5 Linkage, Recombination, and Eukaryotic Gene Mapping

Alfred Henry Sturtevant, an early geneticist, developed the first genetic map. *[Institute Archives, California Institute of Technology.]*

Alfred Sturtevant and the First Genetic Map

In 1909, Thomas Hunt Morgan taught the introduction to zoology class at Columbia University. Seated in the lecture hall were sophomore Alfred Henry Sturtevant and freshman Calvin Bridges. Sturtevant and Bridges were excited by Morgan's teaching style and intrigued by his interest in biological problems. They asked Morgan if they could work in his laboratory and, the following year, both young men were given desks in the "Fly Room," Morgan's research laboratory where the study of *Drosophila* genetics was in its infancy (see pp. 75–76 in Chapter 4). Sturtevant, Bridges, and Morgan's other research students virtually lived in the laboratory, raising fruit flies, designing experiments, and discussing their results.

In the course of their research, Morgan and his students observed that some pairs of genes did not segregate randomly according to Mendel's principle of independent assortment but instead tended to be inherited together. Morgan suggested that possibly the genes were located on the same chromosome and thus traveled together during meiosis. He further proposed that closely linked genes—those that are rarely shuffled by recombination—lie close together on the same chromosome, whereas loosely linked genes—those more frequently shuffled by recombination—lie farther apart.

One day in 1911, Sturtevant and Morgan were discussing independent assortment when, suddenly, Sturtevant had a flash of inspiration: variation in the strength of linkage indicated how genes are positioned along a chromosome, providing a way of mapping genes. Sturtevant went home and, neglecting his undergraduate homework, spent most of the night working out the first genetic map (**Figure 5.1**). Sturtevant's first chromosome map was remarkably accurate, and it established the basic methodology used today for mapping genes.

Sturtevant went on to become a leading geneticist. His research included gene mapping and basic mechanisms of inheritance in *Drosophila*, cytology, embryology, and evolution. Sturtevant's career was deeply influenced by his early years in the Fly Room, where Morgan's unique personality and the close quarters combined to stimulate intellectual excitement and the free exchange of ideas.

5.1 Sturtevant's map included five genes on the X chromosome of *Drosophila*. The genes are yellow body (*y*), white eyes (*w*), vermilion eyes (*v*), miniature wings (*m*), and rudimentary wings (*r*).

This chapter explores the inheritance of genes located on the same chromosome. These linked genes do not strictly obey Mendel's principle of independent assortment; rather, they tend to be inherited together. This tendency requires a new approach to understanding their inheritance and predicting the types of offspring produced. A critical piece of information necessary for predicting the results of these crosses is the arrangement of the genes on the chromosomes; thus, it will be necessary to think about the relation between genes and chromosomes. A key to understanding the inheritance of linked genes is to make the conceptual connection between the genotypes in a cross and the behavior of chromosomes in meiosis.

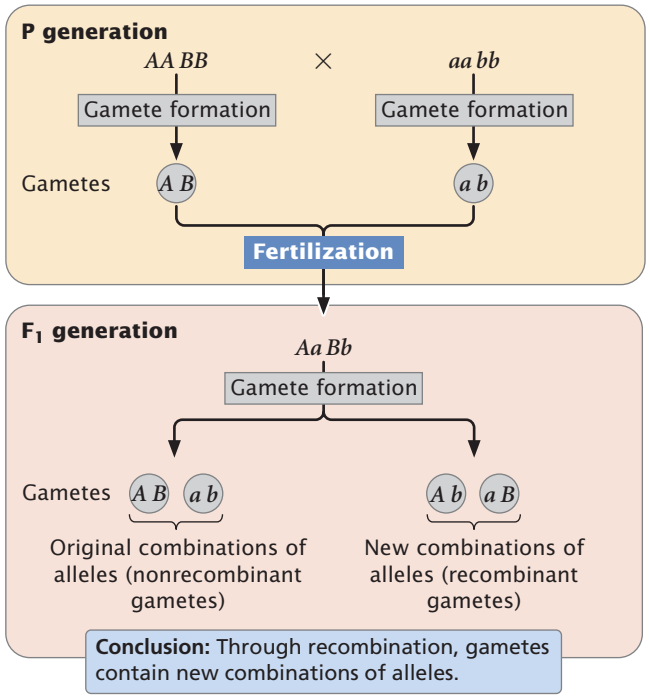
We will begin our exploration of linkage by first comparing the inheritance of two linked genes with the inheritance of two genes that assort independently. We will then examine how crossing over breaks up linked genes. This knowledge of linkage and recombination will be used for predicting the results of genetic crosses in which genes are linked and for mapping genes. Later in the chapter, we will focus on physical methods of determining the chromosomal locations of genes.

5.1 Linked Genes Do Not Assort Independently

Chapter 3 introduced Mendel's principles of segregation and independent assortment. Let's take a moment to review these two important concepts. The principle of segregation states that each individual diploid organism possesses two alleles at a locus that separate in meiosis, with one allele going into each gamete. The principle of independent assortment provides additional information about the process of segregation: it tells us that, in the process of separation, the two alleles at a locus act independently of alleles at other loci.

The independent separation of alleles results in *recombination*, the sorting of alleles into new combinations. Consider a cross between individuals homozygous for two different pairs of alleles: *AA BB* × *aa bb*. The first parent, *AA BB*, produces gametes with alleles *A B*, and the second parent, *aa bb*, produces gametes with the alleles *a b*, resulting in F_1 progeny with genotype *Aa Bb* (Figure 5.2). Recombination means that, when one of the F_1 progeny reproduces, the combination of alleles in its gametes may differ from the combinations in the gametes from its parents. In other words, the F_1 may produce gametes with alleles *A b* or *a B* in addition to gametes with *A B* or *a b*.

Mendel derived his principles of segregation and independent assortment by observing the progeny of genetic crosses, but he had no idea of what biological processes produced these phenomena. In 1903, Walter Sutton proposed a biological basis for Mendel's principles, called the chromosome theory of heredity (see Chapter 3). This theory holds that genes are found on chromosomes. Let's restate Mendel's two principles in relation to the chromosome theory of heredity. The principle of segregation states that a diploid organism possesses two alleles for a trait, each of which is

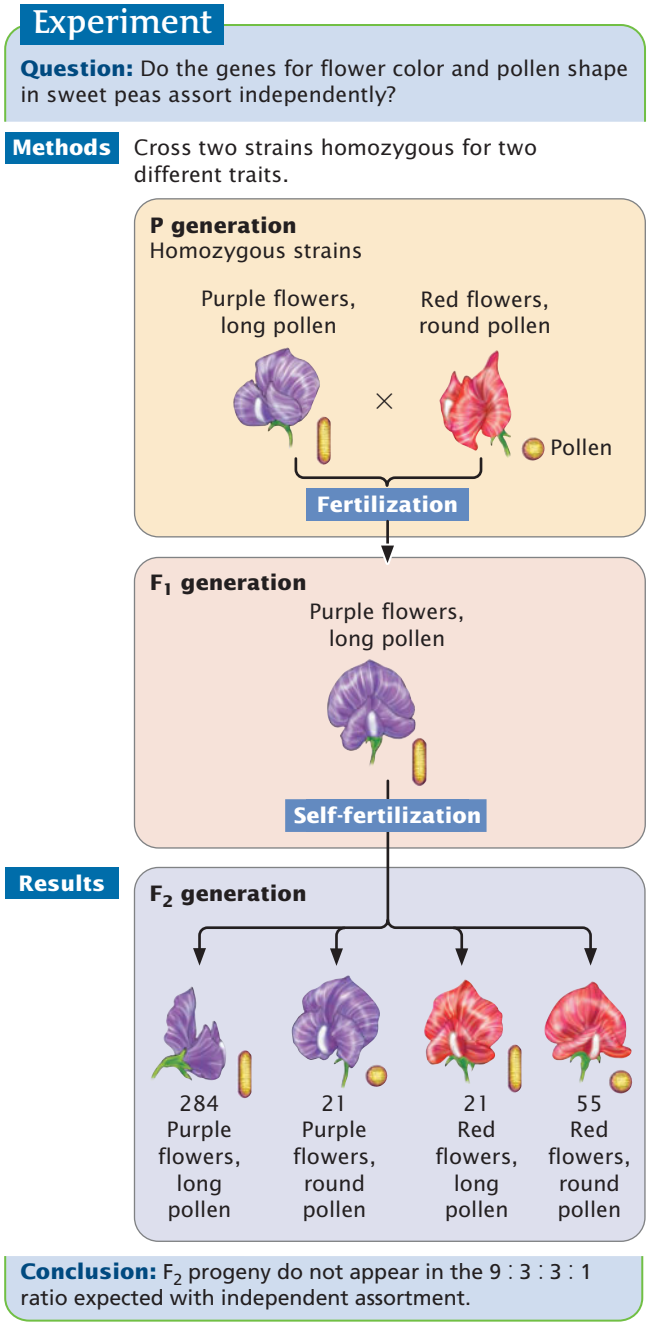


5.2 Recombination is the sorting of alleles into new combinations.

located at the same position, or locus, on each of the two homologous chromosomes. These chromosomes segregate in meiosis, with each gamete receiving one homolog. The principle of independent assortment states that, in meiosis, each pair of homologous chromosomes assorts independently of other homologous pairs. With this new perspective, it is easy to see that the number of chromosomes in most organisms is limited and that there are certain to be more genes than chromosomes; so some genes must be present on the same chromosome and should not assort independently. Genes located close together on the same chromosome are called **linked genes** and belong to the same **linkage group**. Linked genes travel together during meiosis, eventually arriving at the same destination (the same gamete), and are not expected to assort independently.

All of the characteristics examined by Mendel in peas did display independent assortment and, after the rediscovery of Mendel's work, the first genetic characteristics studied in other organisms also seemed to assort independently. How could genes be carried on a limited number of chromosomes and yet assort independently?

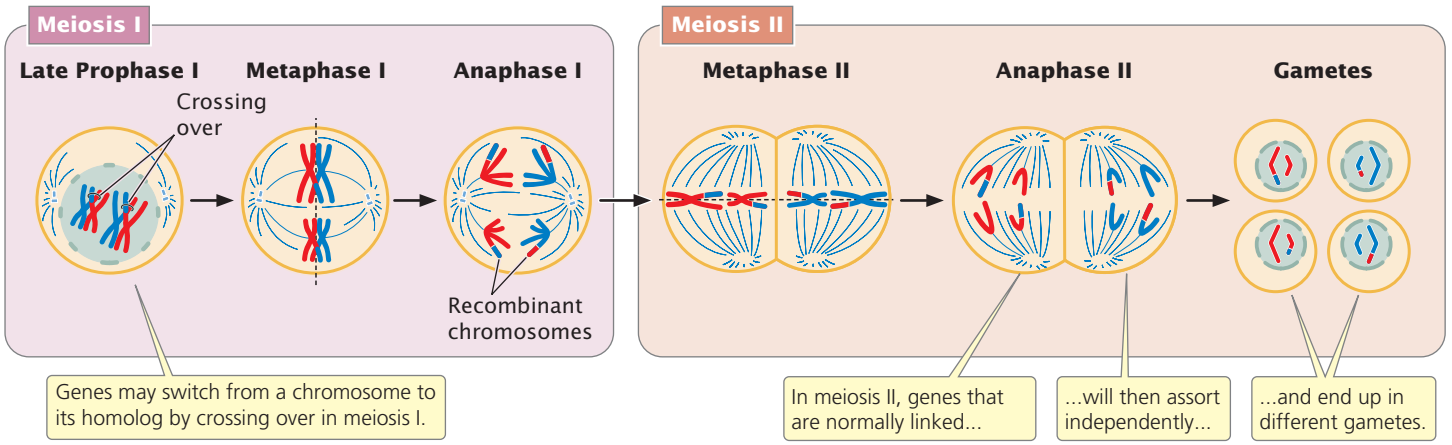
The apparent inconsistency between the principle of independent assortment and the chromosome theory of heredity soon disappeared as biologists began finding genetic characteristics that did not assort independently. One of the first cases was reported in sweet peas by William Bateson, Edith Rebecca Saunders, and Reginald C. Punnett in 1905. They crossed a homozygous strain of peas having purple flowers and long pollen grains with a homozygous strain having red flowers and round pollen grains. All the F_1 had purple flowers and long pollen grains, indicating that purple was dominant over red and long was dominant over round. When they intercrossed the F_1, the resulting F_2 progeny did not appear in the 9 : 3 : 3 : 1 ratio expected with independent assortment (Figure 5.3). An excess of F_2 plants had purple flowers and long pollen or red flowers and round pollen (the parental phenotypes). Although Bateson, Saunders, and Punnett were unable to explain these results, we now know that the two loci that they examined lie close together on the same chromosome and therefore do not assort independently.



5.3 Nonindependent assortment of flower color and pollen shape in sweet peas.

5.2 Linked Genes Segregate Together and Crossing Over Produces Recombination Between Them

Genes that are close together on the same chromosome usually segregate as a unit and are therefore inherited together. However, genes occasionally switch from one homologous chromosome to the other through the process of crossing over (see Chapter 2), as illustrated in Figure 5.4. Crossing over results in recombination; it breaks up the associations of genes that are close together on the same chromosome. Linkage and crossing over can be seen as processes that have opposite effects: linkage keeps particular genes together, and crossing over mixes them up. In Chapter 4 we considered a number of exceptions and extensions to Mendel's principles of heredity. The concept of linked genes adds a further complication to interpretations of the results of genetic crosses. However, with an understanding of how linkage affects



5.4 Crossing over takes place in meiosis and is responsible for recombination.

heredity, we can analyze crosses for linked genes and successfully predict the types of progeny that will be produced.

Notation for Crosses with Linkage

In analyzing crosses with linked genes, we must know not only the genotypes of the individuals crossed, but also the arrangement of the genes on the chromosomes. To keep track of this arrangement, we introduce a new system of notation for presenting crosses with linked genes. Consider a cross between an individual homozygous for dominant alleles at two linked loci and another individual homozygous for recessive alleles at those loci (*AA BB* × *aa bb*). For linked genes, it's necessary to write out the specific alleles as they are arranged on each of the homologous chromosomes:

$$\frac{A \quad B}{A \quad B} \times \frac{a \quad b}{a \quad b}$$

In this notation, each line represents one of the two homologous chromosomes. Inheriting one chromosome from each parent, the F_1 progeny will have the following genotype:

$$\frac{A \quad B}{a \quad b}$$

Here, the importance of designating the alleles on each chromosome is clear. One chromosome has the two dominant alleles *A* and *B*, whereas the homologous chromosome has the two recessive alleles *a* and *b*. The notation can be simplified by drawing only a single line, with the understanding that genes located on the same side of the line lie on the same chromosome:

$$\frac{A \quad B}{a \quad b}$$

This notation can be simplified further by separating the alleles on each chromosome with a slash: *AB/ab*.

Remember that the two alleles at a locus are always located on different homologous chromosomes and therefore must lie on opposite sides of the line. Consequently, we would *never* write the genotypes as

$$\frac{A \quad a}{B \quad b}$$

because the alleles *A* and *a* can *never* be on the same chromosome.

It is also important to always keep the same order of the genes on both sides of the line; thus, we should *never* write

$$\frac{A \quad B}{b \quad a}$$

because it would imply that alleles *A* and *b* are allelic (at the same locus).

Complete Linkage Compared with Independent Assortment

We will first consider what happens to genes that exhibit complete linkage, meaning that they are located very close together on the same chromosome and do not exhibit crossing over. Genes are rarely completely linked but, by assuming that no crossing over occurs, we can see the effect of linkage more clearly. We will then consider what happens when genes assort independently. Finally, we will consider the results obtained if the genes are linked but exhibit some crossing over.

A testcross reveals the effects of linkage. For example, if a heterozygous individual is test-crossed with a homozygous recessive individual (*Aa Bb* × *aa bb*), the alleles that are present in the gametes contributed by the heterozygous parent will be expressed in the phenotype of the offspring, because the homozygous parent could not contribute dominant alleles that might mask them. Consequently, traits that appear in the progeny reveal which alleles were transmitted by the heterozygous parent.

Consider a pair of linked genes in tomato plants. One pair affects the type of leaf: an allele for mottled leaves (*m*) is recessive to an allele that produces normal leaves (*M*). Nearby on the same chromosome is another locus that determines the height of the plant: an allele for dwarf (*d*) is recessive to an allele for tall (*D*).

Testing for linkage can be done with a testcross, which requires a plant heterozygous for both characteristics. A geneticist might produce this heterozygous plant by crossing a variety of tomato that is homozygous for normal leaves and tall height with a variety that is homozygous for mottled leaves and dwarf height:

$$\text{P} \quad \frac{M\ D}{M\ D} \times \frac{m\ d}{m\ d}$$

$$\downarrow$$

$$\text{F}_1 \quad \frac{M\ D}{m\ d}$$

The geneticist would then use these F_1 heterozygotes in a testcross, crossing them with plants homozygous for mottled leaves and dwarf height:

$$\frac{M\ D}{m\ d} \times \frac{m\ d}{m\ d}$$

The results of this testcross are diagrammed in **Figure 5.5a**. The heterozygote produces two types of gametes: some with the $\frac{M\ D}{}$ chromosome and others with the $\frac{m\ d}{}$ chromosome. Because no crossing over occurs, these gametes are the only types produced by the heterozygote. Notice that these gametes contain only combinations of alleles that were present in the original parents: either the allele for normal leaves together with the allele for tall height (*M* and *D*) or the allele for mottled leaves together with the allele for dwarf height (*m* and *d*). Gametes that contain only original combinations of alleles present in the parents are **nonrecombinant gametes**, or *parental* gametes.

The homozygous parent in the testcross produces only one type of gamete; it contains chromosome $\frac{m\ d}{}$ and pairs with one of the two gametes generated by the heterozygous parent (see Figure 5.5a). Two types of progeny result: half have normal leaves and are tall:

$$\frac{M\ D}{m\ d}$$

and half have mottled leaves and are dwarf:

$$\frac{m\ d}{m\ d}$$

These progeny display the original combinations of traits present in the P generation and are **nonrecombinant progeny**, or *parental* progeny. No new combinations of the two traits, such as normal leaves with dwarf or mottled leaves with tall, appear in the offspring, because the genes affecting the two traits are completely linked and are inherited together. New combinations of traits could arise only if the physical connection between *M* and *D* or between *m* and *d* were broken.

These results are distinctly different from the results that are expected when genes assort independently (**Figure 5.5b**). If the *M* and *D* loci assorted independently, the heterozygous plant (*Mm Dd*) would produce four types of gametes: two nonrecombinant gametes containing the original combinations of alleles (*M D* and *m d*) and two gametes containing new combinations of alleles (*M d* and *m D*). Gametes with new combinations of alleles are called **recombinant gametes**. With independent assortment, nonrecombinant and recombinant gametes are produced in equal proportions. These four types of gametes join with the single type of gamete produced by the homozygous parent of the testcross to produce four kinds of progeny in equal proportions (see Figure 5.5b). The progeny with new combinations of traits formed from recombinant gametes are termed **recombinant progeny**.

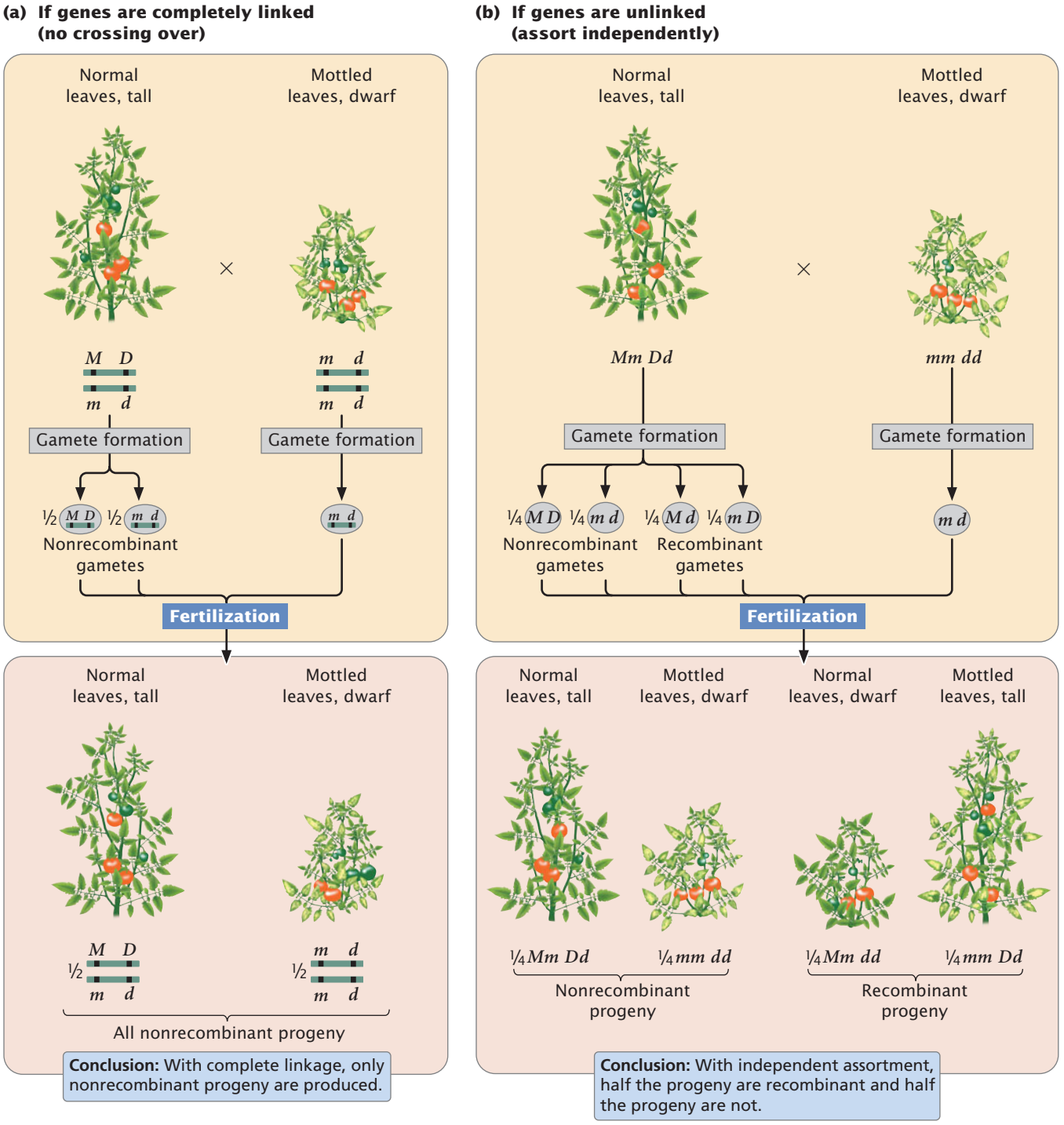
In summary, a testcross in which one of the plants is heterozygous for two completely linked genes yields two types of progeny, each type displaying one of the original combinations of traits present in the P generation. Independent assortment, in contrast, produces progeny in a 1 : 1 : 1 : 1 ratio. That is, there are four types of progeny—two types of recombinant progeny and two types of nonrecombinant progeny in equal proportions.

Crossing Over with Linked Genes

Usually, there is some crossing over between genes that lie on the same chromosome, producing new combinations of traits. Genes that exhibit crossing over are incompletely linked. Let's see how it takes place.

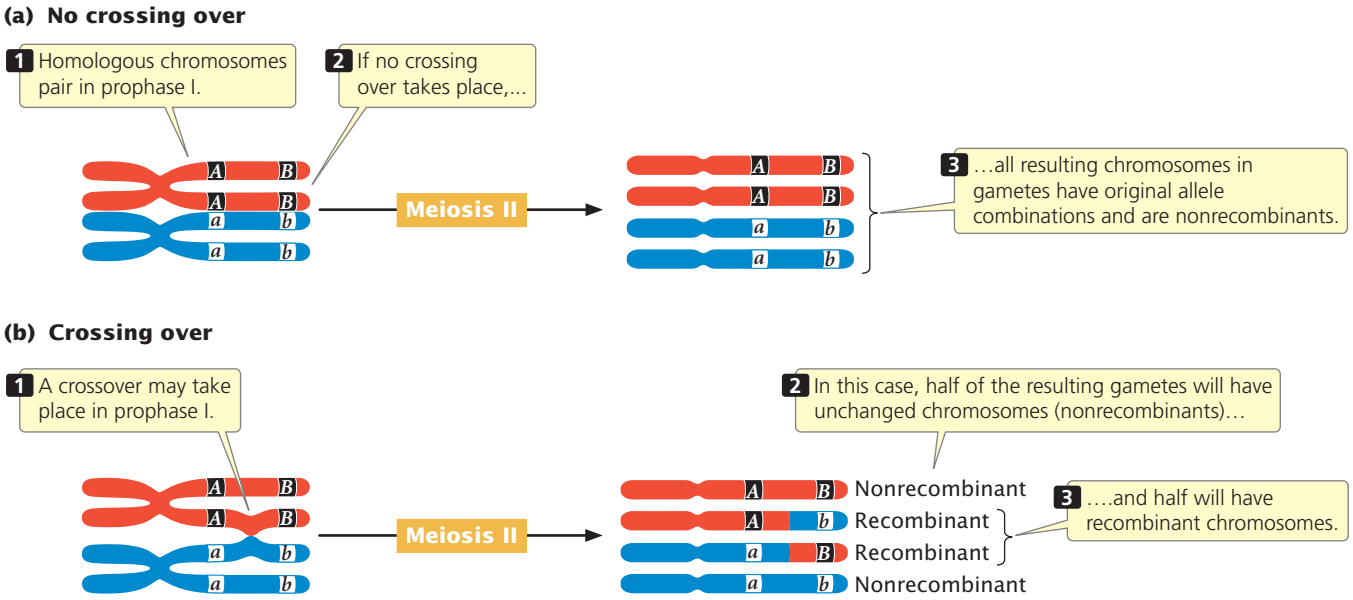
Theory The effect of crossing over on the inheritance of two linked genes is shown in **Figure 5.6**. Crossing over, which takes place in prophase I of meiosis, is the exchange of genetic material between nonsister chromatids (see Figures 2.12 and 2.14). After a single crossover has taken place, the two chromatids that did not participate in crossing over are unchanged; gametes that receive these chromatids are nonrecombinants. The other two chromatids, which did participate in crossing over, now contain new combinations of alleles; gametes that receive these chromatids are recombinants. For each meiosis in which a single crossover takes place, then, two nonrecombinant gametes and two recombinant gametes will be produced. This result is the same as that produced by independent assortment (see Figure 5.5b); so, when crossing over between two loci takes place in every meiosis, it is impossible to determine whether the genes are on the same chromosome and crossing over took place or whether the genes are on different chromosomes.

For closely linked genes, crossing over does not take place in every meiosis. In meioses in which there is no



5.5 A testcross reveals the effects of linkage. Results of a testcross for two loci in tomatoes that determine leaf type and plant height.

crossing over, only nonrecombinant gametes are produced. In meioses in which there is a single crossover, half the gametes are recombinants and half are nonrecombinants (because a single crossover affects only two of the four chromatids); so the total percentage of recombinant gametes is always half the percentage of meioses in which crossing over takes place. Even if crossing over between two genes takes place in every meiosis, only 50% of the resulting gametes will be recombinants. Thus, the frequency of recombinant gametes is always half the frequency of crossing over, and the maximum proportion of recombinant gametes is 50%.



5.6 A single crossover produces half nonrecombinant gametes and half recombinant gametes.

Concepts

Linkage between genes causes them to be inherited together and reduces recombination; crossing over breaks up the associations of such genes. In a testcross for two linked genes, each crossover produces two recombinant gametes and two nonrecombinants. The frequency of recombinant gametes is half the frequency of crossing over, and the maximum frequency of recombinant gametes is 50%.

✔ Concept Check 1

For single crossovers, the frequency of recombinant gametes is half the frequency of crossing over because

a. a test cross between a homozygote and heterozygote produces $\frac{1}{2}$ heterozygous and $\frac{1}{2}$ homozygous progeny.

b. the frequency of recombination is always 50%.

c. each crossover takes place between only two of the four chromatids of a homologous pair.

d. crossovers occur in about 50% of meioses.

Application Let's apply what we have learned about linkage and recombination to a cross between tomato plants that differ in the genes that encode leaf type and plant height. Assume now that these genes are linked and that some crossing over takes place between them. Suppose a geneticist carried out the testcross outlined earlier:

$$\frac{M \quad D}{m \quad d} \times \frac{m \quad d}{m \quad d}$$

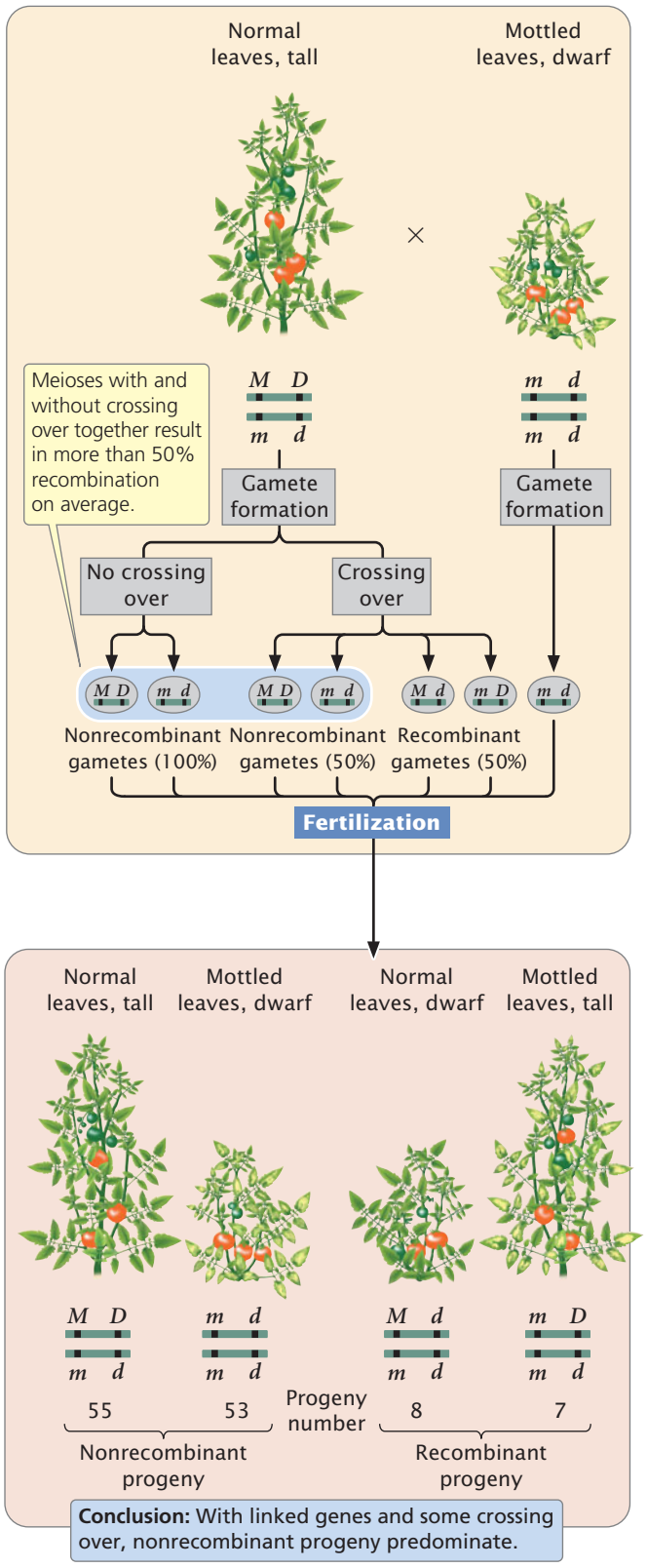
When crossing over takes place between the genes for leaf type and height, two of the four gametes produced will be recombinants. When there is no crossing over, all four resulting gametes will be nonrecombinants. Thus, over all meioses, the majority of gametes will be nonrecombinants. These gametes then unite with gametes produced by the homozygous recessive parent, which contain only the recessive alleles, resulting in mostly nonrecombinant progeny and a few recombinant progeny (**Figure 5.7**). In this cross, we see that 55 of the testcross progeny have normal leaves and are tall and 53 have mottled leaves and are dwarf. These plants are the nonrecombinant progeny, containing the original combinations of traits that were present in the parents. Of the 123 progeny, 15 have new combinations of traits that were not seen in the parents: 8 are normal leaved and dwarf, and 7 are mottle leaved and tall. These plants are the recombinant progeny.

The results of a cross such as the one illustrated in Figure 5.7 reveal several things. A testcross for two independently assorting genes is expected to produce a 1 : 1 : 1 : 1 phenotypic ratio in the progeny. The progeny of this cross clearly do not exhibit such a ratio; so we might suspect that the genes are not assorting independently. When linked genes undergo some crossing over, the result is mostly nonrecombinant progeny and fewer recombinant progeny. This result is what we observe among the progeny of the testcross illustrated in Figure 5.7; so we conclude that the two genes show evidence of linkage with some crossing over.

Calculating Recombination Frequency

The percentage of recombinant progeny produced in a cross is called the **recombination frequency**, which is calculated as follows:

$$\text{recombinant frequency} = \frac{\text{number of recombinant progeny}}{\text{total number of progeny}} \times 100\%$$



5.7 Crossing over between linked genes produces nonrecombinant and recombinant offspring.
In this testcross, genes are linked and there is some crossing over.

In the testcross shown in Figure 5.7, 15 progeny exhibit new combinations of traits; so the recombination frequency is:

$$\frac{8 + 7}{55 + 53 + 8 + 7} \times 100\% = \frac{15}{123} \times 100\% = 12.2\%$$

Thus, 12.2% of the progeny exhibit new combinations of traits resulting from crossing over. The recombination frequency can also be expressed as a decimal fraction (0.122).

Coupling and Repulsion

In crosses for linked genes, the arrangement of alleles on the homologous chromosomes is critical in determining the outcome of the cross. For example, consider the inheritance of two genes in the Australian blowfly, *Lucilia cuprina*. In this species, one locus determines the color of the thorax: a purple thorax (p) is recessive to the normal green thorax (p^+). A second locus determines the color of the puparium: a black puparium (b) is recessive to the normal brown puparium (b^+). These loci are located close together on the chromosome. Suppose we test cross a fly that is heterozygous at both loci with a fly that is homozygous recessive at both. Because these genes are linked, there are two possible arrangements on the chromosomes of the heterozygous progeny fly. The dominant alleles for green thorax (p^+) and brown puparium (b^+) might reside on the same chromosome, and the recessive alleles for purple thorax (p) and black puparium (b) might reside on the other homologous chromosome:

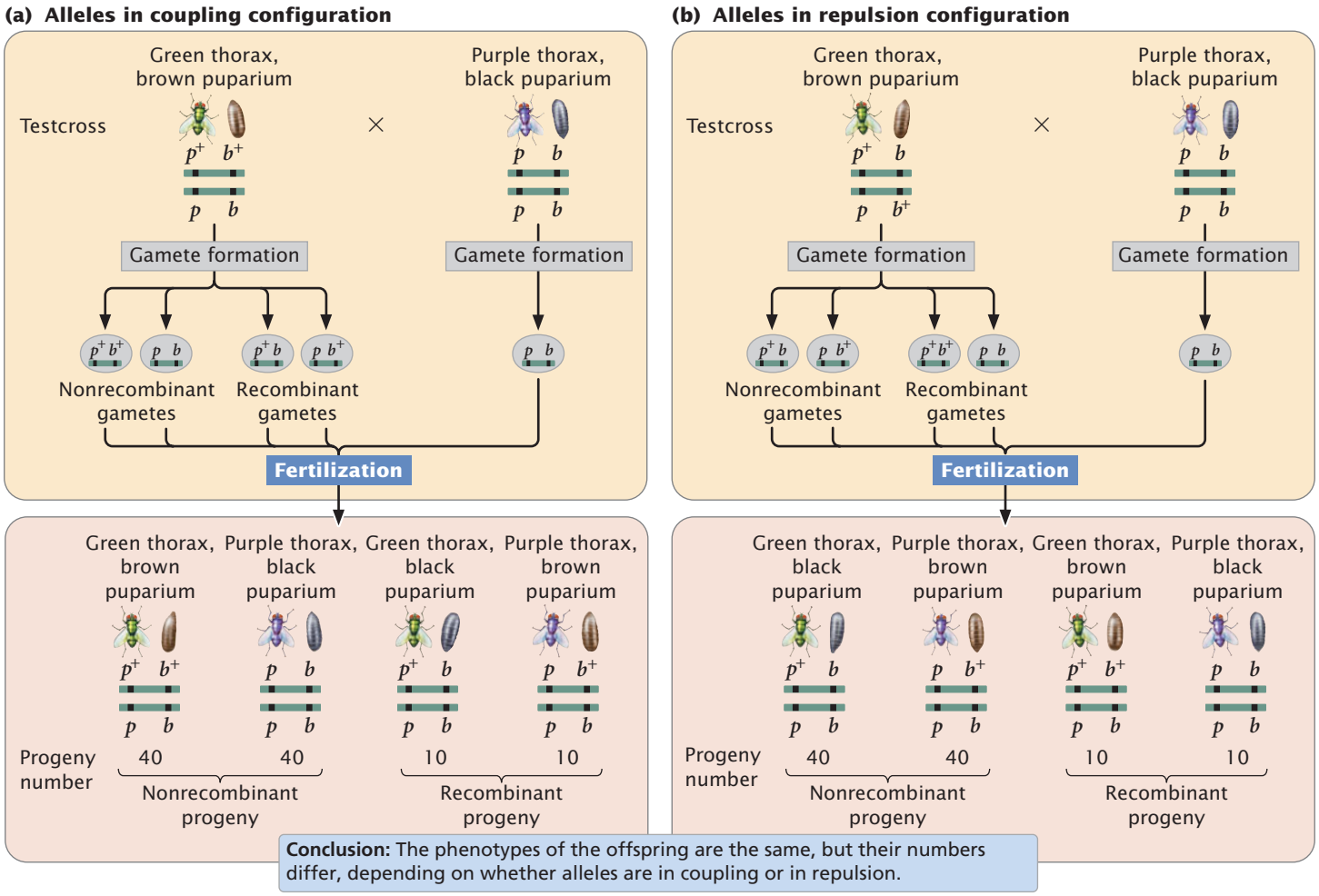
$$\frac{p^+ \quad b^+}{p \quad b}$$

This arrangement, in which wild-type alleles are found on one chromosome and mutant alleles are found on the other chromosome, is referred to as the **coupling** or **cis configuration**. Alternatively, one chromosome might bear the alleles for green thorax (p^+) and black puparium (b), and the other chromosome would carry the alleles for purple thorax (p) and brown puparium (b^+):

$$\frac{p^+ \quad b}{p \quad b^+}$$

This arrangement, in which each chromosome contains one wild-type and one mutant allele, is called the **repulsion** or **trans configuration**. Whether the alleles in the heterozygous parent are in coupling or repulsion determines which phenotypes will be most common among the progeny of a testcross.

When the alleles are in the coupling configuration, the most numerous progeny types are those with green thorax and brown puparium and those with purple thorax and black puparium (Figure 5.8a); but, when the alleles of the heterozygous parent are in repulsion, the most numerous progeny types are those with green thorax and black puparium and those with purple thorax and brown puparium (Figure 5.8b). Notice that the genotypes of the parents in Figure 5.8a and b are the same ($p^+p\ b^+b \times pp\ bb$) and that the dramatic dif-



5.8 The arrangement (coupling or repulsion) of linked genes on a chromosome affects the results of a testcross. Linked loci in the Australian blowfly, *Luciliá cuprina*, determine the color of the thorax and that of the puparium.

ference in the phenotypic ratios of the progeny in the two crosses results entirely from the configuration—coupling or repulsion—of the chromosomes. It is essential to know the arrangement of the alleles on the chromosomes to accurately predict the outcome of crosses in which genes are linked.

Concepts

In a cross, the arrangement of linked alleles on the chromosomes is critical for determining the outcome. When two wild-type alleles are on one homologous chromosome and two mutant alleles are on the other, they are in the coupling configuration; when each chromosome contains one wild-type allele and one mutant allele, the alleles are in repulsion.

✔ Concept Check 2

The following testcross produces the progeny shown: *Aa Bb* × *aa bb* → 10 *Aa Bb*, 40 *Aa bb*, 40 *aa Bb*, 10 *aa bb*. What is the percent recombination between the *A* and *B* loci? Were the genes in the *Aa Bb* parent in coupling or in repulsion?

Connecting Concepts

Relating Independent Assortment, Linkage, and Crossing Over

We have now considered three situations concerning genes at different loci. First, the genes may be located on different chromosomes; in this case, they exhibit independent assortment and combine randomly when gametes are formed. An individual heterozygous at two loci (*Aa Bb*) produces four types of gametes (*A B*, *a b*, *A b*, and *a B*) in equal proportions: two types of nonrecombinants and two types of recombinants.

Second, the genes may be completely linked—meaning that they're on the same chromosome and lie so close together that crossing over between them is rare. In this case, the genes do not recombine. An individual heterozygous for two closely linked genes in the coupling configuration

$$\frac{A \quad B}{a \quad b}$$

produces only the nonrecombinant gametes containing alleles *A B* or *a b*. The alleles do not assort into new combinations such as *A b* or *a B*.

The third situation, incomplete linkage, is intermediate between the two extremes of independent assortment and complete linkage. Here, the genes are physically linked on the same chromosome, which prevents independent assortment. However, occasional crossovers break up the linkage and allow the genes to recombine. With incomplete linkage, an individual heterozygous at two loci produces four types of gametes—two types of recombinants and two types of nonrecombinants—but the nonrecombinants are produced more frequently than the recombinants because crossing over does not take place in every meiosis.

Earlier in the chapter, the term recombination was defined as the sorting of alleles into new combinations. We can now distinguish between two types of recombination that differ in the mechanism that generates these new combinations of alleles.

Interchromosomal recombination is between genes on *different* chromosomes. It arises from independent assortment—the random segregation of chromosomes in anaphase I of meiosis. This is the kind of recombination that Mendel discovered while studying dihybrid crosses. **Intrachromosomal recombination** is between genes located on the *same* chromosome. It arises from crossing over—the exchange of genetic material in prophase I of meiosis. Both types of recombination produce new allele combinations in the gametes; so they cannot be distinguished by examining the types of gametes produced. Nevertheless, they can often be distinguished by the *frequencies* of types of gametes: interchromosomal recombination produces 50% nonrecombinant gametes and 50% recombinant gametes, whereas intrachromosomal recombination frequently produces fewer than 50% recombinant gametes. However, when the genes are very far apart on the same chromosome, they assort independently, as if they were on different chromosomes. In this case, intrachromosomal recombination also produces 50% recombinant gametes. Intrachromosomal recombination of genes that lie far apart on the same chromosome and interchromosomal recombination are genetically indistinguishable.

Predicting the Outcomes of Crosses with Linked Genes

Knowing the arrangement of alleles on a chromosome allows us to predict the types of progeny that will result from a cross entailing linked genes and to determine which of these types will be the most numerous. Determining the *proportions* of the types of offspring requires an additional piece of information—the recombination frequency. The recombination frequency provides us with information about how often the alleles in the gametes appear in new combinations and allows us to predict the proportions of offspring phenotypes that will result from a specific cross with linked genes.

In cucumbers, smooth fruit (t) is recessive to warty fruit (T) and glossy fruit (d) is recessive to dull fruit (D). Geneticists have determined that these two genes exhibit a recombination frequency of 16%. Suppose we cross a plant homozygous for warty and dull fruit with a plant homozygous for smooth and glossy fruit and then carry out a testcross by using the F_1:

$$\frac{T\ \ D}{t\ \ d} \times \frac{t\ \ d}{t\ \ d}$$

What types and proportions of progeny will result from this testcross?

Four types of gametes will be produced by the heterozygous parent, as shown in **Figure 5.9**: two types of nonrecombinant gametes ($\underline{T\ \ D}$ and $\underline{t\ \ d}$) and two types of recombinant gametes ($\underline{T\ \ d}$ and $\underline{t\ \ D}$). The recombination frequency tells us that 16% of the gametes produced by the heterozygous parent will be recombinants. Because there are two types of recombinant gametes, each should arise with a frequency of $^{16\%}/_{2} = 8\%$ This frequency can also be represented as a probability of 0.08. All the other gametes will be nonrecombinants; so they should arise with a frequency of 100% − 16% = 84%. Because there are two types of nonrecombinant gametes, each should arise with a frequency of $^{84\%}/_{2} = 42\%$ (or 0.42). The other parent in the testcross is homozygous and therefore produces only a single type of gamete ($\underline{t\ \ d}$) with a frequency of 100% (or 1.00).

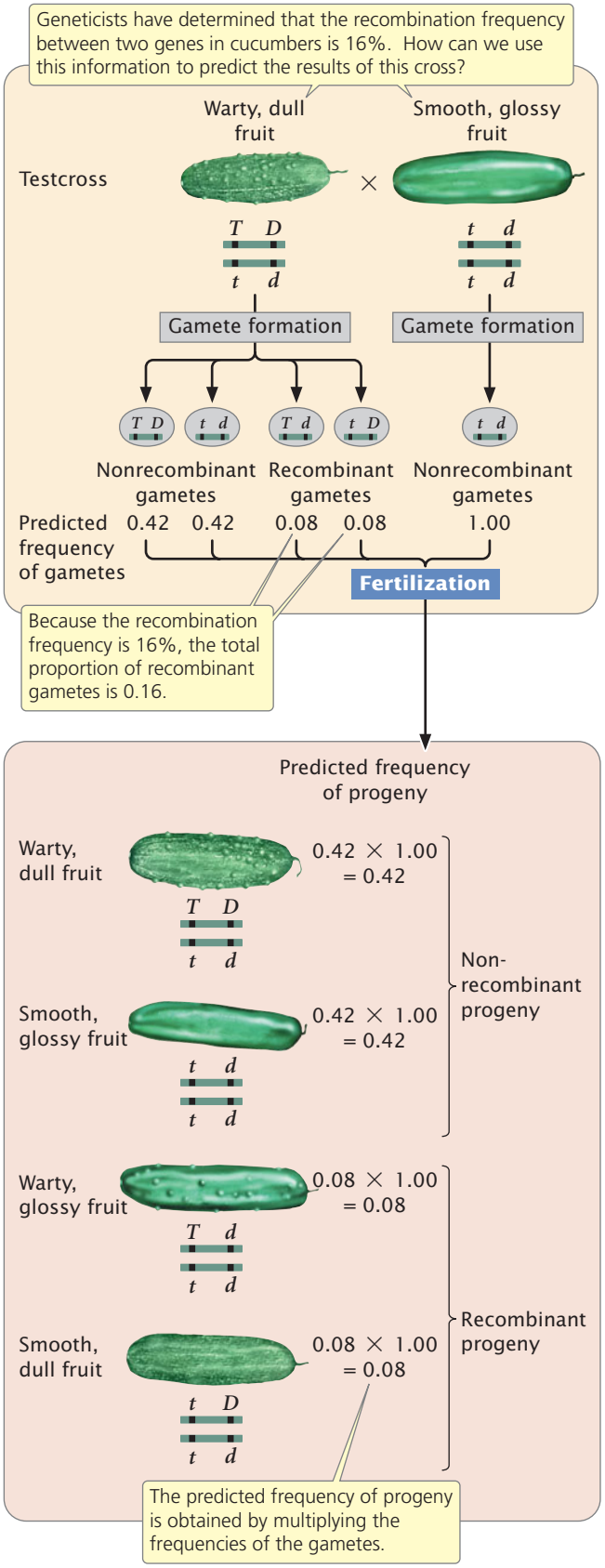
The progeny of the cross result from the union of two gametes, producing four types of progeny (see Figure 5.9). The expected proportion of each type can be determined by using the multiplication rule, multiplying together the probability of each uniting gamete. Testcross progeny with warty and dull fruit

$$\frac{T\ \ D}{t\ \ d}$$

appear with a frequency of 0.42 (the probability of inheriting a gamete with chromosome $\underline{T\ \ D}$ from the heterozygous parent) × 1.00 (the probability of inheriting a gamete with chromosome $\underline{t\ \ d}$ from the recessive parent) = 0.42. The proportions of the other types of F_2 progeny can be calculated in a similar manner (see Figure 5.9). This method can be used for predicting the outcome of any cross with linked genes for which the recombination frequency is known.

Testing for Independent Assortment

In some crosses, the genes are obviously linked because there are clearly more nonrecombinant progeny than recombinant progeny. In other crosses, the difference between independent assortment and linkage is not so obvious. For example, suppose we did a testcross for two pairs of genes, such as *Aa Bb* × *aa bb*, and observed the following numbers of progeny: 54 *Aa Bb*, 56 *aa bb*, 42 *Aa bb*, and 48 *aa Bb*. Is this outcome the 1 : 1 : 1 : 1 ratio we would expect if *A* and *B* assorted independently? Not exactly, but it's pretty close. Perhaps these genes assorted independently and chance produced the slight deviations between the observed numbers and the expected 1 : 1 : 1 : 1 ratio. Alternatively, the genes might be linked, with considerable crossing over taking place between them, and so the number of nonrecombinants is only slightly greater than the number of recombinants. How do we distinguish between the role of chance and the role of linkage in producing deviations from the results expected with independent assortment?



5.9 The recombination frequency allows a prediction of the proportions of offspring expected for a cross entailing linked genes.

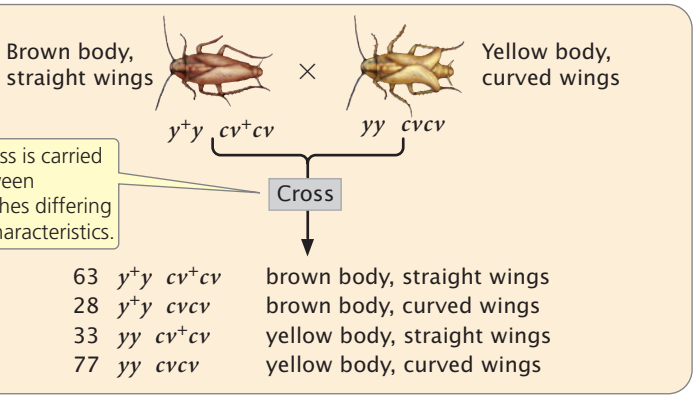
We encountered a similar problem in crosses in which genes were unlinked—the problem of distinguishing between deviations due to chance and those due to other factors. We addressed this problem (in Chapter 3) with the goodness-of-fit chi-square test, which helps us evaluate the likelihood that chance alone is responsible for deviations between the numbers of progeny that we observed and the numbers that we expected by applying the principles of inheritance. Here, we are interested in a different question: Is the inheritance of alleles at one locus independent of the inheritance of alleles at a second locus? If the answer to this question is yes, then the genes are assorting independently; if the answer is no, then the genes are probably linked.

A possible way to test for independent assortment is to calculate the expected probability of each progeny type, assuming independent assortment, and then use the goodness-of-fit chi-square test to evaluate whether the observed numbers deviate significantly from the expected numbers. With independent assortment, we expect $^1/_4$ of each phenotype: $^1/_4$ *Aa Bb*, $^1/_4$ *aa bb*, $^1/_4$ *Aa bb*, and $^1/_4$ *aa Bb*. This expected probability of each genotype is based on the multiplication rule of probability, which we considered in Chapter 3. For example, if the probability of *Aa* is $^1/_2$ and the probability of *Bb* is $^1/_2$, then the probability of *Aa Bb* is $^1/_2 \times {}^1/_2 = {}^1/_4$. In this calculation, we are making two assumptions: (1) the probability of each single-locus genotype is $^1/_2$ and (2) genotypes at the two loci are inherited independently ($^1/_2 \times {}^1/_2 = {}^1/_4$).

One problem with this approach is that a significant chi-square test can result from a violation of either assumption. If the genes are linked, then the inheritances of genotypes at the two loci are not independent (assumption 2), and we will get a significant deviation between observed and expected numbers. But we can also get a significant deviation if the probability of each single-locus genotype is not $^1/_2$ (assumption 1), even when the genotypes are assorting independently. We may obtain a significant deviation, for example, if individuals with one genotype have a lower probability of surviving or the penetrance of a genotype is not 100%. We could test both assumptions by conducting a series of chi-square tests, first testing the inheritance of genotypes at each locus separately (assumption 1) and then testing for independent assortment (assumption 2). However, a faster method is to test for independence in genotypes with a *chi-square test of independence*.

The chi-square test of independence allows us to evaluate whether the segregation of alleles at one locus is independent of the segregation of alleles at another locus, without making any assumption about the probability of single-locus genotypes. To illustrate this analysis, we will examine results from a cross between German cockroaches, in which yellow body (*y*) is recessive to brown body (y^+) and curved wings (*cv*) are recessive to straight wings (cv^+). A testcross ($y^+y\ cv^+cv \times yy\ cvcv$) produced the progeny shown in **Figure 5.10a**.

(a)



(b) Contingency table

2 To test for independent assortment of alleles encoding the two traits, a table is constructed...

3 ...with genotypes for one locus along the top...

4 ...and genotypes for the other locus along the left side.

5 Numbers of each genotype are placed in the table cells, and the row totals, column totals, and grand total are computed.

Segregation of y^+ and y

Segregation of cv^+ and cv	y^+y	yy	Row totals
cv^+cv	63	33	96
$cv\ cv$	28	77	105
Column totals	91	110	201
			Grand total

(c)

Genotype	Number observed	Number expected $\left(\frac{\text{row total} \times \text{column total}}{\text{grand total}}\right)$
$y^+y\ cv^+cv$	63	$\frac{96 \times 91}{201} = 43.46$
$y^+y\ cvcv$	28	$\frac{105 \times 91}{201} = 47.54$
$yy\ cv^+cv$	33	$\frac{96 \times 110}{201} = 52.46$
$yy\ cvcv$	77	$\frac{105 \times 110}{201} = 57.46$

6 The expected numbers of progeny, assuming independent assortment, are calculated.

(d)

7 A chi-square value is calculated.

$$\chi^2 = \Sigma \frac{(\text{observed} - \text{expected})^2}{\text{expected}}$$

$$\chi^2 = \frac{(63 - 43.46)^2}{43.46} + \frac{(28 - 47.54)^2}{47.54} + \frac{(33 - 52.54)^2}{52.54} + \frac{(77 - 57.46)^2}{57.46}$$

$$\chi^2 = 8.79 + 8.03 + 7.27 + 6.64$$

$$\chi^2 = 30.73$$

(e)

df = (number of rows − 1) × (number of columns − 1)

$df = (2 - 1) \times (2 - 1) = 1 \times 1 = 1$

$P < 0.005$

8 The probability is less than 0.005, indicating that the *difference* between numbers of observed and expected progeny is probably not due to chance.

Conclusion: The genes for body color and type of wing are not assorting independently and must be linked.

To carry out the chi-square test of independence, we first construct a table of the observed numbers of progeny, somewhat like a Punnett square, except, here, we put the genotypes that result from the segregation of alleles at one locus along the top and the genotypes that result from the segregation of alleles at the other locus along the side (**Figure 5.10b**). Next, we compute the total for each row, the total for each column, and the grand total (the sum of all row totals or the sum of all column totals, which should be the same). These totals will be used to compute the expected values for the chi-square test of independence.

Now, we compute the values expected if the segregation of alleles at the *y* locus is independent of the segregation of alleles at the *cv* locus. If the segregation of alleles at each locus is independent, then the proportion of progeny with y^+y and yy genotypes should be the same for cockroaches with genotype cv^+cv and for cockroaches with genotype $cvcv$. The converse is also true; the proportions of progeny with cv^+cv and $cvcv$ genotypes should be the same for cockroaches with genotype y^+y and for cockroaches with genotype yy. With the assumption that the alleles at the two loci segregate independently, the expected number for each cell of the table can be computed by using the following formula:

$$\text{expected number} = \frac{\text{row total} \times \text{column total}}{\text{grand total}}$$

For the cell of the table corresponding to genotype $y^+y\ cv^+cv$ (the upper-left-hand cell of the table in Figure 5.10b) the expected number is:

$$\frac{96\ (\text{row total}) \times 91\ (\text{column total})}{201\ (\text{grand total})} = \frac{8736}{201} = 43.46$$

With the use of this method, the expected numbers for each cell are given in **Figure 5.10c**.

We now calculate a chi-square value by using the same formula that we used for the goodness-of-fit chi-square test in Chapter 3:

$$x^2 = \sum \frac{(\text{observed} - \text{expected})^2}{\text{expected}}$$

With the observed and expected numbers of cockroaches from the testcross, the calculated chi-square value is 30.73 (**Figure 5.10d**).

To determine the probability associated with this chi-square value, we need the degrees of freedom. Recall from Chapter 3 that the degrees of freedom are the number of ways in which the observed classes are

5.10 A chi-square test of independence can be used to determine if genes at two loci are assorting independently.

free to vary from the expected values. In general, for the chi-square test of independence, the degrees of freedom equal the number of rows in the table minus 1 multiplied by the number of columns in the table minus 1 (**Figure 5.10e**), or

$$df = (\text{number of rows} - 1) \times (\text{number of columns} - 1)$$

In our example, there were two rows and two columns, and so the degrees of freedom are:

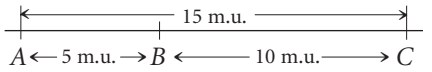
$$df = (2 - 1) \times (2 - 1) = 1 \times 1 = 1$$

Therefore, our calculated chi-square value is 30.73, with 1 degree of freedom. We can use Table 3.4 to find the associated probability. Looking at Table 3.4, we find our calculated chi-square value is larger than the largest chi-square value given for 1 degree of freedom, which has a probability of 0.005. Thus, our calculated chi-square value has a probability less than 0.005. This very small probability indicates that the genotypes are not in the proportions that we would expect if independent assortment were taking place. Our conclusion, then, is that these genes are not assorting independently and must be linked. As is the case for the goodness-of-fit chi-square test, geneticists generally consider that any chi-square value for the test of independence with a probability less than 0.05 is significantly different from the expected values and is therefore evidence that the genes are not assorting independently.

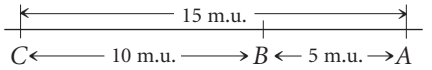
Gene Mapping with Recombination Frequencies

Morgan and his students developed the idea that physical distances between genes on a chromosome are related to the rates of recombination. They hypothesized that crossover events occur more or less at random up and down the chromosome and that two genes that lie far apart are more likely to undergo a crossover than are two genes that lie close together. They proposed that recombination frequencies could provide a convenient way to determine the order of genes along a chromosome and would give estimates of the relative distances between the genes. Chromosome maps calculated by using the genetic phenomenon of recombination are called **genetic maps**. In contrast, chromosome maps calculated by using physical distances along the chromosome (often expressed as numbers of base pairs) are called **physical maps**.

Distances on genetic maps are measured in **map units** (abbreviated m.u.); one map unit equals 1% recombination. Map units are also called **centiMorgans** (cM), in honor of Thomas Hunt Morgan; 100 centiMorgans equals one **Morgan**. Genetic distances measured with recombination rates are approximately additive: if the distance from gene *A* to gene *B* is 5 m.u., the distance from gene *B* to gene *C* is 10 m.u., and the distance from gene *A* to gene *C* is 15 m.u., then gene *B* must be located between genes *A* and *C*. On the basis of the map distances just given, we can draw a simple genetic map for genes *A*, *B*, and *C*, as shown here:



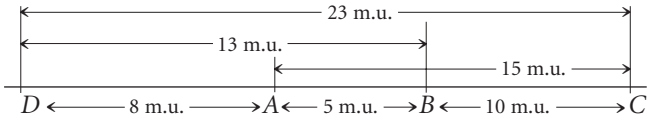
We could just as plausibly draw this map with *C* on the left and *A* on the right:



Both maps are correct and equivalent because, with information about the relative positions of only three genes, the most that we can determine is which gene lies in the middle. If we obtained distances to an additional gene, then we could position *A* and *C* relative to that gene. An additional gene *D*, examined through genetic crosses, might yield the following recombination frequencies:

Gene pair	Recombination frequency (%)
A and *D*	8
B and *D*	13
C and *D*	23

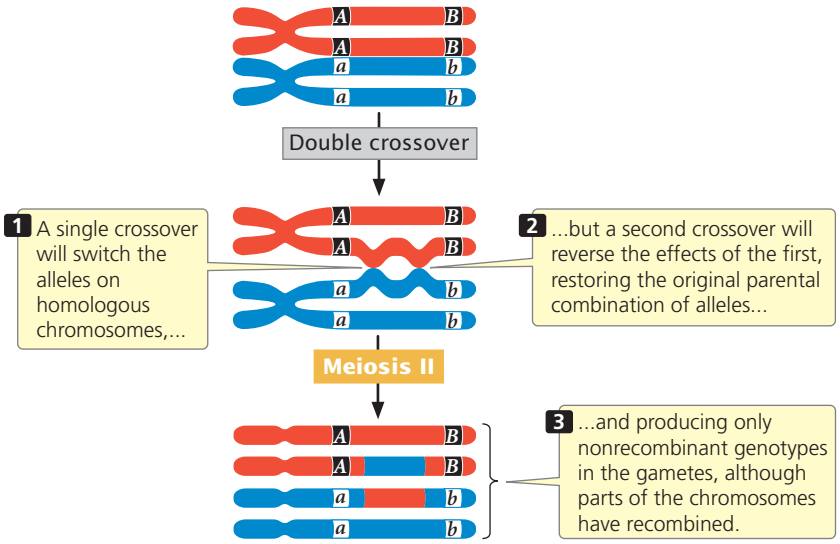
Notice that *C* and *D* exhibit the greatest amount of recombination; therefore, *C* and *D* must be farthest apart, with genes *A* and *B* between them. Using the recombination frequencies and remembering that 1 m.u. = 1% recombination, we can now add *D* to our map:



By doing a series of crosses between pairs of genes, we can construct genetic maps showing the linkage arrangements of a number of genes.

Two points should be emphasized about constructing chromosome maps from recombination frequencies. First, recall that we cannot distinguish between genes on different chromosomes and genes located far apart on the same chromosome. If genes exhibit 50% recombination, the most that can be said about them is that they belong to different groups of linked genes (different linkage groups), either on different chromosomes or far apart on the same chromosome.

The second point is that a testcross for two genes that are relatively far apart on the same chromosome tends to underestimate the true physical distance, because the cross does not reveal double crossovers that might take place between the two genes (**Figure 5.11**). A double crossover arises when two separate crossover events take place between two loci. (For now, we will consider only double crossovers that occur between two of the four chromatids of a homologus pair—a two-strand double crossover. Double crossovers entailing three and four chromatids will be considered later.) Whereas a single crossover produces combinations of alleles that were



5.11 A two-strand double crossover between two linked genes produces only nonrecombinant gametes.

not present on the original parental chromosomes, a second crossover between the same two genes reverses the effects of the first, thus restoring the original parental combination of alleles (see Figure 5.11). Two-strand double crossovers produce only nonrecombinant gametes, and so we cannot distinguish between the progeny produced by double crossovers and the progeny produced when there is no crossing over at all. As we shall see in the next section, we can detect double crossovers if we examine a third gene that lies between the two crossovers. Because double crossovers between two genes go undetected, map distances will be underestimated whenever double crossovers take place. Double crossovers are more frequent between genes that are far apart; therefore genetic maps based on short distances are usually more accurate than those based on longer distances.

Concepts

A genetic map provides the order of the genes on a chromosome and the approximate distances from one gene to another based on recombination frequencies. In genetic maps, 1% recombination equals 1 map unit, or 1 centiMorgan. Double crossovers between two genes go undetected; so map distances between distant genes tend to underestimate genetic distances.

✔ Concept Check 3

How does a genetic map differ from a physical map?

Constructing a Genetic Map with Two-Point Testcrosses

Genetic maps can be constructed by conducting a series of testcrosses between pairs of genes and examining the recombination frequencies between them. A testcross between two genes is called a **two-point testcross** or a two-point cross for short. Suppose that we carried out a series of two-point crosses for four genes, *a*, *b*, *c*, and *d*, and obtained the following recombination frequencies:

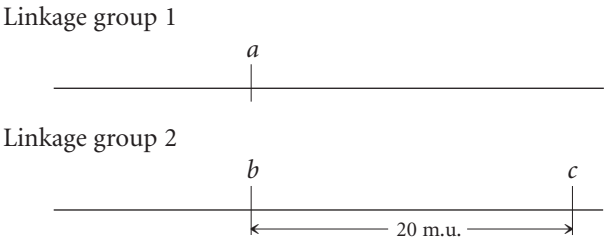
Gene loci in testcross	Recombination frequency (%)
a and *b*	50
a and *c*	50
a and *d*	50
b and *c*	20
b and *d*	10
c and *d*	28

We can begin constructing a genetic map for these genes by considering the recombination frequencies for each pair of genes. The recombination frequency between *a* and *b* is 50%, which is the recombination frequency expected with independent assortment. Therefore, genes *a* and *b* may either be on different chromosomes or be very far apart on the same chromosome; so we will place them in different linkage groups with the understanding that they may or may not be on the same chromosome:

Linkage group 1

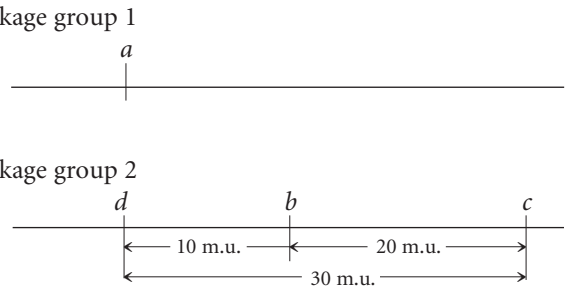
Linkage group 2

The recombination frequency between *a* and *c* is 50%, indicating that they, too, are in different linkage groups. The recombination frequency between *b* and *c* is 20%; so these genes are linked and separated by 20 map units:



The recombination frequency between *a* and *d* is 50%, indicating that these genes belong to different linkage groups, whereas genes *b* and *d* are linked, with a recombination frequency of 10%. To decide whether gene *d* is 10 m.u. to the left or to the right of gene *b*, we must consult the *c*-to-*d* distance. If gene *d* is 10 m.u. to the left of gene *b*, then the distance between *d* and *c* should be 20 m.u. + 10 m.u. = 30 m.u. This distance will be only approximate because any double crossovers between the two genes will be missed and the map distance will be underestimated. If, on the other hand, gene *d* lies to the right of gene *b*, then the distance between gene *d* and gene *c* will be much shorter, approximately 20 m.u. − 10 m.u. = 10 m.u.

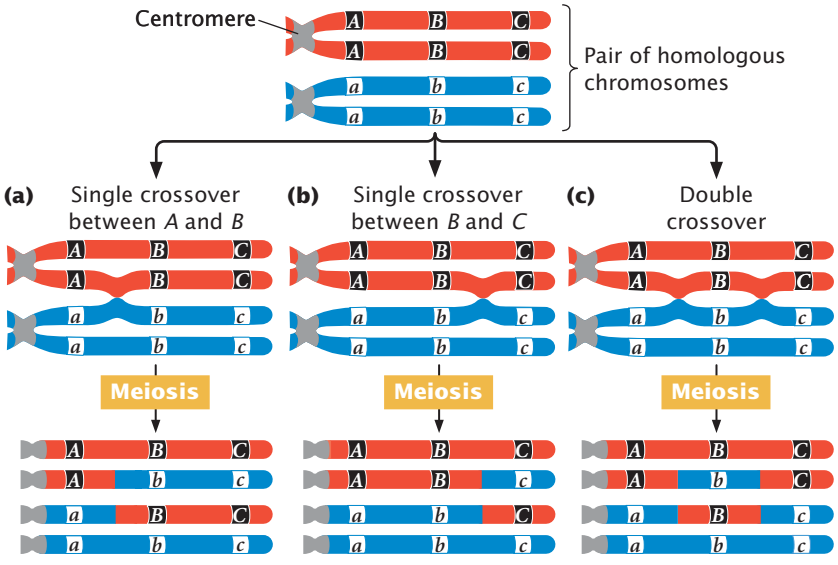
By examining the recombination frequency between *c* and *d*, we can distinguish between these two possibilities. The recombination frequency between *c* and *d* is 28%; so gene *d* must lie to the left of gene *b*. Notice that the sum of the recombination between *d* and *b* (10%) and between *b* and *c* (20%) is greater than the recombination between *d* and *c* (28%). As already discussed, this discrepancy arises because double crossovers between the two outer genes go undetected, causing an underestimation of the true map distance. The genetic map of these genes is now complete:



5.3 A Three-Point Testcross Can Be Used to Map Three Linked Genes

Genetic maps can be constructed from a series of testcrosses for pairs of genes, but this approach is not particularly efficient, because numerous two-point crosses must be carried out to establish the order of the genes and because double crossovers are missed. A more efficient mapping technique is a testcross for three genes—a **three-point testcross**, or three-point cross. With a three-point cross, the order of the three genes can be established in a single set of progeny and some double crossovers can usually be detected, providing more accurate map distances.

Consider what happens when crossing over takes place among three hypothetical linked genes. Figure 5.12 illustrates a pair of homologous chromosomes from an individual that is heterozygous at three loci (*Aa Bb Cc*). Notice that the genes are in the coupling configuration; that is, all the dominant alleles are on one chromosome ($\underline{A\ \ B\ \ C}$) and all the recessive alleles are on the other chromosome



Conclusion: Recombinant chromosomes resulting from the double crossover have only the middle gene altered.

5.12 **Three types of crossovers can take place among three linked loci.**

($\frac{a\ \ b\ \ c}{}$). Three types of crossover events can take place between these three genes: two types of single crossovers (see Figure 5.12a and b) and a double crossover (see Figure 5.12c). In each type of crossover, two of the resulting chromosomes are recombinants and two are nonrecombinants.

Notice that, in the recombinant chromosomes resulting from the double crossover, the outer two alleles are the same as in the nonrecombinants, but the middle allele is different. This result provides us with an important clue about the order of the genes. In progeny that result from a double crossover, only the middle allele should differ from the alleles present in the nonrecombinant progeny.

Constructing a Genetic Map with the Three-Point Testcross

To examine gene mapping with a three-point testcross, we will consider three recessive mutations in the fruit fly *Drosophila melanogaster*. In this species, scarlet eyes (*st*) are recessive to red eyes (st^+), ebony body color (*e*) is recessive to gray body color (e^+), and spineless (*ss*)—that is, the presence of small bristles—is recessive to normal bristles (ss^+). All three mutations are linked and located on the third chromosome.

We will refer to these three loci as *st*, *e*, and *ss*, but keep in mind that either the recessive alleles (*st*, *e*, and *ss*) or the dominant alleles (st^+, e^+, and ss^+) may be present at each locus. So, when we say that there are 10 m.u. between *st* and *ss*, we mean that there are 10 m.u. between the loci at which these mutations occur; we could just as easily say that there are 10 m.u. between st^+ and ss^+.

To map these genes, we need to determine their order on the chromosome and the genetic distances between them. First, we must set up a three-point testcross, a cross between a fly heterozygous at all three loci and a fly homozygous for recessive alleles at all three loci. To produce flies heterozygous for all three loci, we might cross a stock of flies that are homozygous for normal alleles at all three loci with flies that are homozygous for recessive alleles at all three loci:

$$\text{P}\quad \frac{st^+\ e^+\ ss^+}{st^+\ e^+\ ss^+} \times \frac{st\ \ e\ \ ss}{st\ \ e\ \ ss}$$

$$\downarrow$$

$$\text{F}_1\quad \frac{st^+\ e^+\ ss^+}{st\ \ e\ \ ss}$$

The order of the genes has been arbitrarily assigned because, at this point, we do not know which is the middle gene. Additionally, the alleles in these heterozygotes are in coupling configuration (because all the wild-type dominant alleles were inherited from one parent and all the recessive mutations from the other parent), although the testcross can also be done with alleles in repulsion.

In the three-point testcross, we cross the F_1 heterozygotes with flies that are homozygous for all three recessive mutations. In many organisms, it makes no difference whether the heterozygous parent in the testcross is male or female (provided that the genes are autosomal) but, in *Drosophila*, no crossing over takes place in males. Because crossing over in the heterozygous parent is essential for determining recombination frequencies, the heterozygous flies in our testcross must be female. So we mate female F_1 flies that are heterozygous for all three traits with male flies that are homozygous for all the recessive traits:

$$\frac{st^+\ e^+\ ss^+}{st\ \ e\ \ ss}\ \text{Female} \times \frac{st\ \ e\ \ ss}{st\ \ e\ \ ss}\ \text{Male}$$

The progeny produced from this cross are listed in **Figure 5.13**. For each locus, two classes of progeny are produced: progeny that are heterozygous, displaying the dominant trait, and progeny that are homozygous, displaying the recessive trait. With two classes of progeny possible for each of the three loci, there will be $2^3 = 8$ classes of phenotypes possible in the progeny. In this example, all eight phenotypic classes are present but, in some three-point crosses, one or more of the phenotypes may be missing if the number of progeny is limited. Nevertheless, the absence of a particular class can provide important information about which combination of traits is least frequent and, ultimately, about the order of the genes, as we will see.

To map the genes, we need information about where and how often crossing over has taken place. In the homozygous recessive parent, the two alleles at each locus are the same, and so crossing over will have no effect on the types of gametes produced; with or without crossing over, all gametes from this parent have a chromosome with three recessive alleles ($\frac{st\ \ e\ \ ss}{}$). In contrast, the heterozygous parent has different alleles on its two chromosomes, and so crossing over can be detected. The information that we need for mapping, therefore, comes entirely from the gametes produced by the heterozygous parent. Because chromosomes contributed by the homozygous parent carry only recessive alleles, whatever alleles are present on the chromosome contributed by the heterozygous parent will be expressed in the progeny.

As a shortcut, we often do not write out the complete genotypes of the testcross progeny, listing instead only the alleles expressed in the phenotype, which are the alleles inherited from the heterozygous parent. This convention is used in the discussion that follows.

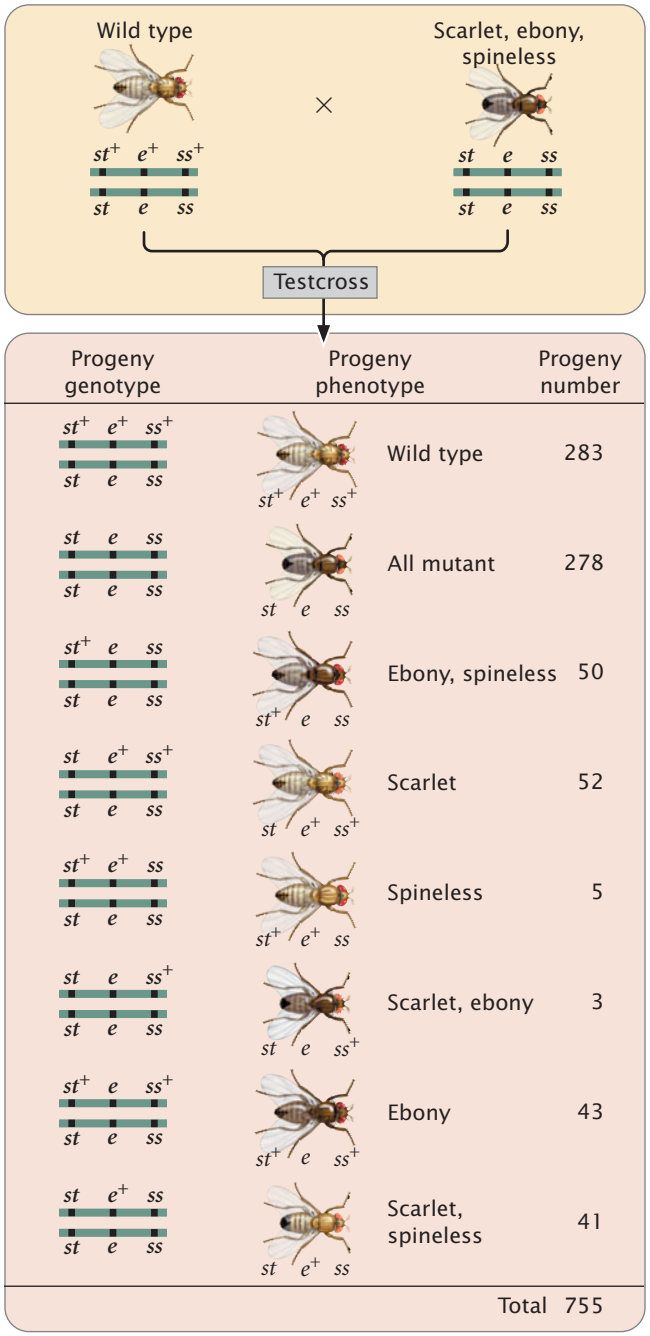
Concepts

To map genes, information about the location and number of crossovers in the gametes that produced the progeny of a cross is needed. An efficient way to obtain this information is to use a three-point testcross, in which an individual heterozygous at three linked loci is crossed with an individual that is homozygous recessive at the three loci.

✔ Concept Check 4

Write the genotypes of all recombinant and nonrecombinant progeny expected from the following three-point cross:

$$\frac{m^+\ p^+\ s^+}{m\ \ p\ \ s} \times \frac{m\ \ p\ \ s}{m\ \ p\ \ s}$$



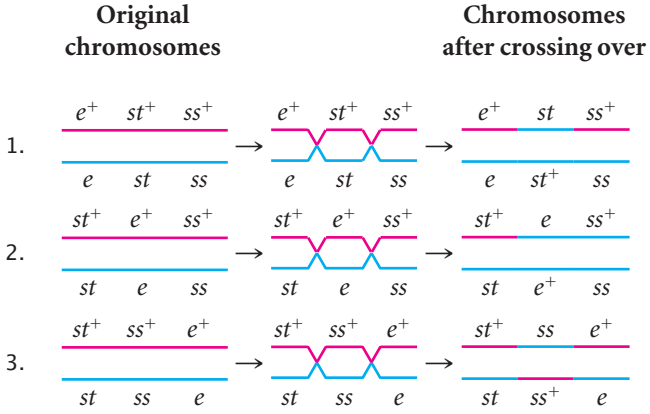
5.13 The results of a three-point testcross can be used to map linked genes. In this three-point testcross of *Drosophila melanogaster*, the recessive mutations scarlet eyes (*st*), ebony body color (*e*), and spineless bristles (*ss*) are at three linked loci. The order of the loci has been designated arbitrarily, as has the sex of the progeny flies.

Determining the gene order The first task in mapping the genes is to determine their order on the chromosome. In Figure 5.13, we arbitrarily listed the loci in the order *st*, *e*, *ss*, but we had no way of knowing which of the three loci was between the other two. We can now identify the middle locus by examining the double-crossover progeny.

First, determine which progeny are the nonrecombinants; they will be the two most-numerous classes of progeny. (Even if crossing over takes place in every meiosis, the nonrecombinants will constitute at least 50% of the progeny.) Among the progeny of the testcross in Figure 5.13, the most numerous are those with all three dominant traits (st^+ e^+ ss^+) and those with all three recessive traits (*st* *e* *ss*).

Next, identify the double-crossover progeny. These progeny should always have the two least-numerous phenotypes, because the probability of a double crossover is always less than the probability of a single crossover. The least-common progeny among those listed in Figure 5.13 are progeny with spineless bristles (st^+ e^+ *ss*) and progeny with scarlet eyes and ebony body (*st* *e* ss^+); so they are the double-crossover progeny.

Three orders of genes on the chromosome are possible: the eye-color locus could be in the middle (*e* *st* *ss*), the body-color locus could be in the middle (*st* *e* *ss*), or the bristle locus could be in the middle (*st* *ss* *e*). To determine which gene is in the middle, we can draw the chromosomes of the heterozygous parent with all three possible gene orders and then see if a double crossover produces the combination of genes observed in the double-crossover progeny. The three possible gene orders and the types of progeny produced by their double crossovers are:



The only gene order that produces chromosomes with the set of alleles observed in the least-numerous progeny or double crossovers (st^+ e^+ *ss* and *st* *e* ss^+ in Figure 5.13) is the one in which the *ss* locus for bristles lies in the middle (gene-order 3). Therefore, this order (*st* *ss* *e*) must be the correct sequence of genes on the chromosome.

With a little practice, we can quickly determine which locus is in the middle without writing out all the gene orders. The phenotypes of the progeny are expressions of the alleles inherited from the heterozygous parent. Recall that, when we looked at the results of double crossovers (see Figure 5.12), only the alleles at the middle locus differed from the nonrecombinants. If we compare the nonrecombinant progeny with double-crossover progeny, they should differ only in alleles of the middle locus.

Let's compare the alleles in the double-crossover progeny $\frac{st^+\ e^+\ ss}{st^+\ e^+\ ss^+}$ with those in the nonrecombinant progeny. We see that both have an allele for red eyes (st^+) and both have an allele for gray body (e^+), but the nonrecombinants have an allele for normal bristles (ss^+), whereas the double crossovers have an allele for spineless bristles (ss). Because the bristle locus is the only one that differs, it must lie in the middle. We would obtain the same results if we compared the other class of double-crossover progeny ($\frac{st\ e\ ss^+}{\ }$) with other nonrecombinant progeny ($\frac{st\ e\ ss}{\ }$). Again, the only locus that differs is the one for bristles. Don't forget that the nonrecombinants and the double crossovers should differ only at one locus; if they differ at two loci, the wrong classes of progeny are being compared.

Concepts

To determine the middle locus in a three-point cross, compare the double-crossover progeny with the nonrecombinant progeny. The double crossovers will be the two least-common classes of phenotypes; the nonrecombinants will be the two most-common classes of phenotypes. The double-crossover progeny should have the same alleles as the nonrecombinant types at two loci and different alleles at the locus in the middle.

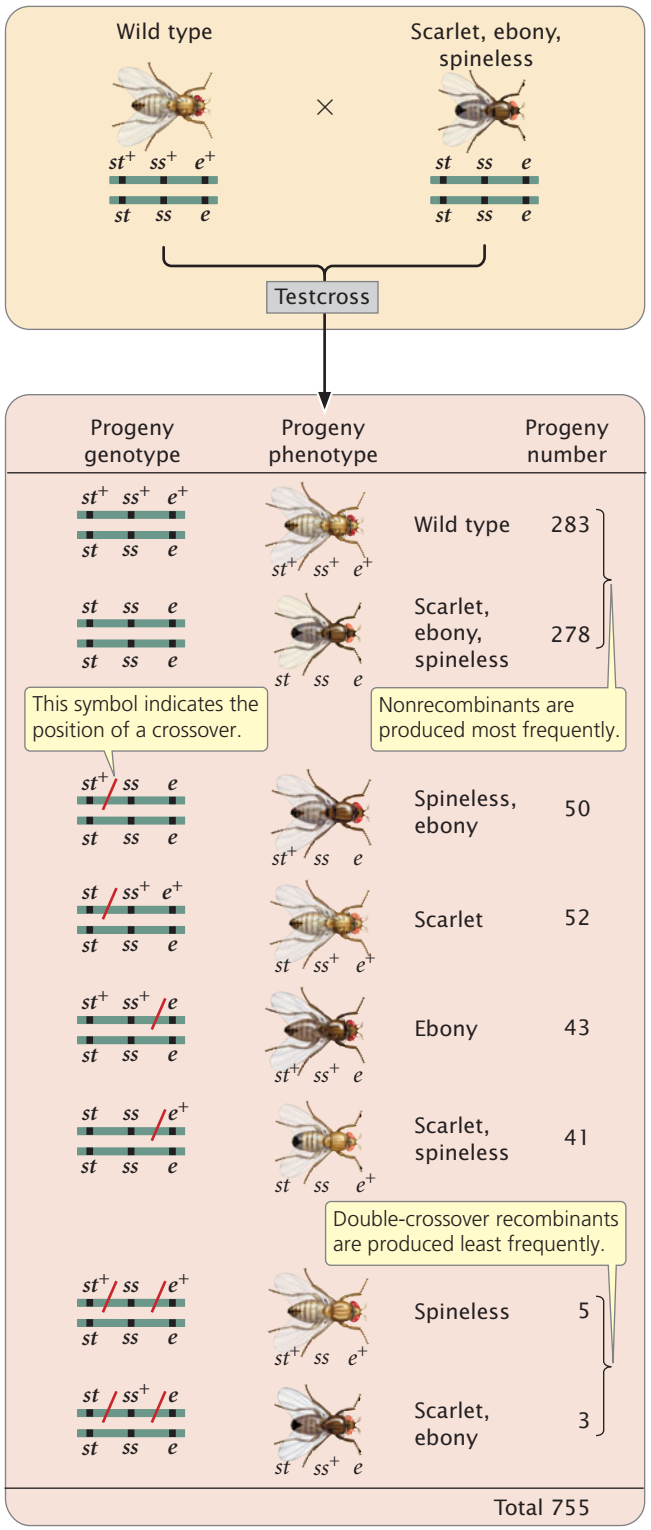
✔ Concept Check 5

A three-point test cross is carried out between three linked genes. The resulting nonrecombinant progeny are $s^+\ r^+\ c^+$ and $s\ r\ c$ and the double-crossover progeny are $s\ r\ c^+$ and $s^+\ r^+\ c$. Which is the middle locus?

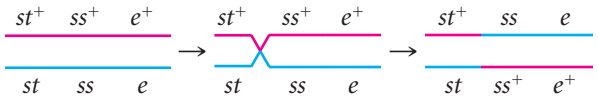
Determining the locations of crossovers When we know the correct order of the loci on the chromosome, we should rewrite the phenotypes of the testcross progeny in Figure 5.13 with the alleles in the correct order so that we can determine where crossovers have taken place (**Figure 5.14**).

Among the eight classes of progeny, we have already identified two classes as nonrecombinants ($\frac{st^+\ ss^+\ e^+}{\ }$ and $\frac{st\ ss\ e}{\ }$) and two classes as double crossovers ($\frac{st^+\ ss\ e^+}{\ }$ and $\frac{st\ ss^+\ e}{\ }$). The other four classes include progeny that resulted from a chromosome that underwent a single crossover: two underwent single crossovers between *st* and *ss*, and two underwent single crossovers between *ss* and *e*.

To determine where the crossovers took place in these progeny, compare the alleles found in the single-crossover progeny with those found in the nonrecombinants, just as we did for the double crossovers. For example, consider progeny with chromosome $\frac{st^+\ ss\ e}{\ }$. The first allele (st^+) came from the nonrecombinant chromosome $\frac{st^+\ ss^+\ e^+}{\ }$ and the other two alleles (*ss* and *e*) must have come from the other nonrecombinant chromosome $\frac{st\ ss\ e}{\ }$ through crossing over:



5.14 Writing the results of a three-point testcross with the loci in the correct order allows the locations of crossovers to be determined. These results are from the testcross illustrated in Figure 5.13, with the loci shown in the correct order. The location of a crossover is indicated by a slash (/). The sex of the progeny flies has been designated arbitrarily.



This same crossover also produces the $st\ ss^+\ e^+$ progeny.

This method can also be used to determine the location of crossing over in the other two types of single-crossover progeny. Crossing between *ss* and *e* produces $st^+\ ss^+\ e$ and $st\ ss\ e^+$ chromosomes:

We now know the locations of all the crossovers. Their locations are marked with a slash in Figure 5.14.

Calculating the recombination frequencies Next, we can determine the map distances, which are based on the frequencies of recombination. We calculate recombination frequency by adding up all of the recombinant progeny, dividing this number by the total number of progeny from the cross, and multiplying the number obtained by 100%. To determine the map distances accurately, we must include all crossovers (both single and double) that take place between two genes.

Recombinant progeny that possess a chromosome that underwent crossing over between the eye-color locus (*st*) and the bristle locus (*ss*) include the single crossovers ($st^+\ /\ ss\ e$ and $st\ /\ ss^+\ e^+$) and the two double crossovers ($st^+\ /\ ss\ /\ e^+$ and $st\ /\ ss^+\ /\ e$); see Figure 5.14. There are a total of 755 progeny; so the recombination frequency between *ss* and *st* is:

$$st\text{–}ss \text{ recombination frequency} = \frac{(50 + 52 + 5 + 3)}{755} \times 100\% = 14.6\%$$

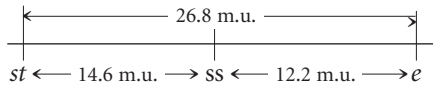
The distance between the *st* and *ss* loci can be expressed as 14.6 m.u.

The map distance between the bristle locus (*ss*) and the body locus (*e*) is determined in the same manner. The recombinant progeny that possess a crossover between *ss* and *e* are the single crossovers $st^+\ ss^+\ /\ e$ and $st\ ss\ /\ e^+$ and the double crossovers $st^+\ /\ ss\ /\ e^+$ and $st\ /\ ss^+\ /\ e$. The recombination frequency is:

$$st\text{–}e \text{ recombination frequency} = \frac{(43 + 41 + 5 + 3)}{755} \times 100\% = 12.2\%$$

Thus, the map distance between *ss* and *e* is 12.2 m.u.

Finally, calculate the map distance between the outer two loci, *st* and *e*. This map distance can be obtained by summing the map distances between *st* and *ss* and between *ss* and *e* (14.6 m.u. + 12.2 m.u. = 26.8 m.u.). We can now use the map distances to draw a map of the three genes on the chromosome:



A genetic map of *D. melanogaster* is illustrated in **Figure 5.15.**

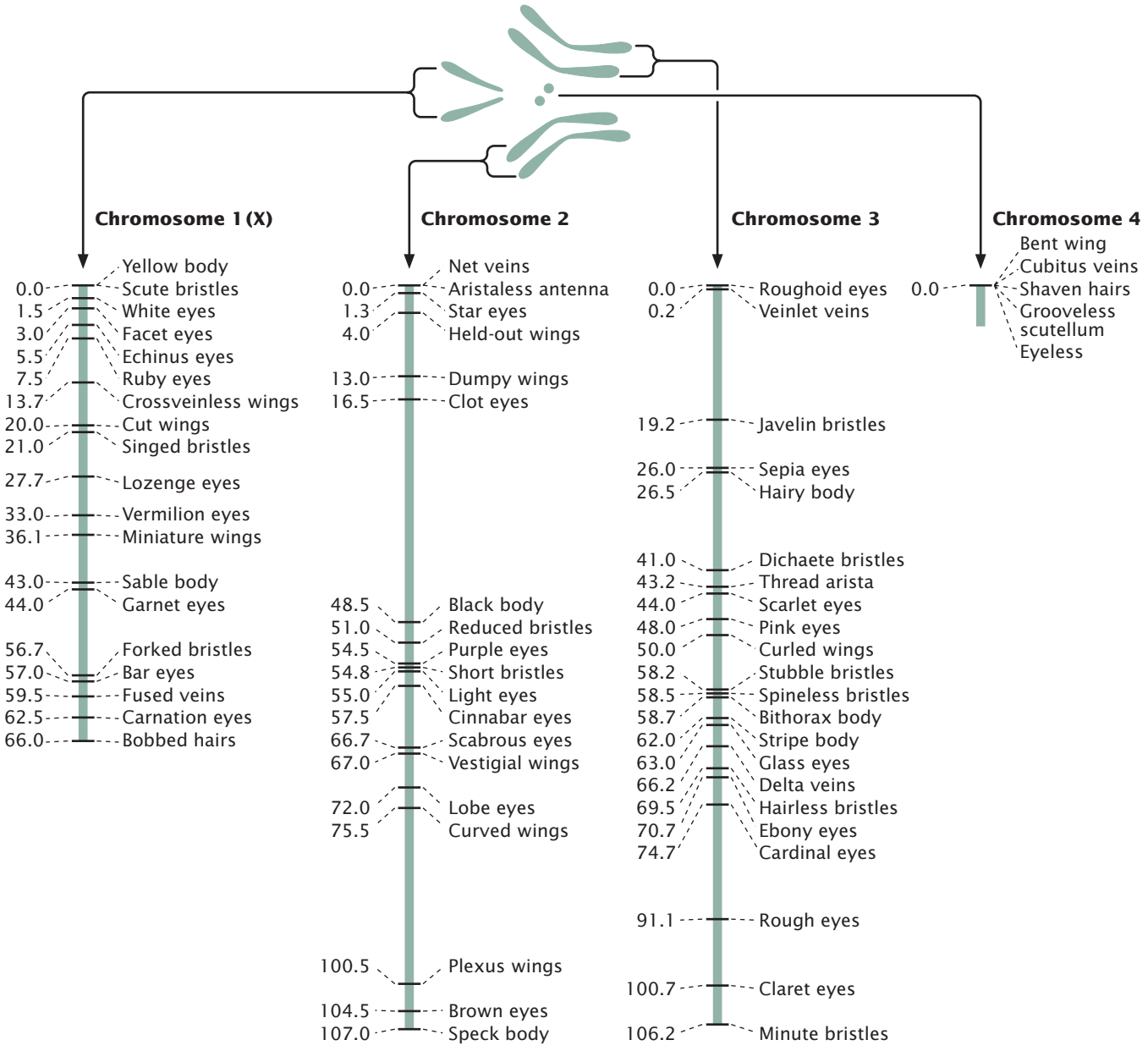
Interference and coefficient of coincidence Map distances give us information not only about the distances that separate genes, but also about the proportions of recombinant and nonrecombinant gametes that will be produced in a cross. For example, knowing that genes *st* and *ss* on the third chromosome of *D. melanogaster* are separated by a distance of 14.6 m.u. tells us that 14.6% of the gametes produced by a fly heterozygous at these two loci will be recombinants. Similarly, 12.2% of the gametes from a fly heterozygous for *ss* and *e* will be recombinants.

Theoretically, we should be able to calculate the proportion of double-recombinant gametes by using the multiplication rule of probability (see Chapter 3), which states that the probability of two independent events occurring together is calculated by multiplying the probabilities of the independent events. Applying this principle, we should find that the proportion (probability) of gametes with double crossovers between *st* and *e* is equal to the probability of recombination between *st* and *ss* multiplied by the probability of recombination between *ss* and *e*, or $0.146 \times 0.122 = 0.0178$. Multiplying this probability by the total number of progeny gives us the *expected* number of double-crossover progeny from the cross: $0.0178 \times 755 = 13.4$. Only 8 double crossovers—considerably fewer than the 13 expected—were observed in the progeny of the cross (see Figure 5.14).

This phenomenon is common in eukaryotic organisms. The calculation assumes that each crossover event is independent and that the occurrence of one crossover does not influence the occurrence of another. But crossovers are frequently *not* independent events: the occurrence of one crossover tends to inhibit additional crossovers in the same region of the chromosome, and so double crossovers are less frequent than expected.

The degree to which one crossover interferes with additional crossovers in the same region is termed the **interference**. To calculate the interference, we first determine the **coefficient of coincidence**, which is the ratio of observed double crossovers to expected double crossovers:

$$\text{coefficient of coincidence} = \frac{\text{number of observed double crossovers}}{\text{number of expected double crossovers}}$$



5.15 *Drosophila melanogaster* has four linkage groups corresponding to its four pairs of chromosomes. Distances between genes within a linkage group are in map units.

For the loci that we mapped on the third chromosome of *D. melanogaster* (see Figure 5.14), we find that the

$$\text{coefficient of coincidence} = \frac{5 + 3}{0.146 \times 0.122 \times 755} = \frac{8}{13.4} = 0.6$$

which indicates that we are actually observing only 60% of the double crossovers that we expected on the basis of the single-crossover frequencies. The interference is calculated as

$$\text{interference} = 1 - \text{coefficient of coincidence}$$

So the interference for our three-point cross is:

$$\text{interference} = 1 - 0.6 = 0.4$$

This value of interference tells us that 40% of the double-crossover progeny expected will not be observed, because of interference. When interference is complete and no double-crossover progeny are observed, the coefficient of coincidence is 0 and the interference is 1.

Sometimes a crossover *increases* the probability of another crossover taking place nearby and we see *more* double-crossover progeny than expected. In this case, the coefficient of coincidence is greater than 1 and the interference is negative.

Concepts

The coefficient of coincidence equals the number of double crossovers observed divided by the number of double crossovers expected on the basis of the single-crossover frequencies. The interference equals 1 – the coefficient of coincidence; it indicates the degree to which one crossover interferes with additional crossovers.

✔ Concept Check 6

In analyzing the results of a three-point testcross, a student determines that the interference is −0.23. What does this negative interference value indicate?

a. Fewer double crossovers took place than expected on the basis of single-crossover frequencies.
b. More double crossovers took place than expected on the basis of single-crossover frequencies.
c. Fewer single crossovers took place than expected.
d. A crossover in one region interferes with additional crossovers in the same region.

Connecting Concepts

Stepping Through the Three-Point Cross

We have now examined the three-point cross in considerable detail and have seen how the information derived from the cross can be used to map a series of three linked genes. Let's briefly review the steps required to map genes from a three-point cross.

1. **Write out the phenotypes and numbers of progeny produced in the three-point cross**. The progeny phenotypes will be easier to interpret if you use allelic symbols for the traits (such as st^+ e^+ ss).
2. **Write out the genotypes of the original parents used to produce the triply heterozygous individual** in the testcross and, if known, the arrangement (coupling or repulsion) of the alleles on their chromosomes.
3. **Determine which phenotypic classes among the progeny are the nonrecombinants and which are the double crossovers**. The nonrecombinants will be the two most-common phenotypes; double crossovers will be the two least-common phenotypes.
4. **Determine which locus lies in the middle**. Compare the alleles present in the double crossovers with those present in the nonrecombinants; each class of double crossovers should be like one of the nonrecombinants for two loci and should differ for one locus. The locus that differs is the middle one.
5. **Rewrite the phenotypes with genes in correct order**.
6. **Determine where crossovers must have taken place to give rise to the progeny phenotypes.** To do so, compare each phenotype with the phenotype of the nonrecombinant progeny.
7. **Determine the recombination frequencies**. Add the numbers of the progeny that possess a chromosome with a crossover between a pair of loci. Add the double crossovers to this number. Divide this sum by the total number of progeny from the cross, and multiply by 100%; the result is the recombination frequency between the loci, which is the same as the map distance.
8. **Draw a map of the three loci**. Indicate which locus lies in the middle, and indicate the distances between them.
9. **Determine the coefficient of coincidence and the interference**. The coefficient of coincidence is the number of observed double-crossover progeny divided by the number of expected double-crossover progeny. The expected number can be obtained by multiplying the product of the two single-recombination probabilities by the total number of progeny in the cross.

Worked Problem

In *D. melanogaster*, cherub wings (*ch*), black body (*b*), and cinnabar eyes (*cn*) result from recessive alleles that are all located on chromosome 2. A homozygous wild-type fly was mated with a cherub, black, and cinnabar fly, and the resulting F_1 females were test-crossed with cherub, black, and cinnabar males. The following progeny were produced from the testcross:

ch	b^+	cn	105
ch^+	b^+	cn^+	750
ch^+	b	cn	40
ch^+	b^+	cn	4
ch	b	cn	753
ch	b^+	cn^+	41
ch^+	b	cn^+	102
ch	b	cn^+	5
Total			1800

a. Determine the linear order of the genes on the chromosome (which gene is in the middle).

b. Calculate the recombinant distances between the three loci.

c. Determine the coefficient of coincidence and the interference for these three loci.

• Solution

a. We can represent the crosses in this problem as follows:

P $\quad \dfrac{ch^+\ b^+\ cn^+}{ch^+\ b^+\ cn^+} \times \dfrac{ch\ \ b\ \ cn}{ch\ \ b\ \ cn}$

$\downarrow$

$F_1 \quad \dfrac{ch^+\ b^+\ cn^+}{ch\ \ b\ \ cn}$

Testcross $\quad \dfrac{ch^+\ b^+\ cn^+}{ch\ \ b\ \ cn} \times \dfrac{ch\ \ b\ \ cn}{ch\ \ b\ \ cn}$

Note that we do not know, at this point, the order of the genes; we have arbitrarily put *b* in the middle.

The next step is to determine which of the testcross progeny are nonrecombinants and which are double crossovers. The nonrecombinants should be the most-frequent phenotype; so they must be the progeny with phenotypes encoded by $\frac{ch^+ \quad b^+ \quad cn^+}{}$ and $\frac{ch \quad b \quad cn}{}$. These genotypes are consistent with the genotypes of the parents, given earlier. The double crossovers are the least-frequent phenotypes and are encoded by $\frac{ch^+ \quad b^+ \quad cn}{}$ and $\frac{ch \quad b \quad cn^+}{}$.

We can determine the gene order by comparing the alleles present in the double crossovers with those present in the nonrecombinants. The double-crossover progeny should be like one of the nonrecombinants at two loci and unlike it at one locus; the allele that differs should be in the middle. Compare the double-crossover progeny $\frac{ch \quad b \quad cn^+}{}$ with the nonrecombinant $\frac{ch \quad b \quad cn}{}$. Both have cherub wings (*ch*) and black body (*b*), but the double-crossover progeny have wild-type eyes (cn^+), whereas the nonrecombinants have cinnabar eyes (*cn*). The locus that determines cinnabar eyes must be in the middle.

b. To calculate the recombination frequencies among the genes, we first write the phenotypes of the progeny with the genes encoding them in the correct order. We have already identified the nonrecombinant and double-crossover progeny; so the other four progeny types must have resulted from single crossovers. To determine *where* single crossovers took place, we compare the alleles found in the single-crossover progeny with those in the nonrecombinants. Crossing over must have taken place where the alleles switch from those found in one nonrecombinant to those found in the other nonrecombinant. The locations of the crossovers are indicated with a slash:

$ch \quad cn \;/\; b^+$	105	single crossover
$ch^+ \quad cn^+ \quad b^+$	750	nonrecombinant
$ch^+ \;/\; cn \quad b$	40	single crossover
$ch^+ \;/\; cn \;/\; b^+$	4	double crossover
$ch \quad cn \quad b$	753	nonrecombinant
$ch \;/\; cn^+ \quad b^+$	41	single crossover
$ch^+ \quad cn^+ \;/\; b$	102	single crossover
$ch \;/\; cn^+ \;/\; b$	5	double crossover
Total	1800	

Next, we determine the recombination frequencies and draw a genetic map:

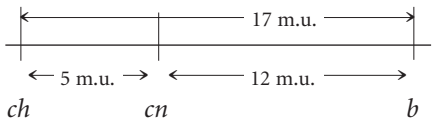
ch–cn recombination frequency =

$$\frac{40 + 4 + 41 + 5}{1800} \times 100\% = 5\%$$

cn–b recombination frequency =

$$\frac{105 + 4 + 102 + 5}{1800} \times 100\% = 12\%$$

ch–b map distance = 5% + 12% = 17%



c. The coefficient of coincidence is the number of observed double crossovers divided by the number of expected double crossovers. The number of expected double crossovers is obtained by multiplying the probability of a crossover between *ch* and *cn* (0.05) × the probability of a crossover between *cn* and *b* (0.12) × the total number of progeny in the cross (1800):

coefficient of coincidence =

$$\frac{4 + 5}{0.05 + 0.12 \times 1800} = 0.84$$

Finally, the interference is equal to 1 – the coefficient of coincidence:

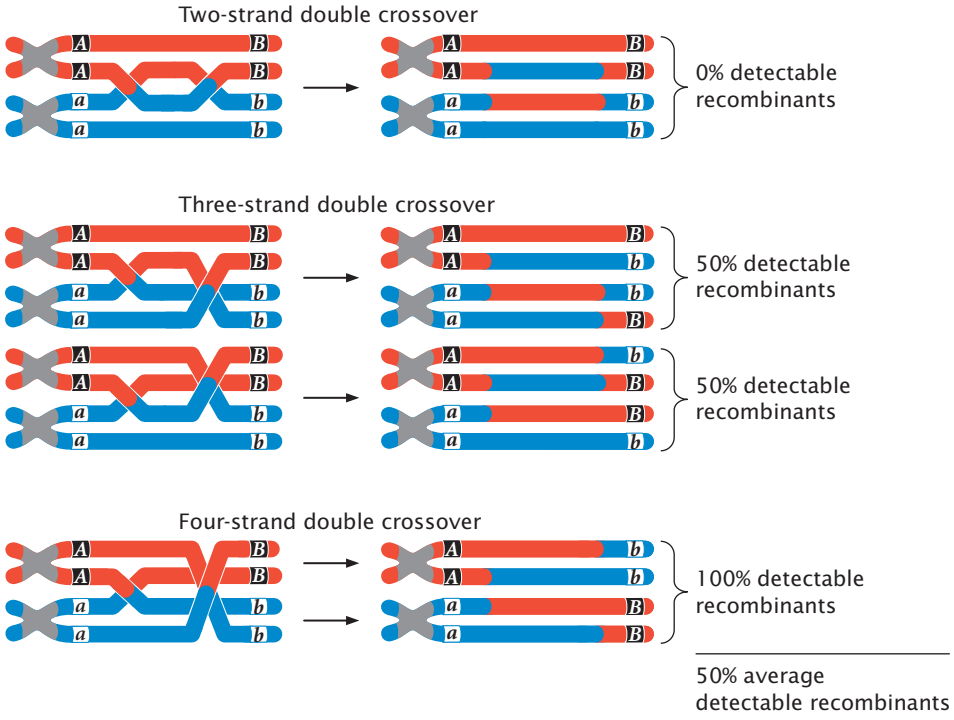
interference = 1 – 0.83 = 0.17

? To increase your skill with three-point crosses, try working Problem 18 at the end of this chapter.

Effect of Multiple Crossovers

So far, we have examined the effects of double crossovers taking place between only two of the four chromatids of a homologous pair. These crossovers are called two-strand crossovers. Double crossovers including three and even four of the chromatids of a homologous pair also may take place (**Figure 5.16**). If we examine only the alleles at loci on either side of both crossover events, two-strand double crossovers result in no new combinations of alleles, and no recombinant gametes are produced (see Figure 5.16). Three-strand double crossovers result in two of the four gametes being recombinant, and four-strand double crossovers result in all four gametes being recombinant. Thus, two-strand double crossovers produce 0% recombination, three-strand double crossovers produce 50% recombination, and four-strand double crossovers produce 100% recombination. The overall result is that all types of double crossovers, taken together, produce an average of 50% recombinant progeny.

As we have seen, two-strand double crossover cause alleles on either side of the crossovers to remain the same and produce no recombinant progeny. Three-strand and four-strand crossovers produce recombinant progeny, but these progeny are the same types produced by single crossovers. Consequently, some multiple crossovers go undetected when the progeny of a genetic cross are observed. Therefore, map distances based on recombination rates will underestimate



5.16 Results of two-, three-, and four-strand double crossovers on recombination between two genes.

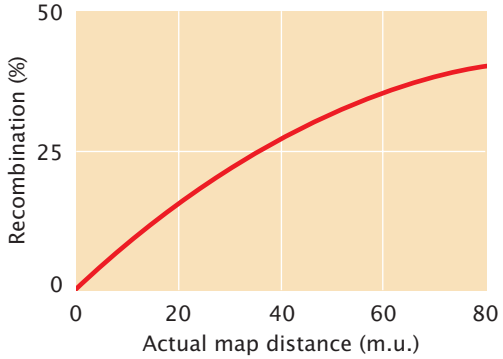
the true physical distances between genes, because some multiple crossovers are not detected among the progeny of a cross. When genes are very close together, multiple crossovers are unlikely, and the distances based on recombination rates accurately correspond to the physical distances on the chromosome. But, as the distance between genes increases, more multiple crossovers are likely, and the discrepancy between genetic distances (based on recombination rates) and physical distances increases. To correct for this discrepancy, geneticists have developed mathematical **mapping functions**, which relate recombination frequencies to actual physical distances between genes (Figure 5.17). Most of these functions are based on the Poisson distribution, which predicts the probability of multiple rare events. With the use of such mapping functions, map distances based on recombination rates can be more accurately estimated.

Mapping with Molecular Markers

For many years, gene mapping was limited in most organisms by the availability of **genetic markers**—that is, variable genes with easily observable phenotypes whose inheritance could be studied. Traditional genetic markers include genes that encode easily observable characteristics such as flower color, seed shape, blood types, and biochemical differences. The paucity of these types of characteristics in many organisms limited mapping efforts.

In the 1980s, new molecular techniques made it possible to examine variations in DNA itself, providing an almost unlimited number of genetic markers that can be used for creating genetic maps and studying linkage relations. The earliest of these molecular markers consisted of restriction fragment length polymorphisms (RFLPs), which are variations in DNA sequence detected by cutting the DNA with restriction enzymes (see Chapter 14). Later, methods were developed for detecting variable numbers of short DNA sequences repeated in tandem, called microsatellites. More recently, DNA sequencing allows the direct detection of individual variations in the DNA nucleotides, called single nucleotide polymorphisms (SNPs; see Chapter 14). All of these methods have expanded the availability of genetic markers and greatly facilitated the creation of genetic maps.

Gene mapping with molecular markers is done essentially in the same manner as mapping performed with traditional phenotypic markers: the cosegregation of two or more markers is studied, and map distances are based on the rates of recombination between markers. These methods and their use in mapping are presented in more detail in Chapter 14.



5.17 Percent recombination underestimates the true physical distance between genes at higher map distances.

Concepts Summary

- Linked genes do not assort independently. In a testcross for two completely linked genes (no crossing over), only nonrecombinant progeny are produced. When two genes assort independently, recombinant progeny and nonrecombinant progeny are produced in equal proportions. When two genes are linked with some crossing over between them, more nonrecombinant progeny than recombinant progeny are produced.
- Recombination frequency is calculated by summing the number of recombinant progeny, dividing by the total number of progeny produced in the cross, and multiplying by 100%. The recombination frequency is half the frequency of crossing over, and the maximum frequency of recombinant gametes is 50%.
- Coupling and repulsion refer to the arrangement of alleles on a chromosome. Whether genes are in coupling configuration or in repulsion determines which combination of phenotypes will be most frequent in the progeny of a testcross.
- Interchromosomal recombination takes place among genes located on different chromosomes through the random segregation of chromosomes in meiosis. Intrachromosomal recombination takes place among genes located on the same chromosome through crossing over.
- A chi-square test of independence can be used to determine if genes are linked.
- Recombination rates can be used to determine the relative order of genes and distances between them on a chromosome. One percent recombination equals one map unit. Maps based on recombination rates are called genetic maps; maps based on physical distances are called physical maps.
- Genetic maps can be constructed by examining recombination rates from a series of two-point crosses or by examining the progeny of a three-point testcross.
- Some multiple crossovers go undetected; thus, genetic maps based on recombination rates underestimate the true physical distances between genes.
- Molecular techniques that allow the detection of variable differences in DNA sequence have greatly facilitated gene mapping.

Important Terms

linked genes (p. 109)
linkage group (p. 109)
nonrecombinant (parental) gamete (p. 111)
nonrecombinant (parental) progeny (p. 111)
recombinant gamete (p. 111)
recombinant progeny (p. 111)
recombination frequency (p. 113)
coupling (cis) configuration (p. 114)
repulsion (trans) configuration (p. 114)
interchromosomal recombination (p. 116)
intrachromosomal recombination (p. 116)
genetic map (p. 119)
physical map (p. 119)
map unit (m.u.) (p. 119)
centiMorgan (p. 119)
Morgan (p. 119)
two-point testcross (p. 120)
three-point testcross (p. 121)
interference (p. 125)
coefficient of coincidence (p. 125)
mapping function (p. 129)
genetic marker (p. 129)

Answers to Concept Checks

1. c

2. 20%, in repulsion

3. Genetic maps are based on rates of recombination; physical maps are based on physical distances.

4. $\frac{m^+p^+s^+}{m\ p\ s}\ \frac{m^+p\ s}{m\ p\ s}\ \frac{m\ p^+s^+}{m\ p\ s}\ \frac{m^+p^+s}{m\ p\ s}\ \frac{m\ p\ s^+}{m\ p\ s}\ \frac{m^+p\ s^+}{m\ p\ s}\ \frac{m\ p^+s}{m\ p\ s}\ \frac{m\ p\ s}{m\ p\ s}$

5. The *c* locus

6. b

Worked Problems

1. In guinea pigs, white coat (*w*) is recessive to black coat (*W*) and wavy hair (*v*) is recessive to straight hair (*V*). A breeder crosses a guinea pig that is homozygous for white coat and wavy hair with a guinea pig that is black with straight hair. The F_1 are then crossed with guinea pigs having white coats and wavy hair in a series of testcrosses. The following progeny are produced from these testcrosses:

black, straight	30
black, wavy	10
white, straight	12
white, wavy	31
Total	83

a. Are the genes that determine coat color and hair type assorting independently? Carry out chi-square tests to test your hypothesis.

b. If the genes are not assorting independently, what is the recombination frequency between them?

• Solution

a. Assuming independent assortment, outline the crosses conducted by the breeder:

P $ww\ vv \times WW\ VV$

↓

F_1 $Ww\ Vv$

↓

Testcross $Ww\ Vv \times ww\ vv$

↓

$Ww\ Vv$	$\frac{1}{4}$ black, straight
$Ww\ vv$	$\frac{1}{4}$ black, wavy
$ww\ Vv$	$\frac{1}{4}$ white, straight
$ww\ vv$	$\frac{1}{4}$ white, wavy

Because a total of 83 progeny were produced in the testcrosses, we expect $\frac{1}{4} \times 83 = 20.75$ of each. The observed numbers of progeny from the testcross (30, 10, 12, 31) do not appear to fit the expected numbers (20.75, 20.75, 20.75, 20.75) well; so independent assortment may not have taken place.

To test the hypothesis, carry out a chi-square test of independence. Construct a table, with the genotypes of the first locus along the top and the genotypes of the second locus along the side. Compute the totals for the rows and columns and the grand total.

	Ww	*ww*	Row totals
Vv	30	12	42
vv	10	31	41
Column totals	40	43	83 ← Grand total

The expected value for each cell of the table is calculated with the formula:

$$\text{expected number} = \frac{\text{row total} \times \text{column total}}{\text{grand total}}$$

Using this formula, we find the expected values (given in parentheses) to be:

	Ww	*ww*	Row totals
Vv	30 (20.24)	12 (21.76)	42
vv	10 (19.76)	31 (21.24)	41
Column totals	40	43	83 ← Grand total

Using these observed and expected numbers, we find the calculated chi-square value to be:

$$\chi^2 = \sum \frac{(\text{observed} - \text{expected})^2}{\text{expected}}$$

$$= \frac{(30-20.24)^2}{20.24} + \frac{(10-19.76)^2}{19.76} + \frac{(12-21.76)^2}{21.76} + \frac{(31-21.24)^2}{21.24}$$

$$= 4.71 + 4.82 + 4.38 + 4.48$$

$$= 18.39$$

The degrees of freedom for the chi-square test of independence are $df =$ (number of rows $-$ 1) $\times$ (number of columns $-$ 1). There are two rows and two columns, so the degrees of freedom are:

$$df = (2 - 1) \times (2 - 1) = 1 \times 1 = 1$$

In Table 3.4, the probability associated with a chi-square value of 18.39 and 1 degree of freedom is less than 0.005, indicating that chance is very unlikely to be responsible for the differences between the observed numbers and the numbers expected with independent assortment. The genes for coat color and hair type have therefore not assorted independently.

b. To determine the recombination frequencies, identify the recombinant progeny. Using the notation for linked genes, write the crosses:

P $\dfrac{w\ \ v}{w\ \ v} \times \dfrac{W\ \ V}{W\ \ V}$

↓

F_1 $\dfrac{W\ \ V}{w\ \ v}$

Testcross $\dfrac{W\ \ V}{w\ \ v} \times \dfrac{w\ \ v}{w\ \ v}$

↓

$\frac{W\ \ V}{w\ \ v}$	30 black, straight (nonrecombinant progeny)
$\frac{w\ \ v}{w\ \ v}$	31 white, wavy (nonrecombinant progeny)
$\frac{W\ \ v}{w\ \ v}$	10 black, wavy (recombinant progeny)
$\frac{w\ \ V}{w\ \ v}$	12 white, straight (recombinant progeny)

The recombination frequency is:

$$\frac{\text{number of recombinant progeny}}{\text{total number progeny}} \times 100\%$$

or

$$\text{recombinant frequency} = \frac{10 + 12}{30 + 31 + 10 + 12} \times 100\%$$

$$= \frac{22}{83} \times 100\% = 26.5$$

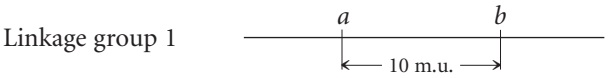
2. A series of two-point crosses were carried out among seven loci (*a*, *b*, *c*, *d*, *e*, *f*, and *g*), producing the following recombination frequencies. Using these recombination frequencies, map the seven loci, showing their linkage groups, the order of the loci in each linkage group, and the distances between the loci of each group:

Loci	Recombination frequency (%)	Loci	Recombination frequency (%)
a and *b*	10	*c* and *d*	50
a and *c*	50	*c* and *e*	8
a and *d*	14	*c* and *f*	50
a and *e*	50	*c* and *g*	12
a and *f*	50	*d* and *e*	50
a and *g*	50	*d* and *f*	50
b and *c*	50	*d* and *g*	50
b and *d*	4	*e* and *f*	50
b and *e*	50	*e* and *g*	18
b and *f*	50	*f* and *g*	50
b and *g*	50		

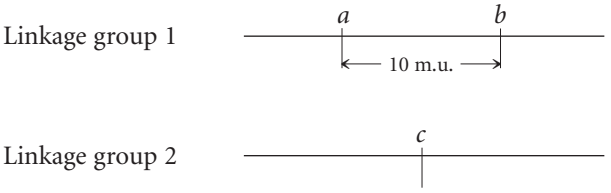
• Solution

To work this problem, remember that 1% recombination equals 1 map unit and a recombination frequency of 50% means that genes at the two loci are assorting independently (located in different linkage groups).

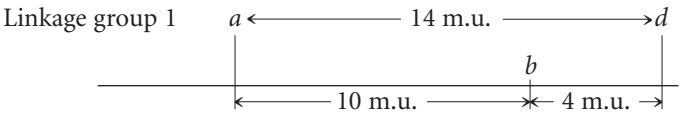
The recombination frequency between *a* and *b* is 10%; so these two loci are in the same linkage group, approximately 10 m.u. apart.



The recombination frequency between *a* and *c* is 50%; so *c* must lie in a second linkage group.

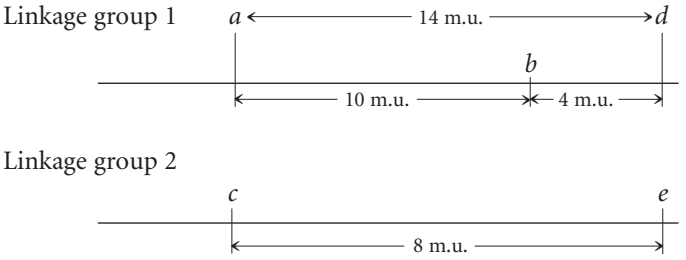


The recombination frequency between *a* and *d* is 14%; so *d* is located in linkage group 1. Is locus *d* 14 m.u. to the right or to the left of gene *a*? If *d* is 14 m.u. to the left of *a*, then the *b*-to-*d* distance should be 10 m.u. + 14 m.u. = 24 m.u. On the other hand, if *d* is to the right of *a*, then the distance between *b* and *d* should be 14 m.u. − 10 m.u. = 4 m.u. The *b*–*d* recombination frequency is 4%; so *d* is 14 m.u. to the right of *a*. The updated map is:

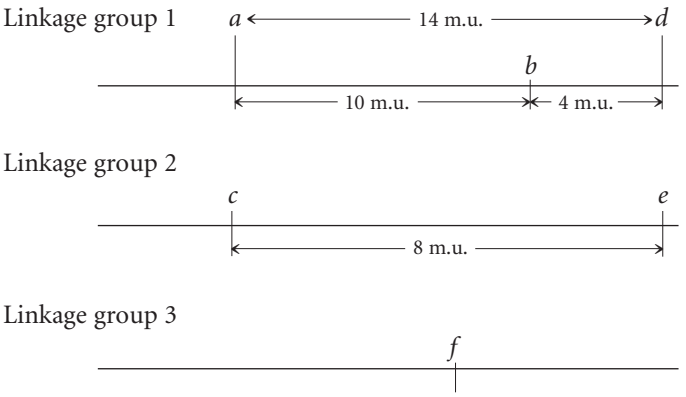


Linkage group 2

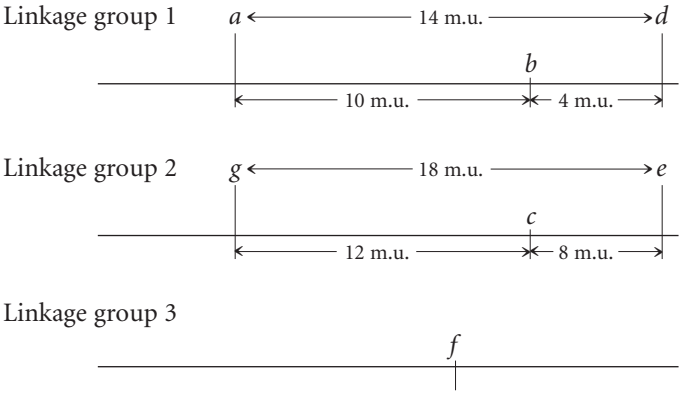
The recombination frequencies between each of loci *a*, *b*, and *d*, and locus *e* are all 50%; so *e* is not in linkage group 1 with *a*, *b*, and *d*. The recombination frequency between *e* and *c* is 8 m.u.; so *e* is in linkage group 2:



There is 50% recombination between *f* and all the other genes; so *f* must belong to a third linkage group:

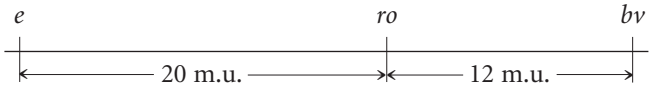


Finally, position locus *g* with respect to the other genes. The recombination frequencies between *g* and loci *a*, *b*, and *d* are all 50%; so *g* is not in linkage group 1. The recombination frequency between *g* and *c* is 12 m.u.; so *g* is a part of linkage group 2. To determine whether *g* is 12 m.u. to the right or left of *c*, consult the *g*–*e* recombination frequency. Because this recombination frequency is 18%, *g* must lie to the left of *c:*



Note that the *g*-to-*e* distance (18 m.u.) is shorter than the sum of the *g*-to-*c* (12 m.u.) and *c*-to-*e* distances (8 m.u.), because of undetectable double crossovers between *g* and *e*.

3. Ebony body color (*e*), rough eyes (*ro*), and brevis bristles (*bv*) are three recessive mutations that occur in fruit flies. The loci for these mutations have been mapped and are separated by the following map distances:



The interference between these genes is 0.4.

A fly with ebony body, rough eyes, and brevis bristles is crossed with a fly that is homozygous for the wild-type traits. The resulting F_1 females are test-crossed with males that have ebony body, rough eyes, and brevis bristles; 1800 progeny are produced. Give the expected numbers of phenotypes in the progeny of the testcross.

• Solution

The crosses are:

P $\quad \dfrac{e^+ \quad ro^+ \quad bv^+}{e^+ \quad ro^+ \quad bv^+} \times \dfrac{e \quad ro \quad bv}{e \quad ro \quad bv}$

$\downarrow$

$F_1 \quad \dfrac{e^+ \quad ro^+ \quad bv^+}{e \quad ro \quad bv}$

$\downarrow$

Testcross $\quad \dfrac{e^+ \quad ro^+ \quad bv^+}{e \quad ro \quad bv} \times \dfrac{e \quad ro \quad bv}{e \quad ro \quad bv}$

In this case, we know that *ro* is the middle locus because the genes have been mapped. Eight classes of progeny will be produced from this cross:

$e^+ \quad ro^+ \quad bv^+$	nonrecombinant
$e \quad ro \quad bv$	nonrecombinant
$e^+ \;/\; ro \quad bv$	single crossover between *e* and *ro*
$e \;/\; ro^+ \quad bv^+$	single crossover between *e* and *ro*
$e^+ \quad ro^+ \;/\; bv$	single crossover between *ro* and *bv*
$e \quad ro \;/\; bv^+$	single crossover between *ro* and *bv*
$e^+ \;/\; ro \;/\; bv^+$	double crossover
$e \;/\; ro^+ \;/\; bv$	double crossover

To determine the numbers of each type, use the map distances, starting with the double crossovers. The expected number of double crossovers is equal to the product of the single-crossover probabilities:

$$\text{expected number of double crossovers} = 0.20 \times 0.12 \times 1800 = 43.2$$

However, some interference occurs; so the observed number of double crossovers will be less than the expected. The interference is 1 – coefficient of coincidence; so the coefficient of coincidence is:

$$\text{coefficient of coincidence} = 1 - \text{interference}$$

The interference is given as 0.4; so the coefficient of coincidence equals 1 – 0.4 = 0.6. Recall that the coefficient of coincidence is:

$$\text{coefficient of coincidence} = \frac{\text{number of observed double crossovers}}{\text{number of expected double crossovers}}$$

Rearranging this equation, we obtain:

$$\text{number of observed double crossovers} = \text{coefficient of coincidence} \times \text{number of expected double crossover}$$

$$\text{number of observed double crossovers} = 0.6 \times 43.2 = 26$$

A total of 26 double crossovers should be observed. Because there are two classes of double crossovers ($e^+ \;/\; ro \;/\; bv^+$ and $e \;/\; ro^+ \;/\; bv$), we expect to observe 13 of each.

Next, we determine the number of single-crossover progeny. The genetic map indicates that the distance between *e* and *ro* is 20 m.u.; so 360 progeny (20% of 1800) are expected to have resulted from recombination between these two loci. Some of them will be single-crossover progeny and some will be double-crossover progeny. We have already determined that the number of double-crossover progeny is 26; so the number of progeny resulting from a single crossover between *e* and *ro* is 360 − 26 = 334, which will be divided equally between the two single-crossover phenotypes ($e \;/\; ro^+ \;/\; bv^+$ and $e^+ \;/\; ro \;/\; bv$).

The distance between *ro* and *bv* is 12 m.u.; so the number of progeny resulting from recombination between these two genes is 0.12 × 1800 = 216. Again, some of these recombinants will be single-crossover progeny and some will be double-crossover progeny. To determine the number of progeny resulting from a single crossover, subtract the double crossovers: 216 − 26 = 190. These single-crossover progeny will be divided between the two single-crossover phenotypes ($e^+ \;/\; ro^+ \;/\; bv$ and $e \;/\; ro \;/\; bv^+$); so there will be $^{190}/_{2} = 95$ of each of these phenotypes. The remaining progeny will be nonrecombinants, and they can be obtained by subtraction: 1800 − 26 − 334 − 190 = 1250; there are two nonrecombinants ($e^+ \quad ro \quad bv^+$ and $e \quad ro \quad bv$); so there will be $^{1250}/_{2} = 625$ of each. The numbers of the various phenotypes are listed here:

$e^+ \quad ro^+ \quad bv^+$	625	nonrecombinant
$e \quad ro \quad bv$	625	nonrecombinant
$e^+ \;/\; ro \quad bv$	167	single crossover between *e* and *ro*
$e \;/\; ro^+ \quad bv^+$	167	single crossover between *e* and *ro*
$e^+ \quad ro^+ \;/\; bv$	95	single crossover between *ro* and *bv*
$e \quad ro \;/\; bv^+$	95	single crossover between *ro* and *bv*
$e^+ \;/\; ro \;/\; bv^+$	13	double crossover
$e \;/\; ro^+ \;/\; bv$	13	double crossover
Total	1800	

Comprehension Questions

Section .5.1

*1. What does the term recombination mean? What are two causes of recombination?

Section 5.2

2. What effect does crossing over have on linkage?

3. Why is the frequency of recombinant gametes always half the frequency of crossing over?

*4. What is the difference between genes in coupling configuration and genes in repulsion? What effect does the arrangement of linked genes (whether they are in coupling configuration or in repulsion) have on the results of a cross?

5. What is the difference between a genetic map and a physical map?

Section 5.3

6. Explain how to determine which of three linked loci is the middle locus from the progeny of a three-point testcross.

*7. What does the interference tell us about the effect of one crossover on another?

Application Questions and Problems

Section 5.2

*8. In the snail *Cepaea nemoralis*, an autosomal allele causing a banded shell (B^B) is recessive to the allele for an unbanded shell (B^O). Genes at a different locus determine the background color of the shell; here, yellow (C^Y) is recessive to brown (C^{Bw}). A banded, yellow snail is crossed with a homozygous brown, unbanded snail. The F_1 are then crossed with banded, yellow snails (a testcross).

a. What will be the results of the testcross if the loci that control banding and color are linked with no crossing over?

b. What will be the results of the testcross if the loci assort independently?

c. What will be the results of the testcross if the loci are linked and 20 m.u. apart?

*9. A geneticist discovers a new mutation in *Drosophila melanogaster* that causes the flies to shake and quiver. She calls this mutation spastic (*sps*) and determines that spastic is due to an autosomal recessive gene. She wants to determine if the gene encoding spastic is linked to the recessive gene for vestigial wings (*vg*). She crosses a fly homozygous for spastic and vestigial traits with a fly homozygous for the wild-type traits and then uses the resulting F_1 females in a testcross. She obtains the following flies from this testcross.

vg^+	sps^+	230
vg	sps	224
vg	sps^+	97
vg^+	sps	99
Total		650

Are the genes that cause vestigial wings and the spastic mutation linked? Do a chi-square test of independence to determine if the genes have assorted independently.

10. In cucumbers, heart-shaped leaves (*hl*) are recessive to normal leaves (*Hl*) and having numerous fruit spines (*ns*) is recessive to having few fruit spines (*Ns*). The genes for leaf shape and for number of spines are located on the same chromosome; findings from mapping experiments indicate that they are 32.6 m.u. apart. A cucumber plant having heart-shaped leaves and numerous spines is crossed with a plant that is homozygous for normal leaves and few spines. The F_1 are crossed with plants that have heart-shaped leaves and numerous spines. What phenotypes and proportions are expected in the progeny of this cross?

*11. In tomatoes, tall (*D*) is dominant over dwarf (*d*) and smooth fruit (*P*) is dominant over pubescent fruit (*p*), which is covered with fine hairs. A farmer has two tall and smooth tomato plants, which we will call plant A and plant B. The farmer crosses plants A and B with the same dwarf and pubescent plant and obtains the following numbers of progeny:

	Progeny of	
	Plant A	**Plant B**
Dd Pp	122	2
Dd pp	6	82
dd Pp	4	82
dd pp	124	4

a. What are the genotypes of plant A and plant B?

b. Are the loci that determine height of the plant and pubescence linked? If so, what is the map distance between them?

c. Explain why different proportions of progeny are produced when plant A and plant B are crossed with the same dwarf pubescent plant.

12. DATA ANALYSIS Daniel McDonald and Nancy Peer determined that eyespot (a clear spot in the center of the eye) in flour beetles is

caused by an X-linked gene (*es*) that is recessive to the allele for the absence of eyespot (es^+). They conducted a series of crosses to determine the distance between the gene for eyespot and a dominant X-linked gene for stripped (*St*), which acted as a recessive lethal (is lethal when homozygous in females or hemizygous in males). The following cross was carried out (D. J. McDonald and N. J. Peer. 1961. *Journal of Heredity* 52:261–264).

$$\text{♀}\ \frac{es^+\ St}{es\ St^+} \times \frac{es\ St^+}{\text{Y}}\ \text{♂}$$

$$\downarrow$$

$\frac{es^+\ St}{es\ St^+}$	1630
$\frac{es\ St^+}{es\ St^+}$	1665
$\frac{es\ St^+}{es\ St^+}$	935
$\frac{es^+\ St^+}{es\ St^+}$	1005
$\frac{es\ St^+}{\text{Y}}$	1661
$\frac{es^+\ St^+}{\text{Y}}$	1024

a. Which progeny are the recombinants and which progeny are the nonrecombinants?

b. Calculate the recombination frequency between *es* and *St*.

c. Are some potential genotypes missing among the progeny of the cross? If so, which ones and why?

13. In German cockroaches, bulging eyes (*bu*) are recessive to normal eyes (bu^+) and curved wings (*cv*) are recessive to straight wings (cv^+). Both traits are encoded by autosomal genes that are linked. A cockroach has genotype $bu^+bu\ cv^+\ cv$ and the genes are in repulsion. Which of the following sets of genes will be found in the most-common gametes produced by this cockroach?

a. $bu^+\ cv^+$

b. $bu\ cv$

c. $bu^+\ bu$

d. $cv^+\ cv$

e. $bu\ cv^+$

Explain your answer.

***14.** In *Drosophila melanogaster*, ebony body (*e*) and rough eyes (*ro*) are encoded by autosomal recessive genes found on chromosome 3; they are separated by 20 m.u. The gene that encodes forked bristles (*f*) is X-linked recessive and assorts independently of *e* and *ro*. Give the phenotypes of progeny and their expected proportions when a female of each of the following genotypes is test-crossed with a male.

a. $\frac{e^+\ ro^+}{e\ \ ro}\ \frac{f^+}{f}$

b. $\frac{e^+\ ro}{e\ \ ro^+}\ \frac{f^+}{f}$

***15.** A series of two-point crosses were carried out among seven loci (*a*, *b*, *c*, *d*, *e*, *f*, and *g*), producing the following recombination frequencies. Map the seven loci, showing their linkage groups, the order of the loci in each linkage group, and the distances between the loci of each group.

Loci	Percent recombination	Loci	Percent recombination
a and *b*	50	*c* and *d*	50
a and *c*	50	*c* and *e*	26
a and *d*	12	*c* and *f*	50
a and *e*	50	*c* and *g*	50
a and *f*	50	*d* and *e*	50
a and *g*	4	*d* and *f*	50
b and *c*	10	*d* and *g*	8
b and *d*	50	*e* and *f*	50
b and *e*	18	*e* and *g*	50
b and *f*	50	*f* and *g*	50
b and *g*	50		

16. DATA ANALYSIS R. W. Allard and W. M. Clement determined recombination rates for a series of genes in lima beans (R. W. Allard and W. M. Clement. 1959. *Journal of Heredity* 50:63−67). The following table lists paired recombination rates for eight of the loci (*D*, *Wl*, *R*, *S*, L_1, *Ms*, *C*, and *G*) that they mapped. On the basis of these data, draw a series of genetic maps for the different linkage groups of the genes, indicating the distances between the genes. Keep in mind that these rates are estimates of the true recombination rates and that some error is associated with each estimate. An asterisk beside a recombination frequency indicates that the recombination frequency is significantly different from 50%.

Recombination Rates (%) among Seven Loci in Lima Beans

	Wl	*R*	*S*	L_1	*Ms*	*C*	*G*
D	2.1*	39.3*	52.4	48.1	53.1	51.4	49.8
Wl		38.0*	47.3	47.7	48.8	50.3	50.4
R			51.9	52.7	54.6	49.3	52.6
S				26.9*	54.9	52.0	48.0
L_1					48.2	45.3	50.4
Ms						14.7*	43.1
C							52.0

*Significantly different from 50%.

Section 5.3

17. DATA ANALYSIS Raymond Popp studied linkage among genes for pink eye (*p*), shaker-1 (*sh*-1), and hemoglobin (*Hb*) in mice (R. A. Popp. 1962. *Journal of Heredity* 53:73–80). He performed a series of test crosses, in which mice heterozygous for pink eye, shaker-1, and hemoglobin 1 and 2 were crossed with

mice that were homozygous for pink eye, shaker-1 and hemoglobin 2.

$$\frac{P\ Sh{-}1\ Hb^1}{p\ sh{-}1\ Hb^2} \times \frac{p\ sh{-}1\ Hb^2}{p\ sh{-}1\ Hb^2}$$

The following progeny were produced.

Progeny genotype	Number
$\frac{p\ sh{-}1\ Hb^2}{p\ sh{-}1\ Hb^2}$	274
$\frac{P\ Sh{-}1\ Hb^1}{p\ sh{-}1\ Hb^2}$	320
$\frac{P\ sh{-}1\ Hb^2}{p\ sh{-}1\ Hb^2}$	57
$\frac{p\ Sh{-}1\ Hb^1}{p\ sh{-}1\ Hb^2}$	45
$\frac{p\ Sh{-}1\ Hb^2}{p\ sh{-}1\ Hb^2}$	6
$\frac{p\ sh{-}1\ Hb^1}{p\ sh{-}1\ Hb^2}$	5
$\frac{p\ Sh{-}1\ Hb^2}{p\ sh{-}1\ Hb^2}$	0
$\frac{P\ sh{-}1\ Hb^1}{p\ sh{-}1\ Hb^2}$	1
Total	708

a. Determine the order of these genes on the chromosome.

b. Calculate the map distances between the genes.

c. Determine the coefficient of coincidence and the interference among these genes.

***18.** Waxy endosperm (*wx*), shrunken endosperm (*sh*), and yellow seedling (*v*) are encoded by three recessive genes in corn that are linked on chromosome 5. A corn plant homozygous for all three recessive alleles is crossed with a plant homozygous for all the dominant alleles. The resulting F_1 are then crossed with a plant homozygous for the recessive alleles in a three-point testcross. The progeny of the testcross are:

wx	*sh*	*V*	87
Wx	*Sh*	*v*	94
Wx	*Sh*	*V*	3,479
wx	*sh*	*v*	3,478
Wx	*sh*	*V*	1,515
wx	*Sh*	*v*	1,531
wx	*Sh*	*V*	292
Wx	*sh*	*v*	280
Total			10,756

a. Determine the order of these genes on the chromosome.

b. Calculate the map distances between the genes.

c. Determine the coefficient of coincidence and the interference among these genes.

19. DATA ANALYSIS Priscilla Lane and Margaret Green studied the linkage relations of three genes affecting coat color in mice: mahogany (*mg*), agouti (*a*), and ragged (*Ra*). They carried out a series of three-point crosses, mating mice that were heterozygous at all three loci with mice that were homozygous for the recessive alleles at these loci (P W. Lane and M. C. Green. 1960. *Journal of Heredity* 51:228–230). The following table lists the results of the test crosses.

Phenotype			Number
a	*Rg*	+	1
+	+	*mg*	1
a	+	+	15
+	*Rg*	*mg*	9
+	+	+	16
a	*Ra*	*mg*	36
a	+	*mg*	76
+	*Ra*	+	69
Total			213

a. Determine the order of the loci that encode mahogany, agouti, and ragged on the chromosome, the map distances between them, and the interference and coefficient of coincidence for these genes.

b. Draw a picture of the two chromosomes in the triply heterozygous mice used in the testcrosses, indicating which of the alleles are present on each chromosome.

***20.** In *Drosophila melanogaster*, black body (*b*) is recessive to gray body (b^+), purple eyes (*pr*) are recessive to red eyes (pr^+), and vestigial wings (*vg*) are recessive to normal wings (vg^+). The loci encoding these traits are linked, with the following map distances:

The interference among these genes is 0.5. A fly with a black body, purple eyes, and vestigial wings is crossed with a fly homozygous for a gray body, red eyes, and normal wings. The female progeny are then crossed with males that have a black body, purple eyes, and vestigial wings. If 1000 progeny are produced from this testcross, what will be the phenotypes and proportions of the progeny?

Challenge Question

Section 5.3

21. DATA ANALYSIS Transferrin is a blood protein that is encoded by the transferrin locus (*Trf*). In house mice the two alleles at this locus (Trf^a and Trf^b) are codominant and encode three types of transferrin:

Genotype	Phenotype
Trf^a/Trf^a	Trf-a
Trf^a/Trf^b	Trf-ab
Trf^b/Trf^b	Trf-b

The dilution locus, found on the same chromosome, determines whether the color of a mouse is diluted or full; an allele for dilution (*d*) is recessive to an allele for full color (d^+):

Genotype	Phenotype
d^+d^+	d^+ (full color)
d^+d	d^+ (full color)
dd	d (dilution)

Donald Shreffler conducted a series of crosses to determine the map distance between the tranferrin locus and the dilution locus (D. C. Shreffler. 1963 *Journal of Heredity* 54:127–129). The table at right presents a series of crosses carried out by Shreffler and the progeny resulting from these crosses.

a. Calculate the recombinant frequency between the *Trf* and the *d* loci by using the pooled data from all the crosses.

b. Which crosses represent recombination in male gamete formation and which crosses represent recombination in female gamete formation?

c. On the basis of your answer to part *b*, calculate the frequency of recombination among male parents and female parents separately.

d. Are the rates of recombination in males and females the same? If not, what might produce the difference?

			Progeny phenotypes				
Cross	♂	♀	d^+ Trf-ab	d^+ Trf-b	d Trf-ab	d Trf-b	Total
1	$\frac{d^+\ Trf^a}{d\ Trf^b}$	$\frac{d\ Trf^b}{d\ Trf^b}$	32	3	6	21	62
2	$\frac{d\ Trf^b}{d\ Trf^b}$	$\frac{d^+\ Trf^a}{d\ Trf^b}$	16	0	2	20	38
3	$\frac{d^+\ Trf^a}{d\ Trf^b}$	$\frac{d\ Trf^b}{d\ Trf^b}$	35	9	4	30	78
4	$\frac{d\ Trf^b}{d\ Trf^b}$	$\frac{d^+\ Trf^a}{d\ Trf^b}$	21	3	2	19	45
5	$\frac{d^+\ Trf\ ^b}{d\ Trf^{\ a}}$	$\frac{d\ Trf\ ^b}{d\ Trf^{\ b}}$	8	29	22	5	64
6	$\frac{d\ Trf^b}{d\ Trf^b}$	$\frac{d^+\ Trf^b}{d\ Trf^a}$	4	14	11	0	29