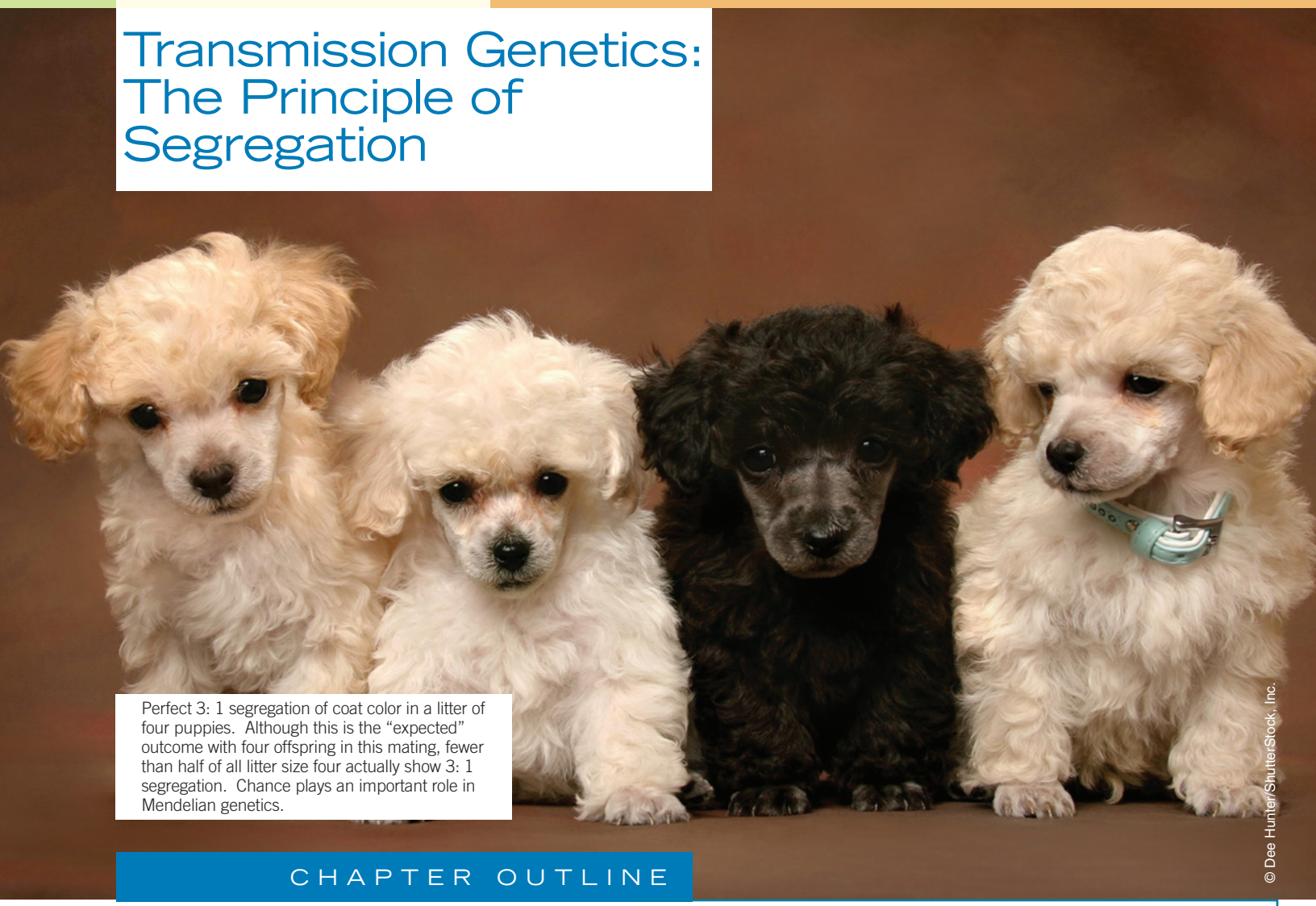
CHAPTER 3

Transmission Genetics: The Principle of Segregation



Perfect 3: 1 segregation of coat color in a litter of four puppies. Although this is the "expected" outcome with four offspring in this mating, fewer than half of all litter size four actually show 3: 1 segregation. Chance plays an important role in Mendelian genetics.

CHAPTER OUTLINE

In this chapter we consider how genes are transmitted from parents to offspring and how this determines the distribution of genotypes and phenotypes among related individuals. The study of the inheritance of traits constitutes **transmission genetics.** This subject is also called **Mendelian genetics** because the underlying principles were first deduced from experiments in garden peas (*Pisum sativum*) carried out in the years 1856–1863 by Gregor Mendel, a monk at the monastery of St. Thomas in the town of Brno (Brünn), in the Czech Republic. He reported his experiments to a local natural history society, published the results and his interpretation in its scientific journal in 1866, and began exchanging letters with one of the leading botanists of the time. His experiments were careful and exceptionally well documented, and his paper contains the first clear exposition of the statistical rules governing the transmission of genes from generation to generation. Nevertheless, Mendel's paper was ignored for 34 years until its significance was finally appreciated.

3.1 Morphological and Molecular Phenotypes

Until the advent of molecular genetics, geneticists dealt mainly with morphological traits, in which the differences between organisms can be expressed in terms of color, shape, or size. Mendel studied seven morphological traits contrasting in seed shape, seed color, flower color, pod shape, and so forth (**FIGURE 3.1**). Perhaps the most widely known example of a contrasting Mendelian trait is round versus wrinkled seeds. As pea seeds dry, they lose water and shrink. Round seeds are round because they shrink uniformly; wrinkled seeds are wrinkled because they shrink irregularly. The wrinkled phenotype is due to the absence of a branched-chain form of starch known as *amylopectin,* which is not synthesized in wrinkled seeds owing to a defect in the enzyme starch-branching enzyme I (SBEI).

The nonmutant, or **wildtype,** allele of the gene for SBEI is designated *W* and the mutant allele *w.* Seeds that are heterozygous *Ww* have only half as much SBEI enzyme as wildtype homozygous *WW* seeds, but this half the normal amount of enzyme produces enough amylopectin for the heterozygous *Ww* seeds to shrink uniformly and remain phenotypically round. Hence, with respect to seed shape, genotypes *WW* and *Ww* have the same phenotype (round). The *W* allele is called the **dominant allele** and *w* is called the **recessive allele.**

The molecular basis of the wrinkled mutation is that the *SBEI* gene has become interrupted by the insertion, into the gene, of a DNA

IN THIS SMALL garden plot adjacent to the monastery of St. Thomas, Gregor Mendel grew more than 33,500 pea plants over a period of eight years, including more than 6400 plants in one year alone. He received some help from two fellow monks who assisted in the experiments.

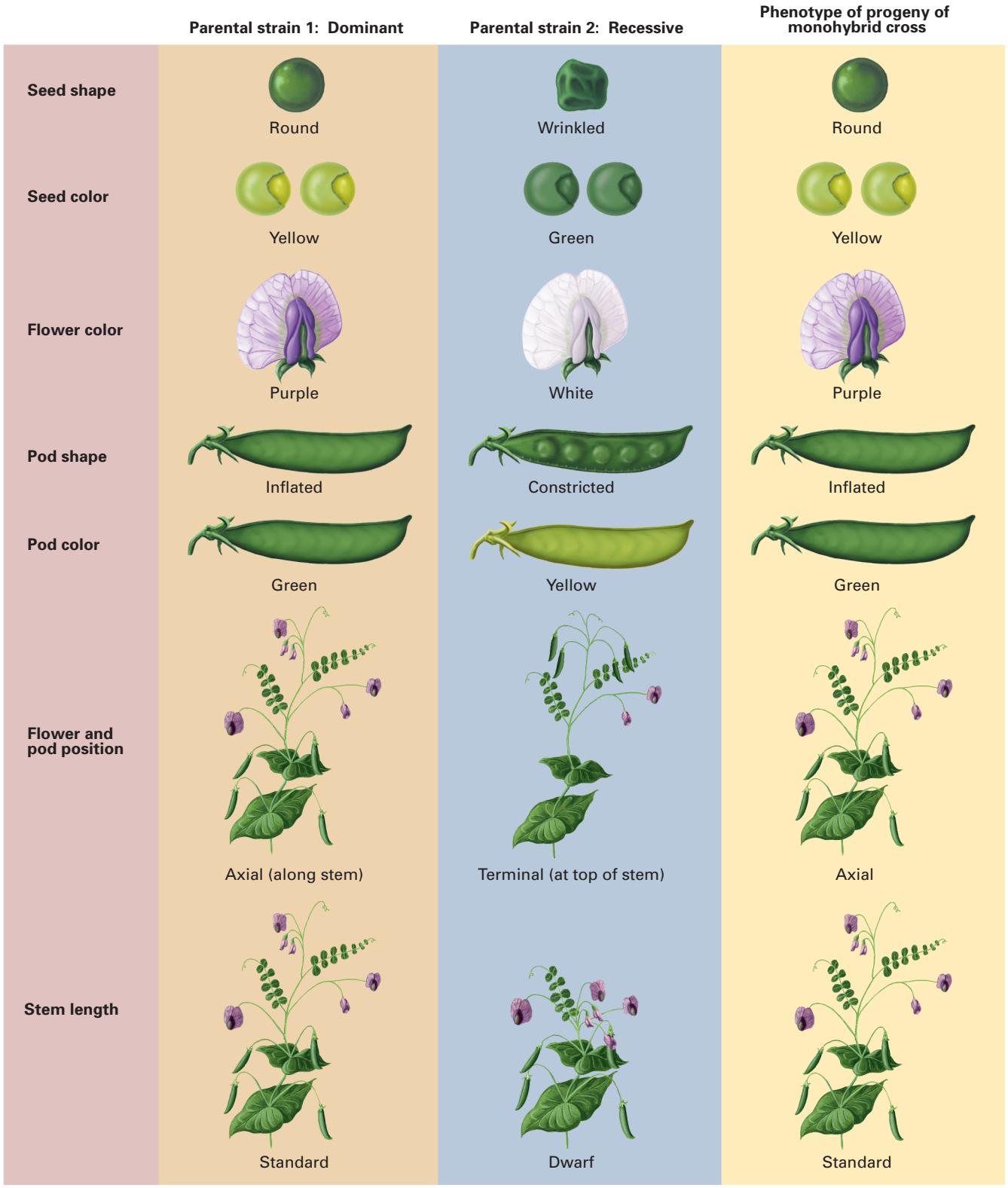


FIGURE 3.1 The seven different traits in peas studied by Mendel. The phenotype shown at the far right is the dominant trait, which appears in the hybrid produced by crossing.

sequence called a **transposable element.** Such DNA sequences are capable of moving (*transposition*) from one location to another within a chromosome or between chromosomes. The molecular mechanism of transposition is discussed in Chapter 14, but for our present purposes it is necessary to know only that transposable elements are present in most genomes, especially the large genomes of eukaryotes, and that many spontaneous mutations result from the insertion of transposable elements into a gene.

FIGURE 3.2 includes a diagram of the DNA structure of the wildtype *W* and mutant *w* alleles and shows the DNA insertion that interrupts the *w* allele. Highlighted are two *Eco*RI restriction sites, present in both alleles, that flank the site of the insertion. The diagram in part C indicates the pattern of bands that would be expected if one were to carry out a Southern blot (Section 2.3) in which genomic DNA was digested to completion with *Eco*RI and then the resulting fragments were separated by electrophoresis and hybridized with a labeled probe complementary to a region shared between the *W* and *w* alleles. The *Eco*RI fragment from the *W* allele would be smaller than that of the *w* allele, because of the inserted DNA in the *w* allele, and thus it would migrate faster than the corresponding fragment from the *w* allele and move to a position closer to the bottom of the gel. Genomic DNA from homozygous *WW* would yield a single, faster-migrating band; that from homozygous *ww* a single, slower-migrating band; and that from heterozygous *Ww* two bands with the same electrophoretic mobilities as those observed in the homozygous genotypes. In Figure 3.2C, the band in the homozygous genotypes is shown as somewhat thicker than those from the heterozygous genotype, because the single band in each homozygous genotype comes from the two copies of whichever allele is homozygous, and thus it contains more DNA than the corresponding DNA in the heterozygous genotype, in which only one copy of each allele is present.

Hence, as illustrated in Figure 3.2C, the RFLP analysis clearly distinguishes between the genotypes *WW, Ww,* and *ww,* because the heterozygous *Ww* genotype exhibits both of the bands observed in the homozygous genotypes. This situation is described by saying that the *W* and *w* alleles are **codominant** with respect to the molecular phenotype. However, as indicated by the seed shapes in Figure 3.2C, *W* is dominant over *w* with respect to morphological phenotype. In the discussion that follows, we use the RFLP analysis of the *W* and *w* alleles to emphasize the importance of molecular phenotypes in modern genetics and to demonstrate experimentally the principles of genetic transmission.

3.2 Segregation of a Single Gene

Mendel selected peas for his experiments for two primary reasons. First, he had access to varieties that differed in contrasting traits, such as round versus wrinkled seeds and yellow versus green seeds. Second, his earlier studies had indicated that peas usually reproduce by self-pollination, in which pollen produced in a flower is used to fertilize the eggs in the same flower. Producing hybrids by cross-pollination (**outcrossing**), re-

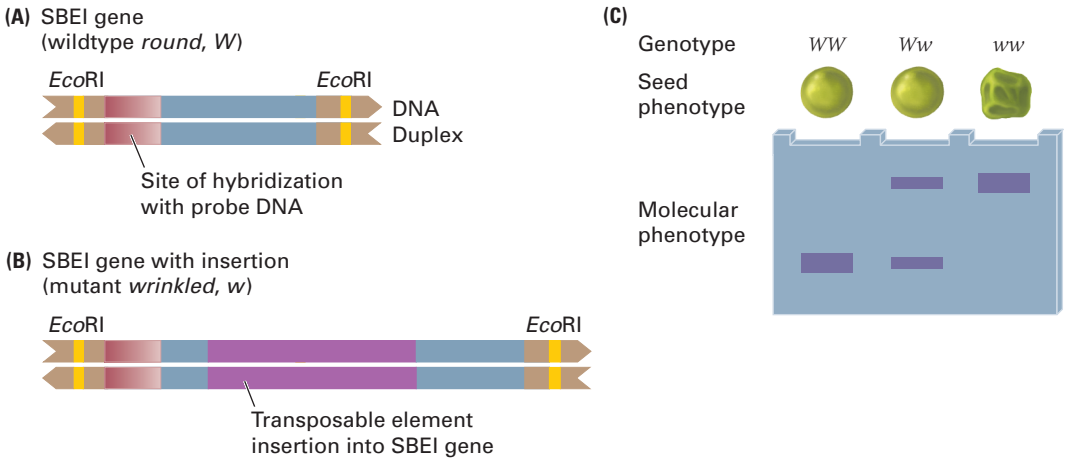


FIGURE 3.2 (A) *W* (round) is an allele of a gene that specifies the amino acid sequence of starch branching enzyme I (SBEI). (B) *w* (wrinkled) is an allele that encodes an inactive form of the enzyme because its DNA sequence is interrupted by the insertion of a transposable element. (C) At the level of the morphological phenotype, *W* is dominant to *w:* Genotypes *WW* and *Ww* have round seeds, whereas genotype *ww* has wrinkled seeds. The molecular difference between the alleles can be detected as a restriction fragment length polymorphism (RFLP) using the enzyme *Eco*RI and a probe that hybridizes at the site shown. At the molecular level, the alleles are codominant: DNA from each genotype yields a different molecular phenotype—a single band differing in size for homozygous *WW* and *ww*, and both bands for heterozygous *Ww*.

quired painstaking surgery on immature flowers to reveal the receptive female structures, excise and discard the undeveloped male pollen-producing structures, and expose the female structures to mature pollen from a a different plant. After the cross-fertilization, each dissected flower was enclosed in a fine mesh bag to prevent stray pollen from accessing the female structures inside. The relatively small space needed to grow each plant, and the relatively large number of progeny that could be obtained, gave him the opportunity, as he says in his paper, to "determine the number of different forms in which hybrid progeny appear" and to "ascertain their numerical interrelationships."

The fact that garden peas are normally self-fertilizing means that in the absence of deliberate outcrossing, plants with contrasting traits are usually homozygous for alternative alleles of a gene affecting the trait; for example, plants with round seeds have genotype *WW* and those with wrinkled seeds genotype *ww*. The homozygous genotypes are indicated experimentally by the observation that hereditary traits in each variety are **true-breeding**, which means that plants produce only progeny like themselves when allowed to self-pollinate normally. Outcrossing between plants that differ in one or more traits creates a **hybrid**. If the parents differ in one, two, or three traits, the hybrid is a *monohybrid, dihybrid,* or *trihybrid.* In keeping track of parents and their hybrid progeny, we say that the parents constitute the **P_1 generation** and their hybrid progeny the **F_1 generation**.

Because garden peas are sexual organisms, each cross can be performed in two ways, depending on which phenotype is present in the female parent and which in the male parent. With round versus wrinkled, for example, the female parent can be round and the male wrinkled, or the reverse. These are called **reciprocal crosses**. Mendel was the first to demonstrate the following important principle:

> The outcome of a genetic cross does not depend on which trait is present in the male and which is present in the female; reciprocal crosses yield the same result.

This principle is illustrated for round and wrinkled seeds in **FIGURE 3.3**. The gel icons show the

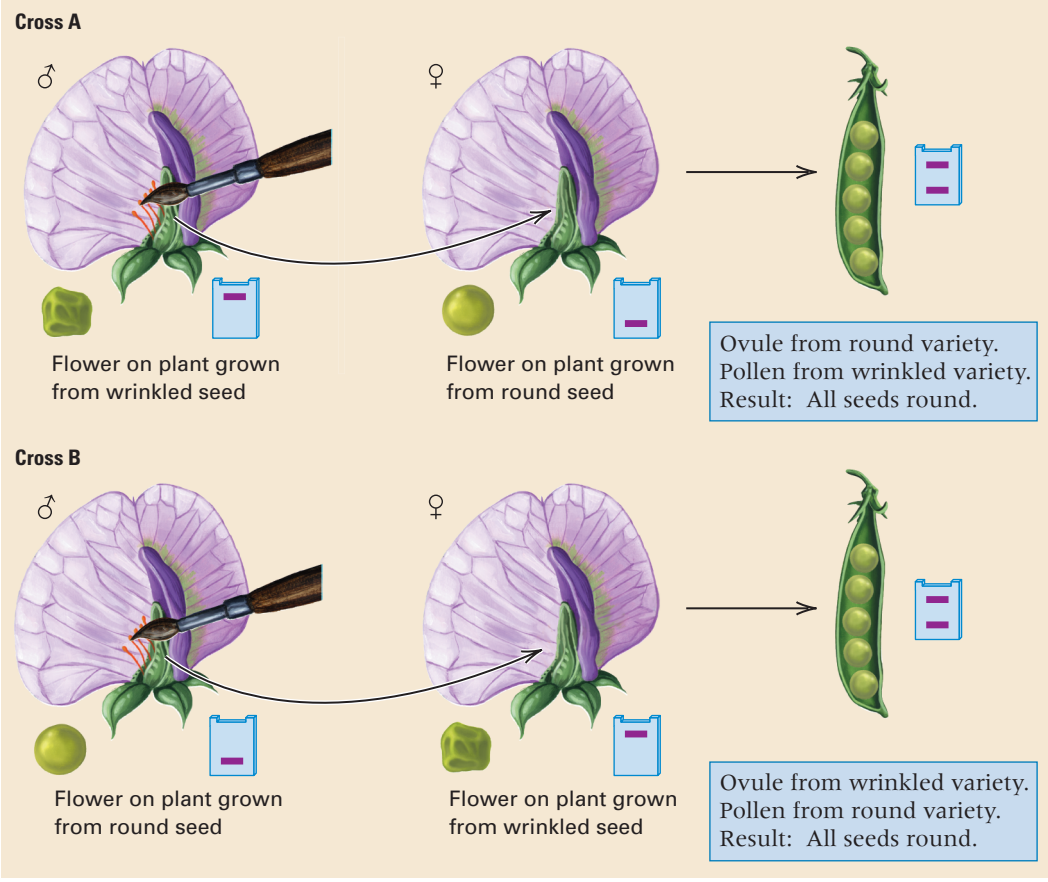


FIGURE 3.3 Morphological and molecular phenotypes showing the equivalence of reciprocal crosses. In this example, the hybrid seeds are round and yield an RFLP pattern with two bands, irrespective of the direction of the cross.

RFLP bands that genomic DNA from each type of seed in these crosses would yield. The genotypes of the crosses and their progeny are as follows:

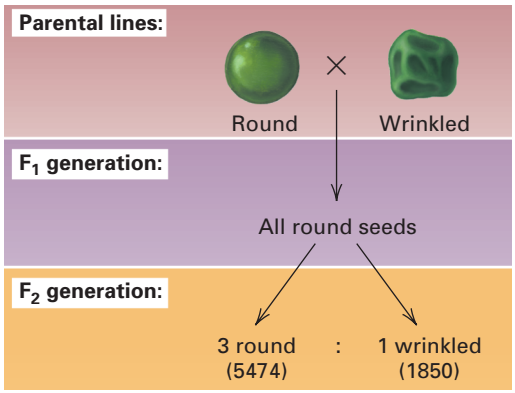
Cross A: WW ♀ × ww ♂ → Ww
Cross B: ww ♀ × WW ♂ → Ww

In both reciprocal crosses, the progeny have the morphological phenotype of round seeds, but as shown by the RFLP diagrams on the right, all progeny genotypes are actually heterozygous *Ww* and therefore different from either parent. The genetic equivalence of reciprocal crosses illustrated in Figure 3.3 is a principle that is quite general in its applicability, but there is an important exception, having to do with sex chromosomes, that will be discussed in Chapter 4.

In the following section we examine a few of Mendel's original experiments in the context of RFLP analysis in order to relate the morphological phenotypes and their ratios to the molecular phenotypes that would be expected.

Phenotypic Ratios in the F_2 Generation

Although the progeny of the crosses in Figure 3.3 have the dominant phenotype of round seeds, the RFLP analysis shows that they are actually heterozygous. The *w* allele is hidden with respect to the morphological phenotype because *w* is recessive to *W*. Nevertheless, the wrinkled phenotype reappears in the next generation when the hybrid progeny are allowed to undergo self-fertilization. For example, if the round F_1 seeds from Cross A in Figure 3.3 were grown into sexually mature plants and allowed to undergo self-fertilization, some of the resulting seeds would be round and others wrinkled. The progeny seeds produced by self-fertilization of the F_1 generation constitute the **F_2 generation.** When Mendel carried out this experiment, in the F_2 generation he observed the results shown in the following diagram:



Note that the ratio 5474 : 1850 is approximately 3 : 1.

A 3 : 1 ratio of dominant : recessive forms in the F_2 progeny is characteristic of simple Mendelian inheritance. Mendel's data demonstrating this point are shown in **TABLE 3.1**. Note that the first two traits in the table (round versus wrinkled seeds and yellow versus green seeds) have many more observations than any of the others. The reason is that these traits can be classified directly in the seeds, whereas the others can be classified only in the mature plants, and Mendel could analyze many more seeds than he could mature plants. The principal observations from the data in Table 3.1 are:

- The F_1 hybrids express only the dominant trait (because the F_1 progeny are heterozygous—for example, *Ww*)
- In the F_2 generation, plants with either the dominant or the recessive trait are present (which means that some F_2 progeny are homozygous—for example, *ww*).
- In the F_2 generation, there are approximately three times as many plants with the dominant phenotype as plants with the recessive phenotype.

The 3 : 1 ratio observed in the F_2 generation is the key to understanding the mechanism of genetic transmission. In the next section we use RFLP analysis of the *W* and *w* alleles to explain why this ratio is produced.

Table 3.1 Results of Mendel's monohybrid experiments

Parental traits	F_1 trait	Number of F_2 progeny	F_2 ratio
Round × wrinkled (seeds)	Round	5474 round, 1850 wrinkled	2.96 : 1
Yellow × green (seeds)	Yellow	6022 yellow, 2001 green	3.01 : 1
Purple × white (flowers)	Purple	705 purple, 224 white	3.15 : 1
Inflated × constricted (pods)	Inflated	882 inflated, 299 constricted	2.95 : 1
Green × yellow (unripe pods)	Green	428 green, 152 yellow	2.82 : 1
Axial × terminal (flower position)	Axial	651 axial, 207 terminal	3.14 : 1
Long × short (stems)	Long	787 long, 277 short	2.84 : 1

The Principle of Segregation

The 3 : 1 ratio can be explained with reference to **FIGURE 3.4**. This is the heart of Mendelian genetics. You should master it and be able to use it to deduce the progeny types produced in crosses. The diagram illustrates these key features of single-gene inheritance:

1. Genes come in pairs, which means that a cell or individual has *two* copies (alleles) of each gene.
2. For each pair of genes, the alleles may be identical (homozygous *WW* or homozygous *ww*), or they may be different (heterozygous *Ww*).
3. Each reproductive cell (**gamete**) produced by an individual contains only *one* allele of each gene (that is, either *W* or *w*).
4. In the formation of gametes, any particular gamete is equally likely to include either allele (hence, from a heterozygous *Ww* genotype, half the gametes contain *W* and the other half contain *w*).
5. The union of male and female reproductive cells is a random process that reunites the alleles in pairs.

The essential feature of transmission genetics is the separation, technically called **segregation**, in unaltered form, of the two alleles in an individual during the formation of its reproductive cells. Segregation corresponds to points 3 and 4 in the foregoing list. The prin-

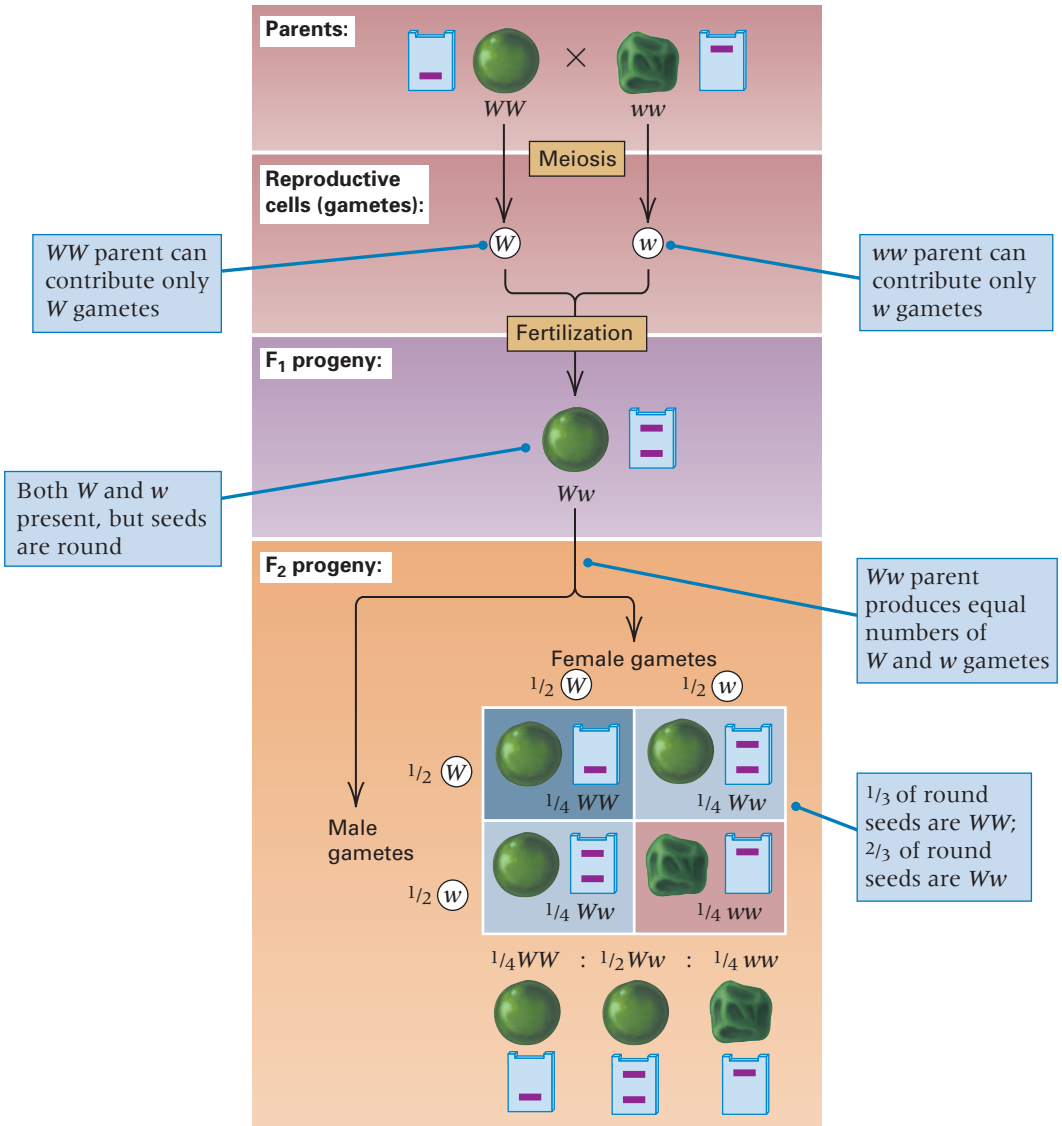


FIGURE 3.4 A diagrammatic explanation of the 3 : 1 ratio of dominant : recessive morphological phenotypes observed in the F_2 generation of a monohybrid cross. The 3 : 1 ratio is observed because of dominance. Note that the ratio of *WW* : *Ww* : *ww* genotypes in the F_2 generation is 1 : 2 : 1, as can be seen from the restriction fragment phenotypes.

connection

What Did Gregor Mendel Think He Discovered?

Gregor Mendel 1866
Monastery of St. Thomas, Brno [then Brünn], Czech Republic
Experiments on Plant Hybrids (original in German)

Mendel's paper is remarkable for its precision and clarity. It is worth reading in its entirety for this reason alone. Although the most important discovery attributed to Mendel is segregation, he never uses this term. His description of segregation is found in the first passage in italics in the excerpt. (All of the italics are reproduced from the original.) In his description of the process, he takes us carefully through the separation of A *and* a *in gametes and their coming together again at random in fertilization. One flaw in the description is Mendel's occasional confusion between genotype and phenotype, which is illustrated by his writing* A *instead of* AA *and* a *instead of* aa *in the display toward the end of the passage. Most early geneticists made no consistent distinction between genotype and phenotype until 1909, when the terms themselves were coined.*

Artificial fertilization undertaken on ornamental plants to obtain new color variants initiated the experiments reported here. The striking regularity with which the same hybrid forms always reappeared whenever fertilization between like species took place suggested further experiments whose task it was to follow that development of hybrids in their progeny. . . . This paper discusses the attempt at such a detailed experiment. . . . Whether the plan by which the individual experiments were set up and carried out was adequate to the assigned task should be decided by a benevolent judgment. . . . [Here the experimental results are described in detail.] Thus experimentation also justifies the assumption *that pea hybrids form germinal and pollen cells that in their composition correspond in equal numbers to all the constant forms resulting from the combination of traits united through fertilization.* The difference of forms among the progeny of hybrids, as well as the ratios in which they are observed, find an adequate explanation in the principle [of segregation] just deduced. The simplest case is given by the series for *one pair of differing traits.* It is shown that this series is described by the expression: $A + 2Aa + a$, in which A and a signify the forms with constant differing traits, and Aa the form hybrid for both. The series contains four individuals in three different terms. In their production, pollen and germinal cells of form A and a participate, on the average, equally in fertilization; therefore each form manifests itself twice, since four individuals are produced. Participating in fertilization are thus:

Pollen cells	$A + A + a + a$
Germinal cells	$A + A + a + a$

It is entirely a matter of chance which of the two kinds of pollen combines with each single germinal cell. However, according to the laws of probability, in an average of many cases it will always happen that every pollen form A and a will unite equally often with every germinal-cell form A and a; therefore, in fertilization, one of the two pollen cells A will meet a germinal cell A, the other a germinal cell a, and equally, one pollen cell a will become associated with a germinal cell A, and the other a.

Pollen cells	A	A	a	a
	↓	⤫		↓
Germinal cells	A	A	a	a

The result of fertilization can be visualized by writing the designations for associated germinal and pollen cells in the form of fractions, pollen cells above the line, germinal cells below. In the case under discussion one obtains

$$\frac{A}{A} + \frac{A}{a} + \frac{a}{A} + \frac{a}{a}$$

In the first and fourth terms germinal and pollen cells are alike; therefore the products of their association must be constant, namely A and a; in the second and third, however, a union of the two differing parental traits takes place again, therefore the forms arising from such fertilizations are absolutely identical with the hybrid from which they derive. *Thus, repeated hybridization takes place.* The striking phenomenon, that hybrids are able to produce, in addition to the two parental types, progeny that resemble themselves is thus explained: Aa and aA both give the same association, Aa, since, as mentioned earlier, it makes no difference to the consequence of fertilization which of the two traits belongs to the pollen and which to the germinal cell. Therefore

$$\frac{A}{A} + \frac{A}{a} + \frac{a}{A} + \frac{a}{a} = A + 2Aa + a$$

This represents the *average* course of self-fertilization of hybrids when two differing traits are associated in them. In individual flowers and individual plants, however, the ratio in which the members of the series are formed may be subject to not insignificant deviations. . . . Thus it was proven experimentally that, in *Pisum,* hybrids form *different kinds* of germinal and pollen cells and that this is the reason for the variability of their offspring.

Whether the plan by which the individual experiments were set up and carried out was adequate to the assigned task should be decided by a benevolent judgment.

Source: G. Mendel, *Verhandlungen des naturforschenden den Vereines in Brünn* 4(1866): 3–47.

ciple of segregation is sometimes called *Mendel's first law.*

> **The Principle of Segregation:** In the formation of gametes, the paired hereditary determinants separate (segregate) in such a way that each gamete is equally likely to contain either member of the pair.

Another key feature of transmission genetics is that the hereditary determinants are present as pairs in both the parental organisms and the progeny organisms but as single copies in the reproductive cells. This feature corresponds to points 1 and 5 in the foregoing list.

Figure 3.4 illustrates the biological mechanism underlying the important Mendelian ratios in the F_2 generation of 3 : 1 for phenotypes and 1 : 2 : 1 for genotypes. To understand these ratios, consider first the parental generation in which the original cross is *WW* × *ww*. The sex of the parents is not stated because reciprocal crosses yield identical results. (There is, however, a convention in genetics that unless otherwise specified, crosses are given with the female parent listed first.) In the original cross, the *WW* parent produces only *W*-containing gametes, whereas the *ww* parent produces only *w*-containing gametes. Segregation still takes place in the homozygous genotypes as well as in the heterozygous genotype, even though all of the gametes carry the same type of allele (*W* from homozygous *WW* and *w* from homozygous *ww*). When the *W*-bearing and *w*-bearing gametes come together in fertilization, the hybrid genotype is heterozygous *Ww*, which is shown by the bands in the gel icon next to the F_1 progeny. With regard to seed shape, the hybrid *Ww* seeds are round because *W* is dominant over *w*.

When the heterozygous F_1 progeny form gametes, segregation implies that half the gametes will contain the *W* allele and the other half will contain the *w* allele. These gametes come together at random when an F_1 individual is self-fertilized or when two F_1 individuals are crossed. The result of random fertilization can be deduced from the sort of cross-multiplication square shown at the bottom of the figure, in which the female gametes and their frequencies are arrayed across the top margin and the male gametes and their frequencies along the left-hand margin. This calculating device is widely used in genetics and is called a **Punnett square** after its inventor Reginald C. Punnett (1875–1967). The Punnett square in Figure 3.4 shows that random combinations of the F_1 gametes result in an F_2 generation with the genotypic composition 1/4 *WW*, 1/2 *Ww*, and 1/4 *ww*. This can be confirmed by the RFLP banding patterns in the gel icons because of the codominance of *W* and *w* with respect to the molecular phenotype. But because *W* is dominant over *w* with respect to the morphological phenotoype, the *WW* and *Ww* genotypes have round seeds and the *ww* genotypes have wrinkled seeds, yielding the phenotypic ratio of round : wrinkled seeds of 3 : 1. Hence, it is a combination of segregation, random union of gametes, and dominance that results in the 3 : 1 ratio.

The ratio of F_2 genotypes is as important as the ratio of F_2 phenotypes. The Punnett square in Figure 3.4 also shows that the ratio of *WW* : *Ww* : *ww* genotypes is 1 : 2 : 1, which can be confirmed directly by the RFLP analysis.

Verification of Segregation

The round seeds in Figure 3.4 conceal a genotypic ratio of 1 *WW* : 2 *Ww*. To say the same thing in another way, among the F_2 seeds that are round (or, more generally, among organisms that show the dominant morphological phenotype), 1/3 are homozygous (in this example, *WW*) and 2/3 are heterozygous (in this example, *Ww*). This conclusion is obvious from the RFLP patterns in Figure 3.4, but it is not at all obvious from the morphological phenotypes. Unless you knew something about genetics already, it would be a very bold hypothesis, because it implies that two organisms with the same morphological phenotype (in this case round seeds) might nevertheless differ in molecular phenotype and in genotype.

Yet this is exactly what Mendel proposed. But how could this hypothesis be tested experimentally? He realized that it could be tested via self-fertilization of the F_2 plants. With self-fertilization, plants grown from the homozygous *WW* genotypes should be true-breeding for round seeds, whereas those from the heterozygous *Ww* genotypes should yield round and wrinkled seeds in the ratio of 3 : 1. On the other hand, the plants grown from wrinkled seeds should be true-breeding for wrinkled because these plants are homozygous *ww*. The results Mendel obtained are summarized in **FIGURE 3.5**. As predicted from the genetic hypothesis, plants grown from F_2 wrinkled seeds were true breeding for wrinkled seeds, yielding only wrinkled seeds in the F_3 generation. But some of the plants grown from round seeds showed evidence of segregation. Among 565 plants grown from round F_2 seeds, 372 plants produced both round and wrinkled seeds in a proportion very close to 3 : 1, whereas the re-

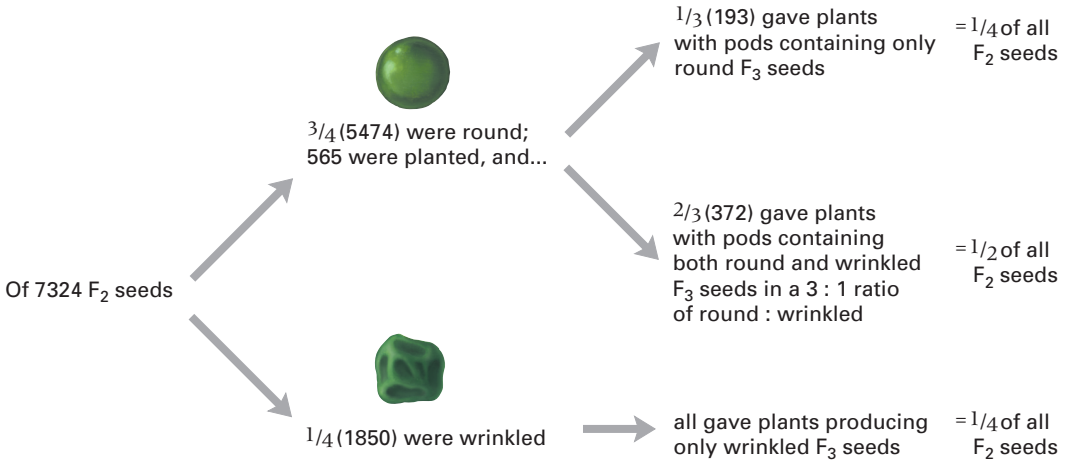


FIGURE 3.5 Summary of F_2 phenotypes and the progeny produced by self-fertilization.

maining 193 plants produced only round seeds in the F_3 generation. The ratio 193 : 372 equals 1 : 1.93, which is very close to the ratio 1 : 2 of *WW* : *Ww* genotypes predicted from the genetic hypothesis.

An important feature of the homozygous round and homozygous wrinkled seeds produced in the F_2 and F_3 generations is that the phenotypes are exactly the same as those observed in the original parents in the P_1 generation. This makes sense in terms of DNA, because the DNA of each allele remains unaltered unless a new mutation happens to occur. Mendel described this result in a letter by saying that in the progeny of crosses, "the two parental traits appear, separated and unchanged, and there is nothing to indicate that one of them has either inherited or taken over anything from the other." From this finding, he concluded that the hereditary determinants for the traits in the parental lines were transmitted as two different elements that retain their purity in the hybrids. In other words, the hereditary determinants do not "mix" or "contaminate" each other. In modern terminology, this means that, with rare but important exceptions, genes are transmitted unchanged from generation to generation.

The Testcross and the Backcross

Another straightforward way of testing the genetic hypothesis in Figure 3.4 is by means of a **testcross,** a cross between an organism that is heterozygous for one or more genes (for example, *Ww*) and an organism that is homozygous for the recessive allele (for example, *ww*). The result of such a testcross is shown in **FIGURE 3.6**. Because the heterozygous parent is expected to produce *W* and *w* gametes in equal numbers, whereas the homozygous recessive produces only *w* gametes, the expected progeny are 1/2 with the genotype *Ww* and 1/2 with the genotype *ww*. The former have the dominant phenotype because *W* is dominant over *w*, and the latter have the recessive phenotype. A testcross is often extremely useful in genetic analysis:

> In a testcross, the phenotypes of the progeny reveal the relative frequencies of the different gametes produced by the heterozygous parent, because the recessive parent contributes only recessive alleles.

Mendel carried out a series of testcrosses with various traits. The results are shown in **TABLE 3.2**. In all cases, the ratio of phenotypes among the test-

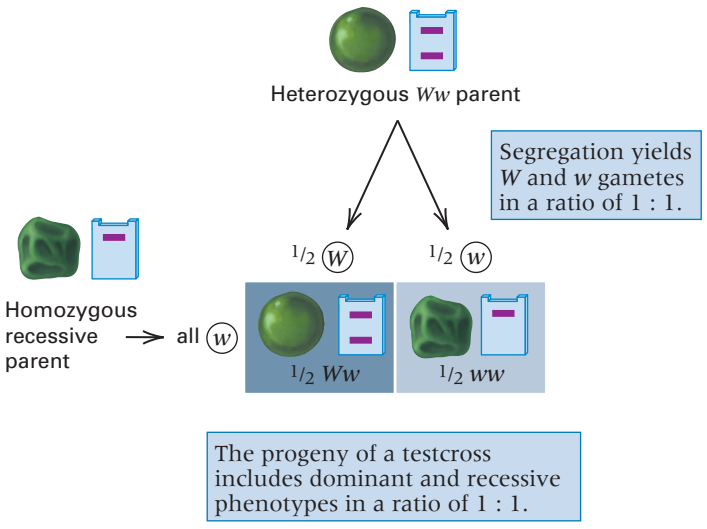


FIGURE 3.6 In a testcross of a *Ww* heterozygous parent with a *ww* homozygous recessive, the progeny are *Ww* and *ww* in the ratio of 1 : 1. A testcross shows the result of segregation.

Table 3.2 Mendel's testcross results

Testcross (F_1 heterozygote × homozygous recessive)	Progeny from testcross	Ratio
Round × wrinkled seeds	193 round, 192 wrinkled	1.01 : 1
Yellow × green seeds	196 yellow, 189 green	1.04 : 1
Purple × white flowers	85 purple, 81 white	1.05 : 1
Long × short stems	87 long, 79 short	1.10 : 1

cross progeny is very close to the 1 : 1 ratio expected from segregation of the alleles in the heterozygous parent.

Another valuable type of cross is a **backcross**, in which hybrid organisms are crossed with one of the parental genotypes. Backcrosses are commonly used by geneticists and by plant and animal breeders, as we will see in later chapters. Note that the testcrosses in Table 3.2 are also backcrosses, because in each case, the F_1 heterozygous parent came from a cross between the homozygous dominant and the homozygous recessive.

3.3 Segregation of Two or More Genes

The results of many genetic crosses depend on the segregation of the alleles of two or more genes. The genes may be in different chromosomes or in the same chromosome. Although in this section we consider the case of genes that are in two different chromosomes, the same principles apply to genes that are in the same chromosome but are so far distant from each other that they segregate independently. The case of *linkage* of genes in the same chromosome is examined in Chapter 5.

To illustrate the principles, we consider again a cross between homozygous genotypes, but in this case homozygous for the alleles of two genes. A specific example is a true-breeding variety of garden peas with seeds that are wrinkled and green (genotype *ww gg*) versus a variety with seeds that are round and yellow (genotype *WW GG*). As suggested by the use of uppercase and lowercase symbols for the alleles, the dominant alleles are *W* and *G*, the recessive alleles *w* and *g*. The mutant gene responsible for Mendel's green seeds has been identified. It is an inborn error in metabolism that blocks the metabolic pathway for breaking down the green pigment chlorophyll. Homozygous mutant seeds are unable to break down their chlorophyll and therefore they stay green, whereas wildtype seeds do break down their chlorophyll and turn yellow like the leaves of certain trees in the autumn. The gene is officially named *staygreen*.

A cross of *WW GG* plants with *ww gg* plants yields F_1 seeds with the genotype *Ww Gg*, which are phenotypically round and yellow because of the dominance relations. When the F_1 seeds are grown into mature plants and self-fertilized, the F_2 progeny show the result of simultaneous segregation of the *W, w* allele pair and the *G, g* allele pair. When Mendel perfomed this cross, he obtained the following numbers of F_2 seeds:

Round, yellow	315
Round, green	108
Wrinkled, yellow	101
Wrinkled, green	32
Total	556

In these data, the first thing to be noted is the expected 3 : 1 ratio for each trait considered separately. The ratio of round : wrinkled (pooling across yellow and green) is

$$(315 + 108) : (101 + 32)$$
$$= 423 : 133$$
$$= 3.18 : 1$$

And the ratio of yellow : green (pooling across round and wrinkled) is

$$(315 + 101) : (108 + 32)$$
$$= 416 : 140$$
$$= 2.97 : 1$$

Both of these ratios are in satisfactory agreement with 3 : 1. (Testing for goodness of fit to a predicted ratio is described in Chapter 4.) Furthermore, in the F_2 progeny of the dihybrid cross, the separate 3 : 1 ratios for the two traits were combined at random. With random combinations, as shown in **FIGURE 3.7**, among the 3/4 of the progeny that are round, 3/4 will be yellow and 1/4 green; similarly, among the 1/4 of the progeny that are wrinkled, 3/4 will be yellow and 1/4 green. The overall proportions of round yellow to round green to wrinkled yellow to wrinkled green are therefore expected to be

$$\begin{array}{cccccccccc} & 3/4 \times 3/4 & : & 3/4 \times 1/4 & : & 1/4 \times 3/4 & : & 1/4 \times 1/4 \\ = & 9/16 & : & 3/16 & : & 3/16 & : & 1/16 \end{array}$$

The observed ratio of 315 : 108 : 101 : 32 equals 9.84 : 3.38 : 3.16 : 1, which is a satisfactory fit

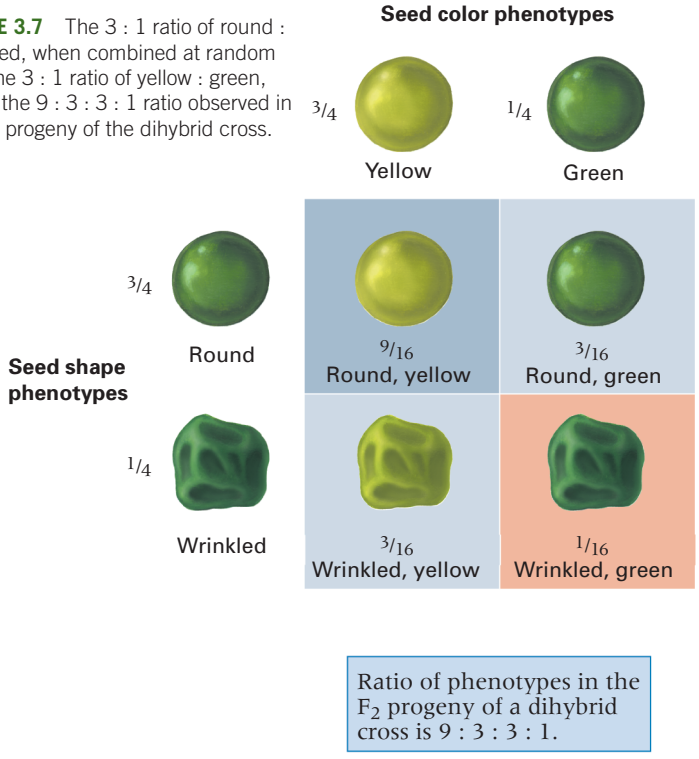


FIGURE 3.7 The 3 : 1 ratio of round : wrinkled, when combined at random with the 3 : 1 ratio of yellow : green, yields the 9 : 3 : 3 : 1 ratio observed in the F_2 progeny of the dihybrid cross.

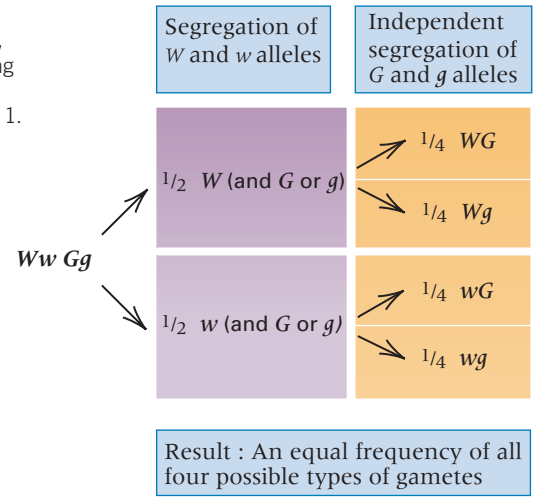


FIGURE 3.8 Independent segregation of the *W*, *w* and *G*, *g* allele pairs means that among each of the *W* and *w* gametic classes, the ratio of *G* : *g* is 1 : 1. Likewise, among each of the *G* and *g* gametic classes, the ratio of *W* : *w* is 1 : 1.

to the 9 : 3 : 3 : 1 ratio expected from the Punnett square in Figure 3.7.

The Principle of Independent Assortment

The independent segregation of the *W*, *w* and *G*, *g* allele pairs is illustrated in **FIGURE 3.8**. What independence means is that if a gamete contains *W*, it is equally likely to contain *G* or *g*; and if a gamete contains *w*, it is equally likely to contain *G* or *g*. The implication is that the four gametes are formed in equal frequencies:

$$1/4\ W\,G \quad 1/4\ W\,g \quad 1/4\ w\,G \quad 1/4\ w\,g$$

The result of independent assortment when the four types of gametes combine at random to form the zygotes of the next generation is shown in **FIGURE 3.9**. Note that the expected ratio of phenotypes among the F_2 progeny is

$$9 : 3 : 3 : 1$$

However, as the Punnett square also shows, the ratio of genotypes in the F_2 generation is more complex; it is

$$1 : 2 : 1 : 2 : 4 : 2 : 1 : 2 : 1$$

The reason for this ratio is shown in **FIGURE 3.10**. Among seeds that have the *WW* genotype, the ratio of

$$GG : Gg : gg \text{ equals } 1 : 2 : 1$$

Among seeds that have the *Ww* genotype, the ratio is

$$2 : 4 : 2$$

(which is a 1 : 2 : 1 ratio multiplied by 2 because there are twice as many *Ww* genotypes as either *WW* or *ww*). And among seeds that have the *ww* genotype, the ratio of

$$GG : Gg : gg \quad \text{equals} \quad 1 : 2 : 1$$

The phenotypes of the seeds are shown beneath the genotypes, and the combined phenotypic ratio is

$$9 : 3 : 3 : 1$$

The principle of independent segregation of two pairs of alleles in different chromosomes (or located sufficiently far apart in the same chromosome) has come to be known as the principle of independent assortment. It is also sometimes referred to as *Mendel's second law.*

> **The Principle of Independent Assortment:** Segregation of the members of any pair of alleles is independent of the segregation of other pairs in the formation of reproductive cells.

Although the principle of independent assortment is fundamental in Mendelian genetics, the phenomenon of linkage, caused by proximity of genes in the same chromosome, is an important exception.

The Testcross with Unlinked Genes

Genes that show independent assortment are said to be **unlinked.** The hypothesis of independent assortment can be tested directly in a testcross with the double homozygous recessive:

$$Ww\ Gg \times ww\ gg$$

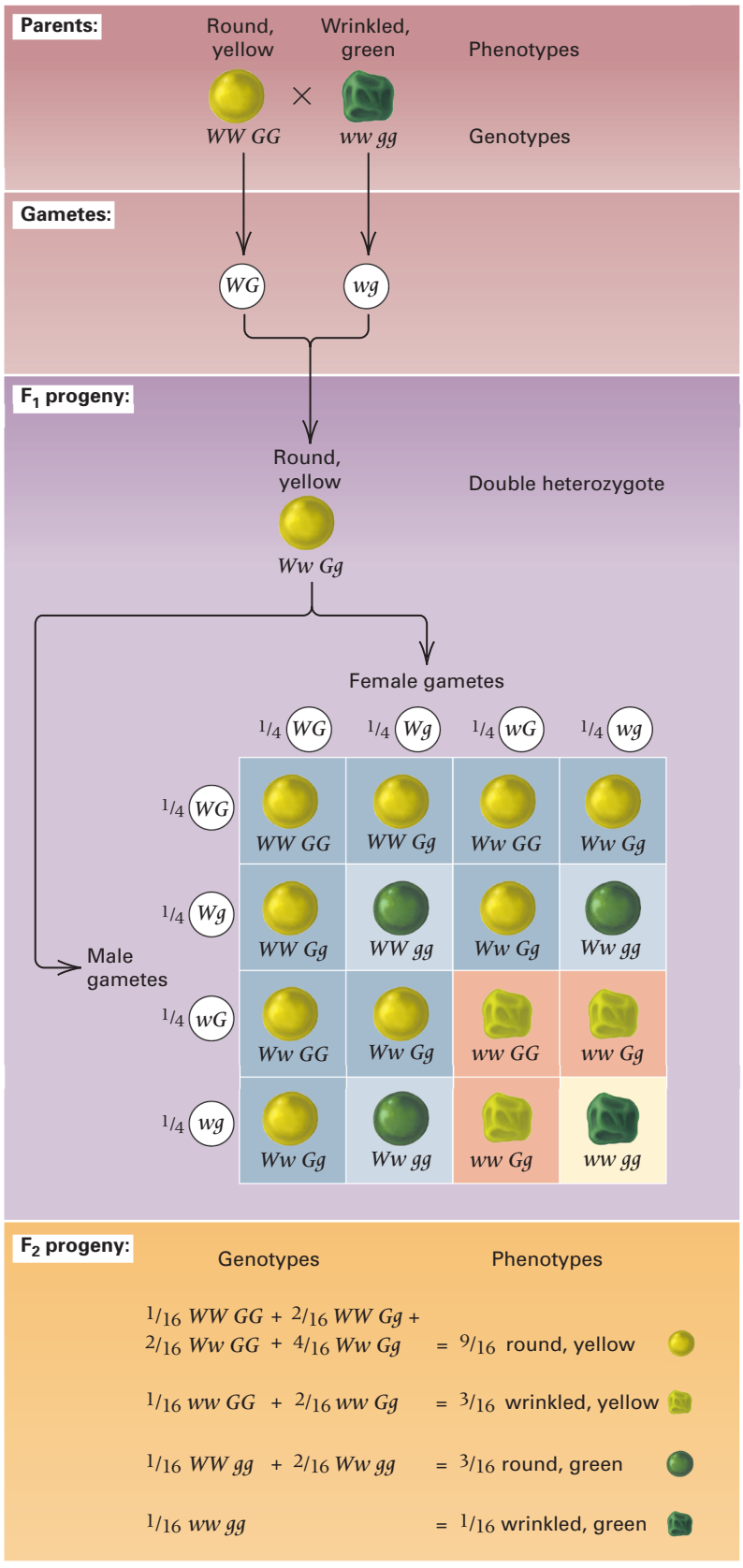


FIGURE 3.9 Independent assortment is the biological basis for the 9 : 3 : 3 : 1 ratio of F_2 phenotypes resulting from a dihybrid cross.

The result of the testcross is shown in **FIGURE 3.11**. Because plants with doubly heterozygous genotypes produce four types of gametes—*W G*, *W g*, *w G*, and *w g*—in equal frequencies, whereas the plants with *ww gg* genotypes produce only *w g* gametes, the possible progeny genotypes are *Ww Gg*, *Ww gg*, *ww Gg*, and *ww gg*, and these are expected in equal frequencies. Because of the dominance relations—*W* over *w* and *G* over *g*—the progeny phenotypes are expected to be round yellow, round green, wrinkled yellow, and wrinkled green in a ratio of

$$1 : 1 : 1 : 1$$

As with the one-gene testcross, in a two-gene testcross the ratio of progeny phenotypes is a direct demonstration of the ratio of gametes produced by the doubly heterozygous parent. In the actual cross, Mendel obtained 55 round yellow, 51 round green, 49 wrinkled yellow, and 53 wrinkled green, which is in good agreement with the predicted 1 : 1 : 1 : 1 ratio.

Three or More Genes

A Punnett square of the type in Figure 3.9 is a nice way to show the logic behind the genotype and phenotype frequencies for two genes that undergo independent assortment. As a method for solving problems, however, a Punnett square is not very efficient. In working a problem, especially when time is limited as during an exam, drawing and filling out the whole square takes too long. Another approach is less graphical but quicker. For example, the expected frequency of any genotype with independent assortment for two genes can be obtained by picking out the corresponding term in the expression

$$(\tfrac{1}{4}\ WW + \tfrac{1}{2}\ Ww + \tfrac{1}{4}\ ww) \times (\tfrac{1}{4}\ GG + \tfrac{1}{2}\ Gg + \tfrac{1}{4}\ gg)$$

Likewise, the expected frequency of any phenotype is given by the corresponding term in the expression

$$(\tfrac{3}{4}\ \text{round} + \tfrac{1}{4}\ \text{wrinkled}) \times (\tfrac{3}{4}\ \text{yellow} + \tfrac{1}{4}\ \text{green})$$

For example, the expected frequency of *Ww gg* genoypes in the F_2 generation is given by $\frac{1}{2} \times \frac{1}{4} = \frac{1}{8}$, and the expected frequency of round, green phenotypes is given by $\frac{3}{4} \times \frac{1}{4} = \frac{3}{16}$.

The time saved by use of the algebraic expressions is small when the cross has only two genes, but it is already considerable with three genes that show independent assortment. Consider an example of a trihybrid cross with the allele pairs (*W*, *w*), (*G*, *g*), and (*P*, *p*), where *P*

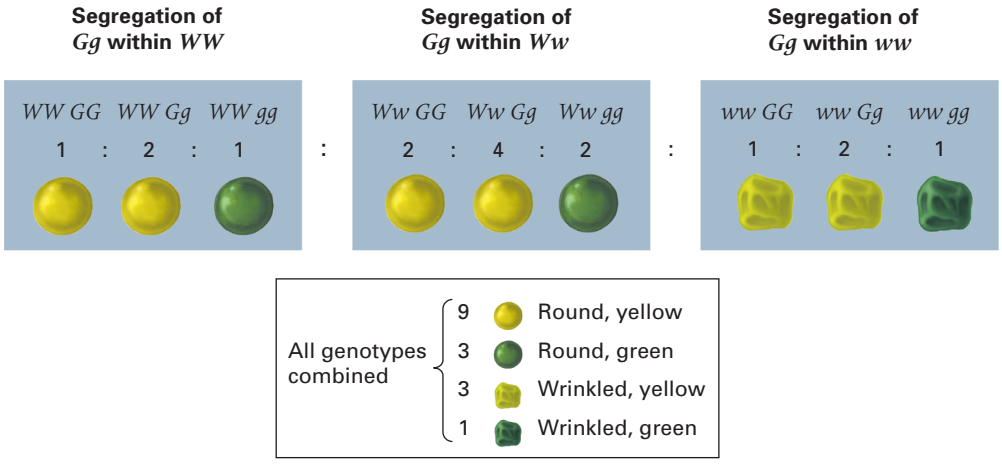


FIGURE 3.10 Genotypes and phenotypes of the F_2 progeny of the dihybrid cross for seed shape and seed color.

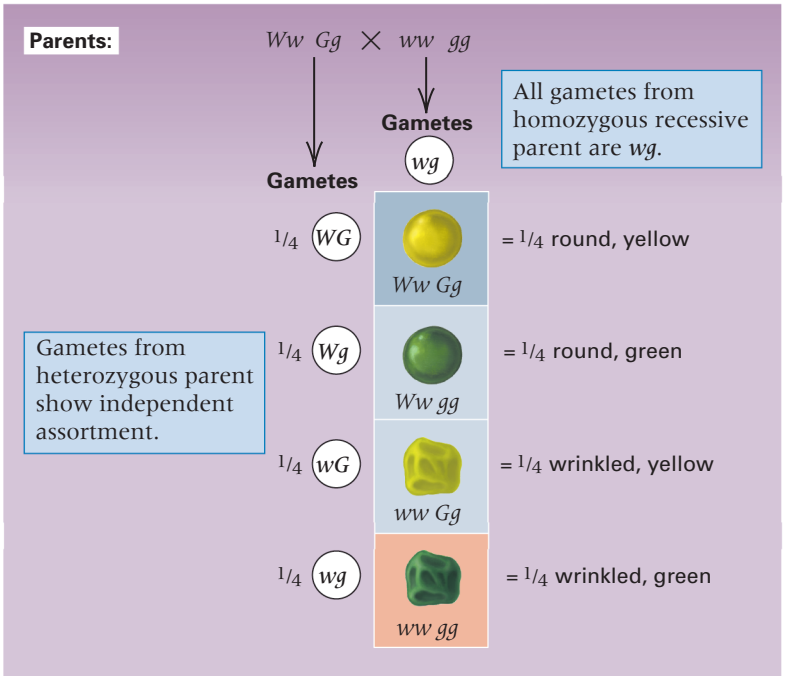


FIGURE 3.11 Genotypes and phenotypes resulting from a testcross of the *Ww Gg* double heterozygote.

is a dominant allele for purple flowers and *p* is a recessive mutation for white flowers. In this case, the Punnett square is a cube containing 64 cells that is tricky to draw and tedious to complete. The genotype frequencies nevertheless correspond to terms in the expression

$$(\tfrac{1}{4}\ WW + \tfrac{1}{2}\ Ww + \tfrac{1}{4}\ ww) \times (\tfrac{1}{4}\ GG + \tfrac{1}{2}\ Gg + \tfrac{1}{4}\ gg) \times (\tfrac{1}{4}\ PP + \tfrac{1}{2}\ Pp + \tfrac{1}{4}\ pp)$$

and the phenotype frequencies to terms in

$$(\tfrac{3}{4}\ \text{round} + \tfrac{1}{4}\ \text{wrinkled}) \times (\tfrac{3}{4}\ \text{yellow} + \tfrac{1}{4}\ \text{green}) \times (\tfrac{3}{4}\ \text{purple} + \tfrac{1}{4}\ \text{white})$$

If you wish to know the probability that an offspring genotype is *Ww Gg Pp*, for example, this can be obtained by multiplying the frequencies $\frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} = \frac{1}{8}$. Likewise, if you wish to know the probability that offspring phenotype is round, yellow, purple, this can be calculated as $\frac{3}{4} \times \frac{3}{4} \times \frac{3}{4} = \frac{24}{64}$.

Mendel actually carried out this three-gene cross. The results for the phenotypes are shown in FIGURE 3.12. They conform well to the ratio of 27 : 9 : 9 : 9 : 3 : 3 : 3 : 1 expected from independent assortment. The various progeny types represent $3^3 = 27$ different genotypes. Mendel did a testcross for each offspring having the dominant phenotype for one or more traits in order to determine whether its genotype was homozygous or heterozygous for each of the genes. This effort alone required 632 testcrosses. Little wonder that he complained that "of all the experiments, [this one] required the most time and effort."

Now that we have emphasized the efficiency and utility of a more abstract approach to genetic calculations than the Punnett square, we are in a position to explain a little more formally the foundations of probability as they relate to Mendelian genetics and to take such reasoning an important step further.

3.4 Probability in Genetic Analysis

As you already know, Mendel's laws of genetic transmission are fundamentally laws of chance (probability). He surpassed any of his contemporaries in understanding that his principles of inheritance accounted for the different types of progeny he observed as well as for the ratios in which they were found. No discoveries in genetics made since Mendel's time have undermined the fundamental role of chance in heredity that he was the first to recognize.

To fully understand Mendelian genetics, we therefore need to understand the elementary

$(3/4\ W\ +\ 1/4\ ww)\ \times\ (3/4\ G-\ +\ 1/4\ gg)\ \times\ (3/4\ P-\ +\ 1/4\ pp)$

			Observed number	Expected number
27/64	*W*− *G*− *P*−	Round, yellow, purple	269	270
9/64	*W*− *G*− *pp*	Round, yellow, white	98	90
9/64	*W*− *gg P*−	Round, green, purple	86	90
9/64	*wwG*− *P*−	Wrinkled, yellow, purple	88	90
3/64	*W*− *gg pp*	Round, green, white	27	30
3/64	*wwG*− *pp*	Wrinkled, yellow, white	34	30
3/64	*ww gg P*−	Wrinkled, green, purple	30	30
1/64	*ww gg pp*	Wrinkled, green, white	7	10

- For any one gene, the ratio of phenotypes is 48 : 16 = 3 : 1
- For any pair of genes, the ratio of phenotypes is 36 : 12 : 12 : 4 = 9 : 3 : 3 : 1

FIGURE 3.12 With independent assortment, the expected ratio of phenotypes in a trihybrid cross is obtained by multiplying the three independent 3 : 1 ratios of he dominant and recessive phenotypes. A dash used in a genotype symbol indicates that either the dominant or the recessive allele is present; for example, *W*− refers collectively to the genotypes *WW* and *Ww*. (The expected numbers total 640 rather than 639 because of round-off error.)

principles of probability. Every problem in probability begins with an experiment, which may be real or imaginary. In genetics, the experiment is typically a cross. Associated with the experiment is a set of possible outcomes of the experiment. In genetics the possible outcomes are typically genotypes or phenotypes. The possible outcomes are called *elementary outcomes*. They are elementary outcomes in the sense that none of them can be reduced to combinations of the others. In our applications of probability, the number of elementary outcomes is often relatively small, or in any case can be enumerated. The principles of probability also can deal with conceptual experiments in which there are an infinite number of elementary outcomes, but this requires some technicalities that are beyond the scope of this book.

Each elementary outcome is assigned a *probability* that is proportional to its likelihood of occurrence. In principle the probabilities can be assigned arbitrarily. There are only two rules. First, the probability of each elementary outcome must be a nonnegative number between 0 and 1, and may actually equal 0 or 1. An elementary outcome with a probability of 0 cannot occur, and one with a probability of 1 *must* occur. The second rule is that the sum of the probabilities of all the elementary outcomes must equal 1. This rule assures that some one of the elementary outcome must occur. These two rules also handle an annoying question often asked in regard to a coin toss: What happens if it lands on its edge? The answer is that this elementary outcome is assigned a probability of 0, and so we need not bother with it.

Here it will be helpful to consider a specific example. Consider the conceptual experiment of self-fertilization of the F_1 progeny of a cross between pea plants homozygous for the round *W* and yellow *G* alleles with plant homozygous for *w* and *g*. Recall that the *W*, *w* and *G*, *g* allele pairs undergo independent assortment, and so there are 16 elementary outcomes. These are shown in **FIGURE 3.13A**. Each of the elementary outcomes is equally likely, and so the probability of each elementary outcome is assigned a value of $1/16$. Note that the progeny genotype *Ww Gg* is listed four times. This is because there are four possible combinations of parental gametes that can yield the progeny genotype *Ww Gg* (see Figure 3.9).

The enumeration of the elementary outcomes and their probabilities constitute what is often called the *sample space* of the probability problem. For the progeny from self-fertilization of *Ww Gg* the sample space is shown in Figure 3.13A. This is the sample space in which all probability considerations regarding this conceptual experiment take place.

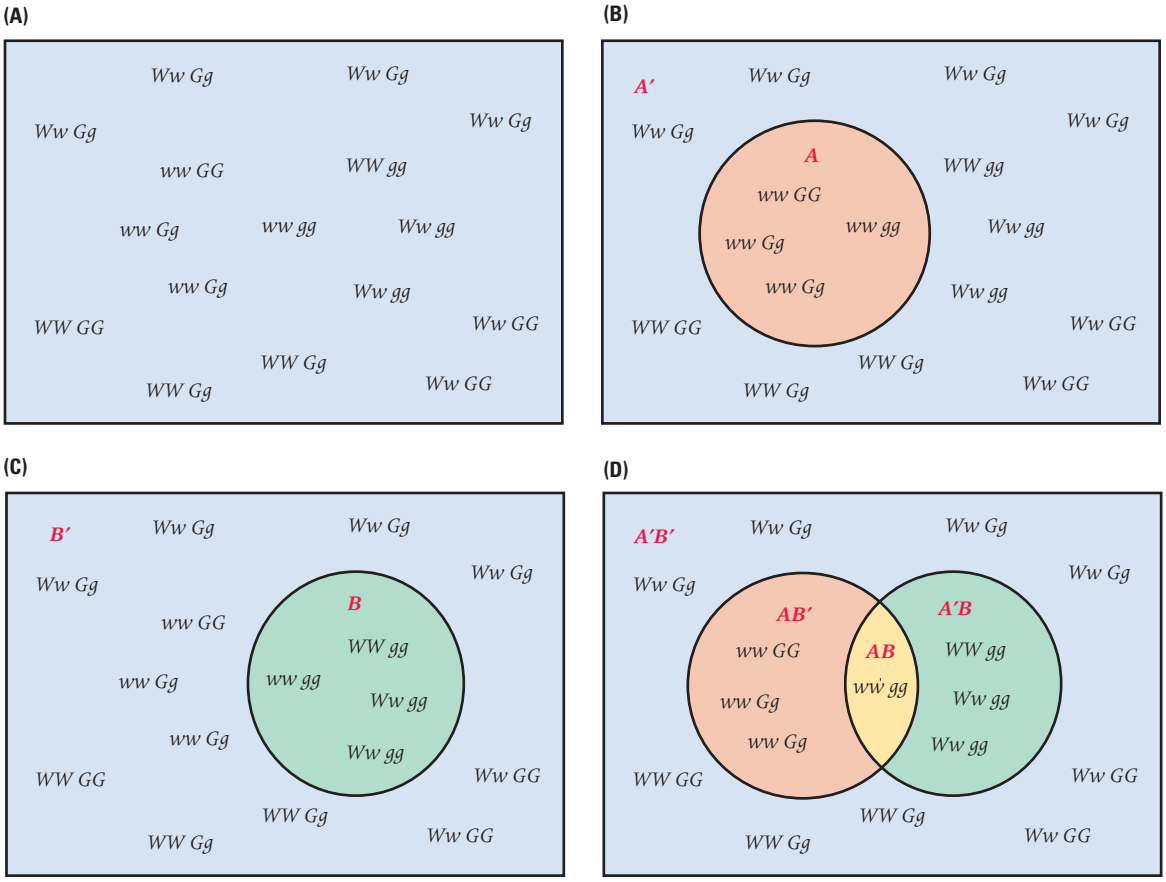


FIGURE 3.13 (A) Sample space for the possible progeny in a cross for the allele pairs *W*, *w* for round versus wrinkled seeds and *G*, *g* for yellow versus green seeds. (B) The event *A* includes all genotypes whose phenotype is wrinkled. (C) The event *B* includes all genotypes whose phenotype is yellow. (D) The union and intersection (yellow) of *A* and *B*.

Elementary Outcomes and Events

Any combination of elementary outcomes constitutes an **event** in the sample space. In Figure 3.13B, the circle labeled *A* includes four elementary outcomes defining the event *A* as "the offspring genotype is *ww*." The event A could be defined equivalently as "the offspring phenotype is wrinkled." The alternative ways of defining *A* in words shows how subsets of elementary outcomes can relate genotypes and phenotypes. Corresponding to every event is a *probability* of that event, in this case symbolized Pr{*A*}. A fundamental principle of probability is that:

> The **probability** of any event equals the sum of the probabilities of all the elementary outcomes included in the event.

In the example in Figure 3.13B, therefore, Pr{*A*} = $^4/_{16}$ because the event *A* consists of four elementary outcomes (possible progeny genotypes), each of which has probability of $^1/_{16}$. An event may include no elementary outcomes, in which case its probability is 0; or it may include all elementary outcomes, in which case its probability is 1.

The elementary outcomes outside the circle defining event *A* in Figure 3.13B define another event that we have denoted *A′*. This event consists of all elementary outcomes not present in *A*, and so in other words it consists of all progeny whose genotype is not *ww*, or equivalently it consists of progeny whose phenotype is round. The event *A′* is called the *complement* of *A*, and it is also variously denoted as A^C, *A*, or not-*A*. The probability of *A′* again equals the sum of the probabilities of the elementary outcomes that constitute *A′*, and so in this case Pr{*A′*} = $^{12}/_{16}$.

Events also can be composed of other events. To see how, consider the event denoted *B* in Figure 3.13C. Event *B* is defined as all progeny from the cross whose genotype is *gg*, which could also be defined as all progeny whose seeds are green. Because *B* includes four elementary outcomes, its probability is Pr{*B*} = $^4/_{16}$. Again there is a complementary event *B′* defined as all elementary outcomes not included in *B*; or to define *B′* in another way, all progeny whose genotype is either *GG* or *Gg*. There are 12 elementary outcomes in *B′*, and so Pr{*B′*} = $^{12}/_{16}$.

Look again at Figures 3.13B and C and consider the event consisting of elementary outcomes that are included in A or included in B or both. This event is called the **union** of A and B. We will denote the union of A and B as $A + B$, but in many probability textbooks you may find it symbolized as $A \cup B$, where the symbol $\cup$ is pronounced "cup." The event $A + B$ consists of seven elementary outcomes in which the progeny genotype is either *ww*, *gg*, or both, and so $Pr\{A + B\} = 7/16$.

Another important way in which two events A and B can be combined is called the **intersection** of A and B, and it consists of all the elementary outcomes that are included in both A and B. An example is shown in Figure 3.13D, where the intersection of A and B is shaded yellow. We will denote the intersection of A and B as AB, but in probability textbooks the intersection is often represented as $A \cap B$, where the symbol $\cap$ is pronounced "cap."

Probability of the Union of Events

As with all events, the probability of the union of events $A + B$ equals the sum of the probabilities of the elementary outcomes that are included in A or B or both. Likewise, the probability of the intersection of events AB equals the sum of the probabilities of the elementary outcomes that are included in both A and B. In general, an equation for the probability of the union of A and B is

$$\Pr\{A + B\} = \Pr\{A\} + \Pr\{B\} - \Pr\{AB\} \qquad (1)$$

The reason for subtracting $\Pr\{AB\}$ will become evident by looking at Figure 3.13D. Because the progeny genotype *ww gg* is included in both A and B, when the probabilities of the elementary outcomes in A are added to those in B, the genotype *ww gg* is included twice, and to correct for the overcounting the probability of this outcome must be subtracted from the total. Because the elementary outcomes counted twice are exactly those that are present in both A and B, they constitute the intersection of A and B. Hence, $\Pr\{AB\}$ is the quantity that is subtracted in Equation 1.

An important special case of Equation 1 pertains when A and B do not overlap, that is, when they have no elementary outcomes in common. In this case A and B are said to be **mutually exclusive** (or *disjoint*), and so $\Pr\{AB\} = 0$. Therefore, when A and B are mutually exclusive, Equation 1 becomes

$$\Pr\{A + B\} = \Pr\{A\} + \Pr\{B\} \qquad (2)$$

This equation is sometimes called the **addition rule** for mutually exclusive events. Events are mutually exclusive when the occurrence of one event excludes the occurrence of the other. In other words, mutually exclusive events are mutually incompatible in the sense that no elementary outcome can be present in both. An example, again based on the experiment of crossing *Ww Gg* × *Ww Gg*, is shown in **FIGURE 3.14**. Here we have defined events A^* and B^* in such a way that they are modified versions of A and B that do not overlap. In words, the event A^* may be defined as "the progeny phenotype is wrinkled but not green," and the event B^* may be defined as "the progeny phenotype is green but not wrinkled." Defined in this way, A^* and B^* are mutually exclusive, and therefore Equation 2 has the implication that $\Pr\{A^* + B^*\} = \Pr\{A^*\} + \Pr\{B^*\} = 6/16$.

Probability of the Intersection of Events

Returning again to events A and B in Figure 3.13D, consider the probability of the intersection AB. This event consists of all elementary outcomes that are included in both A and B, which in this example consists of one genotype only, namely *ww gg*. The probability of AB is therefore $\Pr\{AB\} = 1/16$. This is a special case in which the events are **independent**, which means that knowing whether or not the event A occurs tells you nothing about whether or not the event B occurs. The probability of the joint occurrence of independent events has the special property that it is proportional to the product of the probabilities of the individual events. In the example in Figure 3.13D we have already seen that $\Pr\{A\} = 4/16$ and that $\Pr\{B\} = 4/16$, and so in this case, $\Pr\{AB\} = 4/16 \times 4/16 = 16/256 = 1/16$. The events A and B in this example illustrate an important general principle: When the events A and B are independent, their probability is given by

$$\Pr\{AB\} = \Pr\{A\}\Pr\{B\} \qquad (3)$$

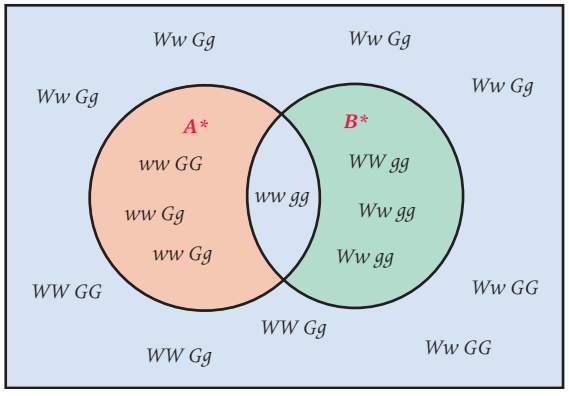


FIGURE 3.14 The events A^* and B^* are defined in such a way that they each exclude genotypes whose phenotype is wrinkled and green. A^* and B^* do not overlap and hence are mutually exclusive.

This equation is sometimes called the **multiplication rule** for independent events. The choice of terms for "independent events" and "independent assortment" is not fortuitous, because independent assortment means that knowing the genotype of an offspring for the *W*, *w* allele pair tells you nothing about the genotype for the *G*, *g* allele pair. This principle is illustrated in **FIGURE 3.15A**. Another situation in genetics in which independence is the rule is shown in part B, which deals with successive offspring from a cross. Successive offspring are independent because the genotypes of early progeny have no influence on the probabilities in later progeny. The independence of successive offspring contradicts the widespread belief that in each human familiy, the ratio of girls to boys must "even out" at approximately 1:1. According to this reasoning, a family with four girls would be more likely to have a boy the next time around. But this belief is supported neither by theory nor by actual data. The data indicate that human families are equally likely to have a girl or a boy on any birth, irrespective of the sex distribution in previous births. Although statistics guarantees that the sex ratio will balance out when averaged over a very large number of cases, this does not imply that it will equalize in any individual case. To be concrete, among families in which there are five children, those consisting of five boys balance those consisting of five girls, yielding an overall sex ratio of 1 : 1; nevertheless, both of these sex ratios are unusual.

In anticipation of thinking about non-independent events, consider that every elementary outcome in event *B* in Figure 3.13D must either be in *A* or in *A′*. In particular, the offspring genotype *ww gg* that is included in *B* is also in *A*, whereas all the other genotypes included in *B* are in *A′*. This principle is true in general. Any event *B* can always be written as the union of two other events, *BA* and *BA′*. Furthermore, *BA* and *BA′* are mutually exclusive, because no elementary outcome can be in both *A* and *A′*. Equation 2 therefore implies that

$$\Pr\{B\} = \Pr\{BA\} + \Pr\{BA'\} \qquad (4)$$

Since $\Pr\{BA\} = 1/16$ and $\Pr\{BA'\} = 3/16$, it follows that $\Pr\{B\} = 1/16 + 3/16 = 4/16$. As we shall see in the next section, Equation 4 has many important applications.

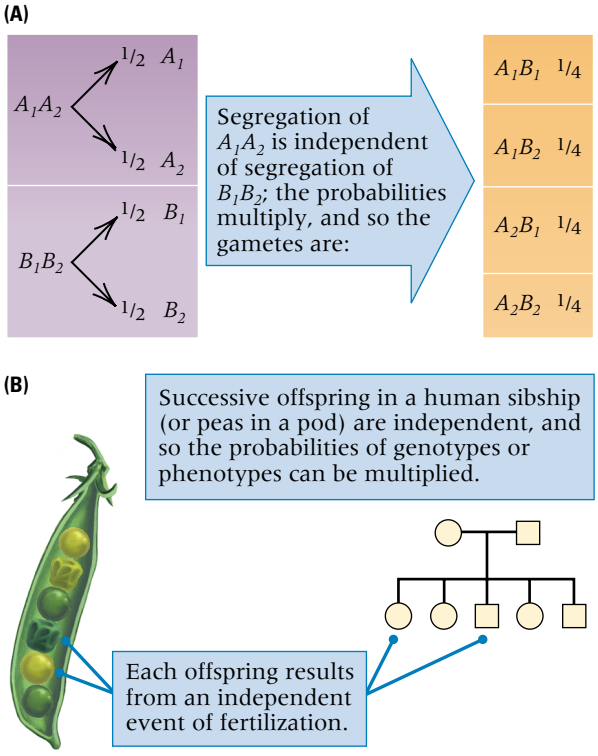


FIGURE 3.15 In genetics, two important types of independence are (A) independent segregation of alleles that show independent assortment and (B) independent fertilizations resulting in successive offspring.

Conditional Probability

In some situations there may be partial information about the outcome of an experiment that can affect the chance that an event has occurred. The known information is called *prior information*, and it is taken into account by means of the concept of conditional probability. The *conditional probability* that an event *A* has occurred, given prior information that another event *B* has occurred, is denoted by the symbol $\Pr\{A \mid B\}$, where the event $A \mid B$ is pronounced as "*A* given *B*." To take a specific example, consider the events *A* and *B* in **FIGURE 3.16**. Here the event *A* can be described in words as "an offspring is wrinked," and event *B* as "an offspring is either wrinkled or green or both." In this example, if we had prior knowledge that *B* had occurred it would change our assessment of the probability that *A* had occurred. The reason is simple: given that *B* has occurred eliminates nine elementary outcomes from the sample space. The prior information that *B* has occurred therefore defines a new experiment and sample space because elementary outcomes that are included in *B′*, the complement of *B*, are not allowed. Given that *B* has occurred, the probability of any event is equal to its relative likelihood within *B*. Because *B* includes seven elementary outcomes and each is equally likely, each of these elementary outcomes has a probability of 1/16. Hence

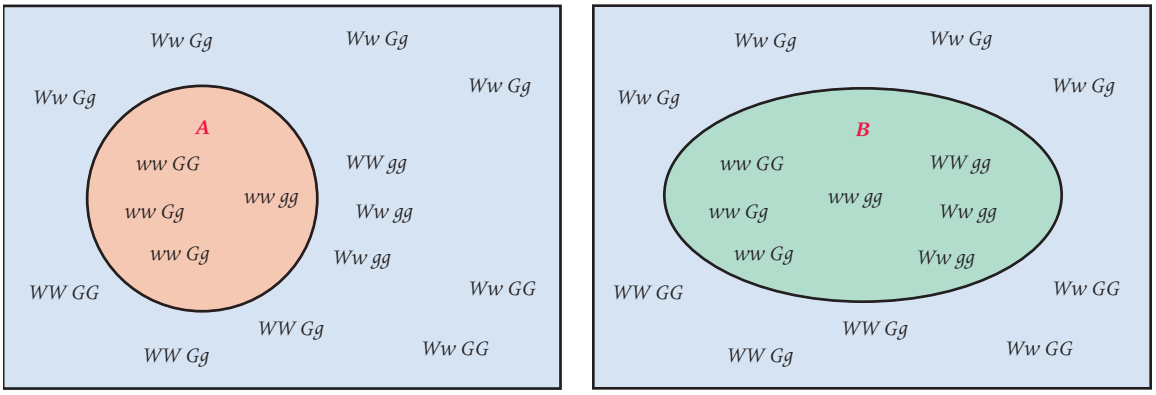


FIGURE 3.16 In this example the elementary outcomes in event *A* are completely included among those in event *B*. Knowing that the event *B* has occurred gives important information about whether the event *A* has occurred.

the probability that *A* occurs, given that *B* has occurred, is $\Pr\{A \mid B\} = 3/7$.

These ideas may become clearer in the technical definition of conditional probability. The **conditional probability** $\Pr\{A \mid B\}$ is *defined* as

$$\Pr\{A \mid B\} = \frac{\Pr\{AB\}}{\Pr\{B\}} \qquad (5)$$

provided that $\Pr\{B\} \neq 0$. The denominator in this equation automatically adjusts the sample space for the fact that only elementary outcomes included in *B* are allowed. For the events in Figure 3.16, for example, $\Pr\{AB\} = 3/16$ and $\Pr\{B\} = 7/16$, and so Equation 5 says that $\Pr\{A \mid B\} = (3/16)/(7/16) = 3/7$, which we calculated earlier simply by redefining the probability of each of the six elementary outcomes in *B* as $1/7$.

To illustrate a more complex application of Equation 5, consider the cross involving the allele pairs (*W, w*), (*G, g*), and (*P, p*) whose F_2 phenotypes are shown in Figure 3.12. Suppose you were asked, "What is the probability that a plant has genotype *Ww Gg Pp*, given that it arose from a round yellow seed and has purple flowers?" Let *A* be the event "the offspring genotype is *Ww Gg Pp*" and *B* be the event "the offspring phenotype is round, yellow, and purple." In the earlier discussion of this cross, we showed that $\Pr\{A\} = 1/2 \times 1/2 \times 1/2 = 1/8$. In this case $\Pr\{AB\} = \Pr\{A\}$, because a plant of genotype *Ww Gg Pp* must have phenotype corresponding to *B*, namely, round, yellow, and purple. Likewise we showed earlier that $\Pr\{B\} = 3/4 \times 3/4 \times 3/4 = 27/64$. Hence, according to Equation 5, the probability that an offspring genotype is *Ww Gg Pp* (event *A*) given that its phenotype is round, yellow, and purple (event *B*) is given by $\Pr\{AB\}/\Pr\{B\} = (1/8)/(27/64) = 8/27$. This answer also can be gotten by mean of a Punnett square for three segregating genes, but this approach requires very much more time and effort.

Equation 5 becomes a particularly powerful tool in probability when combined with Equation 4, in which case its form becomes

$$\Pr\{A \mid B\} = \frac{\Pr\{AB\}}{\Pr\{BA\} + \Pr\{BA'\}} \qquad (6)$$

Note that $\Pr\{AB\} = \Pr\{BA\}$ because the events *AB* and *BA* consist of exactly the same elementary outcomes. To illustrate the utility of Equation 6, consider a conceptual experiment in which we choose one round, green seed from the sample space in Figure 3.13A. What is the probability that the genotype of the seed is *Ww*? To use Equation 6, let *B* be the event that "a randomly chosen green seed is round" and *A* be the event that "a randomly chosen green seed has genotype *Ww*." Then *BA* = *AB* is the event that "a randomly chosen green seed has genotype *Ww* and is round" and *BA'* is the event that "a randomly chosen green seed has genotype *WW* and is round." Mendel's principle of segregation tells us that $\Pr\{BA\} = \Pr\{AB\} = 1/2$, and $\Pr\{BA'\} = 1/4$. Putting all this together in the form of Equation 6 yields

$$\Pr\{A \mid B\} = \frac{1/2}{1/2 + 1/4} = 2/3$$

If you are averse to algebraic equations, application of Equation 6 also can be carried out by means of a table laid out as that shown below, again taking as an example the probability that a randomly chosen green, round seed has genotype *Ww*:

		B
A	1/2	1/2
A'	1/4	1/4
Sum		3/4

On the left are the probabilities that a randomly chosen green seed has genotype *Ww*

(event A) or WW (event A'), and the column headed B contains entries for the probability that a randomly chosen seed is Ww and round (event BA) or WW and round (event BA'). The sum at the bottom is the sum of the entries in column B. To obtain $\Pr\{A \mid B\}$ from the table, divide the entry for BA by the sum to obtain $(\frac{1}{2})/(\frac{3}{4}) = \frac{2}{3}$, which is what we calculated earlier directly from Equation 6.

These calculations may seem like an unnecessarily elaborate way to deduce the elementary principle that, among F_2 seeds that are round, $\frac{2}{3}$ are Ww and $\frac{1}{3}$ are WW, but we shall see in the next section that this type of formal reasoning allows much more complex types of probability problems to be solved.

Bayes Theorem

The wonderful insight of conditional probability is credited to the mind of Thomas Bayes (1702–1761), a Presbyterian minister in Tunbridge Wells, England. His ideas were summarized in his *Essay Towards Solving a Problem in the Doctrine of Chances*, published in 1763, two years after his death. What is now known as **Bayes theorem** can be written in several forms, including the forms in Equations 5 and 6. Many sources give Bayes theorem in another form in which the numerator and denominator in Equation 6 are expressed more fully. This can be done using the expressions in **TABLE 3.3**. These expressions yield a more complete form of Bayes theorem, but they results in an equation that appears much more formidable than Equation 6. Note that the layout of Table 3.3 is the same as that used earlier, and in particular the B column consists of the entries $\Pr\{BA\}$ and $\Pr\{BA'\}$. These entries have also been converted into statements of conditional probability using Equation 5, which can be written as $\Pr\{AB\} = \Pr\{B\}\Pr\{A \mid B\}$ merely by multiplying both sides by $\Pr\{B\}$. The roles of A and B in this expression can be interchanged by relabeling A as B and B as A. It then follows that $\Pr\{BA\} = \Pr\{A\}\Pr\{B \mid A\}$ and also that $\Pr\{BA'\} = \Pr\{A'\}\Pr\{B \mid A'\}$. These expressions correspond to those on the right hand side of Table 3.3. The sum of column B in Table 3.3 still equals $\Pr\{B\}$, which when expressed as the sum of its terms equals

$$\Pr\{B\} = \Pr\{A\}\Pr\{B \mid A\} + \Pr\{A'\}\Pr\{B \mid A'\}$$

Substituting $\Pr\{A\}\Pr\{B \mid A\}$ into the numerator of Equation 6 and the above expression into the denominator yields an alternative form of Bayes theorem:

$$\Pr\{A \mid B\} = \frac{\Pr\{A\}\Pr\{B \mid A\}}{\Pr\{A\}\Pr\{B \mid A\} + \Pr\{A'\}\Pr\{B \mid A'\}} \quad (7)$$

Equation 7 looks extremely complex, but it is merely another way of writing Equation 5 or Equation 6.

We will illustrate application of Equation (7) using the tabular layout of Table 3.3 in connection with the situation shown in **FIGURE 3.17**. In this example, suppose that a round seed (seed 1) was chosen at random from among the green seeds in Figure 3.19A. Suppose that we wish to know whether its genotype is Ww or WW and so carry out a testcross. We therefore grow a plant from this seed and fertilize the ovules with pollen from a plant grown from the wrinkled green seed (seed 2). Expecting a large number of progeny seeds, we hope to be able to specify the genotype of seed 1 because if it is Ww there will be $\frac{1}{4}$ wrinkled seeds among the progeny. Unfortunately, the plant grown from seed 1 was attacked by vicious pea weevils leaving one soli-

Table 3.3 The logical framework of Bayes theorem

	B
A $\Pr\{A\}$	$\Pr\{BA\} = \Pr\{A\}\Pr\{B \mid A\}$
A' $\Pr\{A'\}$	$\Pr\{BA'\} = \Pr\{A'\}\Pr\{B \mid A'\}$
Sum	$\Pr\{B\}$

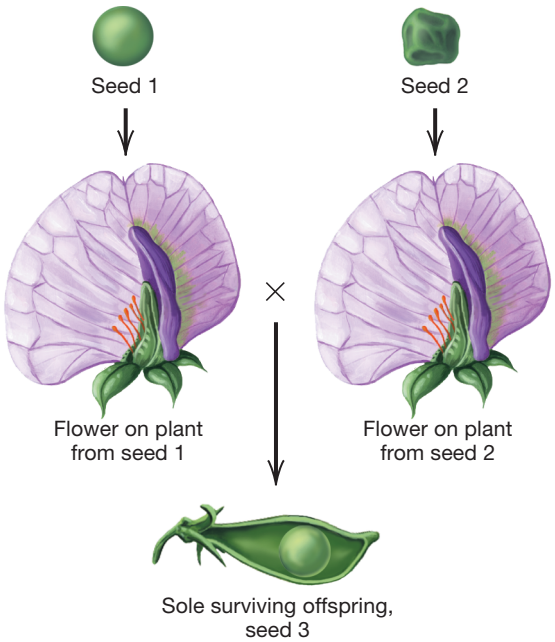


FIGURE 3.17 A testcross of a plant grown from a single round, green seed taken from the F_2 progeny of a dihybrid cross. Only one offspring seed from the textcross survives, but this contributes data that bear on the probability that seed 1 has genotype WW versus Ww.

tary, stunted pod containing only one forlorn offspring, namely seed 3, which proved to be round. Based on this finding, what is the probability that the seed 1 was of genotype *Ww*?

We can solve this problem using Bayes theorem in the form of Table 3.3. The relevant calculations are shown below. The event *A* is "the seed 1 has genotype *Ww*" and event *A′* is "the seed 1 has genotype *WW*." We have already seen that the probability that a seed known to be round had genotype *Ww* is $\frac{2}{3}$ and the probability that it has genotype *WW* is $\frac{1}{3}$. These are the entries in the

		B
A	2/3	(2/3)(1/2)
A′	1/3	(1/3)(1)
Sum		2/3

left-hand column. In column *B*, the entry on the top is $\Pr\{A\}\{\Pr\{B \mid A\}$, where *B* is the event "the seed 3 is round." Likewise the entry $(\frac{1}{3})(1)$ is $\Pr\{A'\}\{\Pr\{B \mid A'\}$. The sum at the bottom is $\Pr\{B\}$, and hence the required conditional probability $\Pr\{A \mid B\} = [(\frac{2}{3})(\frac{1}{2})]/(\frac{2}{3}) = \frac{1}{2}$. Note that this is quite different from the probability $\frac{2}{3}$ that we would have deduced if the seed 3 had not survived.

3.5 Human Pedigree Analysis

Large deviations from expected genetic ratios are often found in individual human families and in domesticated large animals because of the relatively small number of progeny. The effects of segregation are nevertheless evident upon examination of the phenotypes among several generations of related individuals. A diagram of a family tree showing the phenotype of each individual among a group of relatives is a **pedigree.** In this section we introduce basic concepts in pedigree analysis.

Characteristics of Dominant and Recessive Inheritance

FIGURE 3.18 defines the standard symbols used in depicting human pedigrees. Females are represented by circles and males by squares. (A diamond is used if the sex is unknown—as, for example, in a miscarriage.) Persons with the phenotype of interest are indicated by colored or shaded symbols. For recessive alleles, heterozygous carriers are sometimes depicted with half-filled symbols. A mating between a female and a male is indicated by joining their symbols with a horizontal line, which is connected vertically to a second horizontal line, below, that connects the symbols for their offspring. The offspring within a sibship, called **siblings** or **sibs,** are represented from left to right in order of their birth.

A pedigree for the trait *Huntington disease,* caused by a dominant mutation, is shown in **FIGURE 3.19**. The numbers in the pedigree are for convenience in referring to particular persons. The successive generations are designated by Roman numerals. Within any generation, all of the persons are numbered consecutively from left to right. The pedigree starts with the

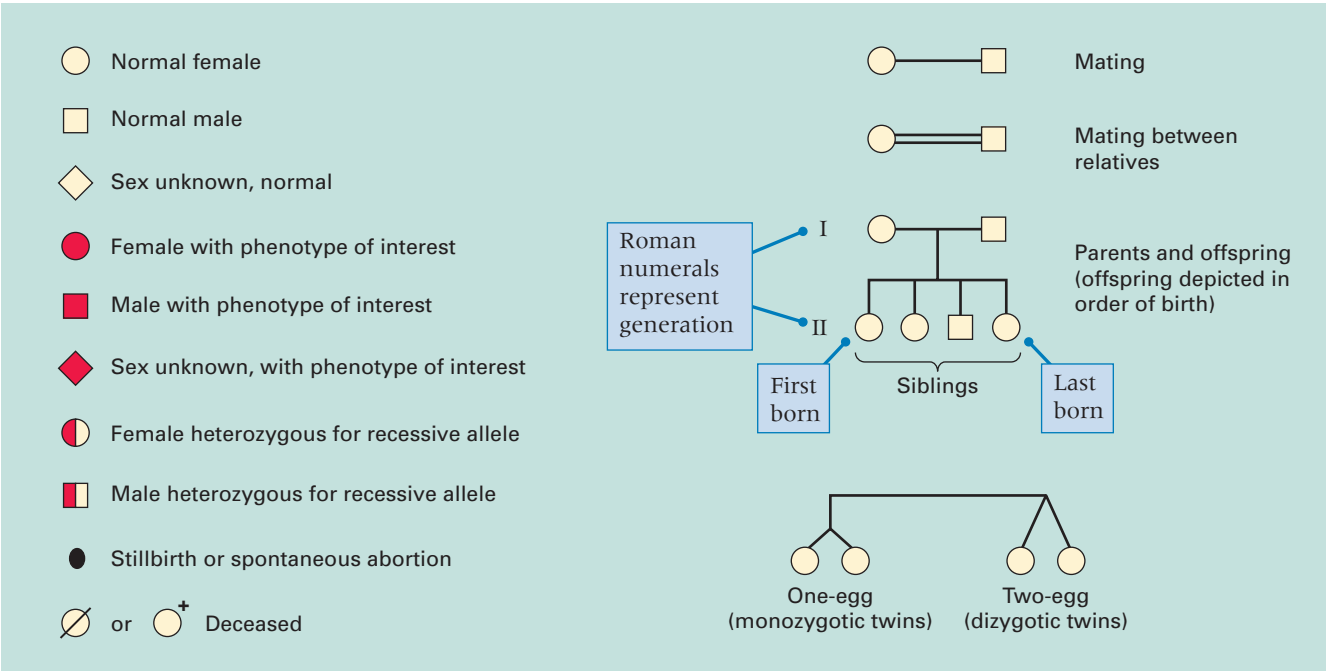


FIGURE 3.18 Conventional symbols used in depicting human pedigrees.

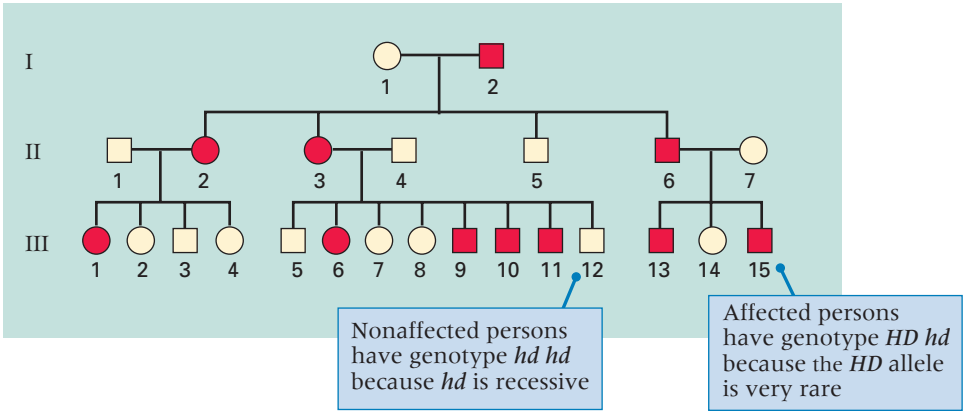


FIGURE 3.19 Pedigree of a human family showing the inheritance of the dominant gene for Huntington disease. Females and males are represented by circles and squares, respectively. Red symbols indicate persons affected with the disease.

woman I-1 and the man I-2. The man has Huntington disease, which is a progressive nerve degeneration that usually begins about middle age. It results in severe physical and mental disability and then death. The dominant allele, *HD*, that causes Huntington disease is rare. All affected persons in the pedigree have the heterozygous genotype *HD hd*, whereas nonaffected persons have the homozygous normal genotype *hd hd*. The disease has complete penetrance. The **penetrance** of a genetic disorder is the proportion of individuals with the at-risk genotype who actually express the trait; *complete penetrance* means the trait is expressed in 100 percent of persons with that genotype. The pedigree demonstrates the following characteristic features of inheritance due to a rare dominant allele with complete penetrance.

1. Females and males are equally likely to be affected.
2. Affected offspring have one affected parent (except for rare new mutations), and the affected parent is equally likely to be the mother or the father.
3. On average, half of the individuals in sibships with an affected parent are affected.

A pedigree for a trait due to a homozygous recessive allele is shown in **FIGURE 3.20**. The trait is *albinism*, absence of pigment in the skin, hair, and iris of the eyes. This pedigree illustrates characteristics of inheritance due to a rare recessive allele with complete penetrance:

1. Females and males are equally likely to be affected.
2. Affected individuals, if they reproduce, usually have unaffected progeny.
3. Most affected individuals have unaffected parents.
4. The parents of affected individuals are often relatives.
5. Among siblings of affected individuals, the proportion affected is approximately 25 percent.

With rare recessive inheritance, the mates of homozygous affected persons are usually homozygous for the normal allele, so all of the offspring will be heterozygous and not affected. Heterozygous **carriers** of the mutant allele are considerably more common than homozygous affected individuals, because it is more likely that a person will inherit only one copy of a rare mutant allele than two copies. Most homozygous recessive genotypes therefore result from matings between carriers (heterozygous × heterozygous), in which each offspring has a 1/4 chance of being affected. Another important feature of rare recessive inheritance is that the parents of affected individuals are often related

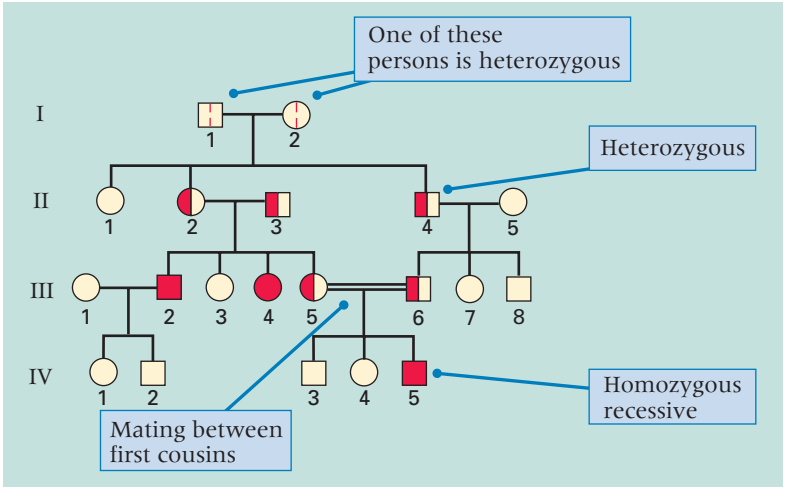


FIGURE 3.20 Pedigree of albinism. With recessive inheritance, affected persons (filled symbols) often have unaffected parents. The double horizontal line indicates a mating between relatives—in this case, first cousins.

(**consanguineous**). A mating between relatives is indicated with a double line connecting the partners, as for the first-cousin mating in Figure 3.20. Mating between relatives is important for recessive alleles to become homozygous, because when a recessive allele is rare, it is more likely to become homozygous through inheritance from a common ancestor than from parents who are completely unrelated. The reason is that the carrier of a rare allele may have many descendants who are also carriers. If two of these carriers should mate with each other (for example, in a first-cousin mating), then the hidden recessive allele can become homozygous with a probability of 1/4. Mating between relatives constitutes *inbreeding*, and the consequences of inbreeding are discussed further in Chapter 17. Because an affected individual indicates that the parents are heterozygous carriers, the expected proportion of affected individuals among the siblings is approximately 25 percent, but the exact value depends on the details of how affected individuals are identified and included in the database.

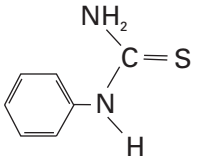
Most Human Genetic Variation is Not "Bad"

Before the advent of molecular methods, there were many practical obstacles to the study of human genetics. With the exception of traits such as the ABO and other blood groups, few traits showing simple Mendelian inheritance were known. Most of these were associated with genetic diseases, and these presented special challenges:

- Most genes that cause simple Mendelian genetic diseases are rare, so they are observed in only a small number of families.
- Many genes for simple Mendelian diseases are recessive, so they are not detected in heterozygous genotypes.
- The number of offspring per human family is relatively small, so segregation cannot usually be detected in single sibships.
- The human geneticist cannot perform testcrosses, backcrosses, or other experimental matings.

This book includes numerous examples of how molecular genetics has revolutionized the study of human genetics. For example, in Chapter 2 we discussed the prevalence of single-nucleotide polymorphisms (SNPs) and copy-number polymorphisms (CNPs) in the human genome. Only a very small number of these common polymorphisms are associated with diseases, and even those serve as genetic risk factors that interact with other risk factors, including other genes and the environment. Although most SNPs and CNPs have no adverse effects, both types of variation show simple Mendelian inheritance, which is why human geneticists have come to rely on them for family and population studies.

But a few, seemingly harmless, simple-Mendelian traits had been detected in the premolecular era. One of the best known was associated with the ability to taste a chemical substance known as **phenylthiocarbamide (PTC)**, which has the molecular structure shown here.



PTC is an artificial chemical created by an industrial chemist in the early 1930s. The taste polymorphism was discovered one day when he carelessly released a cloud of powdered PTC into the air. The PTC powder didn't bother the chemist at all, but his lab mate loudly complained about the bitter taste it left in his mouth. Out of curiosity, the chemist started to test family and friends for their ability to taste PTC, and he recruited a geneticist who began to study the situation. It was soon shown from family studies that the ability to taste PTC is a trait inherited as a simple Mendelian dominant. In European populations about 70 percent of the people are tasters and 30 percent are nontasters, but these proportions differ greatly among ethnic groups. Among people of African or Asian origin the frequency of tasters is about 90 percent, whereas among Australian aborigines it is only about 50 percent.

The ability to taste PTC is quantitative, however. The most sensitive tasters can taste concentrations as low as 0.001 millimolar (mM) whereas the most insensitive nontasters fail to detect concentrations as high as 10 mM. For classifying individuals as "tasters" or "nontasters" an arbitrary cutoff is employed, typically at a concentration of 0.2 mM PTC. Most of the variation in tasting ability between tasters and nontasters is due to the major taster polymorphism, but there are also differences due to other genes, gender, and probably environmental factors. The result of the other variables is that about 5 percent of the heterozygous tasters get classified as "nontaster" and at least 5 percent of the homozygous nontasters become classified as "tasters."

The molecular basis of the taster polymorphism is now known to reside in a taste receptor

protein known as hTAS2R38. There are several alleles of the gene, but the most common forms of the protein differ by three amino acids at scattered positions along the protein. The allelic forms are known as *PAV* and *AVI*, because the three key amino acids in the PAV protein are proline, alanine, and valine, whereas these three positions in the AVI protein are occupied by alanine, valine, and isoleucine. The PAV form is the one that confers the ability to taste PTC.

When you think about it, a polymorphism in PTC tasting makes no sense. PTC is a completely artificial chemical synthesized in the laboratory, so why should there be a polymorphism in the ability to taste it? A clue comes from the observation that the chemical structure of PTC resembles a large and heterogeneous class of molecules called *glucosinolates*. These are compounds synthesized by some plants, including some human food plants, and their synthesis likely evolved as a chemical defense against plant-eating insects. Among the plants that produce glucosinolates is one singled out by former President George H. W. Bush, who in 1989 removed broccoli from the White House menu, proclaiming: "I do not like broccoli. And I haven't liked it since I was a little kid and my mother made me eat it. And I'm President of the United States and I'm not going to eat any more broccoli!" In good-humored protest, broccoli growers throughout the country sent him tons of the stuff. Ironically, 17 years after Bush Sr.'s broccoli boycott, new studies showed that individuals carrying the PAV form of the hTAS2R38 taste receptor do, in fact, find broccoli to be significantly more bitter tasting than individuals homozygous for the allele encoding AVI form. Tasters also report a greater perceived bitterness for collard greens, turnips, rutabagas, and horseradish.

Molecular Markers in Human Pedigrees

Because techniques for manipulating DNA allow direct access to the DNA, modern genetic studies of human pedigrees are carried out primarily using genetic markers present in the DNA itself, rather than through the phenotypes produced by mutant genes. Various types of DNA polymorphisms were discussed in Chapter 2, along with the methods by which they are detected and studied. An example of a DNA polymorphism segregating in a three-generation human pedigree is shown in **FIGURE 3.21**. The type of polymorphism is a *simple sequence repeat polymorphism (SSRP),* in which each allele differs in size according to the number of copies it contains of a short DNA sequence repeated in tandem. The differences in size are detected by electrophoresis after amplification of the region by PCR. SSRPs usually have many codominant alleles, and the majority of individuals are heterozygous for two different alleles. In the example in Figure 3.21, each of the parents is heterozygous, as are all of the children.

Six alleles are depicted in Figure 3.21, denoted by A_1 through A_6. In the gel, the numbers of the bands correspond to the subscripts of the alleles. The mating in generation II is between two heterozygous genotypes: $A_4A_6 \times A_1A_3$. Because of segregation in each parent, four genotypes are possible among the offspring (A_4A_1, A_4A_3, A_6A_1, and A_6A_3); these would conventionally be written with the smaller subscript first, as A_1A_4, A_3A_4, A_1A_6, and A_3A_6. With random fertilization the offspring genotypes are equally likely, as may be verified from a Punnett square for the mating. Figure 3.21 illustrates some of the principal advantages of multiple, codominant alleles for human pedigree analysis:

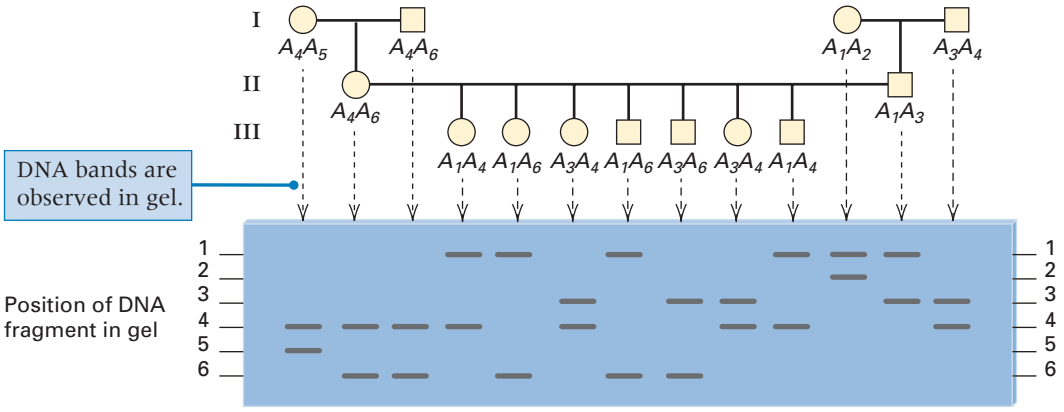


FIGURE 3.21 Human pedigree showing segregation of SSRP alleles. Six alleles (A_1–A_6) are present in the pedigree, but any one person can have only one allele (if homozygous) or two alleles (if heterozygous).

connection

This Land is Your Land

The Huntington's Disease Collaborative Research Group 1993

Comprising 58 authors among 9 institutions

A Novel Gene Containing a Trinucleotide Repeat That Is Expanded and Unstable on Huntington's Disease Chromosomes

Modern genetic research is sometimes carried out by large collaborative groups in a number of research institutions scattered across several countries. This approach is exemplified by the search for the gene responsible for Huntington disease. The search was highly publicized because of the severity of the disease, the late age of onset, and the dominant inheritance. Famed folk singer Woody Guthrie, who wrote "This Land Is Your Land" and other well-known tunes, died of the disease in 1967. When the gene was identified, it turned out to encode a protein (now called huntingtin) of unknown function that is expressed in many cell types throughout the body and not, as expected, exclusively in nervous tissue. Within the coding sequence of this gene is a trinucleotide repeat (5′-CAG-3′) that is repeated in tandem a number of times according to the general formula (5′-CAG-3′)$_n$. Among normal alleles, the number n of repeats ranges from 11 to 34 with an average of 18; among mutant alleles, the number of repeats ranges from 40 to 86. This tandem repeat is genetically unstable in that it can, by some unknown mechanism, increase in copy number ("expand"). In two cases in which a new mutant allele was analyzed, one had increased in repeat number from 36 to 44 and the other from 33 to 49. This is a mutational mechanism that is quite common in some human genetic diseases. The excerpt cites several other examples. The authors also emphasize that their discovery raises important ethical issues, including genetic testing, confidentiality, and informed consent.

Huntington disease (HD) is a progressive neurodegenerative disorder characterized by motor disturbance, cognitive loss, and psychiatric manifestations. It is inherited in an autosomal dominant fashion and affects approximately 1 in 10,000 individuals in most populations of European origin. The hallmark of HD is a distinctive choreic [jerky] movement disorder that typically has a subtle, insidious onset in the fourth to fifth decade of life and gradually worsens over a course of 10 to 20 years until death. . . .The genetic defect causing HD was assigned to chromosome 4 in one of the first successful linkage analyses using DNA markers in humans. Since that time, we have pursued an approach to isolating and characterizing the HD gene based on progressively refining its localization. . . . [We have found that a] 500-kb segment is the most likely site of the genetic defect. [The abbreviation kb stands for kilobase pairs; 1 kb equals 1000 base pairs.] Within this region, we have identified a large gene, spanning approximately 210 kb, that encodes a previously undescribed protein. The reading frame contains a polymorphic $(CAG)_n$ trinucleotide repeat with at least 17 alleles in the normal population, varying from 11 to 34 CAG copies. On HD chromosomes, the length of the trinucleotide repeat is substantially increased. . . . Elongation of a trinucleotide repeat sequence has been implicated previously as the cause of three quite different human disorders, the fragile-X syndrome, myotonic dystrophy, and spino-bulbar muscular atrophy. . . . It can be expected that the capacity to monitor directly the size of the trinucleotide repeat in individuals "at risk" for HD will revolutionize testing for the disorder. . . . We consider it of the utmost importance that the current internationally accepted guidelines and counseling protocols for testing people at risk continue to be observed, and that samples from unaffected relatives should not be tested inadvertently or without full consent. . . . With the mystery of the genetic basis of HD apparently solved, [it opens] the next challenges in the effort to understand and to treat this devastating disorder.

We consider it of the utmost importance that the current internationally accepted guidelines and counseling protocols for testing people at risk continue to be observed, and that samples from unaffected relatives should not be tested inadvertently or without full consent.

Source: The Huntington's Disease Collaborative Research Group, *Cell* 72(1993): 971–983

1. Heterozygous genotypes can be distinguished from homozygous genotypes.
2. Many individuals in the population are heterozygous, and so many matings are informative in regard to segregation.
3. Each segregating genetic marker yields up to four distinguishable offspring genotypes.

3.6 Incomplete Dominance and Epistasis

Dominance and codominance are not the only possibilities for pairs of alleles. There are situations of **incomplete dominance,** in which the phenotype of the heterozygous genotype is intermediate between the phenotypes of the ho-

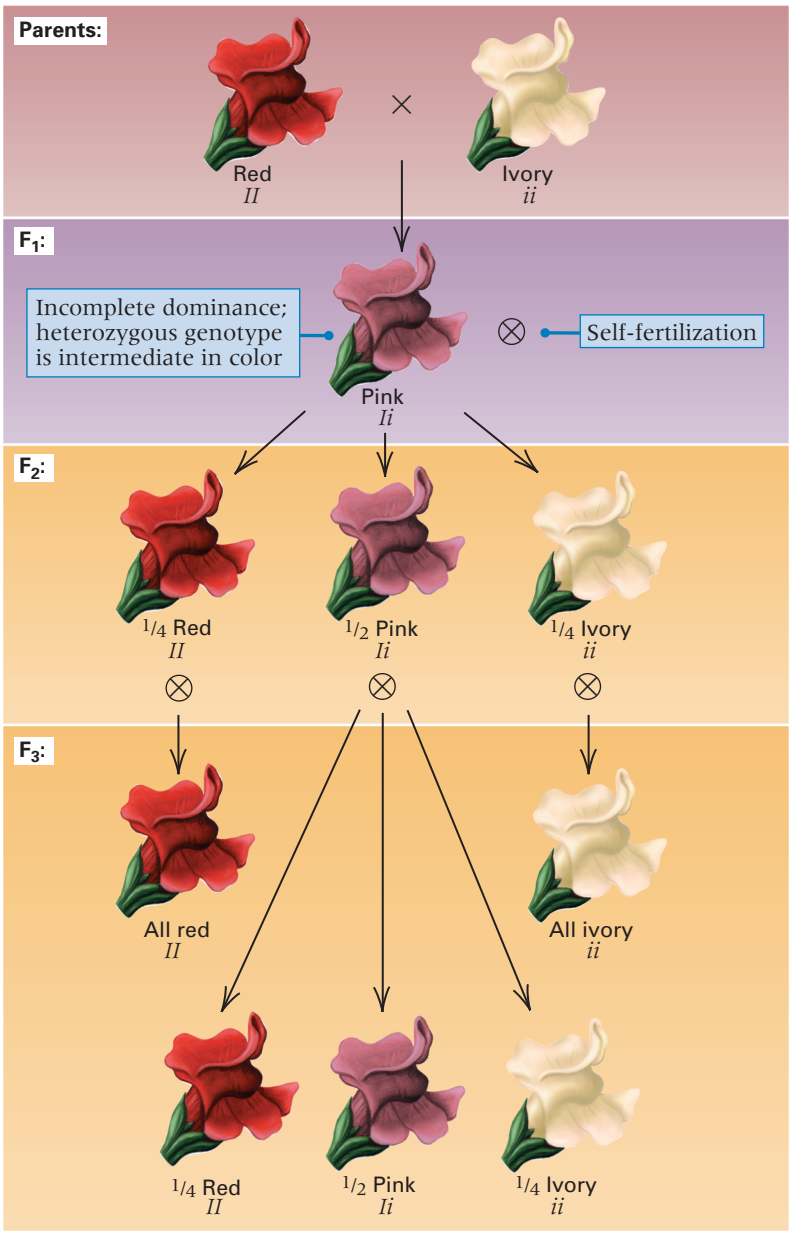


FIGURE 3.22 Red versus white flower color in snapdragons shows incomplete dominance.

mozygous genotypes. A classic example of incomplete dominance concerns flower color in the snapdragon *Antirrhinum majus* (**FIGURE 3.22**). In wildtype flowers, a red type of anthocyanin pigment is formed by a sequence of enzymatic reactions. A wildtype enzyme, encoded by the *I* allele, is limiting to the rate of the overall reaction, so the amount of red pigment is determined by the amount of enzyme that the *I* allele produces. The alternative *i* allele codes for an inactive enzyme, and *ii* flowers are ivory in color. Because the amount of the critical enzyme is reduced in *Ii* heterozygotes, the amount of red pigment in the flowers is reduced also, and the effect of the dilution is to make the flowers pink. A cross between plants differing in flower color therefore gives direct phenotypic evidence of segregation (Figure 3.22). The cross *II* (red) × *ii* (ivory) yields F_1 plants with genotype *Ii* and pink flowers. In the F_2 progeny obtained by self-pollination of the F_1 hybrids, one experiment resulted in 22 plants with red flowers, 52 with pink flowers, and 23 with ivory flowers, which fits the expected ratio of 1 : 2 : 1.

Multiple Alleles

The occurrence of **multiple alleles** is exemplified by the alleles A_1–A_6 of the SSRP marker in the human pedigree in Figure 3.21. Multiple alleles are relatively common in natural populations and, as in this example, can be detected most easily by molecular methods. In the DNA of a gene, each nucleotide can be A, T, G, or C, so a gene of *n* nucleotides can theoretically mutate at any of the positions to any of the three other nucleotides. The number of possible single-nucleotide differences in a gene of length *n* is therefore 3 × *n*. If $n = 5000$, for example, there are potentially 15,000 alleles (not counting any of the possibilities with more than one nucleotide substitution). Most of the potential alleles do not actually exist at any one time. Some are absent because they did not occur, others did occur but were eliminated by chance or because they were harmful, and still others are present but at such a low frequency that they remain undetected. Nevertheless, at the level of DNA sequence, most genes in most natural populations have multiple alleles, all of which can be considered "wildtype." Multiple wildtype alleles are useful in such applications as DNA typing because two unrelated people are unlikely to have the same genotype, especially if several different loci, each with multiple alleles, are examined. Many harmful mutations also exist in multiple forms. Recall from Chapter 1 that more than 400 mutant forms of the phenylalanine hydroxylase gene have been identified in patients with phenylketonuria. The alleles A_1–A_6 in Figure 3.21 also illustrate that although a *population* of organisms may contain any number of alleles, any particular organism or cell may carry no more than two, and any gamete may carry no more than one.

In some cases, the multiple alleles of a gene exist merely by chance and reflect the history of mutations that have taken place in the population and the dissemination of these mutations among population subgroups by migration and interbreeding. In other cases, there are biological mechanisms that favor the maintenance of

a large number of alleles. For example, genes that control self-sterility in certain flowering plants can have large numbers of allelic types. This type of self-sterility is found in species of red clover that grow wild in many pastures. The self-sterility genes prevent self-fertilization because a pollen grain can undergo pollen tube growth and fertilization only if it contains a self-sterility allele different from either of the alleles present in the flower on which it lands. In other words, a pollen grain containing an allele already present in a flower will not function on that flower. Because all pollen grains produced by a plant must contain one of the self-sterility alleles present in the plant, pollen cannot function on the same plant that produced it, and self-fertilization cannot take place. Under these conditions, any plant with a new allele has a selective advantage, because pollen that contains the new allele can fertilize all flowers except those on the same plant. Through evolution, populations of red clover have accumulated hundreds of alleles of the self-sterility gene, many of which have been isolated and their DNA sequences determined. Many of the alleles differ at multiple nucleotide sites, which implies that the alleles in the population are very old.

Human ABO Blood Groups

In a multiple allelic series, there may be different dominance relationships between different pairs of alleles. An example is found in the human ABO blood groups, which are determined by three alleles denoted I^A, I^B, and I^O. (Actually, there are two slightly different variants of the I^A allele.) The blood group of any person may be A, B, AB, or O, depending on the type of polysaccharide (polymer of sugars) present on the surface of red blood cells. One of two different polysaccharides, A or B, can be formed from a precursor molecule that is modified by the enzyme product of the I^A or the I^B allele. The gene product is a glycosyl transferase enzyme that attaches a sugar unit to the precursor (**FIGURE 3.23**). The I^A or the I^B alleles encode different forms of the enzyme with replacements at four amino acid sites; these alter

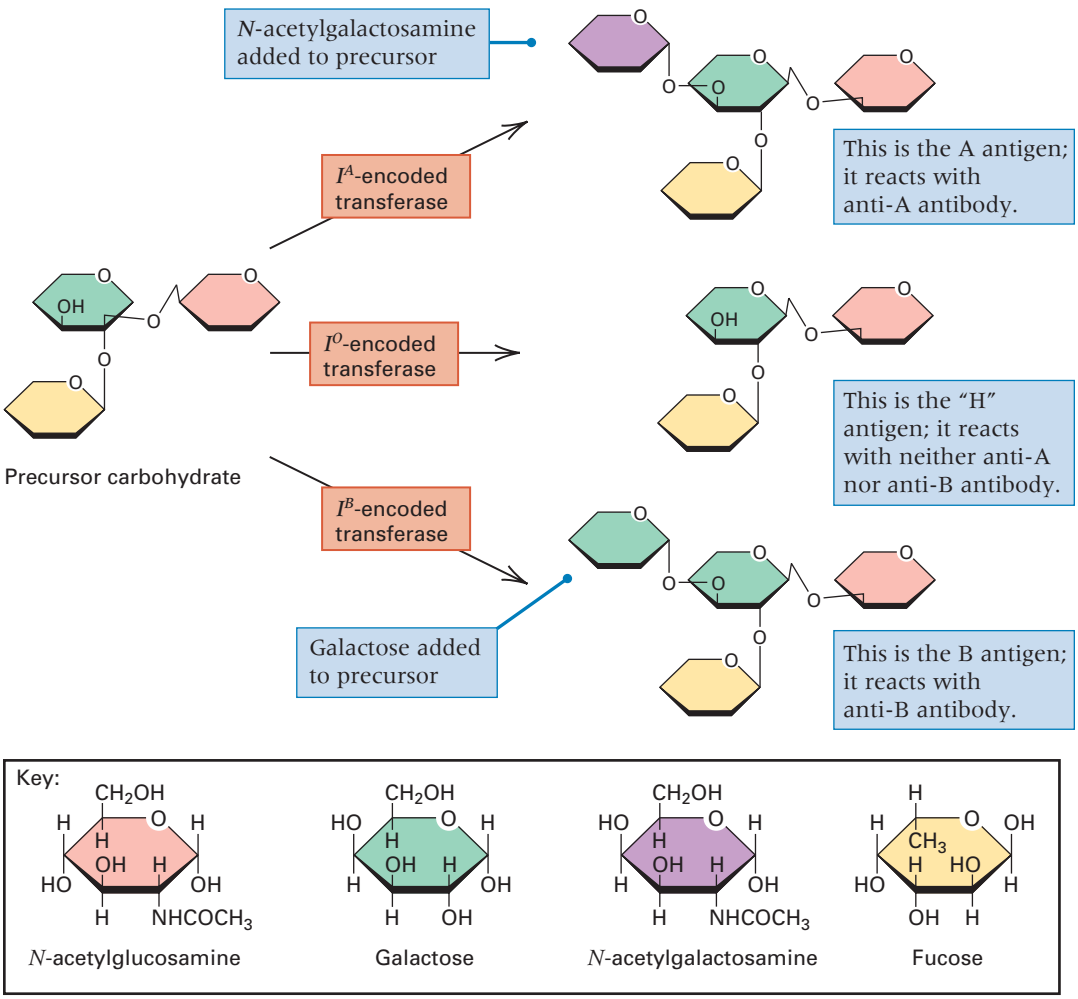


FIGURE 3.23 The ABO antigens on the surface of human red blood cells are carbohydrates.

the substrate specificity so that each enzyme attaches a different sugar. People of genotype I^AI^A produce red blood cells having only the A polysaccharide and are said to have blood type A. Those of genotype I^BI^B have red blood cells with only the B polysaccharide and have blood type B. Heterozygous I^AI^B people have red cells with both the A and the B polysaccharides and have blood type AB. The third allele, I^O, encodes an enzymatically inactive protein that leaves the precursor unchanged; neither the A nor the B type of polysaccharide is produced. Homozygous I^OI^O persons therefore lack both the A and the B polysaccharides and are said to have blood type O.

In this multiple allelic series, the I^A and I^B alleles are codominant: The heterozygous genotype has the characteristics of both homozygous genotypes—the presence of both the A and the B carbohydrate on the red blood cells. On the other hand, the I^O allele is recessive to both I^A and I^B. Hence, heterozygous I^AI^O genotypes produce the A polysaccharide and have blood type A, and heterozygous I^BI^O genotypes produce the B polysaccharide and have blood type B. The genotypes and phenotypes of the ABO blood group system are summarized in the first three columns of **TABLE 3.4**.

ABO blood groups are important in medicine because of the need for blood transfusions. A crucial feature of the ABO system is that most human blood contains antibodies to either the A or the B polysaccharide. An **antibody** is a protein made by the immune system in response to a stimulating molecule called an **antigen** and is capable of binding to the antigen. An antibody is usually specific in that it recognizes only one antigen. Some antibodies combine with antigen and form large molecular aggregates that may precipitate.

Antibodies act to defend against invading viruses and bacteria. Although antibodies do not normally form without prior stimulation by the antigen, people capable of producing anti-A and anti-B antibodies do produce them. Production of these antibodies may be stimulated by antigens that are similar to polysaccharides A and B and that are present on the surfaces of many common bacteria. However, a mechanism called *tolerance* prevents an organism from producing antibodies against its own antigens. This mechanism ensures that A antigen or B antigen elicits antibody production only in people whose own red blood cells do not contain A or B, respectively. The end result:

> People of blood type O make both anti-A and anti-B antibodies, those of blood type A make anti-B antibodies, those of blood type B make anti-A antibodies, and those of blood type AB make neither type of antibody.

The antibodies found in the blood fluid of people with each of the ABO blood types are shown in the fourth column in Table 3.4. The clinical significance of the ABO blood groups is that transfusion of blood containing A or B red-cell antigens into persons who make antibodies against them results in an agglutination reaction in which the donor red blood cells are clumped. In this reaction, the anti-A antibody agglutinates red blood cells of either blood type A or blood type AB, because both carry the A antigen. Similarly, anti-B antibody agglutinates red blood cells of either blood type B or blood type AB. When the blood cells agglutinate, many blood vessels are blocked, and the recipient of the transfusion goes into shock and may die. Incompatibility in the other direction, in which the donor blood contains antibodies against the recipient's red blood cells, is usually acceptable because the donor's antibodies are diluted so rapidly that clumping is avoided. The types of compatible blood transfusions are shown in the last two columns of Table 3.4. Note that a person of blood type AB can receive blood from a person of any other ABO type; type AB is called

Table 3.4 Genetic control of the human ABO blood groups

Genotype	Antigens present on red blood cells	ABO blood group phenotype	Antibodies present in blood fluid	Blood types that can be tolerated in transfusion	Blood types that can accept blood for transfusion
I^AI^A	A	Type A	Anti-B	A & O	A & AB
I^AI^O	A	Type A	Anti-B	A & O	A & AB
I^BI^B	B	Type B	Anti-A	B & O	B & AB
I^BI^O	B	Type B	Anti-A	B & O	B & AB
I^AI^B	A & B	Type AB	Neither anti-A nor anti-B	A, B, AB & O	AB only
I^OI^O	Neither A nor B	Type O	Anti-A & anti-B	O only	A, B, AB & O

a *universal recipient.* Conversely, a person of blood type O can donate blood to a person of any ABO type; type O is called a *universal donor.*

Epistasis

In Chapter 1 we saw how the products of several genes may be necessary to carry out all the steps in a biochemical pathway. In genetic crosses in which two mutations that affect different steps in a single pathway are both segregating, the typical F_2 ratio of 9 : 3 : 3 : 1 is not observed. Gene interaction that perturbs the normal Mendelian ratios is known as **epistasis.** One type of epistasis is illustrated by the interaction of the *C, c* and *P, p* allele pairs affecting flower coloration in peas. These genes encode enzymes in the biochemical pathway for the synthesis of anthocyanin pigment, and the production of anthocyanin requires the presence of at least one wildtype dominant allele of each gene. The proper way to represent this situation genetically is to write the required genotype as

$$C\text{—}\ P\text{—}$$

where each dash is a "blank" that may be filled with either allele of the gene. Hence *C*— comprises the genotypes *CC* and *Cc*, and likewise *P*— comprises the genotypes *PP* and *Pp*. All four genotypes are included in the symbol *C*— *P*—, and only these genotypes, have purple flowers.

FIGURE 3.24 shows a cross between the homozygous recessive genotypes *CC pp* and *cc PP*. The phenotype of the flowers in the F_1 generation is the wildtype purple. Why? Because the *C* allele is dominant to *c* and the *P* allele is dominant to *p*, the F_1 plant is a double heterozygote of genotype *Cc Pp* and therefore has purple flowers. The result is nevertheless strange, because the original cross involved two homozygous recessive mutants, each of which had white flowers. Once we have been told that the mutant *c* and the mutant *p* alleles are in different genes, the finding of wildtype flowers in the F_1 generation is logical. But what if we did not know whether the mutant alleles were in different genes? What if both mutants were picked up in a mutant screen, or discovered in different laboratories, and we did not know whether the *c* allele and the *p* allele were alleles of the same gene or alleles of different genes? The answer is that the phenotype of the F_1 progeny in a cross like that in Figure 3.24 tells you. If the phenotype of the F_1 progeny is wildtype, as in Figure 3.24, this observation tells you that *c* and *p* are alleles of *different* genes. On the other hand, if the phenotype of the F_1 progeny is mutant (in this case a plant with white flowers), this result tells you that *c* and *p* are alleles of the *same* gene. In the latter case it would be best to change their names in such a way as to signify their allelism, such as $c = c_1$ and $p = c_2$, because then the genotype of the F_1 plant would be written as c_1/c_2 and it would be obvious from the genotype symbol that the phenotype is mutant, because each allele of the gene is mutant.

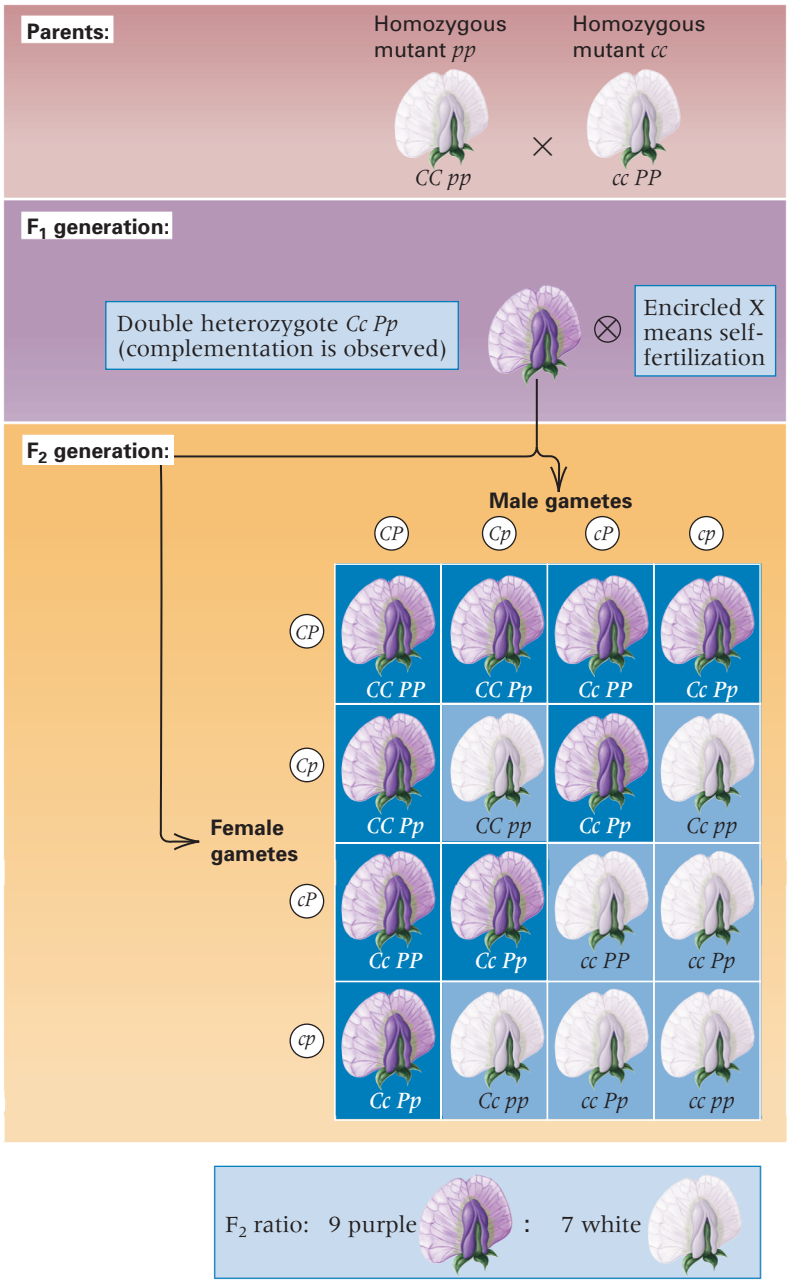


FIGURE 3.24 Epistasis in the determination of flower color in peas. Formation of the purple pigment requires the dominant allele of both the *C* and *P* genes. With this type of epistasis, the dihybrid F_2 ratio is modified to 9 purple : 7 white.

This discussion will hopefully remind you of the Beadle-Tatum experiments in Chapter 1

(Section 1.4) and in particular their use of the *complementation test* to determine whether two nutritional mutants in *Neurospora* were, or were not, mutants of the same gene. The principle in Figure 3.24 is exactly the same. The difference is that Beadle and Tatum used *Neurospora* heterokaryons, in which mutant alleles are brought together by forming hybrid filaments with two types of haploid mutant nuclei, whereas in diploid organisms like the peas in Figure 3.24 the hybrid nuclei are created by crossing two homozygous recessive mutants. The main point is that complementation tests are also used in sexual diploid organisms to identify which recessive mutations are alleles and which are not, and the phenotype of the F_1 generation in Figure 3.24 illustrates the principle.

Whereas the F_1 generation in Figure 3.24 illustrates complementation, the F_2 generation illustrates one type of epistasis. Suppose that plants of the F_1 generation are self-fertilized (indicated by the encircled cross sign), and assume that the (*C*, *c*) and (*P*, *p*) allele pairs undergo independent assortment. The Punnett square in Figure 3.24 gives the genotypes and phenotypes of the F_2 generation. Because only the *C*— *P*— progeny have purple flowers, the ratio of purple flowers to white flowers in the F_2 generation is 9 : 7. The epistasis does not change the result of independent segregation, it merely conceals the fact that the underlying ratio of the genotypes *C*— *P*— : *C*— *pp* : *cc P*— : *cc pp* is 9 : 3 : 3 : 1.

For a trait determined by the interaction of two genes, each with a dominant allele, there are only a limited number of ways in which the 9 : 3 : 3 : 1 dihybrid ratio can be modified. The possibilities are illustrated in **FIGURE 3.25**. In the absence of epistasis, the F_2 ratio of phenotypes is 9 : 3 : 3 : 1. In each row, different colors indicate different phenotypes. For example, in the modified ratio at the bottom, the phenotypes of the "3 : 3 : 1" classes are indistinguishable, resulting in a 9 : 7 ratio. This is the ratio observed in the segregation of the *C*, *c* and *P*, *p* alleles in Figure 3.24, with its 9 : 7 ratio of purple to white flowers. Taking all the possible modified ratios in Figure 3.25 together, there are nine possible dihybrid ratios when both genes show complete dominance. Examples of each of the modified ratios are known. The types of epistasis that result in these modified ratios are illustrated in the following examples, which are taken from a variety of organisms. Other examples can be found in the problems at the end of the chapter.

9 : 7 is observed when a homozygous recessive mutation in either or both of two different genes results in the same mutant phenotype, as in Figure 3.24.

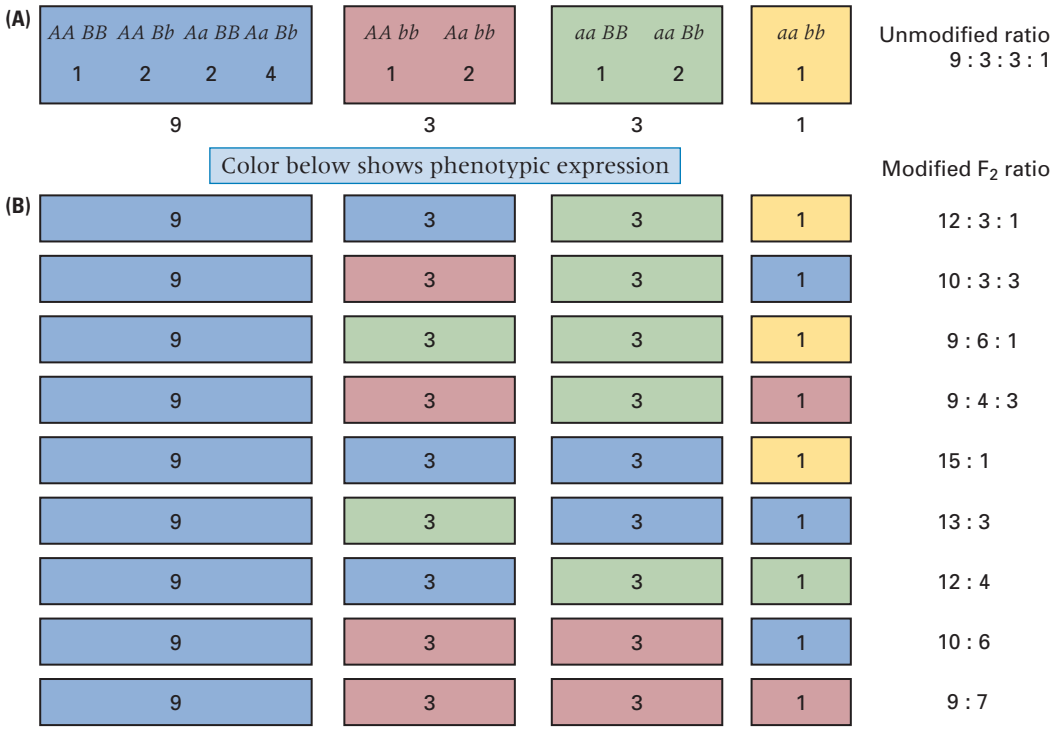


FIGURE 3.25 Modified F_2 dihybrid ratios. In each row, different colors indicate different phenotypes.

12 : 3 : 1 results when the presence of a dominant allele at one locus masks the genotype at a different locus, such as the *A*— genotype rendering the *B*— and *bb* genotypes indistinguishable. For example, in genetic study of the color of the hull in oat seeds, a variety with white hulls was crossed with a variety with black hulls. The F_1 hybrid seeds had black hulls. Among 560 progeny in the F_2 generation, the hull phenotypes observed were 418 black, 106 gray, and 36 white. The ratio of phenotypes is 11.6 : 3.9 : 1, or very nearly 12 : 3 : 1. A genetic hypothesis to explain these results is that the black-hull phenotype is due to the presence of a dominant allele *A* and the gray-hull phenotype is due to another dominant allele *B* whose effect is apparent only in *aa* genotypes. On the basis of this hypothesis, the original varieties had genotypes *aa bb* (white) and *AA BB* (black). The F_1 has genotype *Aa Bb* (black). If the *A, a* allele pair and the *B, b* allele pair undergo independent assortment, then the F_2 generation is expected to have the genotypic and phenotypic composition 9/16 *A*— *B*— (black hull), 3/16 *A*— *bb* (black hull), 3/16 *aa B*— (gray hull), 1/16 *aa bb* (white hull), or 12 : 3 : 1.

13 : 3 is illustrated by the difference between White Leghorn chickens (genotype *CC II*) and White Wyandotte chickens (genotype *cc ii*). Both breeds have white feathers because the *C* allele is necessary for colored feathers, but the *I* allele in White Leghorns is a dominant inhibitor of feather coloration. The F_1 generation of a dihybrid cross between these breeds has the genotype *Cc Ii*, which is expressed as white feathers because of the inhibitory effects of the *I* allele. In the F_2 generation, only the *C*— *ii* genotype has colored feathers, so there is a 13 : 3 ratio of white : colored.

9 : 4 : 3 is observed when homozygosity for a recessive allele with respect to one gene masks the expression of the genotype of a different gene. For example, if the *aa* genotype has the same phenotype regardless of whether the genotype is *B*— or *bb*, then the 9 : 4 : 3 ratio results. As an example, in mice the grayish "agouti" coat color results from a horizontal band of yellow pigment just beneath the tip of each hair. The agouti pattern is due to a dominant allele *A*, and in *aa* animals the coat color is black. A second dominant allele, *C*, is necessary for the formation of hair pigments of any kind, and *cc* animals are albino (white). In a cross of *AA CC* (agouti) × *aa cc* (albino), the F_1 progeny are *Aa Cc* and phenotypically agouti. Crosses between F_1 males and females produce F_2 progeny in the proportions 9/16 *A*— *C*— (agouti), 3/16 *A*— *cc* (albino), 3/16 *aa C*— (black), 1/16 *aa cc* (albino), or 9 agouti : 4 albino : 3 black.

9 : 6 : 1 implies that homozygosity for either of two recessive alleles yields the same phenotype but that the phenotype of the double homozygote is different. In Duroc–Jersey pigs, red coat color requires the presence of two dominant alleles *R* and *S*. Pigs of genotype *R*— *ss* and *rr S*— have sandy-colored coats, and *rr ss* pigs are white. The F_2 ratio is therefore 9/16 *R*— *S*— (red), 3/16 *R*— *ss* (sandy), 3/16 *rr S*— (sandy), 1/16 *rr ss* (white), or 9 red : 6 sandy : 1 white.

CHAPTER SUMMARY

- Inherited traits are determined by the genes present in the reproductive cells united in fertilization.
- Genes are usually inherited in pairs: one from the mother and one from the father.
- The genes in a pair may differ in DNA sequence and in their effect on the expression of a particular inherited trait.
- The maternally and paternally inherited genes are not changed by being together in the same organism.
- In the formation of reproductive cells, the paired genes separate again into different cells.
- Random combinations of reproductive cells containing different genes result in Mendel's ratios of traits appearing among the progeny.
- The ratios actually observed for any trait are determined by the types of dominance and gene interaction.

REVIEW THE BASICS

- What is the principle of segregation, and how is this principle demonstrated in the results of a single-gene (monohybrid) cross?
- What is the principle of independent assortment, and how is this principle demonstrated in the results of a two-gene (dihybrid) cross?
- Explain why random union of male and female gametes is necessary for Mendelian segregation and independent assortment to be observed in the progeny of a cross.
- What is the difference between mutually exclusive events and independent events? How are the probabilities of these two types of events combined? Give two examples of genetic events that are mutually exclusive and two examples of genetic events that are independent.
- When two pairs of alleles show independent assortment, under what conditions will a 9 : 3 : 3 : 1 ratio of phenotypes in the F_2 generation *not* be observed?
- Explain the following statement: "Among the F_2 progeny of a dihybrid cross, the ratio of genotypes is 1 : 2 : 1, but among the progeny that express the dominant phenotype, the ratio of genotypes is 1 : 2."
- What are the principal features of human pedigrees in which a rare dominant allele is segregating? In which a rare recessive allele is segregating?
- What is a mutant screen and how is it used in genetic analysis?
- Explain the statement: "In genetics, a gene is identified experimentally by a set of mutant alleles that fail to show complementation." What is complementation? How does a complementation test enable a geneticist to determine whether two different mutations are or are not mutations in the same gene?
- What does it mean to say that epistasis results in a "modified dihybrid F_2 ratio?" Give two examples of a modified dihybrid F_2 ratio, and explain the gene interactions that result in the modified ratio.

GUIDE TO PROBLEM SOLVING

Problem 1 Complete the table by inserting 0, 1/4, 1/2, or 1 for the probability of each genotype of the progeny from each type of mating. For which mating are the parents identical in genotype but the progeny as variable in the genotype as they can be for a single locus? For which mating are the parents as different as they can be for a single locus but the progeny identical to each other and different from either parent?

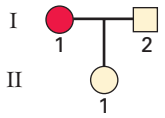
	Progeny genotypes		
Mating	***AA***	***Aa***	***aa***
AA × *AA*			
AA × *Aa*			
AA × *aa*			
Aa × *Aa*			
Aa × *aa*			
aa × *aa*			

Answer This kind of table is fundamental to being able to solve almost any type of quantitative problem in transmission genetics.

	Progeny genotypes		
Mating	***AA***	***Aa***	***aa***
AA × *AA*	1	0	0
AA × *Aa*	1/2	1/2	0
AA × *aa*	0	1	0
Aa × *Aa*	1/4	1/2	1/4
Aa × *aa*	0	1/2	1/2
aa × *aa*	0	0	1

The mating for which the parents are identical in genotype, but the progeny are as variable in genotype as they can be for a single locus, is *Aa* × *Aa*. The mating for which the parents are as different as they can be for a single locus, but the progeny identical to each other and different from either parent, is *AA* × *aa*

Problem 2 The pedigree below shows the inheritance of a rare, simple Mendelian dominant mutation with penetrance equal to 1/3. (A penetrance of 1/3 means that 1/3 of the individuals with the mutant genotype actually express the mutant phenotype.) The woman I-1 necessarily has the genotype *Mm*, where *M* is the mutant allele. Because the mutant allele is rare, you may assume that the male I-2 has the genotype *mm*. Individual II-1 is not affected. What is the probability that II-1 has the genotype *Mm*?



Answer This is a typical problem that makes use of Bayes' theorem. Let *A* be the event that individual II-1 has genotype *Mm*, and let *B* be the event that individual II-1 is not affected. Then event *A'* is therefore the event that individual II-1 has the genotype *mm*. Because I-2 has genotype *Mm*, then Pr{A} = 1/2 and also Pr{*A'*} = 1/2. Now we can apply Bayes' theorem:

$$\Pr\{A \mid B\} = \frac{\Pr\{B \mid A\}\Pr\{A\}}{\Pr\{B \mid A\}\Pr\{A\} + \Pr\{B \mid A'\}\Pr\{A'\}}$$

where Pr{$B \mid A$} is the probability that an individual of genotype *Mm* is not affected and Pr{$B \mid A'$} is the probability that an individual of genotype *m*m is not affected. Because the penetrance is 1/3, then Pr{$B \mid A$} = 2/3. Because the genotype *mm* is never affected, P{$B \mid A'$} = 1. Putting all this together, we obtain the answer we were seeking:

$$\Pr\{A \mid B\} = \frac{(2/3)(1/2)}{(2/3)(1/2) + (1)(1/2)} = 2/5$$

Another approach avoids the machinery of Bayes' theorem. Because individual I-1 has genotype *Mm*, the possible offspring of I-1 are of three types: (1) *Mm* and affected with probability 1/2 × 1/3 = 1/6, (2) *Mm* and not affected with probability 1/2 × 2/3 = 1/3, and (3) *mm* and not affected with probability 1/2. Because we know that II-1 is not affected, possibility (1) can be ruled out, and so the probability that II-1 has genotype *Mm*, given that she is not affected, is given by (1/3)/[(1/3 + 1/2)] = 2/5, which agrees with the answer above.

GUIDE TO PROBLEM SOLVING, CONT.

Problem 3 In domesticated chickens, a dominant allele *C* is required for colored feathers, but a dominant allele *I* of an unlinked gene is an inhibitor of color that overrides the effects of *C*. White Leghorns have genotype *CC II* whereas White Wyandottes have genotype *cc ii*. Both breeds are white, but for different reasons. In the F_2 generation of a cross between White Leghorns and White Wyandottes:

(a) What is the phenotypic ratio of white : colored?

(b) Among F_2 chicks that are white, what is the proportion of *Cc Ii*?

Answer

(a) The initial cross of *CC II* × *cc ii* will yield F_1 chickens with genotype *Cc Ii*. Because the genes are not linked, the offspring in the F_2 generation will have a ratio of the genotypes 9 *C– I–* : 3 *C– ii* : 3 *cc I–* : 1 *cc ii*. Only the chickens with genotype *C– ii* will be colored because *C* is required for colored feathers but *I* inhibits the effect of *C*. This type of interaction of genes, or *epistasis*, will lead to a phenotypic ratio of 13 white : 3 colored.

(b) In any dihybrid cross with unlinked genes, the probability that an F_2 progeny is heterozygous for both genes equals (1/2) × (1/2) = 1/4. Because only 13/16 of the chicks in this example are white, the proportion of *Cc Ii* among the white chicks is (1/4)/(13/16) = 4/13. You can also get this result by counting squares in the Punnett square for a dihybrid cross: A total of 13 genotypes yield white chicks, and among these four are heterozygous for both genes.

Problem 4 In Duroc-Jersey pigs, animals with genotype *R– S–* are red, those with genotype *R– ss* or *rr S–* are sandy colored, and those with the genotype *rr ss* are white. The genes show independent assortment. In the F_2 generation of a cross between the genotypes *RR ss* and *rr SS*:

(a) What is the phenotypic ratio of red : nonred?

(b) Among F_2 piglets that are red, what is the proportion of *Rr Ss*?

Answer

(a) The initial cross *RR ss* × *rr SS* will yield F_1 pigs with genotype *Rr Ss*. The progeny of the F_2 generation will have the genotypic ratio 9 *R– S–* : 3 *R– ss* : 3 *rr S–* : 1 *rr ss*. Only the pigs with genotype *R– S–* will be red, while all the others will be nonred (some sandy and some white, in the ratio of 6 : 1). This type of interaction of two genes (epistasis) will lead to a phenotypic ratio of 9 red : 7 nonred.

(b) 4/9. The proportion of pigs with the genotype *Rr Ss* is the frequency of those heterozygous for both genes, which equals 1/2 × 1/2 = 1/4. Because only 9/16 of the pigs are red, the proportion of *Rr Ss* among them is (1/2 × 1/2)/(9/16) = 4/9. You can also get this answer by counting squares in the Punnett square for a dihybrid cross: A total of 9 genotypes yield red pigs, among which 4 are heterozygous for both genes.

ANALYSIS AND APPLICATIONS

3.1 In the cross *Aa Bb Cc Dd* × *Aa Bb Cc Dd*, in which all genes undergo independent assortment, what proportion of offspring are expected to be heterozygous for all four genes?

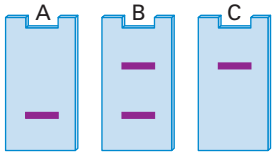
3.2 Consider a gene with four alleles, A_1, A_2, A_3, and A_4. In a cross between A_1A_2 and A_3A_4, what is the probability that a particular offspring inherits either A_1 or A_3 or both?

3.3 Assuming equal numbers of boys and girls, if a mating has already produced a girl, what is the probability that the next child will be a boy? If a mating has already produced two girls, what is the probability that the next child will be a boy? On what type of probability argument do you base your answers?

3.4 A cross is carried out between genotypes *Aa BB Cc dd Ee* and *Aa Bb cc DD Ee*. How many genotypes of progeny are possible?

3.5 An individual of genotype *AA Bb Cc DD Ee* is testcrossed. Assuming that the loci undergo independent assortment, what fraction of the progeny are expected to have the genotype *Aa Bb Cc Dd Ee*?

3.6 Homozygous pea plants with round seeds are crossed to homozygous pea plants with wrinkled seeds. The F_1 progeny undergo self-fertilization. A single round seed is chosen at random from the F_2 progeny, and its DNA examined by electrophoresis as described in the text. What is the probability that the observed gel pattern will be B?



3.7 Assume that the trihybrid cross *MM SS tt* × *mm ss TT* is made in a plant species in which *M* and *S* are dominant but there is no dominance between *T* and *t*. Consider the F_2 progeny from this cross, and assume independent assortment.

(a) How many phenotypic classes are expected?

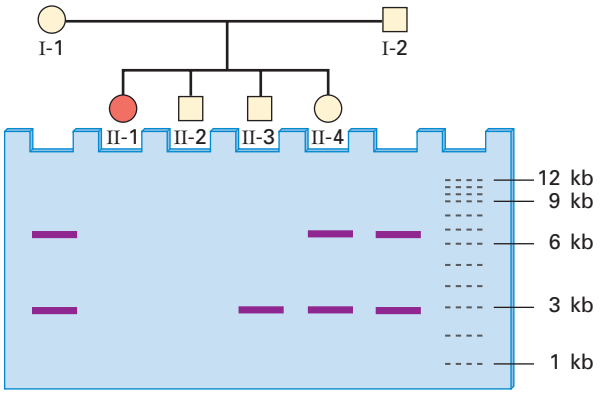
(b) What is the probability of genotype *MM SS tt*?

(c) What proportion would be expected to be homozygous for all genes?

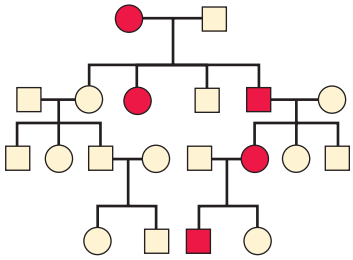
3.8 Shown here is a pedigree and a gel diagram indicating the clinical phenotypes with respect to phenylketonuria and the molecular phenotypes with respect to an RFLP that overlaps the *PAH* gene for phenylalanine hydroxylase. The individual II-1 is affected.

(a) Indicate the expected molecular phenotype of II-1.

(b) Indicate the possible molecular phenotypes of II-2.



3.9 What mode of inheritance is suggested by the following pedigree? Based on this hypothesis, and assuming that the trait is rare and has complete penetrance, what are the possible genotypes of all individuals in this pedigree?

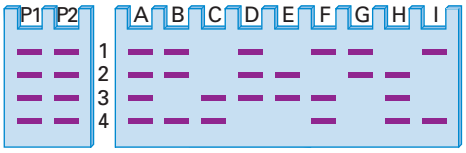


3.10 Huntington disease is a rare neurodegenerative human disease determined by a dominant allele, *HD*. The disorder is usually manifested after the age of *45*. A young man has learned that his father has developed the disease.

(a) What is the probability that the young man will later develop the disorder?

(b) What is the probability that a child of the young man carries the *HD* allele?

3.11 The gel diagram below shows the banding patterns observed among the progeny of a cross between double heterozygous genotypes shown at the left (P1 and P2). On the right are shown the patterns of bands observed among the progeny. The bands are numbered 1–4.



(a) Identify which pairs of numbered bands correspond to the two segregating pairs of alleles.

(b) Assuming that the two pairs of alleles undergo independent assortment, what is the probability that an offspring of the cross shows the banding pattern in lane D?

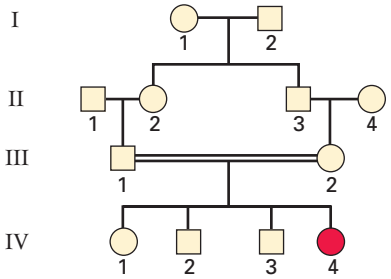
3.12 Assume that the trait in the accompanying pedigree is due to simple Mendelian inheritance.

(a) Is it likely to be due to a dominant allele or a recessive allele? Explain.

(b) What is the meaning of the double horizontal line connecting III-1 with III-2?

(c) What is the biological relationship between III-1 and III-2?

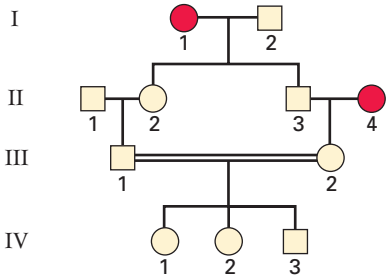
(d) If the allele responsible for the condition is rare, what are the most likely genotypes of all of the persons in the pedigree in generations I, II, and III? (Use *A* and *a* for the dominant and recessive alleles, respectively.)



3.13 In *Drosophila*, the mutant allele bw^{dts} causing brown eyes (normal eyes are red) is temperature sensitive. In flies reared at 29°C the mutant allele is dominant, but in flies reared at 22°C the mutant allele is recessive. In a cross of $bw^{dts}/+ \times bw^{dts}/+$, where the + sign denotes the wildtype allele of bw^{dts}, what is the expected ratio of brown-eyed flies to red-eyed flies if the progeny are reared at 29°C? At 22°C?

3.14 Pedigree analysis tells you that a particular parent may have the genotype *AA BB* or *AA Bb*, each with the same probability. Assuming independent assortment, what is the probability of this parent to produce an *A b* gamete? What is the probability of the parent to produce an *A B* gamete?

3.15 The pedigree shown here is for a rare autosomal recessive trait with complete penetrance. You may assume that no one the pedigree has the recessive allele unless that person inherits it from either I-1 or II-4 or both.

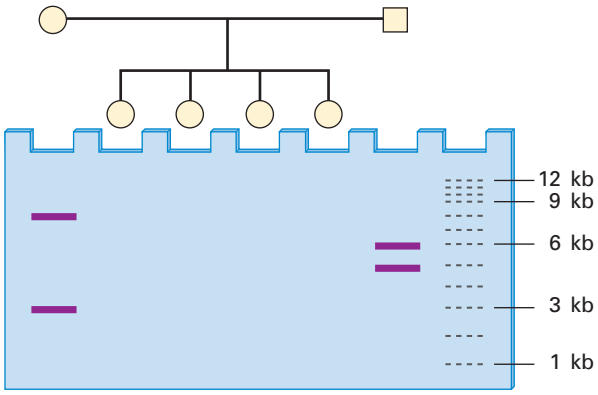


(a) Ignoring the sibship in generation IV, what is the probability that both parents in the first cousin mating are heterozygous?

(b) Taking the sibship in generation IV into account, what is the probability that both parents in the first cousin mating are heterozygous?

3.16 A clinical test for a certain disease is 100 percent accurate in individuals who are affected but also yields a false positive result in 0.2 percent of healthy individuals. If the proportion of affected individuals in a population is 0.002, what is overall probability that an individual selected at random will yield a positive test result?

3.17 The accompanying pedigree and gel diagram shows the phenotypes of the parents for an RFLP that has multiple alleles. What are the possible phenotypes of the progeny, and in what proportions are they expected?



3.18 The pollen from round pea seeds obtained from the F_2 generation of a cross between a true-breeding strain with round seeds and a true-breeding strain with wrinkled seeds was collected and used *en masse* to fertilize plants from the true-breeding wrinkled strain. What fraction of the progeny is expected to have wrinkled seeds? (Assume equal fertility among all genotypes.)

3.19 Heterozygous *Cp cp* chickens express a condition called creeper, in which the leg and wing bones are shorter than normal (*cp cp*). The dominant *Cp* allele is lethal when homozygous. Two alleles of an independently segregating gene determine white (*W*−) versus yellow (*ww*) skin color. From matings between chickens heterozygous for both of these genes, what phenotypic classes will be represented among the viable progeny, and what are their expected relative frequencies?

3.20 The F_2 progeny from a particular cross exhibit a modified dihybrid ratio of 9 : 7 (instead of 9 : 3 : 3 : 1). What phenotypic ratio would be expected from a testcross of the F_1 progeny?

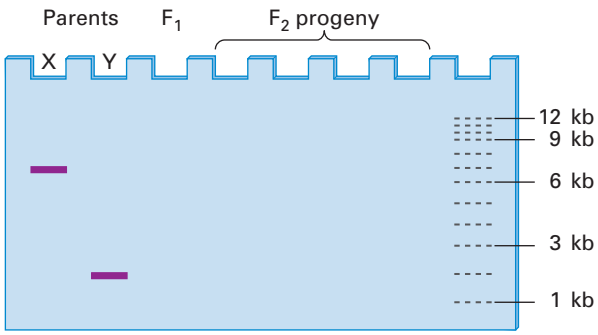
3.21 In the mating $Aa \times Aa$, what is the smallest number of offspring, *n*, for which the probability of at least one *aa* offspring exceeds 95 percent?

3.22 A woman is affected with a trait due to a dominant mutant allele that shows 50% penetrance. If she has a child, what is the probability that it will be affected?

3.23 The pattern of coat coloration in dogs is determined by the alleles of a single gene, with *S* (solid) being dominant over *s* (spotted). Black coat color is determined by the dominant allele *A* of a second gene, and homozygous recessive *aa* animals are tan. A female having a solid tan coat is mated with a male having a solid black coat and produces a litter of six pups. The phenotypes of the pups are 2 solid tan, 2 solid black, 1 spotted tan, and 1 spotted black. What are the genotypes of the parents?

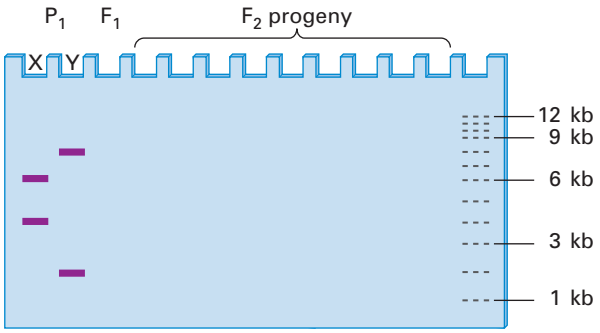
3.24 Consider a phenotype for which the allele *N* is dominant to the allele *n*. A mating $Nn \times Nn$ is carried out, and one individual with the dominant phenotype is chosen at random. This individual is testcrossed and the mating yields four offspring, each with the dominant phenotype. What is the probability that the parent with the dominant phenotype has the genotype *Nn*?

3.25 Some polymorphisms can be identified using oligonucleotides with randomly chosen sequences as primers. These are known and RAPD polymorphisms, where RAPD stands for *randomly amplified polymorphic DNA*. Typically, a RAPD polymorphism results from a site in which only some of the chromosomes in a population will bind with the primers and yield a band, whereas the other chromosomes will not bind with the primers and so not yield a band. DNA from an individual who is heterozygous for the site will yield a band. The gel diagram shown here includes the phenotype of two parents (X and Y) with respect to two RAPD polymorphisms corresponding to different sites in the genome. Each parent is homozygous for the site associated with the RAPD band its DNA exhibits. The two RAPD polymorphisms result from amplification of different sites in the genome that undergo independent assortment. In the gel diagram, indicate the expected phenotype of the F_1 progeny as well as all possible phenotypes of the F_2 progeny along with their expected proportions.



3.26 The accompanying gel diagram shows the phenotype of two parents (X and Y), each homozygous for two RFLPs that undergo independent assortment. Parent X has genotype $A_1A_1\,B_1B_1$, where the A_1 allele yields a band of 4 kb and the B_1 allele yields a band of 6 kb, and parent Y has genotype $A_2A_2\,B_2B_2$, where the A_2 allele yields a band of 8 kb and the B_2 allele yields a band of 2 kb. Show the expected phenotype

of the F_1 progeny as well as all possible phenotypes of the F_2 progeny along with their expected proportions.



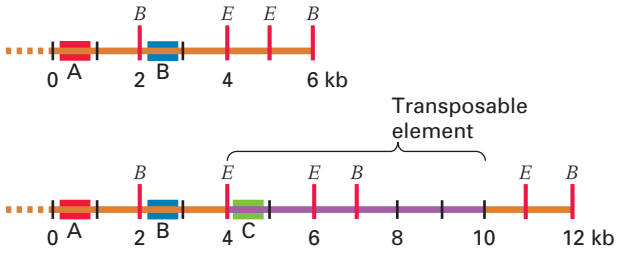
3.27 Complementation tests of the recessive mutant genes *a* through *f* produced the data in the accompanying matrix. The circles represent missing data. Assuming that all of the missing mutant combinations would yield data consistent with the entries that are known, complete the table by filling each circle with a + or − as needed.

	a	*b*	*c*	*d*	*e*	*f*
a	○	+	−	○	+	○
b	○	○	○	○	−	
c	○	−	○	○		
d	○	○	○			
e	○	+				
f	○					

3.28 In plants, certain mutant genes are known that affect the ability of gametes to participate in fertilization. Suppose that an allele *A* is such a mutation and that pollen cells bearing the *A* allele are only half as likely to survive and participate in fertilization as pollen cells bearing the *a* allele. Complete the Punnett square for the F_2 generation in a monohybrid cross. What is the expected ratio of *AA* : *Aa* : *aa* plants in the F_2 generation?

CHALLENGE PROBLEMS

Challenge Problem 1 Diagrammed here is DNA from a wildtype gene (top) and a mutant allele (bottom) that has an insertion of a transposable element that inactivates the gene. The transposable element is present in many copies scattered throughout the genome. The symbols *B* and *E* represent the positions of restriction sites for *Bam*HI and *Eco*RI, respectively, and the rectangles show sites of hybridization with each of three probes (A, B, and C) that are available. The dots at the left indicate that the nearest site of either *Bam*HI or *Eco*RI cleavage is very far to the left of the region shown. Explain which probe and which single restriction enzyme you would use for RFLP analysis to identify both alleles. Also, explain why any other choices would be unsuitable.



Challenge Problem 2 Meiotic drive is an unusual phenomenon in which two alleles do not show Mendelian segregation from the heterozygous genotype. Examples are known from mammals, insects, fungi, and other organisms. The usual mechanism is one in which both types of gametes are formed, but one of them fails to function normally. The excess of the driving allele over the other can range from a small amount to nearly 100 percent. Suppose that *D* is an allele showing meiotic drive against its alternative allele *d*, and suppose that *Dd* heterozygotes produce functional *D*-bearing and *d*-bearing gametes in the proportions 3/4 : 1/4. In the mating *Dd* × *Dd*,

(a) What are the expected proportions of *DD*, *Dd*, and *dd* genotypes?

(b) If *D* is dominant, what are the expected proportions of *D*− and *dd* phenotypes?

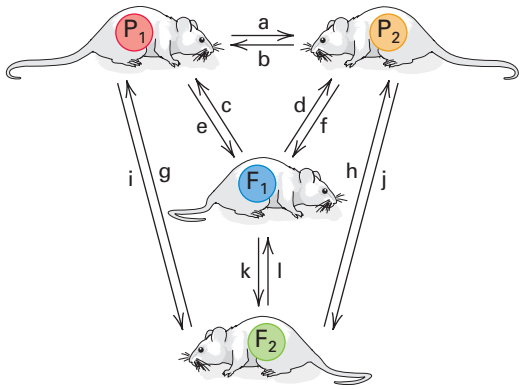
(c) Among the *D*− phenotypes, what is the ratio of *DD* : *Dd*?

(d) Answer parts (a) through (c), assuming that the meiotic drive takes place in only one sex.

Challenge Problem 3 The accompanying table summarizes the effect of inherited tissue antigens on the acceptance or rejection of transplanted tissues, such as skin grafts, in mammals. The tissue antigens are determined in a codominant fashion, so that tissue taken from a donor of genotype *Aa* carries both the A and the a antigen. In the table, the + sign means that a graft of donor tissue is accepted by the recipient, and the − sign means that a graft of donor tissue is rejected by the recipient. The rule is: *Any graft will be rejected whenever the donor tissue contains an antigen not present in the recipient.* In other words, any transplant will be accepted if, and only if, the donor tissue does *not* contain an antigen different from any already present in the recipient.

		Donor		
		AA	*Aa*	*aa*
Recipient	*AA*	+	−	−
	Aa	+	+	+
	aa	−	−	+

The diagram illustrated here shows all possible skin grafts between inbred (homozygous) strains of mice (P_1 and P_2) and their F_1 and F_2 progeny. Assume that the inbred lines P_1 and P_2 differ in only one tissue-compatibility gene. For each of the arrows, what is the probability of acceptance of a graft in which the donor is an animal chosen at random from the population shown at the base of the arrow and the recipient is an animal chosen at random from the population indicated by the arrowhead?



GENETICS *on the web*

GeNETics on the Web will introduce you to some of the most important sites for finding genetics information on the Internet. To explore these sites, visit the Jones and Bartlett companion site to accompany Genetics: Analysis of Genes and Genomes, Seventh Edition at http://biology.jbpub.com/book/genetics.

There you will find a chapter-by-chapter list of highlighted keywords. When you select one of the keywords, you will be linked to a Web site containing information related to that keyword.