that promote cell growth and division. When proto-oncogenes are mutated or mis-expressed in cancer cells, they are known as oncogenes.

8. Tumor suppressor genes normally regulate cell-cycle checkpoints and apoptosis. When tumor suppressor genes are mutated or inactivated, cells cannot correct DNA damage. This leads to accumulation of mutations that may cause cancer.

9. Inherited mutations in cancer-susceptibility genes are not sufficient to trigger cancer. A second somatic mutation, in the other copy of the gene, is necessary to trigger tumorigenesis. In addition, mutations in other cancer-related genes are necessary for the development of hereditary cancers.

10. RNA and DNA tumor viruses contribute to cancers by stimulating cells to proliferate, introducing new oncogenes, interfering with the cell’s normal tumor suppressor gene products, or stimulating the expression of a cell’s proto-oncogenes.

11. Environmental agents such as chemicals, radiation, viruses, and chronic infections contribute to the development of cancer. The most significant environmental factors that affect human cancers are tobacco smoke, diet, and natural radiation.

**KEY TERMS**

- acute transforming retroviruses, 369
- aflatoxin, 370
- apoptosis, 363
- basal lamina, 367
- benign tumor, 359
- Burkitt’s lymphoma, 359
- carcinogens, 360
- caspases, 363
- cell adhesion molecules, 367
- cell proliferation, 359
- chronic myelogenous leukemia (CML), 360
- cyclin-dependent kinases (CDKs), 362
- cyclins, 362
- Epstein-Barr virus, 369
- extracellular matrix, 367
- familial adenomatous polyposis (FAP), 368
- G1/S checkpoint, 362
- G2/M checkpoint, 362
- hepatitis B virus, 369
- hereditary nonpolyposis colorectal cancer (HNPCC), 361
- human T-cell leukemia virus (HTLV-1), 369
- loss of heterozygosity, 368
- M checkpoint, 362
- malignant tumors, 359
- metalloproteinases, 367
- metastases, 359
- metastasis, 359
- mutator phenotype, 360
- oncogene, 364
- papilloma viruses (HPV 16, 18), 369
- p53 gene, 365
- Philadelphia chromosome, 360
- polyp, 368
- protein kinase, 360
- proto-oncogenes, 364
- provirus, 369
- programmed cell death, 363
- ras gene family, 365
- RB1 (retinoblastoma 1), 366
- retinoblastoma, 366
- retinoblastoma protein (pRB), 366
- retroviruses, 369
- reverse transcriptase, 369
- Rous sarcoma virus (RSV), 369
- signal transduction, 362
- tissue inhibitors of metalloproteinases (TIMPs), 367
- transformed, 368
- tumorigenesis, 360
- tumor suppressor genes, 364

**AU:** this term is not bold in chapter text. Pls advise
In disorders such as retinoblastoma, a mutation in one allele of the RB1 gene can be inherited from the germ line, causing an autosomal dominant predisposition to the development of eye tumors. To develop tumors, a somatic mutation in the second copy of the RB1 gene is necessary, indicating that the mutation itself acts as a recessive trait. Given that the first mutation can be inherited, in what ways can a second mutational event occur?

**Solution:** In considering how this second mutation arises, we must look at several types of mutational events, including changes in nucleotide sequence and events that involve whole chromosomes or chromosome parts. Retinoblastoma results when both copies of the RB1 locus are lost or inactivated. With this in mind, you must first list the phenomena that can result in a mutational loss or the inactivation of a gene.

One way the second RB1 mutation can occur is by a nucleotide alteration that converts the remaining normal RB1 allele to a mutant form. This alteration can occur through a nucleotide substitution or by a frameshift mutation caused by the insertion or deletion of nucleotides during replication. A second mechanism involves the loss of the chromosome carrying the normal allele. This event would take place during mitosis, resulting in chromosome 13 monosomy and leaving the mutant copy of the gene as the only RB1 allele. This mechanism does not necessarily involve loss of the entire chromosome; deletion of the long arm (RB1 is on 13q) or an interstitial deletion involving the RB1 locus and some surrounding material would have the same result. Alternatively, a chromosome aberration involving loss of the normal copy of the RB1 gene might be followed by duplication of the chromosome carrying the mutant allele. Two copies of chromosome 13 would be restored to the cell, but the normal RB1 allele would not be present. Finally, a recombination event followed by chromosome segregation could produce a homozygous combination of mutant RB1 alleles.

2. Proto-oncogenes can be converted to oncogenes in a number of different ways. In some cases, the proto-oncogene itself becomes amplified up to hundreds of times in a cancer cell. An example is the cyclin D1 gene, which is amplified in some cancers. In other cases, the proto-oncogene may be mutated in a limited number of specific ways leading to alterations in the gene product’s structure. The ras gene is an example of a proto-oncogene that becomes oncogenic after suffering point mutations in specific regions of the gene. Explain why these two proto-oncogenes (cyclin D1 and ras) undergo such different alterations in order to convert them into oncogenes.

**Solution:** The first step to solving this question is to understand the normal functions of these proto-oncogenes and to think about how either amplification or mutation would affect each of these functions.

The cyclin D1 protein regulates progression of the cell cycle from G1 into S phase, by binding to CDK4 and activating this kinase. The cyclin D1/CDK4 complex phosphorylates a number of proteins including pRB, which in turn activates other proteins in a cascade that results in transcription of genes whose products are necessary for DNA replication in S phase. The simplest way to increase the activity of cyclin D1 would be to increase the number of cyclin D1 molecules available for binding to the cell’s endogenous CDK4 molecules. This can be accomplished by several mechanisms, including amplification of the cyclin D1 gene.

In contrast, a point mutation in the cyclin D1 gene would most likely interfere with the ability of the cyclin D1 protein to bind to CDK4; hence, mutations within the gene would probably repress cell-cycle progression rather than stimulate it.

The ras gene product is a signal transduction protein that operates as an on/off switch in response to external stimulation by growth factors. It does so by binding either GTP (the on state) or GDP (the off state). Oncogenic mutations in the ras gene occur in specific regions that alter the ability of the Ras protein to exchange GDP for GTP. Oncogenic Ras proteins are locked in the on conformation, bound to GTP. In this way, they constantly stimulate the cell to divide. An amplification of the ras gene would simply provide more molecules of normal Ras protein, which would still be capable of on/off regulation. Hence, simple amplification of ras would be less likely to be oncogenic.

3. Explain why many oncogenic viruses contain genes whose products interact with tumor suppressor proteins.

**Solution:** In order to answer this question, it is useful to consider what viruses try to accomplish in cells and what the roles of tumor suppressors are in normal cells.

The goal of oncogenic viruses is not to cause cancer but to maximize the potential for viral replication. Viruses are relatively simple entities, consisting of only a nucleic acid genome—either RNA or DNA—and a relatively small number of structural and enzymatic proteins. They depend on their host cells for much of the biosynthetic machinery and structural components necessary to replicate their genomes and assemble new virus particles. For example, many viruses require the host cell’s RNA polymerase II enzyme, transcription factors, and ribonucleotide precursors in order to transcribe their viral genes. They also require components of the host cell’s DNA replication machinery to replicate viral genomes. Hence, the ideal host cell for a virus infection is one that is within the cell cycle, preferably in late G1 to early S phase.

Most cells in a higher eukaryote such as humans are quiescent (in G0 phase). In order to stimulate the infected cell to enter the cell cycle and become primed for DNA replication, many viruses contain genes that encode growth-stimulating proteins. As we learned in this chapter, tumor suppressor proteins are those involved in either restraining the cell cycle at checkpoints or in triggering the process of programmed cell death. Both of these functions are inhibitory to the goals of a typical virus; therefore, many viruses have evolved methods to inactivate tumor suppressors. One of the ways in which viral proteins can inactivate tumor suppressor proteins is to bind to them and inhibit their functions.

The tumor suppressors p53 and pRB are common targets of viral regulatory proteins, such as the E6 and E7 proteins of HPV 16 and 18. By inactivating tumor suppressors, these viruses are able to maintain the cell within the cell cycle. For the host, however, this growth stimulation in the absence of functional cell-cycle checkpoints can lead to increased mutation accumulation and possible tumorigenesis.
1. As a genetic counselor, you are asked to assess the risk for a couple with a family history of retinoblastoma who are thinking about having children. Both the husband and wife are phenotypically normal, but the husband has a sister with familial retinoblastoma in both eyes. What is the probability that this couple will have a child with retinoblastoma? Are there any tests that you could recommend to help in this assessment?

2. What events occur in each phase of the cell cycle? Which phase is most variable in length?

3. Where are the major regulatory points in the cell cycle?

4. List the functions of kinases and cyclins, and describe how they interact to cause cells to move through the cell cycle.

5. (a) How does pRB function to keep cells at the G1 checkpoint? (b) How do cells get past the G1 checkpoint to move into S phase?

6. What is the difference between saying that cancer is inherited and saying that the predisposition to cancer is inherited?

7. Although tobacco smoking is responsible for a large number of human cancers, not all smokers develop cancer. Similarly, some people who inherit mutations in the tumor suppressor genes p53 or RB1 never develop cancer. Describe some reasons for these observations.

8. What is apoptosis, and under what circumstances do cells undergo this process?

9. Define tumor suppressor genes. Why is a mutation in a single copy of a tumor suppressor gene expected to behave as a recessive gene?

10. In the Rous sarcoma virus (RSV) genome, the host cell proto-oncogene is converted into an oncogene. How does this conversion occur?

11. Part of the Ras protein is associated with the plasma membrane, and part extends into the cytoplasm. How does the Ras protein transmit a signal from outside the cell into the cytoplasm? What happens in cases where the ras gene is mutated?

12. If a cell suffers damage to its DNA while in S phase, how can this damage be repaired before the cell enters mitosis?

13. Distinguish between oncogenes and proto-oncogenes. In what ways can proto-oncogenes be converted to oncogenes?

14. Of the two classes of genes associated with cancer, tumor suppressor genes and oncogenes, mutations in which group can be considered gain-of-function mutations? In which group are the loss-of-function mutations? Explain.

15. How do translocations such as the Philadelphia chromosome contribute to cancer?

16. Given that cancers can be environmentally induced and that some environmental factors are the result of lifestyle choices such as smoking, sun exposure, and diet, what percentage of the money spent on cancer research do you think should be devoted to research and education on preventing cancer rather than on finding cancer cures?

17. In CML, leukemic blood cells can be distinguished from other cells of the body by the presence of a functional BCR-ABL hybrid protein. Explain how this characteristic provides an opportunity to develop a therapeutic approach to a treatment for CML.

18. How do normal cells protect themselves from accumulating mutations in genes that could lead to cancer? How do cancer cells differ from normal cells in these processes?

19. Describe the difference between an acute transforming virus and a virus that does not cause tumors.

20. Explain how environmental agents such as chemicals and radiation cause cancer.

21. Radiotherapy (treatment with ionizing radiation) is one of the most effective current cancer treatments. It works by damaging DNA and other cellular components. In which ways could radiotherapy control or cure cancer, and why does radiotherapy often have significant side effects?

22. Assume that a young woman in a suspected breast cancer family takes the BRCA1 and BRCA2 genetic tests and receives negative results. That is, she does not test positive for the mutant alleles of BRCA1 or BRCA2. Can she consider herself free of risk for breast cancer?

23. People with a genetic condition known as Li-Fraumeni syndrome inherit one mutant copy of the p53 gene. These people have a high risk of developing a number of different cancers, such as breast cancer, leukemias, bone cancers, adrenocortical tumors, and brain tumors. Explain how mutations in one cancer-related gene can give rise to such a diverse range of tumors.

24. Explain the differences between a benign and malignant tumor.

25. As part of a cancer research project, you have discovered a gene that is mutated in many metastatic tumors. After determining the DNA sequence of this gene, you compare the sequence with those of other genes in the human genome sequence database. Your gene appears to code for an amino acid sequence that resembles sequences found in some serine proteases. Conjecture how your new gene might contribute to the development of highly invasive cancers.

26. A study by Bose and colleagues (1998. Blood 92: 3362–3367) and a previous study by Biernaux and others (1996. Bone Marrow Transplant 17: (Suppl. 3) S45–S47) showed that BCR-ABL fusion gene transcripts can be detected in 25 to 30 percent of healthy adults who do not develop chronic myelogenous leukemia (CML). Explain how these individuals can carry a fusion gene that is transcriptionally active and yet do not develop CML.

27. Those who inherit a mutant allele of the RB1 gene are at risk for developing a bone cancer called osteosarcoma. You suspect that in these cases, osteosarcoma requires a mutation in the second RB1 allele and have cultured some osteosarcoma cells and obtained a cDNA clone of a normal human RB1 gene. A colleague sends you a research paper revealing that a strain of cancer-prone mice develop malignant tumors when injected with osteosarcoma cells, and you obtain these mice. Using these three resources, what experiments would you perform to determine (a) whether osteosarcoma cells produce any pRB protein, and (c) if the addition of a normal RB1 gene will change the cancer-causing potential of osteosarcoma cells?

28. The compound benzo[a]pyrene is found in cigarette smoke. This compound chemically modifies guanine bases in DNA. Such abnormal bases are usually removed by an enzyme that hydrolyzes the base, leaving an apurinic site. If such a site is left unrepaird, an adenine is preferentially inserted across from the apurinic site. In a study of lung cancer patients (Harris, A. 1991. Nature 350: 377–378), tumor cells from 15 out of 25 patients had a G to T transversion in the p53 gene, which has a known role in cancer formation. You are testifying as an expert witness in a court case in which the widow of a man who was a lifelong smoker and died of lung cancer is suing a company for manufacturing the
Problems and Discussion Questions

29. Table 16.5 summarizes some of the data that have been collected on BRCA1 mutations in families with a high incidence of both early-onset breast and ovarian cancer. Table 16.6 shows neutral polymorphisms found in control families (with no increased frequency of breast and ovarian cancer). (a) Note the coding effect of the mutation found in kindred group 2082 in Table 16.5. This results from a single base-pair substitution. Draw the normal double-stranded DNA sequence for this codon (with the 5’ and 3’ ends labeled), and show the sequence of events that generated this mutation, assuming that it resulted from an uncorrected mismatch event during DNA replication. (b) Examine the types of mutations that are listed in Table 16.5 and determine if the BRCA1 gene is likely to be a tumor suppressor gene or an oncogene. (c) Although the mutations in Table 16.5 are clearly deleterious and cause breast cancer in women at very young ages, each of the kindred groups had at least one woman who carried the mutation but lived until age 80 without developing cancer. Name at least two different mechanisms (or variables) that could underlie variation in the expression of a mutant phenotype and propose an explanation for the incomplete penetrance of this mutation. How do these mechanisms or variables relate to this explanation?

30. Examine Table 16.6. (a) What is meant by a neutral polymorphism? (b) What is the significance of this table in the context of examining a family or population for BRCA1 mutations that predispose an individual to cancer? (c) Is the PM2 polymorphism likely to result in a neutral missense mutation or a silent mutation? (d) Answer part (c) for the PM3 polymorphism.

Table 16.5: Predisposing Mutations in BRCA1

<table>
<thead>
<tr>
<th>Kindred</th>
<th>Codon</th>
<th>Nucleotide Change</th>
<th>Coding Effect</th>
<th>Frequency in Control Chromosomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1901</td>
<td>24</td>
<td>*11 bp</td>
<td>Frameshift or splice</td>
<td>0/180</td>
</tr>
<tr>
<td>2082</td>
<td>1313</td>
<td>C → T</td>
<td>Gln → Stop</td>
<td>0/170</td>
</tr>
<tr>
<td>1910</td>
<td>1756</td>
<td>Extra C</td>
<td>Frameshift</td>
<td>0/162</td>
</tr>
<tr>
<td>2099</td>
<td>1775</td>
<td>T → G</td>
<td>Met → Arg</td>
<td>0/120</td>
</tr>
<tr>
<td>2035</td>
<td>NA*</td>
<td>?</td>
<td>Loss of transcript</td>
<td>NA*</td>
</tr>
</tbody>
</table>


*NA indicates not applicable, as the regulatory mutation is inferred, and the position has not been identified.

Table 16.6: Neutral Polymorphisms in BRCA1

<table>
<thead>
<tr>
<th>Name</th>
<th>Codon Location</th>
<th>Base in Codon</th>
<th>Frequency in Control Chromosomes*</th>
</tr>
</thead>
<tbody>
<tr>
<td>PM1</td>
<td>317</td>
<td>2</td>
<td>152 0 10 0</td>
</tr>
<tr>
<td>PM6</td>
<td>878</td>
<td>2</td>
<td>0 55 0 100</td>
</tr>
<tr>
<td>PM7</td>
<td>1190</td>
<td>2</td>
<td>109 0 53 0</td>
</tr>
<tr>
<td>PM2</td>
<td>1443</td>
<td>3</td>
<td>0 115 0 58</td>
</tr>
<tr>
<td>PM3</td>
<td>1619</td>
<td>1</td>
<td>116 0 52 0</td>
</tr>
</tbody>
</table>

*The number of chromosomes with a particular base at the indicated polymorphic site (A, C, G, or T) is shown.