CHAPTER 11

Population Structure and Genetic Drift

This chapter treats two profoundly important facts: natural populations, unlike ideal populations at Hardy-Weinberg equilibrium, are finite in size; and species are structured, by geographic or other factors, so that mating is not entirely random. In finite populations, allele frequencies can fluctuate by chance, a process called **random genetic drift**. Random fluctuations can result in the replacement of old alleles by new ones, resulting in NONADAPTIVE evolution. Genetic drift and natural selection are the two important potential causes of allele substitution,—that is, of evolution within populations.

Adaptations do not result from genetic drift, so this process is not responsible for many of the most interesting features of organisms. Genetic drift nevertheless has many important consequences, some of which can influence the course of adaptation. The topic of genetic drift is intimately related to POPULATION STRUCTURE, especially the subdivision of species into local breeding units (Figure 11.1), which exchange genes to a greater or lesser degree. Such structuring has many important consequences.

Because all populations are finite, alleles at all loci are potentially subject to random genetic drift—but all are not necessarily subject to natural selection. For this reason, and because genetic drift constitutes evolution by chance alone, some evolutionary geneticists feel that genetic drift should be the "null hypothesis" used to explain an evolutionary observation unless there is positive evidence for natural selection or some other factor. This perspective is analogous to the "null hypothesis" in statistics: the hypothesis that the data do not depart from those expected on the basis of chance alone.* According to this view, we should not assume that a characteristic, or a difference between populations or species, is adaptive or evolved by natural selection unless there is evidence for this conclusion.

Genetic drift and population structure are the subjects of some of the most highly refined mathematical models in population genetics—or in all of biology, for that matter. Much of the theory was developed by Sewall Wright, starting in the 1930s, and by Motoo Kimura, starting in the 1950s. We will present the material as verbally as possible, but readers who enjoy mathematics can consult the boxes in this chapter for a taste of the models. (See Hartl and Clark 1997 or Crow and Kimura 1970 for more extensive treatments.) For each topic, we will begin with theoretical considerations, and then consider empirical data.

In our discussion of the theory of genetic drift, we will describe random fluctuations in the frequencies (proportions) of two or more kinds of self-reproducing entities that do not differ *on average* (or differ very little) in reproductive success (fitness). For the purposes of this chapter, the entities are alleles. But exactly the same theory applies to any other self-replicating entities, such as chromosomes, asexually reproducing genotypes, or even different species.

The Theory of Genetic Drift

Sampling error That chance should affect allele frequencies is readily understandable. Imagine, for example, that a single mutation, A_2, appears in a large population that is otherwise A_1. If the population size is stable, each mating pair leaves an average of two progeny that survive to reproductive age. From the single mating $A_1A_1 \times A_1A_2$ (for there is only one copy of A_2), the probability that two surviving progeny are both A_1A_1 is $(1/2)(1/2) = 1/4$, which is the probability that the A_2 gene will be immediately lost. We may assume that pairs vary at random in the number of surviving offspring they leave (0, 1, 2, 3 ...), in which case, as

*For example, if we measure height in several samples of people, the null hypothesis is that the observed means differ only because of random sampling, and that the means of the populations from which the samples were drawn do not differ. A statistical test, such as a *t*-test or analysis of variance, is designed to show whether or not the null hypothesis can be rejected. It will be rejected if the sample means differ more than expected if samples were randomly drawn from a single population.

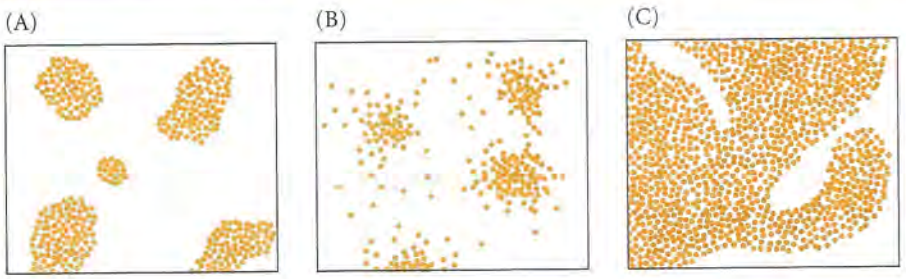


FIGURE 11.1 Some patterns of spatial population structure. Each dot represents an individual. (A) Discrete populations. (B) Perhaps the most common pattern in nature: ill-defined populations between which density is low. (C) A more or less uniform distribution.

the pioneering population geneticist Ronald Fisher (1930) calculated, the probability that A_2 will be lost, averaged over the population, is e^{-1}, or 0.368. In each subsequent generation, there is likewise a probability of loss; Fisher calculated that after 127 generations, the probability is 0.985. These calculations assume that the allele neither increases nor decreases the likelihood of survival of A_1A_2 relative to A_1A_1. However, even if A_1A_2 has a 1 percent advantage, the chance is only 0.027 that A_2 will still be in the population after 127 generations.*

This example illustrates that the frequency of an allele can change (in this instance, to zero from a frequency very near zero) purely by chance: the one or few copies of the A_2 allele might happen not to be included in those gametes that unite into zygotes, or might happen not to be carried by those few newborn individuals that survive to reproductive age. Similarly, if a population carries two (or more) alleles in any frequencies, their frequencies will change from one generation to the next because of random variation in the proportions that unite to form zygotes, variation in the number of offspring produced by carriers of the different alleles, and variation in the number that survive to reproduce. The gene copies in any generation of adult organisms represent a SAMPLE of the gene copies carried by the gametes of the previous generation; and any sample is subject to random variation, or SAMPLING ERROR.

*Fisher (1930) assumed that the number of surviving offspring per pair has a Poisson distribution with a mean of 2. Here, e is the base of natural logarithms, 2.718. You may wonder why Fisher calculated the probability of loss after such an odd number of generations as 127. So do I. Perhaps, since he didn't have a computer or calculator, he got tired of doing arithmetic and quit.

Coalescence The concept of genetic drift is so important that we will develop it by two theoretical approaches. Both theoretical perspectives will recur in later chapters. The first approach is coalescent theory.

Figure 11.2 shows a hypothetical, but realistic, history of lineages of reproducing objects. First, imagine the figure as a lineage of individual asexual organisms such as bacteria. We know from our own experience that not all members of our parents' or grandparents' generations had equal numbers of descendants; some had none. Figure 11.2 diagrams this familiar fact. We note that the individuals in generation t are the progeny of only some of those that existed in the previous generation $(t-1)$: purely by chance, some individuals in generation $t-1$ failed to leave descendants. Likewise, the population at $t-1$ stems from only some of those

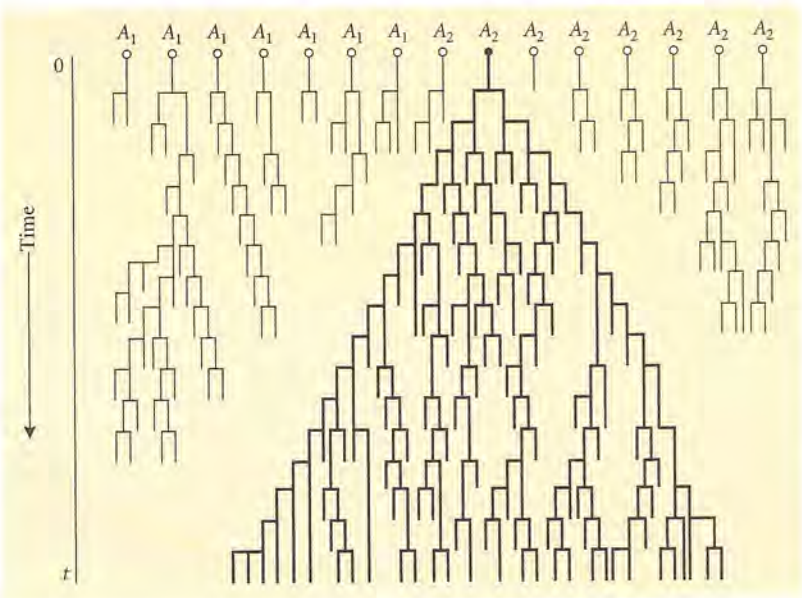


FIGURE 11.2 A possible history of descent of gene copies in a population that begins (at time 0) with 15 copies, representing two alleles. Each gene copy has 0, 1, or 2 descendants in the next generation. Gene copies present at time t are all descended from (coalesce to) a single ancestral copy, which happens to be an A_2 allele (this lineage is shown by the heavier black lines in the figure). Gene lineages descended from all other original gene copies have become extinct. If the failure of gene copies to leave descendants is random, the gene copies at time t could equally likely have descended from any of the other gene copies present at time 0. (After Hartl and Clark 1989.)

individuals that existed in generation $t-2$, and similarly back to the original population, only one member of which has descendants at time t.

Exactly the same model of descent applies to individual genes within a sexually reproducing population as would apply to individuals in an asexually reproducing haploid population; in other words, we can trace the descendants of a gene just as we would trace the descendants of an individual bacterium. Therefore, if we think of the objects in Figure 11.2 as copies of genes at a locus, in either a sexual or an asexual population, Figure 11.2 shows that as time goes on, more and more of the original gene lineages become extinct, so that the population consists of descendants of fewer and fewer of the original gene copies. This implies that *the average degree of relationship among individuals increases* with the passage of time: more and more of them can be traced back to the same common ancestors. In fact, if we look backward rather than forward in time, *all the gene copies* in the population ultimately *are descended from a single ancestral gene copy*, because given long enough, all other gene lineages become extinct. The genealogy of the genes in the present population is said to COALESCE to a single common ancestor. Because that ancestor represents one of the several original alleles, the population, descended entirely from that ancestor, must eventually become monomorphic: one or the other of the original alleles becomes **fixed** (reaches a frequency of 1.00). In our example, all gene copies have descended from a copy of an A_2 allele, but because this is a random process, A_1 might well have been the "lucky" allele if the sequence of random events had been different. If, in the generation that included the single common ancestor of all of today's gene copies, A_1 and A_2 had been equally frequent ($p = q = 0.5$), then it is equally likely that the ancestral gene copy would have been A_1 or A_2; if A_1 had had a frequency of 0.9 in that generation, then the probability is 0.9 that the ancestral gene would have been an A_1 allele. *Our analysis therefore shows that by chance, a population will eventually become monomorphic for one allele or the other, and that the probability that allele* A_1 *will be fixed, rather than another allele, equals the initial frequency of* A_1.

How long will this process take? Suppose the population in Figure 11.2 has a constant size of N gene copies (carried by N individual haploid organisms, or $N/2$ diploid organisms). If we were to pick two copies at random from the current population, the chance that the second came from the same parent copy as the first would be $1/N$. This is the probability that two gene copies coalesce to an ancestor in the previous generation. The probability that the two gene copies come from different parent copies is $1-(1/N)$. By the same reasoning, the probability that those two parents had the same parent is $1/N$, so the probability that the two current copies had the same "grandparent" is $[1-(1/N)]\times 1/N$. Thus the probability that they coalesce two generations back in time is $(1/N)[1-(1/N)]^{(2-1)}$. We can similarly calculate the probability that the two gene copies coalesce 3, 4, or in general G generations back as $(1/N)[1-(1/N)]^{G-1}$. Because N is in the denominator, the smaller the population is, the larger this expression. For example, if the population size is $N = 5$ (i.e., 5 gene copies), the probabilities of common ancestry 1, 2, or 3 generations ago are 0.200, 0.160, and 0.128, whereas for $N = 10$, these probabilities are 0.100, 0.091, and 0.081. The mean of this distribution, the average time back to the common ancestor of random *pairs* of gene copies, can be shown to equal N generations. The mean time back to common ancestry of *all* gene copies in the population is $2N$ generations.*

Thus the coalescence of all gene copies in the current population back to a single ancestral copy is faster, the smaller the population. Viewed from past to present, a single gene copy at some time in the past becomes the ancestor of all gene copies in the population after $G = 2N$ generations on average—i.e., faster in a smaller than in a larger population. If, for the sake of argument, we suppose that each gene copy at some time t generations ago was a different allele or haplotype, (i.e., a distinguishable DNA sequence), then after $2N$ generations we would expect all gene copies to be the same allele or haplotype (assuming no new mutations have occurred during the $2N$ generations). On the other hand, if the population at time t had, let us say, only two alleles, with m copies of A_1 and n copies of A_2 ($m+n=N$), the probability that a copy of A_1 would become the ancestor of all gene copies is $m/N = p$, the allele frequency at time t.

If this process occurs in a large number of independent populations, each of size N, and if allele A_1 had an initial frequency p in each population, then we would expect a fraction p of the populations to become fixed for A_1, and a fraction $1-p$ to become fixed for other alleles such as A_2. Thus the genetic composition of populations diverges by chance.

We have arrived at the following important conclusions about evolution by random genetic drift.

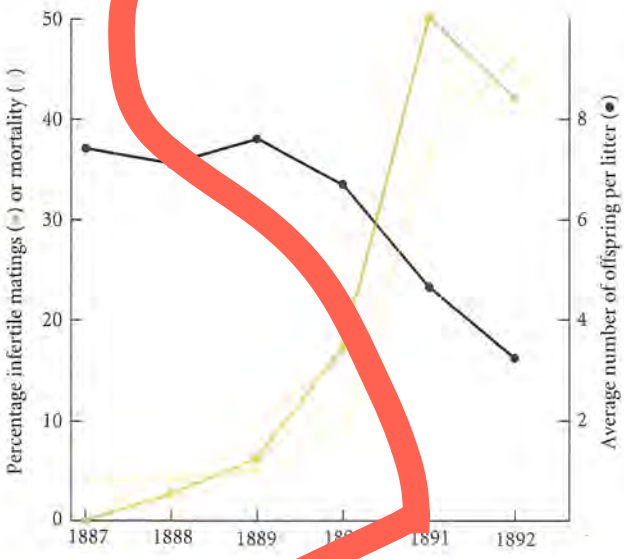
Key points

1. Allele (or haplotype) frequencies fluctuate at random, but eventually one or another allele becomes fixed.
2. Thus, the population eventually loses its genetic variation.
3. Initially similar populations diverge in allele frequency, and may become fixed for different alleles.
4. The probability, at time t, that an allele will eventually become fixed equals the frequency of the allele at that time.
5. The rate at which these events occur is greater, the smaller the population.

Bear in mind that this model, as developed so far, includes only the effects of random genetic drift. Other evolutionary processes—namely, mutation, gene flow, and natural selection—are assumed not to operate. Thus the model, as developed so far, does not describe the evolution of adaptive traits, those that evolve by natural selection. We will incorporate these other evolutionary factors later.

*For a diploid locus, the average time to coalescence for a pair of genes is $2N$ generations, and for all genes in the population is $4N$ generations, which is the expected time to fixation of a newly arisen mutation, i.e., of a single gene copy.

FIGURE 11.15 An early study of inbreeding depression in a laboratory population of rats maintained for about 30 generations of parent-offspring and brother-sister matings. The proportion of infertile matings and the rate of mortality of newborns in the first month of life increased over time. The average number of offspring per litter decreased. (Data from Lerner 1954.)



Relationship between Inbreeding and Genetic Drift

Genetic drift causes changes in allele frequencies, whereas inbreeding alters the proportions of genotypes, but does not alter allele frequencies. Despite this difference, inbreeding and genetic drift are closely related, as we can see from several perspectives. First, imagine that a large, randomly mating population becomes divided into demes, and that mating occurs only within these demes. Within each deme, the probability that two randomly chosen gene copies are identical by descent increases with the passage of generations, as we saw in our discussion of the coalescence of gene lineages (see Figure 11.2), simply because the gene copies in each generation are descended from a smaller number of copies in previous generations. The smaller the population. It can be shown (Box 11.D) that after t generations,

$$F_t = 1 - \left(1 - \frac{1}{2N}\right)^t$$

From the relationship between F and H, $H_t = (1 - 1/2N)^t H_0$. The quantity $(1 - 1/2N)^t$ approaches zero as t becomes large, so ultimately F equals 1 and H equals 0: the population has become completely inbred.

Even if the individuals within each deme mate at random, mates are more closely related to each other than they would be if mating occurred at random throughout the entire metapopulation. Thus, a group of isolated demes is conceptually similar to a population consisting of inbred lines propagated by, say, mating between siblings. In both cases, the population *as a whole* exhibits an excess of all homozygous genotypes and a deficiency of heterozygotes, relative to the Hardy-Weinberg proportions expected if mating occurred at random throughout the population.

Suppose, for example, that the frequencies p and q of alleles A_1 and A_2 are initially 0.5 in a large population that is then subdivided into demes. If A_1 and A_2 each drift to fixation in half of the demes, the overall (or mean) allele frequencies are still 0.5, but there are no heterozygotes: $F = (H_0 - H)/H_0 = (0.5 - 0.0)/0.5 = 1.0$. Less extremely, suppose p has drifted to 0.75 in half of the demes and to 0.25 in the others. Then, if there is random mating *within* demes, we have the following genotype frequencies:

	A_1A_1	A_1A_2	A_2A_2
where $p = 0.75$:	0.5625	0.3750	0.0625
where $p = 0.25$:	0.0625	0.3750	0.5625
Overall (mean):	0.3125	0.3750	0.3125
H-W expectation:	0.25	0.50	0.25 (because $p = 0.5$)

Because p in the population as a whole is $(0.75 + 0.25)/2 = 0.5$, the frequency of heterozygotes *expected* if all individuals mated at random (H_0) would be $2(0.5)(0.5) = 0.5$. Since the observed frequency of heterozygotes is $H = 0.375$, $F = (0.5 - 0.375)/0.5 = 0.25$. In this context, F represents *the probability that two gene copies within a deme are the same, relative to gene copies taken at random from the entire metapopulation.* This value of F, which stems from divergence among demes (or populations of any size) by random genetic drift, was denoted F_{ST} by Sewall Wright, and is sometimes called the **fixation index**. If each deme has an effective size N, then after t generations, $F_{ST} = 1 - (1 - 1/2N)^t$, as we saw earlier.

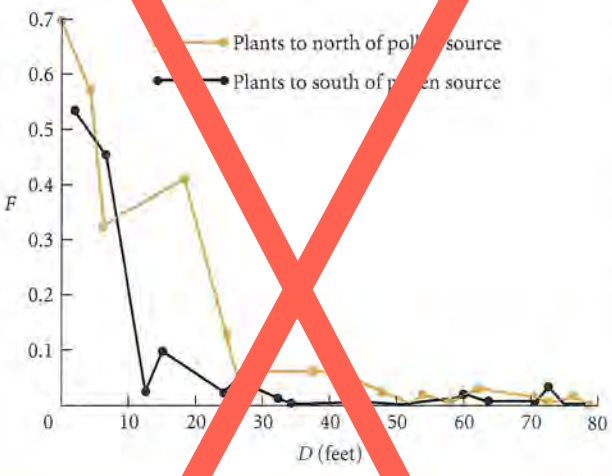
F_{ST} is often used as a measure of the observed variation in allele frequencies among populations (regardless of how the variation may have arisen), and in fact may be calculated from the variance in allele frequencies as $F_{ST} = V_q/[\bar{q}(1 - \bar{q})]$. For example, applying this calculation to the data on the frequency of the *Es-3* allele in house mice in Table 11.2 yields $F_{ST} = 0.0506/(0.318)(0.582) = 0.208$ for small populations and $F_{ST} = 0.0125/(0.372)(0.628) = 0.054$ for large populations.

Gene Flow

Natural populations of a species typically are not completely isolated, but instead exchange genes with one another to a greater or lesser extent. This process is called **gene flow**. *Gene flow, if unopposed by other factors, homogenizes the populations of a species,*—that is, it brings them to the same allele frequencies. Thus conspecific populations differ genetically only if gene flow is sufficiently counterbalanced by the divergent forces of genetic drift or natural selection.

Models of gene flow treat organisms as if they formed either discrete populations (e.g., on islands or in ponds) or continuously distributed populations. In models of dis-

FIGURE 11.17 Gene flow in corn, a wind-pollinated plant. The vertical axis, *F*, gives the proportion of offspring of recessive plants, grown at different distances from a dominant strain, that were fathered by the dominant strain. The curves for plants situated north and south of the pollen source differ because the direction of prevailing winds affects the dispersal of pollen. (After Bateman 1947.)



of about 2 percent per generation, owing chiefly to human destruction of stands of its host plant. Newly planted patches of milkweed were almost all colonized by small numbers of beetles each year, from populations at least several kilometers away. The genetic consequences of these population dynamics are unclear because it is not known whether the colonists at a site are drawn from several source populations (which would tend to reduce genetic differentiation) or from one (which would enhance it).

Indirect estimates An alternative to tracing the movement of genes directly is to infer the level of gene flow from the differences in allele frequency among populations. F_{ST} is one measure of such differences. By rearranging the equation $F_{ST} = 1/[4Nm + 1]$, we might estimate the average number of immigrants into each population per generation as $Nm = (1/F_{ST} - 1)/4$. In doing so, we infer that the more similar the allele frequencies among populations, the higher the rate of gene flow; conversely, strong divergence among populations is taken to indicate that the balance between genetic drift and gene flow is tipped toward genetic drift.

This inference is valid only if two assumptions hold true. First, the alleles for which we calculate F_{ST} must be selectively neutral. F_{ST} would underestimate gene flow if natural selection favored different alleles in different areas, and would overestimate gene flow if selection favored the same allele everywhere. This assumption can be evaluated by the degree of consistency among different loci used to calculate F_{ST}. Genetic drift and gene flow affect all loci the same way, whereas natural selection affects different loci more or less independently. Therefore, if each of a number of polymorphic loci yields about the same value of F_{ST}, it is likely that selection is not strong. Second, allele frequencies must have reached an equilibrium between gene flow and genetic drift. This might not be the case if, for example, the various sites have only recently been colonized, and the populations have not yet had time to differentiate by genetic drift. Their genetic similarity would then lead us to overestimate the rate of gene flow. This assumption can be difficult to evaluate.

The black-tailed prairie dog (*Cynomys ludovicianus*), a social ground squirrel of North American prairies, has a complex population structure. A group of coteries, each with one or two breeding males and several females, makes up a ward; a number of wards, each occupying a patch of suitable habitat, constitute a deme. Ecological studies have shown that males disperse farther than females, and that they move more frequently among coteries within wards than among wards. Ronald Chesser (1983) used electrophoresis to estimate allele frequencies at seven polymorphic enzyme loci in 21 populations of prairie dogs, scattered widely throughout four regions of eastern New Mexico. Most, although not all, of the loci showed fairly similar levels of gene frequency differentiation (F_{ST}), as illustrated in Table 11.4. From the mean F_{ST} values of the several loci, *Nm* was estimated at about 1 animal moving into each coterie from other coteries in the same ward per generation. At higher hierarchical levels, about 4.4 enter each ward, 4.9 enter each deme, and 2.2 enter each population region from other regions. Thus the level of migration even among coteries is low, probably because these social groups tend to exclude outsiders. Gene flow among the larger population units is also low, but nonetheless seems high enough to prevent complete divergence by genetic drift.

Even more extreme subdivision is evident in the pocket gopher *Thomomys bottae* (Patton 1972; Patton and Yang 1977). This burrowing rodent seldom emerges from the soil, and the maximal dispersal distance of marked individuals is only about 900 feet. This species is famous for its localized variation in coloration and other morphological features, which has led taxonomists to name more than 150 subspecies. Moreover, local populations differ more in chromosome configuration than in any other known species of mammal. A study of 21 polymorphic enzyme loci in 825 specimens from 50 localities in the southwestern United States and Mexico also showed extreme geographic differentiation, with different loci displaying different patterns of variation (Figure 11.18). F_{ST}, averaged over these loci, was extraordinarily high at 0.412 across all 50 populations (which might imply $Nm = 0.36$). It was likewise high within smaller regions (e.g., $F_{ST} = 0.198$, $Nm = 1.01$, among 17 localities in Arizona). Patton and Yang found that populations were genetically most different when they were geographically distant and/or segregated by expanses of unsuitable habitat—both factors that would reduce gene flow.

Generally, the magnitude of gene flow estimated from genetic data corresponds fairly well with what we might ex-

If natural selection is defined by differences in survival and reproduction, then selfish genetic elements provide another example of different levels of selection: in these cases, selection among *genes* acts in opposition to selection among *individual organisms*. We note further that selection among genes may not only be harmful to individual organisms, but might also cause the extinction of populations or species.

The Nature of Natural Selection

Definitions of Natural Selection

The examples presented above show that selection can take many forms. Consequently, many definitions of natural selection have been proposed (Endler 1986); it is for this reason that we have not yet provided a definition in this text. Most authors agree that the definition must include the following concepts: some attribute or trait must *vary* among biological entities, and *there must be a consistent relationship*, within a defined context, *between the trait and one or more components of reproductive success*, where "reproductive success" includes both survival (a prerequisite for reproduction) and the reproductive processes themselves. An entity's reproductive success is its **fitness**, defined as the average per capita rate of increase. If the entities are different classes of individual organisms or genes (such as genotypes or alleles), then fitness is usually measured as the mean number of descendants, per individual organism or gene copy, counted as newly produced offspring (fertilized egg or newborns) after one generation. (Thus fitness is usually measured by the entity's rate of increase, R or r, as defined in Chapter 4; see also Chapter 13. Occasionally, it is useful to measure fitness by counting descendants after two generations rather than one; see Chapter 21.)

Some authors treat sexual selection as a process distinct from natural selection, and restrict natural selection to differences in survival and fertility. More commonly, sexual selection is considered a kind of natural selection, and will be so considered in this book.

For selection to exist, there must be average differences in reproductive success among different classes of entities. Evolutionary biologists differ on whether or not the definition of selection should require that the classes differ genetically. Some authors, such as Russell Lande and Stevan Arnold (1983), define selection as a consistent difference in fitness among *phenotypes*, acting within a single generation. Whether or not it alters the frequencies of phenotypes in the next generation depends on whether and how the phenotypic differences are inherited. The change in the population from one generation to another is termed the **response to selection**. Authors who advocate this phenotypic definition distinguish the response, which is solely a matter of inheritance, from differences in survival and reproduction, which constitute selection itself. Thus the experiment on widowbird phenotypes demonstrated selection on tail length, but provided no information on the response to selection, because nothing was discovered about the inheritance of this feature. This definition emphasizes that *selection acts on phenotypes, but may change allele and genotype frequencies if the phenotypes differ in genotype*. Genetically identical members of an asexual clone may differ in phenotype, perhaps because of environmental influences, and in reproductive rate, but the phenotype frequencies in the next generation will not be altered unless the distribution of environmental effects has changed. According to Lande and Arnold, there is selection in this case, but no response to selection, and no evolution.

Biologists have more commonly included the genetic response to selection in the definition. For example, one of the founders of population genetics, Sewall Wright (1969), defined selection as "any process in a population that alters gene frequency in a directed fashion without change of the genetic material (mutation) or introduction from without (immigration)." John Endler (1986) defined natural selection as a process in which (to paraphrase Endler), if a population exhibits (*a*) variation in a trait, (*b*) a consistent relationship between the trait and fitness, and (*c*) inheritance of the trait, then the frequency distribution of the variations (1) will differ among age classes and (2) may differ between generations. Parts *a* and *b* of this definition describe differences among phenotypes; parts 1 and 2 describe their consequences (response to selection), mediated by inheritance (*c*).

Some authors include under natural selection only selection at the level of genes, genotypes, and individual organisms, excluding selection among groups (such as populations or species). However, many authors (e.g., Endler 1986; Sober 1984) include all these entities within natural selection, which then may operate at a variety of levels, ranging from the gene to the species. We shall follow this convention.

For our purposes, we will define natural selection as *any consistent difference in fitness* (i.e., survival and reproduction) *among phenotypically different biological entities*. The entities may be individual genes (which must have some phenotypically variable property if they differ consistently in fitness), groups of genes, individual organisms, populations, or taxa such as species. (Although we have adopted a phenotypic definition, we will almost always discuss the fitness of phenotypes that are inherited to at least some degree, because selection has no evolutionary effect unless there is inheritance.)

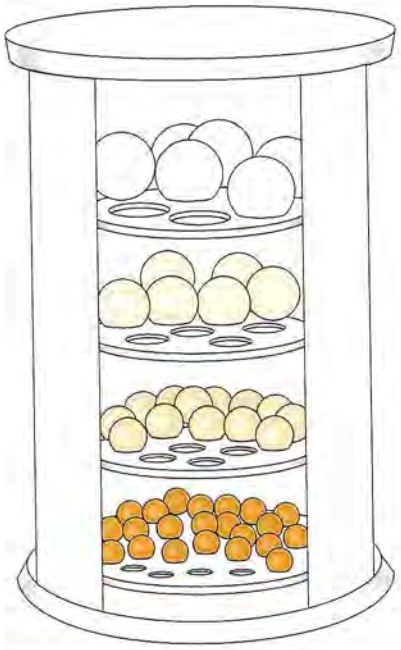
Natural selection and chance A critical feature of our definition is that selection operates only if phenotypes differ *consistently* in fitness. If one neutral allele replaces another in a population by genetic drift (see Chapter 11), the bearers of this allele in this population have had a greater rate of increase than the bearers of the other, but natural selection has not occurred, and the genotypes are not considered different in fitness, because the allele does not consistently confer higher rates of survival or reproduction: in any generation, one of the alleles will increase in frequency in about half of a number of replicate populations, and the other allele will increase in the other half. There is no *average* dif-

ference between them, *no bias* toward increase of one relative to the other. Fitness differences, in contrast, are *average* differences, *biases*, differences in the *probability* of reproductive success. This does not mean, of course, that every individual of a fitter genotype (or phenotype) survives and reproduces prolifically, while every individual of an inferior genotype perishes; a great deal of mortality and variation in reproductive rate occurs independent of—that is, at random with respect to—phenotypic differences. Thus the difference in fitness among phenotypes is the difference that is *not* due to chance, but is *caused* by some characteristic difference between them. Therefore natural selection is the difference in rates of increase among biological entities that is *not* due to chance. *Natural selection is the antithesis of chance.*

If fitness and natural selection are defined by consistent, or average, differences, then we cannot tell whether a difference in reproductive success between two *individuals* is due to chance or to a difference in fitness. We cannot say that one identical twin had lower fitness than the other because she was struck by lightning (Sober 1984), or that the genotype of Pyotr Tchaikovsky, who had no children, was less fit than the genotype of Johann Sebastian Bach, who had many. Fitness cannot be measured for an individual gene, organism, or population, but only as the average of some number of like genes, organisms, or populations. The biologists who performed the experiments described above could ascribe genetic changes to natural selection rather than genetic drift because they observed consistent changes in replicate populations, or measured numerous individuals of each phenotype and found an average difference in survival or mating success.

Selection of and selection for The philosopher of science Elliot Sober (1984) uses a child's "selection toy" (Figure 12.16) to make some useful conceptual points. Balls of several sizes, placed in the top compartment, fall through holes in partitions, the holes in each partition being smaller than in the one above. The toy thus selects small balls. Natural selection may similarly be considered a sieve that selects organisms with a certain body size, mating behavior, or other feature. If the small balls in the toy are green, and the larger ones have other colors, the toy will select small, green balls. Sober emphasizes, however, that we should distinguish *selection of objects* from *selection for properties*. The objects selected are the balls, but they are selected *for* the property of small size; i.e., *because* of their small size. They are not selected for their color because of their color, but nonetheless there is selection *of* green balls. A biological analogy is provided by the hitchhiking effect in the experiment by Atwood and others: in the populations of *E. coli*, there was selection for the capacity for histidine synthesis, but there was selection *for* linked advantageous mutations at other loci not for histidine synthesis itself. The property of histidine synthesis was not the cause of natural selection of genotypes with this property. Thus the evolution of histidine synthesis was an **effect** of natural selection, but not its cause, just as uniform green color is an effect of selection in the toy, but not a cause of the prevalence of green balls.

FIGURE 12.16 A child's toy that selects small balls, which drop through smaller and smaller holes from top to bottom. In this case there is selection *of* darkly shaded balls, which happen to be the smallest, but selection *for* small size. (After Sober 1984.)



These semantic points are more important than they may seem. When we speak of the FUNCTION of a feature, we imply that there has been natural selection *of* organisms with the feature and *of* genes that program it, but *for* the feature itself. We suppose that the feature *caused* its bearers to have higher fitness. The feature may, however, have other *effects*, or consequences, that were not its function, and *for* which there was no selection. For instance, there was selection for an opposable thumb and digital dexterity in early hominids, with the effect, millions of years later, that we can play the piano. There was selection for a large brain and a great capacity for learning, for reasons on which we can only speculate, with the effect that we can do calculus and invent computers. There is selection for more cryptic coloration in a population of guppies exposed to predation, and an *effect* of this might well be a lessening of the likelihood that the population will become extinct, but *avoidance of extinction is not a cause of evolution* in the guppy population. *There has not been selection for avoidance of extinction*; avoidance of extinction is *not the function* of the guppies' coloration, nor, probably, of any feature of any organism.

Levels of Selection

Selection of organisms and groups It is common to read, in student essays and even in professional biological literature, statements to the effect that clams have a high re-

CHAPTER 10

The Origin of Genetic Variation

Evolution cannot occur unless there is genetic variation. As we have seen (in Chapter 9), there is considerable genetic variation within and among populations of most species. We now turn our attention to the processes by which this genetic variation originates.

We will first treat gene mutations, the alterations of individual genes that are so fundamentally important in evolution. The many aspects of this topic occupy much of this chapter. We will then consider recombination and changes in the structure and number of chromosomes, sometimes referred to as chromosomal mutations. Finally, we will note the significance of genetic variation acquired from other populations and species. (These topics are also treated in Chapters 11 and 22.)

The word "*mutation*" refers both to the process of alteration of a gene or chromosome and to the product, the altered state of a gene or chromosome. It is usually clear from the context which is meant.

Gene Mutations

Mutational changes of individual genes are overwhelmingly important in evolution. Many, perhaps most, evolutionary changes in phenotypic characters are attributable to changes in enzymes or other proteins, and thus to changes in the DNA sequences that encode them. At the molecular level, however, many alterations of DNA sequences occur that have slight or no phenotypic consequences.

In a broad sense, a **gene mutation** is *an alteration of a DNA sequence.* Thus, our modern knowledge of the molecular basis of heredity provides a definition in molecular terms. Before the development of molecular genetics, however, a mutation was identified by its effect on a phenotypic character (Box 10.A describes the history of the concept of mutation). That is, a mutation was a newly arisen change in morphology, survival, behavior, or some other property that was inherited and could be mapped (at least in principle) to a specific locus on a chromosome. In practice, many mutations are still discovered, characterized, and named by their phenotypic effects. Thus, we will frequently use the term "*mutation*" to refer to an alteration of a gene from one allele to another in which the alleles are distinguished by their phenotypic effects. However, not all alterations of DNA sequences have phenotypic consequences. Hence, in a molecular context, the term "*mutation*" refers to a change in DNA sequence, independent of whatever phenotypic effect it may have.

Mutations have evolutionary consequences only if they are transmitted to succeeding generations. If a mutation occurs in a somatic cell, it is extinguished with the organism's death in the case of many animals; it may, however, be inherited in certain animals and plants in which the reproductive structures arise from somatic meristems. In those animals in which the germ line is segregated from the soma early in development, a mutation is inherited only if it occurs in a germ line cell. The chance that a gamete will carry a new mutation increases with the number of cell divisions that transpire in the germ line between the mutation event and gametogenesis. In humans, more cell divisions have preceded spermatogenesis than oogenesis in individuals of equal age, and the incidence of new mutations appears to be higher in sperm than in eggs (Crow 1993). An individual may produce many gametes with the same new mutation if the mutation occurred early in the germ line's history, or few if it occurred immediately before gametogenesis.

Both during replication and at other times, DNA is frequently damaged by chemical and physical events, and changes in base pair sequence occur. Many such changes are repaired by DNA polymerase and other "proofreading" enzymes, but some are not. These alterations, or mutations, are considered by most evolutionary biologists to be *errors.* That is, *the process of mutation is thought not to be an adap-*

BOX 10.A

History of the Concept of Mutation

The meaning of "*mutation*," like that of many other words, has evolved. As far back as the seventeenth century, it was used to describe any drastic change in an organism's form, as in the fossil record. Early in the twentieth century, it was given a different meaning by the Dutch botanist Hugo DeVries, who is widely known as a discoverer of Mendel's neglected paper. DeVries was interested in the origin of new species, and thought he had solved this problem when he found discretely different, true-breeding forms among the offspring of the evening primroses (*Oenothera lamarckiana*) in his experimental garden. He termed these "mutations" and concluded that a new species arises by a spontaneous, discrete change in one or more features. To DeVries and his followers, Darwin's theory of natural selection therefore became superfluous, because the mutation process created new species in a single step, in which natural selection and the environment played no role. Moreover, the slight, continuous hereditary variations in characteristics such as size and shape were considered by the "mutationists" to have an entirely different genetic basis from discrete mutations, and to play no role in evolution. (It was later found that the "mutations" or "new species" that DeVries had observed in *Oenothera* were mostly rare recombinations of several genes, produced in a plant with a very unusual system of chromosomes.)

"*Mutation*" underwent a further change in meaning when the pioneering *Drosophila* geneticist Thomas Hunt Morgan, at Columbia University in New York, discovered newly arisen aberrations, such as white-eyed flies, that obeyed Mendelian rules of inheritance. Thus *mutation* came to mean not necessarily the origin of a new species, but a spontaneous alteration of a gene. (Nonetheless, Morgan continued for much of his life to affirm that new species arise by mutation, and that natural selection plays no causal role in evolution.)

If a mutation is an alteration of a single gene, it may be (and usually is) a genetic variant rather than a new species. If the same mutation occurs only rarely, mutation "pressure" will generally not be adequate to transform a species, and something else (such as natural selection) is required to increase the frequency of the mutation in the population. This reasoning, with its emphasis on evolution as a *population-level* process rather than the origin of species as mutant *individuals*, is the foundation of the Evolutionary Synthesis of the 1930s and 1940s (see Chapter 2), in which mutation and natural selection are complementary rather than mutually exclusive ingredients of evolution.

When geneticists came to realize that continuous variation is based on multiple genes that are inherited in the same way as discrete Mendelian factors, it became understood that the mutational process generates both kinds of variation: mutations with small phenotypic effects are the basis of continuous variation, and those with large effects generate discrete variations. Moreover, there is a continuum of effects from very small to quite large. When the molecular nature of the gene was elucidated in the 1950s, mutation could be recognized as an alteration of the base pair sequence of a gene—including base pair changes that have no effect whatever on the phenotype, even on the amino acid sequence of the protein that the gene encodes.

What mutation is *not* is the birth of new organisms utterly unlike their parents. These exist only in science fiction. Dinosaurs or birds do not hatch fully formed from lizard eggs. Some mutations may be monstrous, such as flies with their antennae transformed into legs, but mutations can only alter already immanent developmental processes, and so must have a limited range of possible effects.

tation, but a consequence of unrepaired damage. Both the existence of repair enzymes and the theory of the evolution of mutation rates, discussed in Chapter 21, are the basis for this conclusion.

Mutational changes of DNA sequences are of many kinds.

Point Mutations

The simplest mutation is a substitution of one base pair for another (Figure 10.1). In classic genetics, a mutation that maps to a single gene locus is called a **point mutation**; in modern usage, this term is often restricted to single base pair substitutions. A **transition** is a substitution of a purine for a purine (A↔G) or a pyrimidine for a pyrimidine (C↔T). **Transversions**, of eight possible kinds, are substitutions of purines for pyrimidines or vice versa (A or G↔C or T).

Some base pair changes occur in nontranslated DNA, and have no known phenotypic effect. Mutations in genes that encode ribosomal and transfer RNA potentially affect the function of these gene products. Other base pair changes may result in amino acid substitutions in polypeptides or proteins. These may have little or no effect on the functional properties of the polypeptide, and thus no effect on the phenotype, or they may have substantial consequences. For example, the change from the (RNA) triplet GAA to GUA causes the amino acid valine to be incorporated instead of glutamic acid. This is the mutational event that in humans caused the abnormal β-chain in sickle-cell hemoglobin, which in turn has many pleiotropic effects (see Figure 21 in Chapter 3) and is usually lethal in homozygotes.

Because of the redundancy of the genetic code, many substitutions at the third base position in codons, and quite

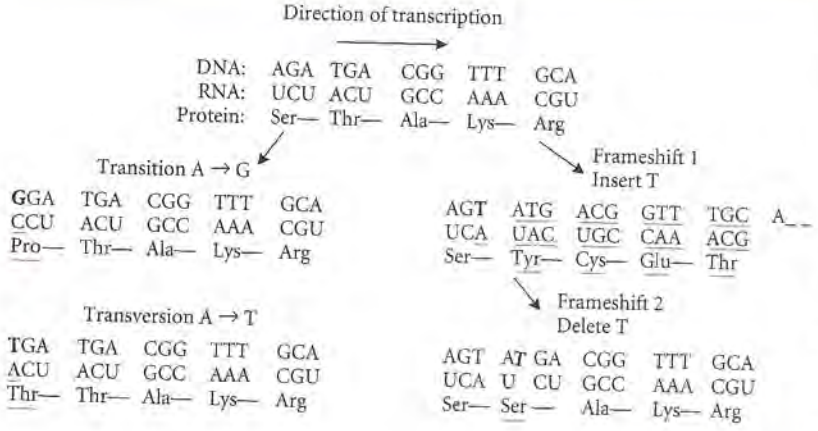


FIGURE 10.1 Examples of kinds of point mutations and their consequences for messenger RNA and amino acid sequences. (Only the transcribed, "sense" strand of the DNA is shown.) At left, transition and transversion mutations at the first base position. At right, a frameshift mutation, caused by the insertion of T between sites 2 and 3, shifts the reading frame so that downstream bases are read in new triplets, altering the amino acid sequence. A second frameshift mutation, a deletion of one base at the fifth site, reestablishes the original reading frame downstream from the site. The encoded amino acids can be found from the code shown in Figure 12 in Chapter 3.

a few at the first base position, are synonymous: they do not alter amino acids. About 24 percent of the possible substitutions in the code are synonymous, but the proportion of synonymous mutations that occur in a species' genome depends on the proportions in which the various codons are represented, as well as any nonrandomness of substitution that may exist.

Three of the triplets in the RNA code are "stop" codons, signaling termination of translation into a polypeptide product. Mutation of an amino acid-encoding triplet into a stop codon will result in an incomplete, usually nonfunctional, gene product. Mutations to termination codons are often found within nonfunctional pseudogenes.

If a single base pair (or more) becomes inserted into or deleted from a DNA sequence, the triplet reading frame is shifted by one nucleotide, so that downstream triplets are translated into different amino acids (Figure 10.1). This is a **frameshift mutation**. The greatly altered gene product is usually nonfunctional.

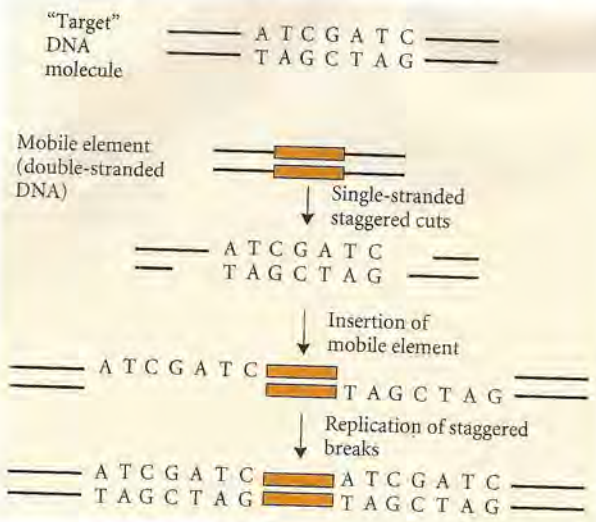
Sequence changes arising from recombination When homologous DNA sequences differ at two or more base pairs, **intragenic recombination** between them can generate new DNA sequences. In molecular terms, intragenic recombination is not mutation, but the new haplotypes might be distinguished as alleles or mutations if they have phenotypic effects. Thus recombination between DNA sequences that code for, say, the amino acid sequence Val-Thr-Arg-Leu and Glu-Thr-Arg-Gly could give rise to the new polypeptide product Val-Thr-Arg-Gly. Precisely this kind of polymorphism was described for the amino acid sequence of the enzyme 6-phosphoglycerate dehydrogenase in the Japanese quail (Ohno et al. 1969). Since then, direct DNA sequencing has revealed many examples of variant haplotypes that apparently arose by intragenic recombination. Kreitman described some instances in his study of sequence variation in the alcohol dehydrogenase gene of *Drosophila melanogaster* (see Chapters 9 and 22).

Recombination appears to be the cause of a peculiar mutational phenomenon called **gene conversion**, which has been studied most extensively in fungi. The gametes of a heterozygote should carry the two alleles (A_1, A_2) in a 1:1 ratio. Occasionally, though, they occur in different ratios, such as 1:3. An A_1 allele has been replaced specifically by an A_2 allele rather than by any of the many other alleles to which it might have mutated: it seems to have been converted into A_2. In some cases gene conversion is unbiased (conversion of A_1 to A_2 is as likely as the converse), but cases of biased gene conversion have been described whereby one allele is preferentially converted to the other. The details of the molecular mechanisms thought to underlie this process need not concern us; suffice it to say that it is believed that a damaged DNA strand of one chromosome is repaired by enzymes that insert bases complementary to the sequence on the undamaged homologous chromosome.

Transposable elements and their effects Until recently, geneticists thought that all genes occupied fixed sites on the chromosomes, except when moved to new positions by inversion or translocation events. This is indeed true of many genes. But in the 1940s, Barbara McClintock described several genes in maize (corn, *Zea mays*) that frequently moved to new sites. Her work was considered a mere curiosity until the 1980s, when it was discovered that apparently all organisms carry in their genomes numerous **transposable elements**: sequences that can move to any of many places in the genome. These DNA sequences carry genes that encode enzymes (transposases) that accomplish the transposition (movement), and sometimes they carry with them other genes near which they had been located. In some cases, the transposable element leaves one "host" gene and becomes inserted elsewhere (conservative transposition). In most cases, a parent element remains in situ, but produces copies that become inserted elsewhere (replicative transposition). The process of insertion generates short (4–12 bp) repeats of the host DNA sequence on either end of the inserted element, called flanking repeats, which are useful for recognizing transposed sequences (Figure 10.2).

The several kinds of transposable elements include

FIGURE 10.2 A model of the origin of flanking repeats in a DNA sequence when a transposable element becomes inserted. The repeats often remain, in whole or in part, if the transposable element is later excised by conservative transposition.



1. *Insertion sequences*, of about 700–2600 bp. Their only functional genes encode transposases, the enzymes that cause transposition. They have been found in bacteria, phage, and maize, among other organisms.
2. *Transposons*, of about 2500–7000 bp. They encode not only transposases, but other functional genes as well. Some plasmids (circular DNA molecules in bacteria that in some cases integrate into the bacterial chromosome and in other cases do not) carry genes that confer resistance to antibiotics and other stresses. Transposons are common in plants, fungi, and animals. There are many kinds; *Drosophila melanogaster* may have as many as 50 to 100. Typically, many copies of a particular kind of transposon are scattered throughout a genome, their locations varying among individuals.
3. *Retroelements.* The traditional view that information flows only from DNA to RNA to protein was changed in the early 1970s by the discovery of **reverse transcription**. The enzyme reverse transcriptase uses RNA as a template for the synthesis of a DNA copy (cDNA). Reverse transcriptase genes are carried by RETROVIRUSES, which are RNA viruses (including the HIV virus that causes AIDS) that invade a cell, make DNA copies of themselves, and insert them into the host genome. These are then transcribed into more RNA virus copies, which infect other cells. RETROPOSONS act similarly except that they do not cross cell boundaries, and spread only by cell division in the host. *Copia* is a retroposon that has been studied extensively in *Drosophila melanogaster.*

Transposable elements have many effects on host genomes, including

1. An increase of total genome size, by replicative transposition.
2. Alteration of expression of host genes. Insertion of a transposable element into the coding region of a host gene can abolish the gene's function. Insertion of a transposable element into the control region of a gene affects its expression; this is the cause of many well-known mutations in *Drosophila*, such as those at the *white* locus, which affect eye color. The promoters carried by a transposable element, which regulate its own transcription, can affect the rate of transcription of nearby host genes as well. The departure of a transposable element is often imprecise, causing the deletion or addition of a few base pairs of the host gene.
3. An increase in the mutation rate of host genes. (We will describe an experiment on mutation rates below.)
4. Chromosome rearrangements in the host genome can result from recombination between two copies of a transposable element located at different sites. This can cause an inversion (a 180° reversal in the orientation of part of a chromosome), or a deletion of the DNA sequence between the transposable elements (Figure 10.3). The deleted material, attached to one copy of the transposable element, can be inserted elsewhere in the genome; the FB transposable elements of *Drosophila* are known to move sequences of hundreds of kilobases. Moreover, the numerous copies of a transposable element promote unequal crossing over, resulting in deletions and duplications of host DNA. For example, a mutation of a human lipoprotein gene, resulting in high cholesterol levels, consists of the deletion of one of the gene's exons, caused by unequal crossing over between repeated sequences distributed throughout the gene's introns (Figure 10.4). (Inversions, deletions, and other chromosome rearrangements are described more fully later in this chapter.)
5. Transposable elements that encode reverse transcriptase sometimes form, and insert into the genome, DNA copies (cDNA) not only of their own RNA, but also of RNA transcripts of the host's genes. A processed RNA transcript lacks the sequences corresponding to the gene's introns, and also lacks the gene's nontranscribed control regions. Therefore a cDNA copy of RNA (a "retrogene") is easily recognized by DNA sequencing: its sequence resembles that of the exons of an ancestral gene located elsewhere in the genome, but it lacks control regions and introns, and its ends correspond precisely to those of the transcribed region of the ancestral gene (Figure 10.5).

Most retrogenes are nonfunctional, partly because they lack control regions. In humans, however, phosphoglycerate kinase is encoded not only by an "ancestral" X-linked gene with introns, but also by an autosomal gene that lacks introns but has a sequence corresponding to the exons of

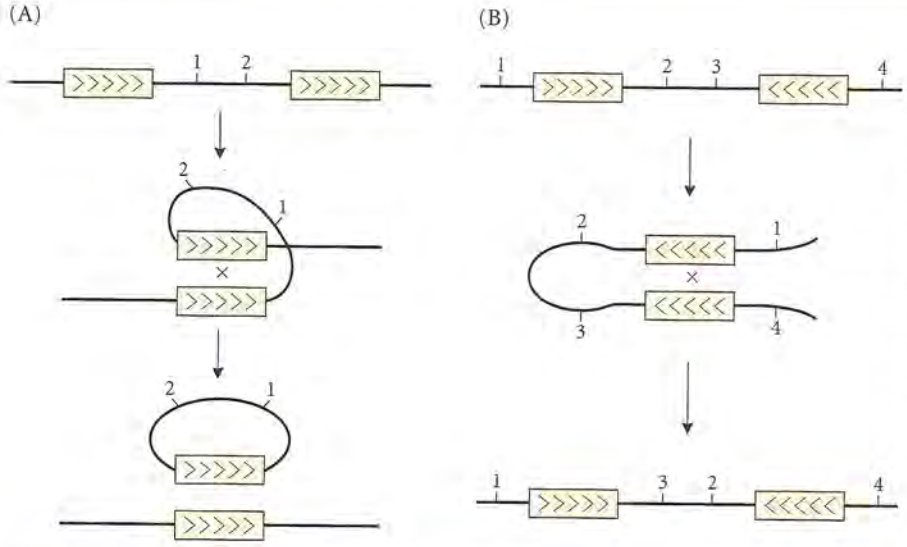


FIGURE 10.3 Recombination between repeated sequences, such as copies of a transposable element, can result in deletions and inversions. The boxes represent repeats, with the polarity of base pair sequence indicated by the arrows within. The numerals represent genetic markers (in different genes or within a single gene). (A) Recombination (×) between two direct repeats (i.e., with the same polarity) excises one repeat and deletes the sequence between the two copies. (B) Recombination between two inverted repeats (with opposite polarity) inverts the sequence between them. (After Lewin 1985.)

the X-linked gene. It is particularly interesting that the X-linked gene is expressed in many tissues, but the autosomal gene is expressed only in the testes, and so has acquired a novel tissue-specific pattern of expression (Li and Graur 1991).

Most retrogenes are **processed pseudogenes** (Figure 10.5), which do not produce functional gene products. They lack sequences corresponding to the ancestral genes' introns, but otherwise show some similarity of sequence. Because they are nonfunctional, however, they accumulate mutations, including termination codons, that are presumably not affected by natural selection, and so their sequence degenerates, diverging from the ancestral gene over time. Mammalian genomes are highly laden with pseudogenes, which may make up as much as 20 percent of the DNA content (Walsh 1985). For example, the hemoglobin gene family includes at least three pseudogenes, and the glyceraldehyde-3-phosphate dehydrogenase sequence is represented by one functional locus and about 20 pseudogenes in humans—and about 200 pseudogenes in the mouse (*Mus musculus*) (Li and Graur 1991). In mammals, a 300-bp sequence called *Alu*, which seems to have been derived by reverse transcription from 7SL RNA, is highly repeated: with more than 500,000 copies, it constitutes about 5 percent of the human genome.

Rate of Mutation

Estimation. Estimates of mutation rates depend on the method used to detect mutations. In classical genetics, a mutation was detected by its phenotypic effects, such as a white vs. red eye in *Drosophila*. Such a mutation, however, might be caused by the alteration of many sites within the locus; moreover, many base pair changes have no phenotypic effect. Thus phenotypically detected rates of mutation underestimate the total mutation rate at a locus. With molecular methods, mutated sequences can be de-

FIGURE 10.4 A mutated low-density-lipoprotein gene in humans, labeled here as FH 626-a, lacks exon 5. It is believed to have arisen by unequal crossing over between two normal gene copies, due to out-of-register pairing between two of the repeated sequences (*Alu* sequences, shown as dark and light shaded boxes) in the introns. The black boxes represent exons. The other product of unequal crossing over, labeled FH-?, has not been found in human populations. (After Hobbs et al. 1986.)