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

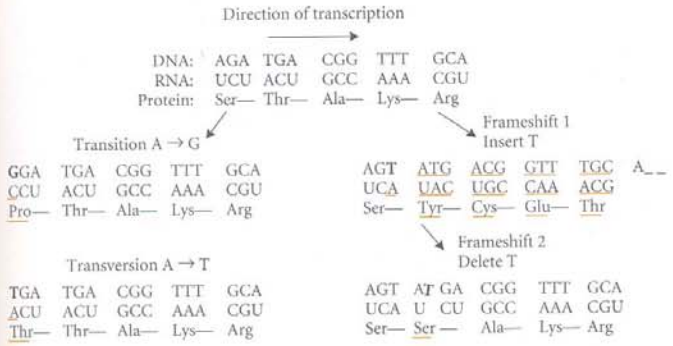


FIGURE 10.1 Examples of kinds of point mutations and their consequences for messenger RNA and amino acid sequences. (Only the transcribed, "sense" strand of the DNA is shown.) At left, transition and transversion mutations at the first base position. At right, a frameshift mutation, caused by the insertion 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.

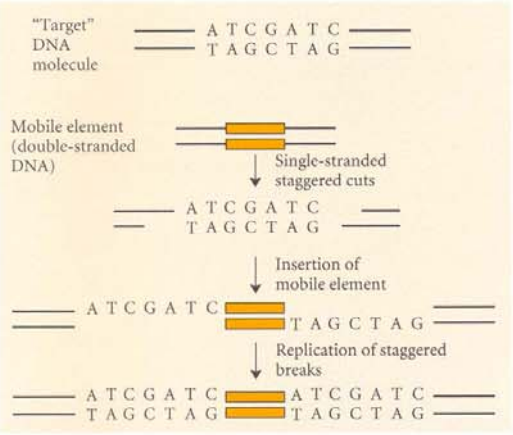
Sequence changes arising from recombination When homologous DNA sequences differ at two or more base pairs, **intragenic recombination** between them can generate new DNA sequences. In molecular terms, intragenic recombination is not mutation, but the new haplotypes might be distinguished as alleles 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

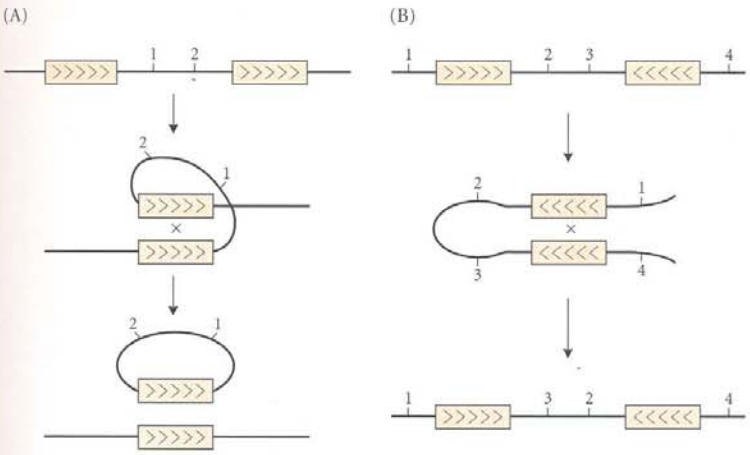


FIGURE 10.3 Recombination between repeated sequences, such as copies of a transposable element, can result in deletions and inversions. The boxes represent repeats, with the polarity of base pair sequence indicated by the arrows within. The numerals represent genetic markers (in different genes or within 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.)

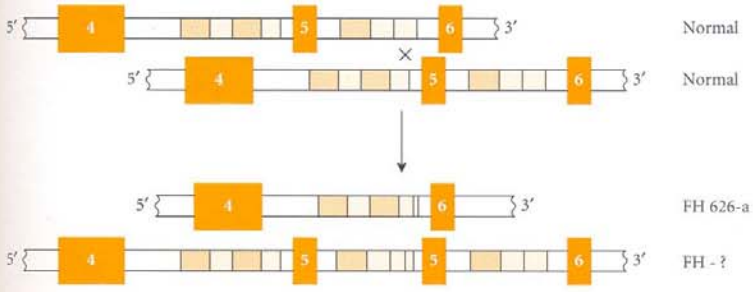
the X-linked gene. It is particularly interesting that the X-linked gene is expressed in many tissues, but the autosomal gene is expressed only in the testes, and so has acquired a novel tissue-specific pattern of expression (Li and Graur 1991).

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.

Rates of Mutation

Estimates Estimates of mutation rates depend on the method used to detect mutations. In classic 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 any 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.)



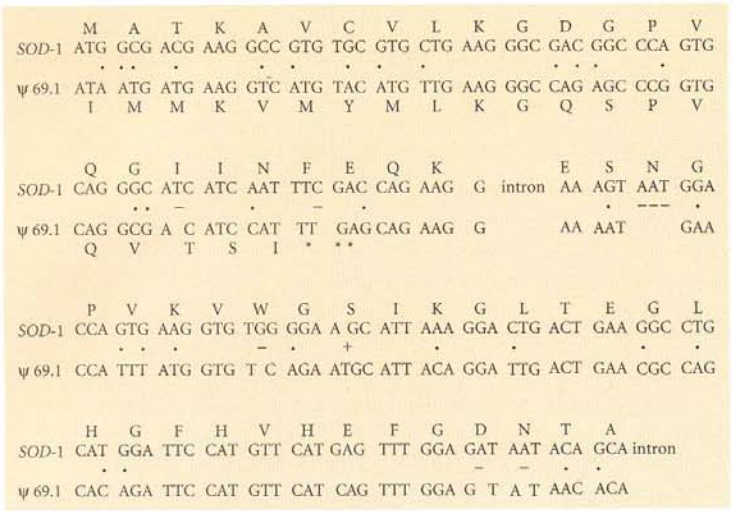


FIGURE 10.5 Comparison of the DNA sequence of the first two exons of the functional Cu/Zn superoxide dismutase gene (*SOD*-1) in humans, and homologous parts of the processed pseudogene (ψ69.1) that arose from *SOD*-1 by reverse transcription from the SOD-1 mRNA. The letters above and below the DNA sequences represent different encoded amino acids, and the dots, dashes, and plus signs between the sequences mark nucleotide substitutions, deletions, and insertions, respectively. Note that the pseudogene precisely lacks the sequence corresponding to the functional gene's intron, and that the deletion of one nucleotide shortly before the intron has generated a chain-termination codon, TGA (marked by asterisks). (After Li and Graur 1991.)

tected directly. These findings are often expressed as mutation rates per base pair.

Rates of mutation are estimated in several ways. A direct method is to count the number of mutations arising in a laboratory stock (which is usually initially homozygous), scoring mutations by their phenotypic effects or by molecular methods. Another, indirect method (Box 10.B) is based on the number of base pair differences between homologous genes in different species, relative to the number of generations that have elapsed since they diverged from their common ancestor. This method depends on a theory of population genetics that is described in Chapter 11.

Direct measures of mutation rate usually count the number of mutations, usually scored by their phenotypic effects, among the offspring of an initially homozygous stock. For example, a bacterial culture can be grown from a single haploid cell, and the rate of origin of mutations conferring resistance to antibiotics can be measured by the number of cells that yield colonies, out of a known number of cells placed on antibiotic-containing medium. This procedure yields the rate of mutation per cell division. In *Drosophila* and other multicellular organisms, the number of chromosomes bearing a particular mutation is scored among a large number of descendants from a stock initially homozygous for a different allele (perhaps the "wild type"). It may be necessary to use special crosses (perhaps using inverted chromosomes; see Figure 6 in Chapter 9) to make the stock homozygous and to make descendant chromosomes homozygous in order to detect recessive mutations. The autosomal chromosomes of N flies represent the $2N$ gametes that formed them, so the mutation rate is calculated as the number of new mutations per gamete per generation. Subsequent crosses, and perhaps mapping of the mutations, may be necessary to determine whether phenotypically similar mutations all occurred at the same locus.

Measured by phenotypic effects, an average locus mutates at a rate of about *10^{-6} to 10^{-5} mutations per gamete per generation* (Tables 10.1, 10.2), an estimate that has been confirmed by studying individual proteins (Neel 1983). The *average mutation rate per base pair*, based mostly on the indirect method of comparing DNA sequences of different species, has been estimated at about 10^{-9}. There is a great deal of variation around these values: mutation rates vary among genes and even among regions within genes, and they can be elevated by transposable elements. Moreover, some genes are known that, in mutant condition ("mutator alleles"), increase the rate of mutation elsewhere in the genome.

Back mutation is mutation of an allele (a "mutant") back to the allele (usually the "wild type") from which it arose. Back mutations are ordinarily described at a phenotypic level. They usually occur at a much lower rate than "forward" mutations from wild type to mutant, presumably because many more substitutions can impair gene function than can restore it. At the molecular level, most phenotypically detected back mutations are not restorations of the original sequence, but instead result from a second amino acid substitution that restores the function that had been altered by the first substitution (Allen and Yanofsky 1963).

Evolutionary implications of mutation rates An average per-locus mutation rate of, say, 10^{-5} (one in 100,000 gametes) is so low that *the rate of change in the frequency of an allele, due to mutation alone, is very low.* Suppose that of two phenotypically distinguishable alleles, A_1 and A_2, with allele frequencies $p\,(=1-q)$ and q, A_1 mutates to A_2 at rate $u=10^{-5}$. In each generation, the frequency of A_2 is increased by $u \times p$ (that is, a fraction u of the A_1 alleles mutate). Denoting the change in the frequency of A_2 from one generation to the next by Δq, we have

$$\Delta q = up = u(1-q).$$

BOX 10.B

Estimating Mutation Rates from Comparisons among Species

In Chapter 11, we will describe the "*neutral theory of molecular evolution*." This theory describes the fate of purely neutral mutations, i.e., those that neither enhance nor lower fitness. One possible fate is that a mutation will become fixed—that is, attain a frequency of 1.0—entirely by chance. The probability that this will occur equals u, the rate at which neutral mutations arise. In each generation, therefore, the probability is u that a mutation that occurred at some time in the past will become fixed. After the passage of t generations, the fraction of mutations that will have become fixed is therefore $x = ut$.

If two species diverged from a common ancestor t generations ago, the expected fraction of fixed mutations in both species is $D = 2ut$, since various mutations have become fixed in both lineages. If the mutations in question are base pair changes, a fraction $D = 2ut$ of the base pairs of a gene should differ between the species, assuming that all base pairs are equally likely to mutate. Thus the average mutation rate per base pair per generation is $u = D/2t$. Thus we can estimate u if we can measure the fraction of base pairs in a gene that differ between two species (D), and if we can estimate the number of generations since they diverged from their common ancestor (t). This requires an estimate of the length of a generation, information from the fossil record on the absolute time at which the common ancestor existed, and an understanding of the phylogenetic relationships among the living and fossilized taxa.

In applying this method to DNA sequence data, it is necessary to assume that most base pair substitutions are neutral, and to correct for the possibility that earlier substitutions at some sites in the gene have been replaced by later substitutions ("multiple hits"). Uncertainty about the time since divergence from the common ancestor is usually the greatest source of error in estimates obtained by this method. Often, the common ancestor has not been identified in the fossil record, and the time estimate is based on the earliest known fossils of either lineage. For example, divergence between a primate and a rodent would date from the late Cretaceous, about 70 Mya, in which the earliest primates are known. Such dates are minimal estimates of divergence time.

The best estimates of mutation rates at the molecular level have been obtained from interspecific comparisons of pseudogenes, other nontranslated sequences, and fourfold degenerate third-base positions (those in which all mutations are synonymous), since these are thought to be least subject to natural selection (although probably not entirely free of it). In comparisons among mammal species, the average rate of nucleotide substitution has been about 3.3–3.5 per nucleotide site per 10^9 years, for a mutation rate of $3.3–3.5 \times 10^{-9}$ per site per year (Li and Graur 1991). If the average generation time were 2 years during the history of the lineages studied, the average rate of mutation per site would be about 1.7×10^{-9} per generation. Comparison of human and chimpanzee sequences yielded an estimate of 1.3×10^{-9} per site per year, assuming divergence 7 Mya. If the average generation time in these lineages has been 15–20 years, the mutation rate is about 2×10^{-8} per generation. The human diploid genome has 6×10^9 nucleotide pairs, so this implies at least 120 new mutations per genome per generation—an astonishingly high number (Crow 1993).

If, for the sake of argument, $q = 0.5$ (A_2 already accounts for half the gene copies), then in the following generation $q' = q + \Delta q = 0.50000495$. At this rate it will take about 70,000 generations to get to $q = 0.75$, and another 70,000 to get to $q = 0.875$. [In fact, each successive halving of the difference between an initial allele frequency and the final equilibrium frequency takes about 70,000 generations (Crow and Kimura 1970).]

With such a low mutation rate per locus, it might seem that mutations occur so rarely that they cannot be impor-

Table 10.1 Comparative spontaneous mutation rates

SPECIES	BASE PAIRS PER GENOME	MUTATION RATE PER BASE PAIR REPLICATION	MUTATION RATE PER GENOME PER GENERATION
Bacteriophage lambda	4.7×10^4	2.4×10^{-8}	0.0001
Bacteriophage T4	1.8×10^5	1.1×10^{-8}	0.0002
Salmonella typhimurium bacteria	3.8×10^6	2.0×10^{-10}	0.0001
Escherichia coli bacteria	3.8×10^6	4.0×10^{-10}	0.0002
Neurospora crassa fungus	4.5×10^7	5.8×10^{-11}	0.0003
Drosophila melanogaster[a]	4.0×10^8	8.4×10^{-11}	0.93

Source: After Drake (1974).

[a]The *Drosophila* values are per diploid genome per generation of flies, not per generation of cells, as in the other species.

Table 10.2 Spontaneous mutation rates of specific genes, detected by phenotypic effects

SPECIES AND LOCUS	MUTATIONS PER 100,000 CELLS OR GAMETES
Escherichia coli	
Streptomycin resistance	0.00004
Resistance to T1 phage	0.003
Arginine independence	0.0004
Salmonella typhimurium	
Tryptophan independence	0.005
Neurospora crassa	
Adenine independence	0.0008–0.029
Drosophila melanogaster	
Yellow body	12
Brown eyes	3
Eyeless	6
Mus musculus (mouse)	
a (coat color)	7.1
c (coat color)	0.97
d (coat color)	1.92
In (coat color)	1.51
Homo sapiens	
Retinoblastinoma	1.2–2.3
Achondroplasia	4.2–14.3
Huntington's chorea	0.5

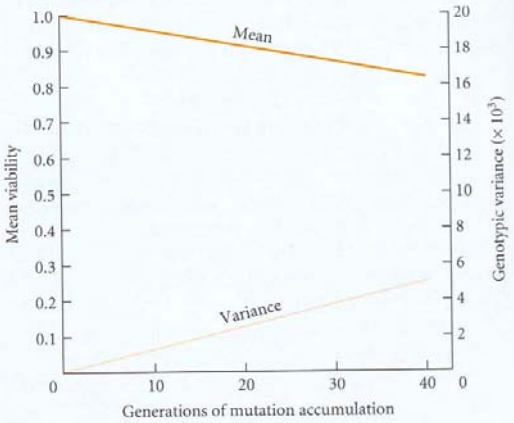
Source: After Dobzhansky (1970)

tant. However, the number of functional genes in a genome is quite large: about 10,000 in *Drosophila* and perhaps 150,000 in humans. This implies that *almost every gamete carries a new, phenotypically detectable mutation somewhere in its genome* (10^{-5} mutations per gene $\times$ 10^5 genes = 1 mutation per haploid genome in humans). So in a population of 500,000 individuals, about one million new mutations arise every generation. If even a tiny fraction of these were advantageous, the amount of new "raw material" for adaptation would be substantial, especially over the course of thousands or millions of years. These figures may be underestimates, for at the molecular level, each human haploid genome may carry about 200 new nucleotide substitutions (Kondrashov and Crow 1993; see Box 10.B). Recall from Chapter 9, furthermore, that many alleles that arose by mutation in the past have high frequencies in natural populations, so the amount of potentially adaptive genetic variation is very high.

Experiments on *Drosophila* have confirmed that the total mutation rate per gamete is quite high. In a heroically large experiment, Terumi Mukai and colleagues (1972) counted more than 1.7 million flies in order to estimate the rate at which the second chromosome accumulates mutations that affect egg-to-adult survival (viability). They used crosses (see Figure 6 in Chapter 9; Box 10.C) in which copies of a single wild chromosome were carried in a heterozygous condition so that deleterious recessive mutations could persist without being eliminated by natural selection. Every ten generations, they made large numbers of these chromosomes homozygous, and measured the proportion of them that reduced viability. The mean viability declined, and the variation among chromosomes increased steadily (Figure 10.6). From the changes in the mean and variance, Mukai et al. calculated a mutation rate of about 0.15 per second chromosome per gamete. This is the sum, over all loci on the chromosome, of mutations that affect viability. Because the second chromosome carries about a third of the genome, the total mutation rate is about 0.5 per gamete. Thus, almost every zygote carries at least one new mutation that affects viability. In a similar experiment, David Houle and colleagues (1992, 1994) confirmed this conclusion, and found that new mutations substantially affect not only viability, but also other aspects of fitness such as reproductive rate.

Because the genetic basis of most phenotypic traits is polygenic, it is important to know how rapidly mutation supplies genetic variance (see Chapter 9) in such characteristics. In order to facilitate comparisons among traits, the variance attributable to new mutations (V_m) is often expressed as a fraction of the environmental variance (V_E) of the trait. One of several methods for estimating this quantity is to propagate a population that has been made entirely homozygous (e.g., by inbreeding), and after several generations to estimate, by the usual methods of quantitative genetics, the genetic variance that has accumulated (Chapters 9, 14). V_m/V_E can also be estimated by artificially selecting a character in an initially homozygous population, for any response to selection must then be based on new mutations.

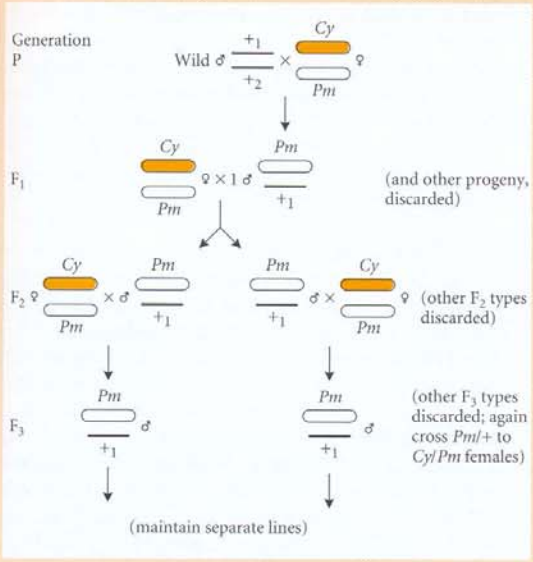
FIGURE 10.6 Effects of the accumulation of spontaneous mutations on the viability (egg-to-adult survival) of *Drosophila melanogaster*. The mean viability of flies made homozygous for chromosomes carrying recessive mutations decreases, and the variation (variance) among chromosomes increases. (After Mukai et al. 1972.)



BOX 10.C

Accumulation of Recessive Mutations

In the crossing scheme used by Mukai, a single wild male fly, with second chromosomes that we may denote $+_1$ and $+_2$, is mated with females of a stock *Cy/Pm* (as in Figure 6 in Chapter 9). Each of the female's two chromosomes has an inversion that prevents crossing over, a recessive lethal gene, and a dominant "marker" gene (*Curly* wing, *Cy*, or *Plum* eye color, *Pm*). A single male F_1 offspring bearing *Pm* (hence $Pm/+_1$ or $Pm/+_2$) is then crossed to *Cy/Pm* females. Thus all wild (+) chromosomes in the F_2 are identical (either $+_1$ or $+_2$). Each of a large number of $Pm/+$ F_2 males is mated to *Cy/Pm* females, and this procedure is repeated in subsequent generations. Note that the + chromosomes, all initially identical, are kept in heterozygous condition (so recessive mutations that affect viability are not expressed) and that the + chromosomes remain intact because they cannot recombine with *Cy* or *Pm* chromosomes.



In each generation, the cross *Cy/Pm* × *Pm/+* yields both females and males of each of four genotypes: *Cy/Pm*, *Cy/+*, *Pm/Pm*, and *Pm/+* (but *Pm/Pm* is lethal and does not appear among the adult progeny). Every ten generations, from each chromosome line, Mukai mated males and females of the *Cy/+* genotypes, yielding *Cy/Cy*, *Cy/+*, and +/+. These should fit a 1:2:1 ratio, but *Cy/Cy* is lethal, so we expect a 2:1 ratio of *Cy/+* to +/+. If the ratio of *Cy/+* to +/+ is greater than 2:1, we can assume that relative viability of the + chromosome in that cross has been reduced. The magnitude of the deviation measures how much viability has been reduced by recessive mutations that have accumulated on this chromosome. Since all the fly lineages started with copies of a single + chromosome, the variation among lineages in their degree of departure from a 2:1 ratio can be used to estimate how many different mutations have occurred.

For bristle number in *Drosophila* and pupal weight in flour beetles (*Tribolium*), V_m/V_E is about 0.1 percent per generation; that is, the genetic variance increases by a factor of 0.001 of the environmental variance, per generation, due to new mutations. Thus, it would take about 500 generations for an initially homozygous population to reach a heritability of 0.5, assuming that none of the new genetic variation were eroded by genetic drift or natural selection. The magnitude of this MUTATIONAL VARIANCE varies somewhat among characters and species (Lynch 1988).

The rate of polygenic mutation can be increased substantially by transposable elements. In *Drosophila melanogaster*, transposable elements called *P* elements cause a syndrome called HYBRID DYSGENESIS (Kidwell 1983; Engels 1983). "P" strains of flies carry *P* elements, and "M" strains do not. "P" strains carry a maternally transmitted cytoplasmic factor that represses transposition of *P* elements, whereas "M" strains lack this factor. In F_1 progeny of the cross ♀M × ♂P, *P* elements transpose freely, causing the "hybrid dysgenesis" syndrome, which includes reduced fertility, recombination in males (crossing over does not normally occur in males), and a high incidence of chromosome breakage and gene mutation.

Trudy Mackay and her colleagues have carried out several experiments (e.g., Mackay et al. 1992) in which they estimated the variance due to mutations in dysgenic flies, compared with that in inbred strains without *P* elements. On average, each transposed *P* element (each "insert") reduced viability by about 12 percent in the homozygous condition and by 5.5 percent in the heterozygous condition. Most interestingly, the genetic variance in bristle number increased by about $V_m/V_E = 0.03$ per generation—about 30 times the rate at which mutation adds genetic variance in "normal" flies. Many of the *P*-induced mutations had only slight effects on bristle number, but a few had rather large effects.

In summary, although any given mutation is a rare event, the rate of origination of new genetic variation in the genome as a whole, and for individual polygenic characters, is appreciable, and can sometimes be quite high.

This conclusion has many important implications that we will discuss in subsequent chapters. The high input of mutations may account for much of the genetic variation within populations (see Chapter 14). Because mutations are more likely to interfere with enzyme function and developmental pathways than to improve them, the steady

input of mutations is likely to cause characters to degenerate in the long term, unless such mutations are removed by natural selection. In the same vein, the input of deleterious mutations lowers fitness, and would ultimately cause the extinction of populations if it were not counteracted by "purifying" natural selection. Selection for characteristics that mitigate the effects of deleterious mutations is thought by some authors to be an important factor in the evolution of recombination and sex (see Chapter 21).

Phenotypic Effects of Mutations

It is useful to make a somewhat arbitrary distinction between the effects of mutations on phenotypic features such as morphology and physiology, and their effects on components of fitness (viability, fertility, developmental rate). The distinction is arbitrary because all differences in fitness are based on physiological, morphological, or behavioral differences. Nevertheless, we can often describe a phenotypic difference, and then ask whether or not it affects fitness. (However, we are often ignorant of the physiological or morphological basis for manifest differences in fitness.)

Effects on the phenotype The phenotypic effects of mutational changes in DNA sequence range from none to drastic. At one extreme, synonymous base pair changes are expected to have no evident phenotypic effect, and this is apparently true also of many amino acid substitutions, which seem not to affect protein function. (Even synonymous mutations, however, can affect fitness, probably because of their effects on the rate of translation of RNA into protein; see Chapter 22.) Mutations of slight effect are exemplified by the mutational increase in variation in bristle number in *Drosophila* described above. In this and many other such instances, the effects of the mutations in aggregate can be measured, but many of the mutations that contribute to the variation have so slight an effect that they cannot be isolated individually for study. Identifiable mutations, however, range from subtle to striking; in *Drosophila*, some have minor effects on features such as eye color and wing veins, and others have large effects, such as reduction of the wings to tiny, useless structures. Similarly drastic mutations are known in humans and in domesticated plants and animals.

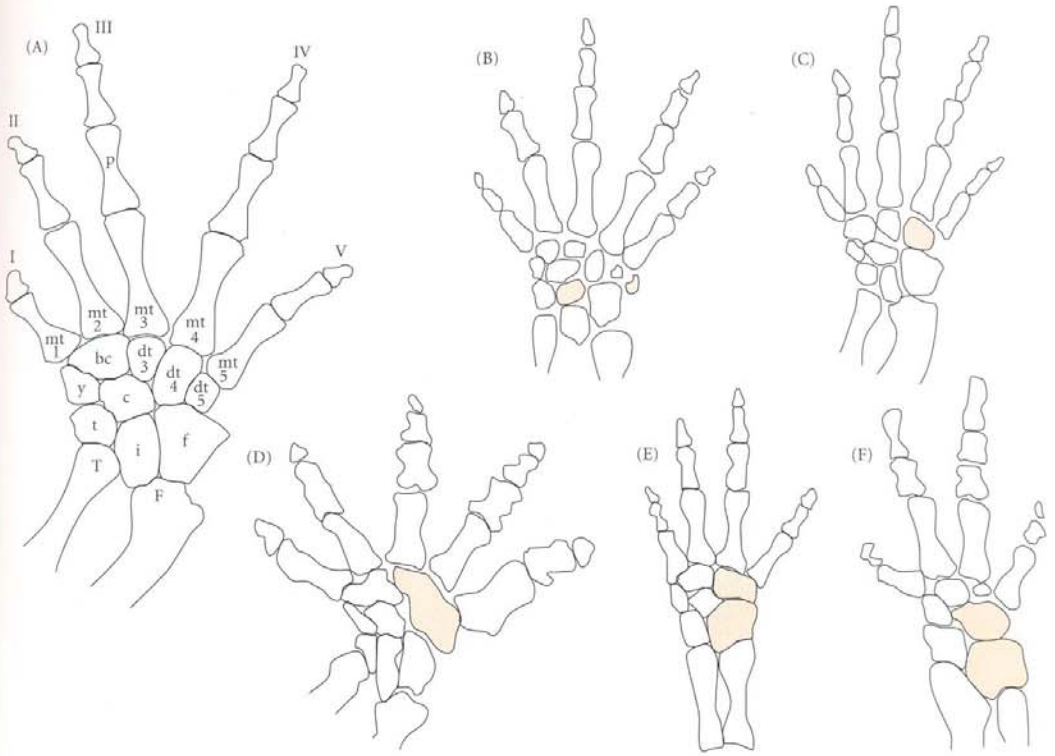
Many mutations affect behavior. For example, in mutants of the *per* (*period*) locus in *Drosophila melanogaster*, a variety of biological rhythms, including circadian activity cycles and the frequency of the wingbeat in the male's courtship behavior, are aberrant (Hall and Kyriacou 1990; Kyriacou 1990). The *yellow* mutation in this species affects not only body color, but also the rate at which males perform certain courtship behaviors, and such males are less successful in mating than wild-type males. A fascinating example of behavioral mutations is illustrated by the honeybee (*Apis mellifera*), in which larvae reside in individual cells of the comb, and are sometimes killed by a bacterial disease that spreads through the hive unless dead larvae are removed. Workers must uncap the cells, then remove the cadavers. Each of these behaviors is abolished by a single dominant mutation, so the normal behavior depends on the bee's being homozygous recessive at two loci. (See Ehrman and Parsons 1981 for more information on behavioral genetics.)

Perhaps the most fascinating of the drastic mutations that affect morphology are HOMEOTIC MUTATIONS, which redirect the development of one part of the body into another. These mutations occur in the genes that determine the basic "body plan" of an organism, conferring a distinct identity on each part of the developing body and regulating the activity of other genes that affect the features of each such part (see Chapters 3, 23). In *Drosophila*, mutations in the *Antennapedia* and *Bithorax* gene complexes transform the identity of certain body segments, so that, for instance, legs may develop in place of antennae, or wings in place of halteres (structures in the Diptera that have been derived in evolution from the second pair of wings).

The limits of mutation It cannot be stressed too strongly that even the most drastic mutations cause alterations of one or more *pre-existing* traits. Mutations with phenotypic effects alter developmental processes, but they cannot alter developmental foundations that do not exist. We may conceive of winged horses and angels, but no mutant horses or humans will ever sprout wings from their shoulders, for the developmental foundations are lacking. (J. B. S. Haldane, one of the founders of population genetics, once quipped that humans could never evolve into a race of angels, because they lack the genetic capacity for both wings and the moral sense.) Thus, some morphologies are highly unlikely, or even impossible, for reasons that we usually do not understand because of our ignorance of developmental pathways. For instance, the numerous ankle bones of a salamander can be organized (by fusions and fissions) in many imaginable ways, but some conceivable patterns have never been found, either as intraspecific variants or as species-typical characters. For example, variations in the foot structure of the newt *Taricha granulosa* almost all occur as typical conditions in other species of salamanders (Figure 10.7), whereas other conceivable variations have not been seen either within *Taricha* or among other species (Shubin et al. 1995). Such data suggest that some variations are more likely to arise and to contribute to evolution than others. Similarly, in laboratory stocks of the green alga *Volvox carteri*, new mutations affecting the relationship between the size and number of germ cells correspond to the typical state of these characters in other species of *Volvox* (Koufopanou and Bell 1991). Because mutations can supply only certain kinds of variation, constrained by existing developmental pathways, some imaginable paths of evolution are closed to a species, or at least are less likely than others.

The likelihood of a given evolutionary change may also be affected by the number of different mutations that generate a particular phenotype. When several or many loci affect the trait (polygeny), evolution can be based on many combinations of mutations, so genetic variation is likely to

FIGURE 10.7 The normal arrangement of tarsal bones in some salamanders. The tarsals lie between the metatarsals (mt1–mt5 in part A) and the tibia (T) and fibula (F). (A) The usual arrangement in the newt *Taricha granulosa*. Almost all variations found in this species are the normal condition in certain other species (B–F). For example, fusion of dt4 and dt5 (dt4 + dt5) is found in *Cynops* (B) and some other salamanders. The other species shown, and their tarsal condition relative to that of *T. granulosa*, shown by the shaded bones, are (C) *Liua* (two extra bones); (D) *Bolitoglossa dolfleini* (dt3 + dt4 + dt5); (E) *Necturus* (dt3 + dt4; f + i); and (F) *Thorius* (dt4 + dt5 + c; f + i). Digit V has been lost in *Necturus* and *Thorius*. (After Shubin et al. 1995.)



arise. Very often, a similar phenotype can be produced by mutation at several different loci, as we know both from mutation in laboratory stocks (e.g., phenotypically similar eye color mutations in *Drosophila*) and from genetic analyses of natural populations. For example, when different copper-tolerant populations of the monkeyflower *Mimulus guttatus* are crossed, the variation in copper tolerance is greater in the F_2 generation than within either parental population, indicating that the populations differ in the loci that confer tolerance (see Cohan 1984 for this and other examples).

Nevertheless, the number of loci providing variation in some polygenic traits is not all that great—perhaps five to ten, with a few loci responsible for much of the variation. In *Drosophila*, for example, some mutations that have small effects on quantitative variation in bristle number are alleles of well-known mutations with large effects (see Chapter 14). Certain phenotypes can apparently be produced by mutation at a very few loci, or perhaps only one. Insecticide resistance offers particularly striking examples. For instance, resistance to dieldrin in different populations of *Drosophila melanogaster* is based on repeated occurrences of the same mutation, which, moreover, is thought (because of its map position) to represent the same mutation that confers resistance in flies (*Lucilia cuprina*, *Musca domestica*) that belong to two other families (ffrench-Constant et al. 1990). A remarkable case is that of organophosphate resistance in the cosmopolitan mosquito *Culex pipiens* (Raymond et al. 1991). Resistance is associated with an electrophoretically distinguishable form of an esterase enzyme that is "overproduced": the locus is repeated (amplified) about 60 times in resistant strains. Unlike other alleles at this locus, which display considerable variation in DNA sequence (based on restriction site mapping), this allele showed identical DNA sequences in resistant mosquitoes from Pakistan, Egypt, Congo, Ivory Coast, California, and Texas. The most reasonable inference is that within the few

decades in which organophosphates have been used for mosquito control and have imposed selection for resistance, a single mutation (i.e., amplification of this locus) has spread by migration across three continents from its site of origin. Even though untold millions of mosquitoes throughout these regions have experienced natural selection for resistance, apparently only one mutational event has proved successful. This may mean that very few genes—perhaps only this one—undergo suitable mutation, and that such a mutation is a very rare event.

In such instances, we can conclude that the supply of rare mutations can limit the capacity of species for adaptation. This may help to explain why species have not become adapted to a broader range of environments, or why, in general, species are not more adaptable than they are (Bradshaw 1991).

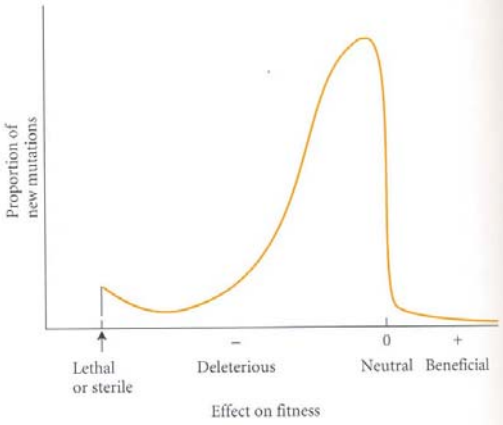
Effects of Mutations on Fitness

The distribution of the effects of new mutations on fitness extends from very positive to very negative (e.g., lethal or sterile). We should like to know what the form of this distribution is: what fraction of mutations are strongly or slightly advantageous, neutral, or slightly or strongly disadvantageous (Figure 10.8). The precise form of the distribution is not known, but it is surely not fixed, for the fitness consequences of many mutations depend on the population's environment and even on its existing genetic constitution.

Undoubtedly many mutations are neutral or nearly so, having very slight effects on fitness (see Chapter 11). However, many do affect fitness, and their *average*, or net, effect is deleterious. This was shown, for example, by the decline in mean fitness as mutations accumulated in Mukai's (1972) experiment, described above (see Figure 10.6). These were spontaneous mutations; new mutations experimentally induced by the mutagenic (mutation-causing) chemical EMS also have a net deleterious effect. Ellen Wijsman (1984) showed that the offspring of *Drosophila melanogaster* that had been exposed to EMS suffered in competition with *Drosophila simulans* to a greater extent than nonmutagenized controls.

In experiments such as those performed by Mukai, it is evident that mutations that affect fitness have occurred, but nothing further is known about which genes have mutated, or what physiological or developmental features have been altered. Conversely, we can study the fitness effects of mutations of known phenotypic effect. These effects often depend on the environment; for example, mutations that slightly alter color pattern might or might not be advantageous, depending on whether or not some environmental factor were to favor a change in this feature. However, we must also recognize that *most mutations have pleiotropic effects*,—i.e., they affect more than one characteristic (see Chapter 3). We noted above, for example, that the *yellow* mutation in *Drosophila* affects not only body color, but also several components of male courtship behavior. The polygenic mutations that generate variation in bristle number in *Drosophila* often reduce larval viability (Kearsey and

FIGURE 10.8 A possible frequency distribution of the effects on fitness of newly arisen mutations. (The real frequency distribution is unknown.) This figure reflects the widespread belief that the great majority of mutations are deleterious or nearly neutral (i.e., with nearly zero effect), and that only a very small proportion are beneficial. The curve rises at the left end because even the worst "superlethal" or "supersterile" mutations cannot cause more than total lethality or sterility.



Barnes 1970; Mackay et al. 1992). Many of the conspicuous mutations studied by *Drosophila* geneticists (as well as those in other organisms) have strongly negative pleiotropic effects on survival and reproduction. Some such mutations appear to lower fitness unconditionally, but others adversely affect viability only under certain conditions (e.g., of temperature) (Dobzhansky 1937). In some cases, the basis of deleterious pleiotropic effects is understood; for example, some mutations that affect *Drosophila* bristles disrupt the development of the nervous system. The developmental disharmonies caused by many major mutations (i.e., those with large phenotypic effects) are cited by those who argue that major mutations play little role in adaptive evolution, which, they claim, is based mostly on the accumulation of many genetic changes, each with only a slight phenotypic effect (Fisher 1930).This is a controversial subject, however, and some authors have argued that mutations with large effects have been important in adaptive evolution (see Chapter 23).

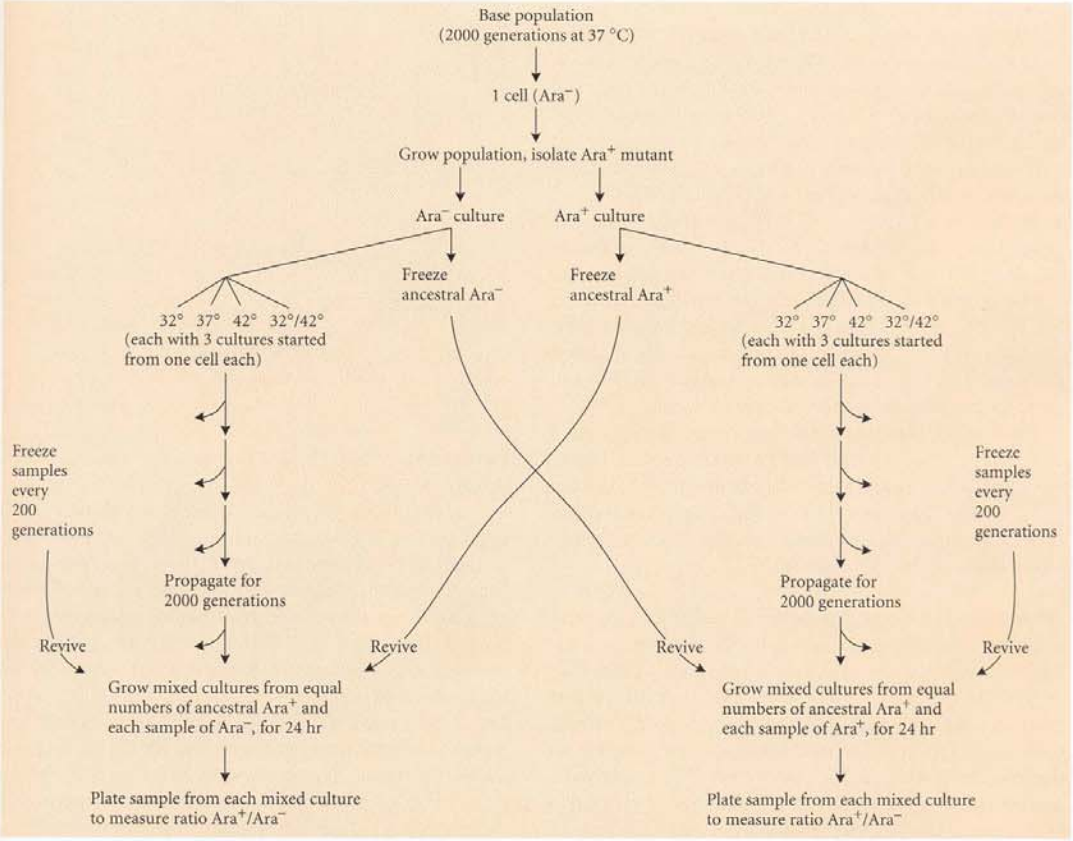
Not all mutations are deleterious. The following examples illustrate experiments in which advantageous mutations have been shown to occur in bacteria. Because of their short generation times and the ease with which huge populations can be cultured, microorganisms such as bacteria and yeasts are exceptionally useful for studying mutation—as well as other aspects of the evolutionary process (Dykhuizen 1990).

Adaptation to temperature Experimental populations of bacteria such as *Escherichia coli* can be kept in a state of

continual population growth (cell division) in several ways, such as transferring a sample to a new vessel of liquid nutrient broth every day. Because bacteria can be frozen (during which time they undergo no genetic change) and later revived, samples taken at different times from an evolving population can be stored, and their fitness can later be directly compared. The fitness of a genotype can be measured by its rate of increase in numbers, relative to that of another genotype with which it competes in the same culture, but which bears a genetic marker and so can be distinguished from it. Suppose, for example, that a culture is begun with equal numbers of genotypes A and B, and that after 24 hours B is twice as abundant as A. If the numbers have grown at a constant rate of cell division, each initial cell has produced 2^x descendants, where x is the number of generations (cell divisions). Thus if genotypes A and B have grown at the respective rates of 2^5 and 2^6 (i.e., B has one more generation per 24 hours), their relative numbers are 32:64, or 1:2. The relative fitnesses of the genotypes—i.e., their relative growth rates—are measured by their rates of cell division per day, namely 5:6 or 1.0:1.2. If the genotypes had the same growth rates—e.g., 2^5—both would increase in number, but their fitnesses would be equal.

Albert Bennett, Richard Lenski, and John Mittler (1992) studied adaptation to temperature in *E. coli*, using replicate cultures derived from a single bacterial cell taken from a culture that had been propagated for 2000 generations at 37°C (Figure 10.9). Among the progeny of this cell, they found a mutant (Ara^+) that forms white colonies on nutrient agar plates containing a certain compound, whereas the Ara^- parent forms red colonies. Thus the relative numbers of Ara^+ and Ara^- bacteria in a culture can be measured by plating out a sample (i.e., spreading a dilute suspension of bacterial cells on a nutrient agar, on which each parent cell forms a separate colony of asexually produced descendants). Ara^+ is only a *marker*: it does not affect fitness, be-

FIGURE 10.9 The design of Bennett et al.'s experiment to monitor adaptation to different temperatures in cultures of *Escherichia coli* that were derived from a single cell, and hence initially lacked genetic variation. The fitness of each culture was measured by allowing it to compete with a genetically marked (*Ara*+ or *Ara*–) sample from the ancestral culture.



cause when *Ara*$^+$ and *Ara*$^-$ were grown together, their relative numbers did not change. Except for the marker mutation, the *Ara*$^+$ and *Ara*$^-$ cells were initially genetically identical, cloned from the same ancestral cell.

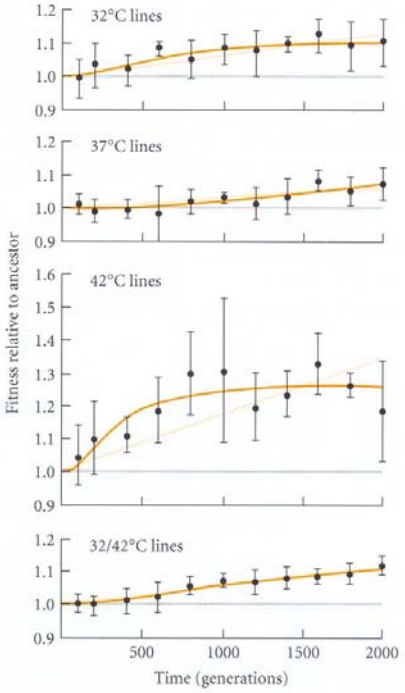
Bennett et al. grew sufficient numbers of *Ara*$^+$ and *Ara*$^-$ cells so that they could freeze a sample of each genotype (the ANCESTORS) and also grow several different populations of each genotype at each of several temperatures: 32°, 37° (the ancestral temperature), 42°, and a regime of daily alternation between 32° and 42°. They grew these populations at these temperatures for 2000 generations. Every 200 generations or so, they removed and froze a sample of each population. They could then determine whether a population had adapted by thawing and competing a sample from, say, the 42° population after 200 generations with a sample from the ancestral population that had not experienced the novel temperature regime. In order to distinguish the two competitors, they grew the 42° *Ara*$^+$ cells with the ancestral *Ara*$^-$ (as well as the converse, 42° *Ara*$^-$ with ancestral *Ara*$^+$, to be sure that the *Ara* allele did not affect the course of adaptation—which it did not). In this manner, the fitness of the bacteria that had been propagated at 42° could be measured, relative to the fitness of the ancestral genotype, by the change in the relative number of the two types over 24 hours.

Figure 10.10 shows some of the results of this experiment. Over the course of 2000 generations, each population increased in fitness (growth rate) relative to the ancestral genotype, when fitness was measured at the temperature at which the population had been maintained. This increase was based on a combination of two factors: new mutations that provided the capacity for more rapid cell division, and an increase in the frequency of these new alleles within the population by natural selection. It is interesting that fitness increased not only under the novel temperatures, but also in the population maintained at the "ancestral" temperature, 37°. The gain in fitness was less at this than at the other temperatures, which is to be expected because the ancestral genotype should already have been adapted to 37°—but some further adaptation nevertheless took place.

All these populations were derived from a single haploid cell. Moreover, this strain of *E. coli* is strictly asexual. Therefore, the adaptation to several temperatures must have been based on new mutations, showing that adaptive mutations do indeed occur. (In Chapter 17, we will describe further elaborations of this experiment.)

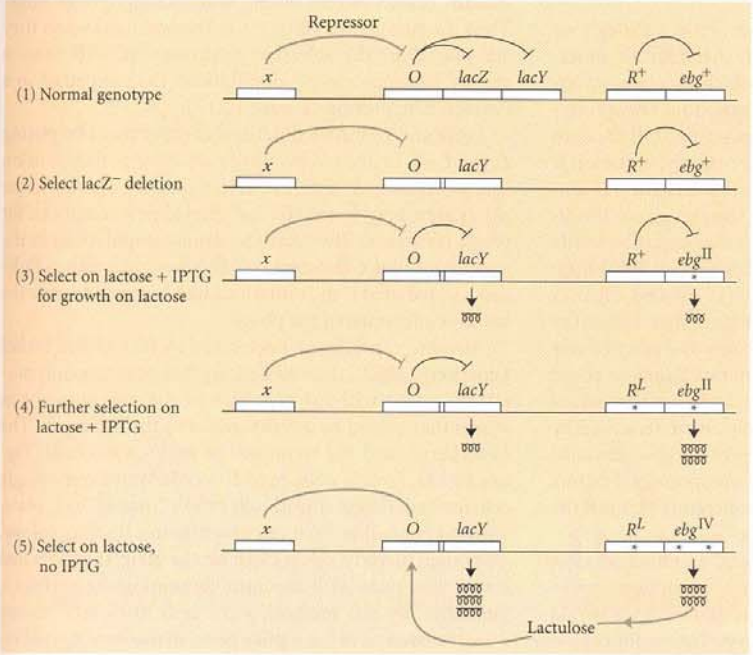
Mutations for novel biochemical abilities Bacteria can be screened for mutations that affect biochemical capacities by placing them on a medium on which that bacterial strain ordinarily cannot grow, such as a medium that lacks an essential amino acid or other nutrient. Whatever colonies do appear on the medium must have grown from the few cells in which mutations occurred that conferred a new biochemical ability. Several investigators have used this technique to select for the evolution of new biochemical pathways (e.g., Clarke 1974; Hall 1982, 1983). The experiments of Barry Hall (1982, 1983) may serve as an example (Figure 10.11).

FIGURE 10.10 Adaptation to temperature in experimental populations of *Escherichia coli* studied by Bennett et al. A population's fitness is its growth rate relative to that of the ancestor, which is set at 1.0 (gray line). Solid circles with vertical bars show the mean fitness of replicate populations and a measure of variation (confidence interval) among them. The colored lines are two statistically calculated expressions of the increase in fitness with time. Because the populations initially lacked genetic variation, the increase in adaptation to different temperatures was due to natural selection acting on new advantageous mutations. (After Bennett et al. 1992.)



The ability of wild-type *Escherichia coli* to use the sugar lactose as an energy source resides in the *lac* operon (see Chapter 3). In this operon, the *lacZ* gene codes for β-galactosidase, which hydrolyzes lactose, and the *lacY* gene codes for a permease, which allows lactose to enter the cell. Transcription of both genes occurs when the operator gene (*O*) is de-repressed, i.e., freed from a repressor protein. When lactose leaks into the cell, it is hydrolyzed by small standing levels of β-galactosidase into glucose and galactose, which are used for energy, and into allolactose, which de-represses the operator, allowing the production of more β-galactosidase and of permease, which enables lactose to enter more freely, thus permitting growth of the cell.

FIGURE 10.11 Diagram of mutational changes in populations of *Escherichia coli* selected for the ability to metabolize lactose in experiments by Hall. Lines ending with crossbars indicate repression of one gene's transcription by another gene. (1) Genes in wild-type *E. coli*. The existence of *R* and *ebg* was not known when the experiment began. (2) A strain with a deletion of *lacZ* (encoding the enzyme that normally metabolizes lactose) is used for the experiments. (3) Cells are placed on a medium containing lactose and IPTG, which enables transcription (symbolized by the curlicue) of the *lacY* gene. This gene encodes a permease. A few cells grow; these prove to have a mutation (^{II}) in *ebg*. (4) Selection for better growth is successful due to the occurrence of a mutation in *R* that enhances the transcription of *ebg*. (5) Further selection, without IPTG, is successful, based on occurrence of another mutation in *ebg*. The new *ebg* enzyme produces lactulose from lactose; lactulose de-represses the operator (*O*), which promotes greater transcription of *lacY*, resulting in greater flux of lactose into the cells.



Hall obtained a strain in which most of the *lacZ* gene had been deleted. This strain, which may be denoted $O^+lacZ^-lacY^+$, cannot synthesize β-galactosidase, and so cannot grow on a medium in which lactose is the sole energy source. On such a medium, moreover, the *lacY* gene remains repressed (so that little lactose can enter the cell) because no allolactose is formed. A chemical additive (IPTG), however, will induce the permease.

Hall selected for the ability to use lactose by putting this strain on a lactose medium with IPTG. He obtained mutants that could grow on lactose. These proved to have mutations in a previously unknown gene, located on the chromosome far from the *lac* operon, which he called *ebg* ("evolved β-galactosidase"). These mutants (ebg^{II}) grew slowly because the *ebg* gene is largely (but not completely) repressed by its regulatory gene (R^+). By selecting for mutants with high growth rates, Hall obtained regulatory mutations (R^L) in which lactose turns off the repression by the regulatory gene. Thus he had obtained a double mutant, $O^+lacZ^-lacY^+ebg^{II}R^L$, which, however, still required IPTG in order for lactose to enter the cell. By growing the double mutant strain on lactose without IPTG, he obtained ebg^{IV}, a second mutation of the *ebg* gene (i.e., $ebg^{II} \rightarrow ebg^{IV}$). The enzyme produced by ebg^{IV} metabolizes lactose to lactulose, which induces the permease gene $lacY^+$.

This genotype ($O^+lacZ^-lacY^+ebg^{IV}R^L$), then, represents the evolution of an entire system of lactose utilization, consisting of changes in enzyme structure enabling hydrolysis of the substrate; alteration of a regulatory gene so that the enzyme can be synthesized in response to the presence of its substrate; and evolution of an enzyme reaction that induces the permease needed for the entry of the substrate. One could not wish for a better demonstration that mutation and selection in concert can give rise to complex adaptations.

Mutation as a Random Process

Mutations occur at random. It is extremely important to understand what this statement does and does not mean. It does not mean that all conceivable mutations are equally likely to occur, because, as we have noted, the develop-

mental foundations for some imaginable transformations do not exist. It does not mean that all loci, or regions within a locus, are equally mutable, for geneticists have described differences in mutation rates, at both the phenotypic and molecular levels, among and within loci (Woodruff et al. 1983; Wolfe et al. 1989). It does not mean that environmental factors cannot influence mutation rates: ultraviolet and other radiation, as well as various chemical mutagens and poor nutrition, do indeed increase rates of mutation.

Mutation *is* random in two senses. First, although we may be able to predict the *probability* that a certain mutation will occur, we cannot predict which of a large number of gene copies will undergo the mutation. The spontaneous process of mutation is stochastic rather than deterministic. Second, and more importantly, mutation is random in the sense that *the chance that a particular mutation will occur is not influenced by whether or not the organism is in an environment in which that mutation would be advantageous.* That is, the environment does not induce adaptive mutations. As Dobzhansky (1970) said, "It may seem a deplorable imperfection of nature that mutability is not restricted to changes that enhance the adaptedness of their carriers. However, only a vitalist Pangloss could imagine that the genes know how and when it is good for them to mutate."* Dobzhansky's point is underscored by our knowledge of the molecular nature of the gene: it is hard to conceive of a mechanism whereby environmental factors could direct the mutation process by dictating that just the right base pair changes should occur.

The argument that adaptively directed mutation does not occur is one of the fundamental tenets of modern evolutionary theory. If "directed mutation" did occur, it would introduce a Lamarckian element into evolution, for organisms would then acquire adaptive hereditary characteristics in response to their environment. The death blow to such "neo-Lamarckian" ideas came in the 1940s and 1950s from experiments with bacteria that appeared to demonstrate that adaptation is a consequence of spontaneous, random mutation followed by natural selection, rather than mutation directed by the environment.

Salvador Luria and Max Delbrück (1943) based an experiment on the following reasoning. Suppose that a large number of bacterial populations, all initially genetically identical, are each grown to size N from a single individual and then placed in a selective environment. If the environment induces an adaptive mutation (one that confers resistance to the selective factor) with some low probability p, the average number of resistant cells in each population will be pN. But as in any probabilistic process, there will be variations among populations in the number of resistant cells. The number of populations with 0, 1, 2 ... k resistant cells will follow a POISSON DISTRIBUTION, in which the variance equals the mean. On the other hand, if the resistance mutations arose spontaneously *before* the populations were placed in the selective environment, any cell that mutated during that time will transmit the mutation to all its descendants, so that some of the populations will include a large number of mutant cells, whereas others will include few or none. Thus the number of resistant cells, measured when they are placed in the selective environment, will show a greater variance among populations than expected in a Poisson distribution (Figure 10.12).

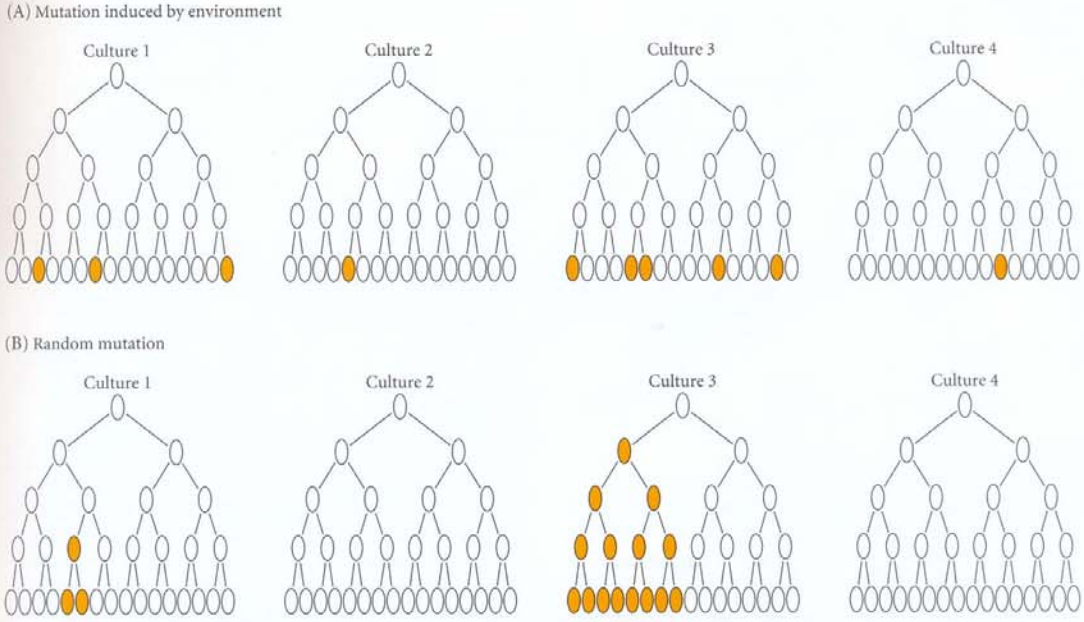
Luria and Delbrück did just this experiment by plating *E. coli* from each of many phage-sensitive populations on agar plates covered with T1 bacteriophage. Colonies on the plates grew only from cells that carried new mutations for phage resistance. The variation among populations in the number of phage-resistant colonies was greater than Poisson, as predicted if the mutations had happened *before* the bacteria encountered the phage.

Another experiment, performed by Joshua and Esther Lederberg (1952), showed directly that advantageous mutations occur without exposure to the environment in which they would be advantageous to the organism. The Lederbergs used the technique of REPLICA PLATING (Figure 10.13). From a culture of *E. coli* derived from a single cell, the Lederbergs spread cells onto a "master" agar plate, without penicillin. Each cell gave rise to a distinct colony. They then placed a velvet cloth on the plate, then touched it to a new plate with medium containing the antibiotic penicillin. By this method, some cells from each colony were transferred to the replica plate, in the same spatial relationships as the colonies from which they had been taken. A few colonies appeared on the replica plate, having grown from penicillin-resistant mutant cells. When all the colonies on the master plate were tested for penicillin resistance, those colonies (and only those colonies) that had been the source of penicillin-resistant cells on the replica plate displayed resistance, showing that the mutations had occurred before the bacteria were exposed to penicillin.

Since these classic experiments, biologists have generally accepted that mutation is adaptively random rather than directed. Recently, however, several investigators have reported results, again with *E. coli*, that at face value seem to suggest that some advantageous mutations might be directed by the environment (Box 10.D). Their interpretations have been challenged by other investigators, who find no compelling evidence for directed mutation in these experiments. The issue is being vigorously debated at this time (Sniegowski and Lenski 1995). It is hard to imagine mechanisms by which mutations for many characteristics, such as morphological features, could be adaptively directed by environmental factors such as natural enemies and a species' social interactions; but a convinc-

*Dobzhansky refers to Dr. Pangloss, a character in Voltaire's satire *Candide*, who despite the most tragic contrary evidence taught that "all is for the best in this best of all possible worlds." Vitalism is the belief, rejected by biologists, that processes in living beings are produced by an immaterial vital force, rather than by purely chemical and physical processes.

FIGURE 10.12 Two hypotheses for the origin of mutations for phage resistance in bacteria. Luria and Delbrück's experiment was designed to determine which hypothesis is correct. In both A and B, the pedigrees of bacteria in several cultures, each derived from a single cell, are shown. Bacteria are exposed to phage in the fourth generation. Shaded cells carry a mutation for phage resistance. It is assumed (and experimentally demonstrated) that a small minority of cells are resistant. (A) Under the hypothesis that exposure to phage directly induces mutations for resistance, the few resistant cells should be distributed at random among the cultures. (The mathematical description of the expected frequency distribution is called the Poisson distribution). (B) Under the hypothesis that mutations for resistance occur at random, before exposure to phage, resistant cells will be "clumped" in certain cultures and absent in all others, because mutations in ancestral cells are inherited by multiple descendants. (After Stent and Calendar 1987.)



ing demonstration that the environment can direct even some mutations, such as those altering biochemical capacities, would constitute a major change in our understanding of evolution.

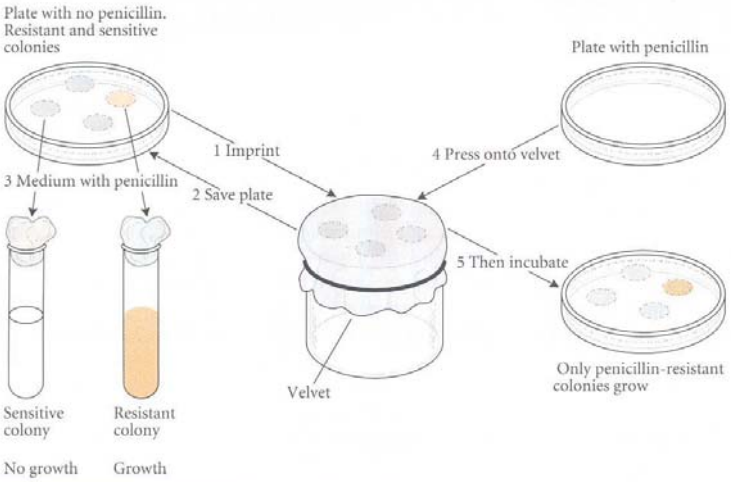
Recombination as a Source of Variation

Release of Genetic Variation

All genetic variation owes its origin ultimately to mutation, but in the short term, a great deal of the genetic variation within populations arises through recombination. Several mechanisms of genetic exchange occur in prokaryotes, but we will restrict our discussion to sexually reproducing eukaryotes, in which genetic variety arises from two processes: the union of genetically different gametes, and the formation of gametes with combinations of alleles different from those that united to form the individual that produces them. The formation, during meiosis, of genetically variable gametes is the consequence of the independent segregation of nonhomologous chromosomes and of crossing over between homologous chromosomes.

The potential genetic variation that can be released by recombination is enormous. If *Drosophila* has about 10,000 functional loci, and if an average individual is heterozygous at 10 percent of its loci (based on estimates from electrophoresis; see Chapter 9), then an individual could conceivably produce 2^{1000}, or approximately 10^{300}, genetically different kinds of gametes, which is immensely greater than Avogadro's number (6.023×10^{23}), the number of molecules per mole. (In actuality, the vast majority of theoretically possible gene combinations will never be formed, because they would require an immense number of crossovers between tightly linked loci.) At a more modest level, if an individual is heterozygous for only one locus on each of five pairs of chromosomes, independent segregation alone generates $2^5 = 32$ allele combinations among its gametes, and mating between two such individuals can give rise to $3^5 = 243$ genotypes among their progeny. (This assumes that both parents have the same heterozygous genotype, say, A_1A_2, at each locus. If each is heterozygous for a different pair of alleles, A_1A_2 and A_3A_4, their progeny can carry $4^5 = 1024$ different genotypes.)

FIGURE 10.13 The method of replica plating, used to show that mutations for penicillin resistance arise spontaneously before exposure to penicillin, rather than being induced by it. We begin (upper left) with an agar plate with numerous colonies of bacteria, each derived from one cell. Part of each colony is transferred to a velvet cloth (i.e., an "imprint") (step 1), and the plate is saved (2). Each colony is tested for penicillin resistance by growing it in medium with penicillin (3). A clean plate with penicillin is pressed onto the velvet (4). This plate is thus seeded with cells from each original colony (5), but only cells from original colonies shown (in step 3) to be penicillin-resistant grow on the new plate. (After Srb et al. 1965.)



If each locus affects a different feature, this represents a great variety of character combinations. If all five loci have equal and additive effects on a single character, such as size, the range of variation among the offspring can greatly exceed the difference between the parents. For instance, if each substitution of + and – alleles in a genotype adds or subtracts one unit of phenotype, two quintuply heterozygous parents (both +–+–+/–+–+–), both of size 20, could have offspring ranging in size from 15 (–/– at all five loci) to 25 (+/+ at all five loci). Each of the two extreme phenotypes, however, would have very low frequency among the offspring (1/1024, assuming independent segregation). By the same token, if + and – alleles have intermediate frequencies (near 0.5) in a population, the majority of individuals will be heterozygous at many loci, hence intermediate in size. Multiple homozygotes, with extreme phenotypes, will segregate only infrequently. (For example, if the frequencies of the + and – alleles are $p = q = 0.5$ at each of five independently segregating loci, the frequency of individuals homozygous +/+ at all loci is $(p^2)^5 = 0.0009766$.)

Some Evidence

The validity of these theoretical constructs has frequently been demonstrated. For example, Figure 7 in Chapter 9 shows results typical of several studies of *Drosophila pseudoobscura* by Theodosius Dobzhansky and his colleagues, who, using crosses by which wild chromosomes can be "extracted" from natural populations (see Figure 6 in Chapter 9), measured the egg-to-adult viability of homozygotes for numerous chromosomes, as well as flies heterozygous for random pairs of these same chromosomes. The variation among homozygotes is greater than among heterozygotes, which is what we would expect if several or many alleles on the chromosome confer extreme phenotypes when homozygous.

In order to judge how much variation is released by recombination, the Dobzhansky team (Spassky et al. 1958) made all possible crosses between homozygotes for only ten different chromosomes. All these chromosomes conferred almost the same, nearly normal viability when homozygous. From the F_1 female offspring, in which crossing over occurred, they then extracted recombinant chromosomes and measured their effect on viability when homozygous (Figure 10.14). Even though the original ten chromosomes differed little, the variance in viability among the recombinant chromosomes was more than 40 percent of the variance among homozygotes for much larger samples of chromosomes from natural populations (as illustrated in Figure 7 in Chapter 9). Thus a single episode of recombination among just ten chromosomes generates a large fraction of a wild population's variability. Some of the recombinant chromosomes were "synthetic lethals," meaning that recombination between two chromosomes that yield normal viability produced chromosomes that were lethal when they were made homozygous. This implies that each of the original chromosomes carried an allele that did not lower viability on its own, but

BOX 10.D

Directed Mutation?

The classic experiments by Luria and Delbrück and by the Lederbergs show that some mutations arise spontaneously, but do not rule out the possibility that some other mutations, under certain conditions, might be environmentally directed. John Cairns and his colleagues (1988) and Barry Hall (1988) reported results suggesting directed mutation for metabolic capacities. In both cases, the investigators scored mutations that occurred not in rapidly growing populations, but in bacteria that were plated on agar with an energy source that the original strain could not use. Cairns et al. plated *lac*$^-$ bacteria, unable to use lactose, on lactose medium, and scored the incidence of *lac*$^+$ mutants that could grow on lactose. They reported that these mutations occurred with a lower variance than if they had been spontaneous (cf. the Luria-Delbrück experiment), and that they occurred only after a long delay (implying that the mutations had not already been carried by the bacteria first plated).

Hall plated bacteria on a medium with salicin, which, he said, could be used for energy only if two mutations occurred in a galactosidase operon. The mutants that grew on salicin had to have experienced both a mutation of the regulatory region (R) and excision of a transposable element (IS) from the structural gene, enabling this operon to function. Hall presented evidence that the IS excision mutations (IS$^-$) occurred before the R mutations, that the IS$^-$ mutations alone could not use salicin, and that the combination of the IS$^-$ and R mutations occurred much more frequently than expected. He suggested that the IS$^-$ excision occurred only on medium with salicin, and that it "serves only to create the potential for a secondary selectively advantageous mutation"—which implies not only directed mutation, but also "anticipatory" mutation.

Several authors have provided criticisms and alternative interpretations of these experiments (see Sniegowski and Lenski 1995 and references therein). For example, the delay in the appearance of mutants able to grow on the medium, as well as the low variance, might be due to an undetected first mutation that provided some capacity for growth, resulting in populations large enough for ordinary spontaneous mutations, conferring a high growth rate, to occur. In the same vein, the high incidence of the double mutant in Hall's experiment might be explained if IS$^-$ excision is stimulated by starvation (rather than by salicin per se), if the IS$^-$ mutants can indeed grow in the presence of salicin, and if spontaneous R mutations then occur in the large resultant populations. Mittler and Lenski (1992) provide evidence for just this hypothesis. So far, it seems that explanations consistent with orthodox neo-Darwinian theory can be found for apparently directed mutation. (For the contrary view, see Thaler 1994.) Most geneticists agree that truly directed mutation should be an explanation of last resort.

did cause death when combined with another allele, at another locus, on the other chromosome. This is a striking example of epistasis (nonadditive interaction among genes).

Erosion of Variation by Recombination

Recombination is a double-edged sword: it both generates and destroys genetic variation. In sexually reproducing populations, genes are transmitted to the next generation, but genotypes are not: they end with organisms' deaths, and are reassembled anew in each generation. Thus an unusual, favorable gene combination may occasionally arise through recombination, but if such individuals mate with other members of the population, it will be lost immediately by the same process. For example, if + and – alleles that affect an additive polygenic trait have frequencies $p = 0.9$, $q = 0.1$ at each of three independently segregating loci, a triply homozygous –/– genotype arises with frequency $(q^2)^3 = 10^{-6}$; but because these few individuals mate with other genotypes, their progeny will be heterozygous, and therefore less phenotypically extreme. In the heterozygous progeny, moreover (such as ++–/– – –), recombination spawns a variety of gametes bearing a mixture of + and – alleles, which unite to yield offspring with intermediate traits. This is true also if the loci are linked, although then the proportions of different gamete types depend on the recombination rate (r) between them.

In Chapter 9, we saw that if two or more loci are initially in a state of *linkage disequilibrium* (in the extreme, only the gamete types *AB* and *ab*, or ++ and – –, are present), then recombination generates the other gene combinations (*Ab* and *aB*, or +– and –+) until the association between alleles has broken down completely (i.e., *linkage equilibrium* is achieved). Thus, for example, the + and – alleles that affect a polygenic trait will be carried by gametes as various +, – mixtures. Consequently, the phenotype of most individuals will be near the mean, and relatively few will have either extreme phenotype. That is, *recombination reduces phenotypic variance*, compared with the variance that might be achieved in its absence. By the same token, it might happen that in a population that is polymorphic for two loci, both double homozygotes (*AABB* and *aabb*) are better adapted than any other genotype; but if mating is random, most of the progeny in each generation will be less well adapted. (For example, if $p = q = 0.5$ at each of two independently segregating loci, the frequency of each of the double homozygotes will be only 0.0625.)

The conflict between natural selection, which tends to increase the frequency of favorable gene combinations,

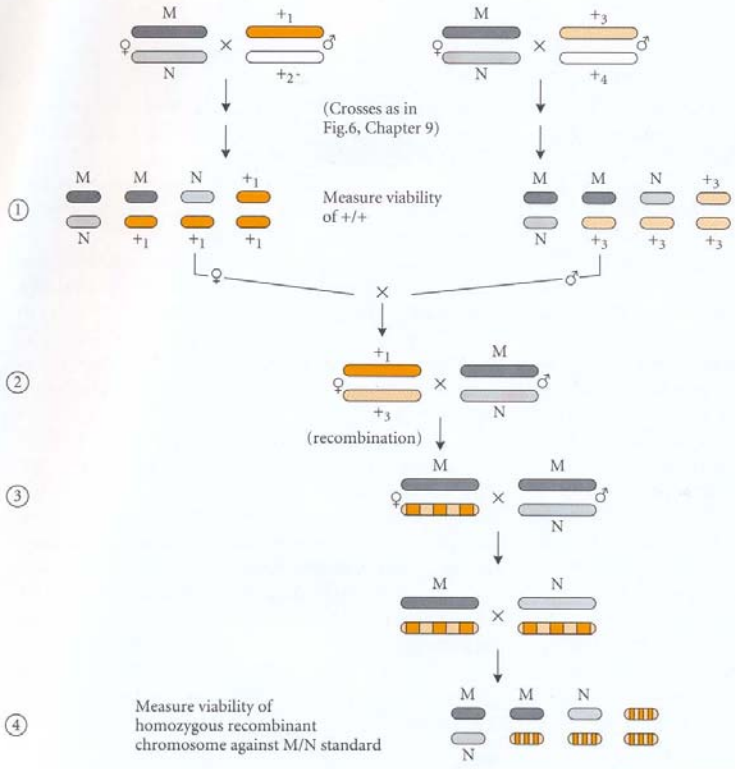


FIGURE 10.14 The method used by Spassky et al. to measure the effect of recombination on variation in egg-to-adult survival (viability) in *Drosophila*. Chromosomes M and N have dominant markers and inversions that prevent crossing over. Two wild chromosomes, say, $+_1$ and $+_3$, are made homozygous (as in Figure 6 in Chapter 9), and their viability is measured by the proportions of wild-type and M/N phenotypes in the progeny (step 1). Heterozygous $+_1/+_3$ females are produced (2). Crossing over between $+_1$ and $+_3$ occurs. Each female is crossed to an M/N male (only one such case is shown), yielding a variety of recombinant chromosomes (shown with dark and light bars) among the progeny. Further crosses (3) bring the recombinant chromosomes into the homozygous state, and the viability of each homozygote is measured (4). Similar crosses are made to obtain recombinants among other wild chromosomes.

and recombination, which tends to destroy them, has many consequences for adaptation, speciation, and the evolution of factors that modify recombination rates, including sexual and asexual reproduction (see Chapters 16 and 21).

Alterations of the Karyotype

An organism's **karyotype** is the description of its complement of chromosomes: their number, size, shape, and internal arrangement. The karyotype may be altered in many ways. In considering these alterations, it is important to bear in mind that the loss of a whole chromosome, or a major part of a chromosome, usually reduces the viability of a gamete or an organism because of the loss of genes. Also, a gamete or organism often is inviable or fails to develop properly if it has an ANEUPLOID, or "unbalanced," chromosome complement—for example, if an otherwise diploid organism has three copies of one of its chromosomes. (For instance, humans with three copies of chromosome 21, a condition known as Down syndrome, have brain and other defects.)

Alterations of the karyotype fall into two major categories: changes in the number of whole sets of chromosomes (*polyploidy*) and *rearrangements* of one or more chromosomes.

Polyploidy

A diploid organism has two entire sets of homologous chromosomes ($2N$); a **polyploid** organism has more than two. (In discussing chromosomes, N refers to the number of different chromosomes in the gametic, or haploid, set, and the numeral refers to the number of representatives of each autosome.) In a diploid species, diploid, or UNREDUCED, gametes are occasionally produced when the reduction division fails to occur in meiosis. (This event can be experimentally induced by inhibitors of cell division such as colchicine.) The union of an unreduced gamete (with $2N$ chromosomes) and a reduced gamete (with N chromosomes) yields a TRIPLOID ($3N$) zygote. Triploids often develop normally, but they are largely sterile because most of their gametes have aneuploid chromosome complements (i.e., at segregation, each daughter cell may receive one copy of certain chromosomes and two of certain others). However, if two unreduced gametes unite, a tetraploid (4N) offspring is formed. Other such unions can form hexaploid (6N), octoploid (8N), or even more highly polyploid genotypes.

In many cases, each set of four homologous chromosomes of a tetraploid is aligned during meiosis into a quartet (quadrivalent), and then may segregate in a balanced (two by two) or unbalanced (one by three) fashion (Figure 10.15). Many ANEUPLOID GAMETES result, so fertility is greatly reduced. In other cases, the four chromosomes align not as a quartet, but as two pairs that segregate normally, resulting in balanced, viable gametes, so that fertility is normal, or nearly so. This would seem to require that the chromosomes be differentiated so that each can recognize and pair with a single homologue rather than with three others.

Many species of plants, and a few species of trout, tree frogs, and other animals, are polyploid (Chapter 16). Polyploidy is common among plants. Many polyploid species appear to have arisen recently from diploid ancestors, and some major groups of plants are believed, because of their high chromosome numbers, to stem from ancient polyploid ancestors. Estimates of the proportion of polyploid angiosperms range from 30 percent (Stebbins 1950) to 50–70 percent (Stace 1989). A substantial minority of recently arisen polyploids has arisen by the union of unreduced gametes of the same species; these are **autopolyploids.** But the majority are **allopolyploids**, which have arisen by hybridization between closely related species.* They present a mixture of, or an intermediate condition between, the characteristics of their parents. In allopolyploids, the parental species' chromosomes apparently are different enough for the chromosomes of each parent to recognize and pair with each other, so that meiosis in an allotetraploid involves normal segregation of pairs rather than quartets of chromosomes (Figure 10.15C). In such a case, the karyotype has $2N$ chromosomes, but N may be based on a smaller ancestral haploid genome of x chromosomes. For example (Anderson 1936), a common blue flag, or iris (*Iris versicolor*), has $2N = 108$ ($N = 54$) chromosomes. It appears to have been derived from hybridization between *Iris virginica*, with $2N = 72$ chromosomes ($N = 36$), and *Iris setosa*, with $2N = 36$ ($N = 18$). The base, or ancestral, chromosome number appears to be $x = 18$, relative to which *I. virginica* has four chromosome sets and *I. versicolor* has six. This is one example of a POLYPLOID COMPLEX, some of which are much more complex than the *Iris* case.

*There is a full spectrum of intermediate cases between auto- and allopolyploidy, and the meiotic behavior of the chromosomes can be very complicated, with some chromosomes forming pairs and others quartets in the same genome. Stebbins (1950) presents an extended discussion of polyploidy.

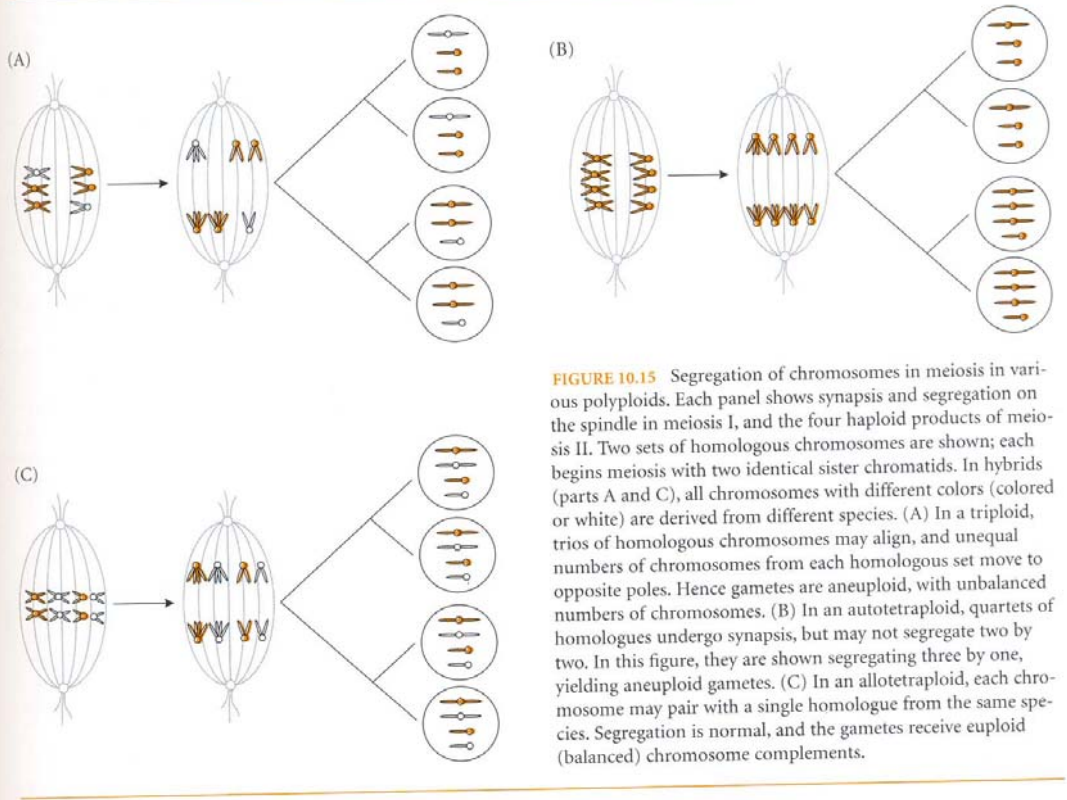


FIGURE 10.15 Segregation of chromosomes in meiosis in various polyploids. Each panel shows synapsis and segregation on the spindle in meiosis I, and the four haploid products of meiosis II. Two sets of homologous chromosomes are shown; each begins meiosis with two identical sister chromatids. In hybrids (parts A and C), all chromosomes with different colors (colored or white) are derived from different species. (A) In a triploid, trios of homologous chromosomes may align, and unequal numbers of chromosomes from each homologous set move to opposite poles. Hence gametes are aneuploid, with unbalanced numbers of chromosomes. (B) In an autotetraploid, quartets of homologues undergo synapsis, but may not segregate two by two. In this figure, they are shown segregating three by one, yielding aneuploid gametes. (C) In an allotetraploid, each chromosome may pair with a single homologue from the same species. Segregation is normal, and the gametes receive euploid (balanced) chromosome complements.

Polyploidy often has many direct effects. Polyploid cells are generally larger and divide more slowly than diploid cells. (This is a specific instance of a very widespread relationship between DNA content and the size and division rate of cells, discussed further in Chapter 22.) Perhaps for this reason, polyploid plants often have thicker leaves, and certain structures (and sometimes the whole plant) are larger. Levels of enzymes, hormones, and other biochemical constituents are often higher, and various physiological functions may differ (Stebbins 1950; Levin 1983b). Some authors have suggested that polyploids may be particularly capable of colonizing stressful environments, possibly because of their physiological vigor.

Many genes in a tetraploid are represented by four, rather than two, copies. If, as in an allotetraploid, they are inherited as two distinct loci, they may undergo independent evolutionary changes and diverge from each other in nucleotide sequence. We will return to this topic in Chapter 22.

Chromosome Rearrangements

Changes in the structure of chromosomes constitute another class of karyotypic alterations. These changes are caused by breaks in chromosomes, followed by rejoining of the pieces in new configurations. Some such changes can affect the pattern of segregation in meiosis, and therefore affect the proportion of viable gametes. Most chromosome rearrangements seem not to have direct effects on morphological or other phenotypic features, but some such effects do occur, because an alteration of gene order may bring certain genes under the influence of the control regions (e.g., the promoters) of other genes, and so alter their expression. Such POSITION EFFECTS have often been described in the laboratory, but it is not certain that they have contributed to evolutionary change.

Individual organisms may be homozygous or heterozygous for a rearranged chromosome, and are referred to as HOMOKARYOTYPES or HETEROKARYOTYPES respectively. It is also useful to distinguish ACROCENTRIC chromosomes, in which the centromere is near one end, from METACENTRIC chromosomes, in which the centromere is somewhere in the middle and separates the chromosome into two arms.

Inversions Consider a segment of a chromosome in which ABCDEFGHIJ denotes a sequence of markers such as genes. If a loop is formed, and breakage and reunion occur at the point of overlap, a new sequence, such as ABI-HGFEDCJ, may be formed. (The inverted sequence is underlined.) Such an **inversion** is PERICENTRIC if it includes the centromere, and PARACENTRIC if it does not. Inversions rearrange gene order, and may also transform metacentric into acrocentric chromosomes, or vice versa.

Because inversion is a rare event, it alters one of the two homologous chromosomes in a cell; if this is a germ line cell, heterokaryotypic daughter cells undergo meiosis. During meiotic synapsis, alignment of genes on the normal and inverted chromosomes requires the formation of a loop, which can sometimes be observed under the microscope (Figure 10.16). Now suppose that in a paracentric inversion, crossing over occurs between loci such as F and G (Figure 10.16). Two of the four strands are affected. One strand lacks certain gene regions (A, B), and also lacks a centromere; it will not migrate to either pole, and is lost. The other affected strand not only lacks some genetic material: it also has two centromeres, so the chromosome breaks when these are pulled to opposite poles. The resulting cells lack certain gene regions, and will not form viable gametes. Consequently, in inversion heterokaryotypes, (1) *fertility may be reduced* because many gametes are inviable; and (2) *recombination is effectively suppressed* because gametes carrying the recombinant chromosomes, which lack some genetic material, are inviable. (If you examine Figure 10.16 carefully, you may deduce that if crossing over occurs at *two* sites within the inversion loop, all the meiotic products are balanced and viable. However, recombination may be fully suppressed if one chromosome carries a series of overlapping inversions.) Crossing over and gamete formation proceed normally in homokaryotypes, whatever gene arrangement they may carry.

Related species often differ in their karyotypes as a consequence of inversions. However, it is rather unusual to find two or more inversions in a polymorphic state within a population because of the reduced fertility of heterokaryotypes: one of the arrangements is rapidly eliminated from the population by natural selection (see Chapter 13). The most prominent exceptions are in *Drosophila* and some other flies (Diptera), in which the unbalanced recombinant chromosomes enter the polar bodies during meiosis, so that female fecundity is not reduced.* It is particularly easy to study inversions in *Drosophila* and some other flies because the larval salivary glands contain giant (polytene) chromosomes that remain in a state of permanent synapsis (so that inversion loops are easily seen), and because these chromosomes display bands, each of which apparently corresponds to a single gene. The banding patterns are as distinct as the computer-scanned bar codes on supermarket products, so an experienced investigator can identify different sequences. **Inversion polymorphisms** are common in *Drosophila*—more than 20 different arrangements of the third chromosome have been described for *Drosophila pseudoobscura*, for example—and they have been extensively studied from both population genetic and phylogenetic points of view (Chapters 8, 13).

Translocations By breakage and reunion, two nonhomologous chromosomes may exchange segments, resulting

* This is true for heterozygous paracentric inversions. Crossing over within a heterozygous pericentric inversion results in genetically unbalanced products of meiosis, in Diptera as in other organisms, and reduces fertility. Male fertility is not reduced in a heterokaryotype, because crossing over does not normally occur in male Diptera.

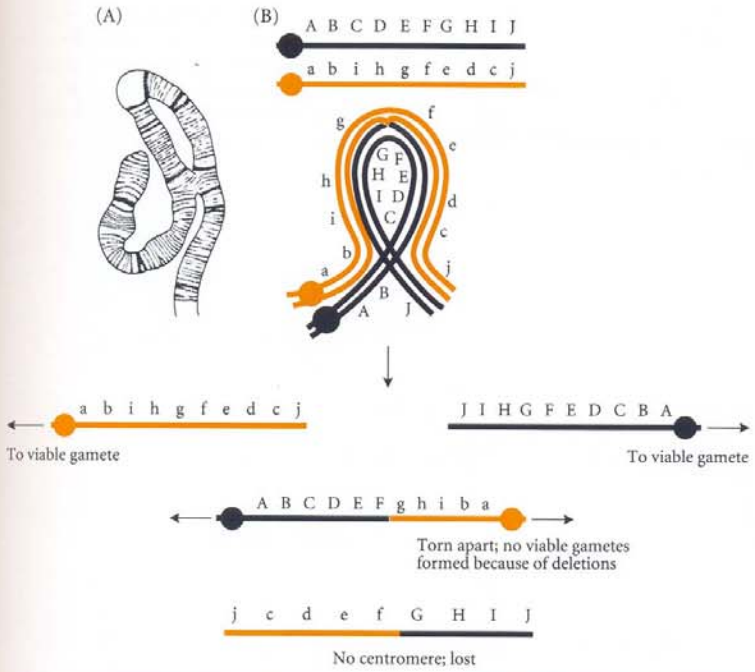


FIGURE 10.16 Chromosome inversions. (A) Synapsed chromosomes in a salivary gland cell of a larval *Drosophila pseudoobscura* heterozygous for *Standard* and *Arrowhead* arrangements. The two homologous chromosomes are so tightly synapsed that they look like a single chromosome. The "bridge" forms the loop shown diagrammatically in part B. Similar synapsis occurs in germ line cells undergoing meiosis. (B) Two homologous chromosomes differing by an inversion of the region C to I, and their configuration in synapsis. Crossing over between two chromatids (between F and G) yields products that lack a centromere or substantial blocks of genes. Because these do not form viable gametes, crossing over appears to be suppressed. Only cells that receive the two chromatids that do not cross over become viable gametes. (After Strickberger 1968.)

in a **reciprocal translocation**. In a translocation heterokaryotype, synapsis of homologous regions results in contorted configurations such as the cross-shaped arrangement in Figure 10.17. In the simple case illustrated, the four chromosomes may segregate pairwise in three different ways, two of which result in gametes with unbalanced gene complements (some genes are lacking). These gametes, or the resulting progeny, are inviable, so the fertility of translocation heterokaryotypes is often reduced by 50 percent or more. Consequently, polymorphism for translocations is rare in natural populations. Nevertheless, related species sometimes differ by translocations, which have the effect of moving groups of genes from one chromosome to another. For example, the Y chromosome of the male *Drosophila miranda* includes a segment that is homologous to part of one of the autosomes of its close relative *D. pseudoobscura* (Dobzhansky 1970).

Fissions and fusions In the simplest form of chromosome **fusion**, two nonhomologous acrocentric chromosomes, in which the centromeres are nearly terminal, may undergo reciprocal translocation near the centromeres so that they are joined into a metacentric chromosome. Conversely, a metacentric may undergo **fission**, becoming two acrocentrics if it undergoes a reciprocal translocation with a minute "donor" chromosome (Figure 10.18). A simple fission heterokaryotype, then, has a metacentric, which we may refer to as AB, with arms that are homologous to two acrocentrics A and B. AB, A, and B together synapse as a "trivalent." Viable gametes and zygotes are formed if AB segregates from A and B, but if homologous chromosomes do not segregate to opposite poles (i.e., if **nondisjunction** occurs) aneuploid gametes result, reducing fertility or zygote viability. The frequency of aneuploid gametes is often about 5–25 percent, and sometimes up to 50 percent (Lande 1979a; see also Chapter 16). Chromosome fusions and fissions often distinguish related species or geographic populations of the same species (Figure 10.19), and are sometimes found as polymorphisms within populations.

Changes in chromosome number Summarizing what we have covered so far, the number of chromosomes may be altered by polyploidy (especially in plants) or by translocations and fusions or fissions of chromosomes. These are the mutational foundations for the evolution of chromosome number. For example, the haploid chromosome number varies among mammals from 3 to 42 (Lande 1979a), and among insects from 1 in a species of ant to about 220 in some butterflies (the highest number known in animals). Related species sometimes differ strikingly in karyotype: in one of the most extreme examples, two very similar species of small deer, *Muntiacus reevesii* and *M. muntjac*, have haploid chromosome numbers of 46 and 3

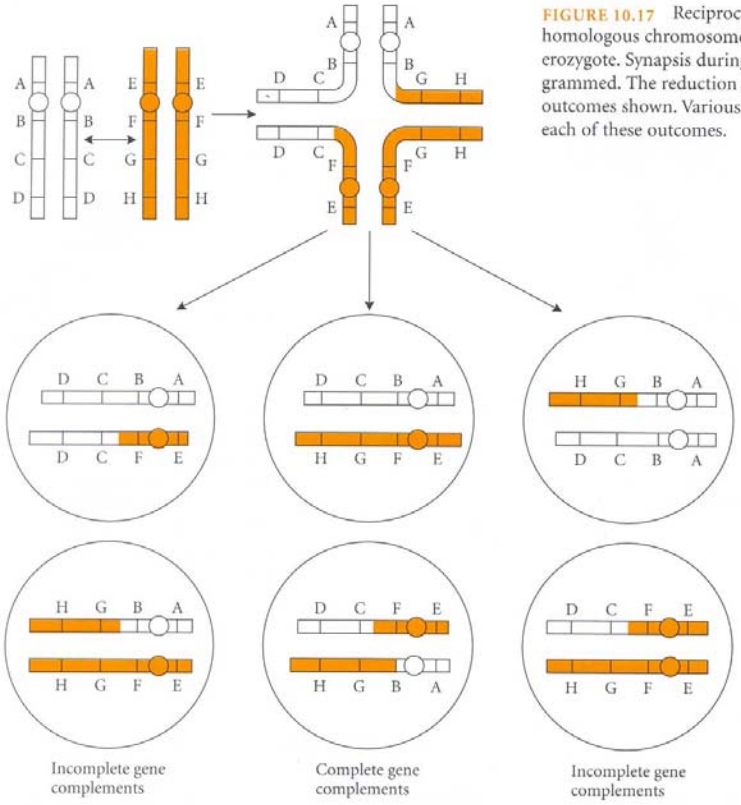


FIGURE 10.17 Reciprocal translocation between two nonhomologous chromosomes gives rise to a translocation heterozygote. Synapsis during meiosis in the heterozygote is diagrammed. The reduction division yields one of the three outcomes shown. Various cells undergoing meiosis will yield each of these outcomes.

or 4 (in different populations) respectively (Figure 10.20). As for all characteristics, the evolution of the karyotype requires not only mutation, but other processes as well (see Chapters 11, 13).

The spontaneous rate of origin of a given class of chromosome rearrangement (e.g., reciprocal translocation) is quite high: about 10^{-4} to 10^{-3} per gamete (Lande 1979a). However, a rearrangement involving breakage at any particular site(s) rarely arises, and is usually considered to be unique.

Duplications and deletions The rearrangements described above, unlike polyploidy, do not greatly alter the amount of genetic material. Several other processes, however, can change the amount of DNA. At this point we will treat only one of these, **unequal crossing over** (unequal exchange). If crossing over occurs between two chromosomes that are not perfectly aligned, a region will be tandemly duplicated on one recombination product and deleted from the other (Figure 10.21). The length of the affected region (e.g., the number of loci) depends on the amount of displace-

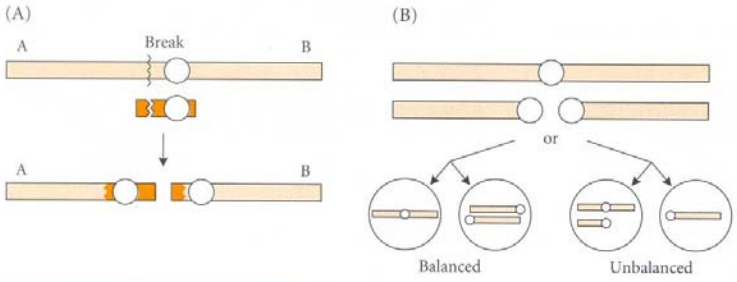


FIGURE 10.18 (A) A simple fission of a metacentric chromosome, with arms A and B, into two acrocentric chromosomes, by translocation with a minute "donor" chromosome. (B) Segregation in meiosis of a fission heterozygote can yield euploid (balanced) or aneuploid (unbalanced) complements of genetic material.

ment of the two misaligned chromosomes. Deletion of a substantial amount of genetic material results in inviable gametes or zygotes, but duplications are sometimes advantageous, perhaps because more gene product is synthesized. Chromosomes with a tandem duplication (ABBC) are even more likely subsequently to engage in unequal crossing over because the duplicate regions can pair out of register,

$$\begin{bmatrix} \text{ABBC}\ldots \\ \ldots\text{ABBC} \end{bmatrix}$$

generating further duplications (ABBBC). Because misalignment occurs rarely, each of the repeated genes is generally inherited as a separate locus, undergoing independent mutations, so that their DNA sequences may diverge over time.* As a result, a great many genes are members of **gene families**: sets of genes that have similar DNA sequences, but differ to some degree in sequence and often in function. For example (Figure 22 in Chapter 3), humans have eight loci that code for the several α-like and β-like hemoglobin polypeptides (as well as several related degenerate sequences that are not translated). These loci differ in their time of expression (embryonic, fetal, or adult). Unequal crossing over is probably one of the processes that have generated the extremely high number of copies of nonfunctional sequences (in satellite DNA) that constitute much of the DNA in most eukaryotes. This

*There also exist processes that may homogenize repeated DNA sequences, giving rise to the phenomenon of "concerted evolution" (Chapter 22).

FIGURE 10.19 Geographic variation in the karyotype of the pocket mouse *Perognathus goldmani* in northwestern Mexico. (A) The distribution of populations with different karyotypes, designated by Greek letters. (B) The proposed sequence of evolutionary changes in karyotype, involving four fusions (e.g., of chromosomes 3 and 4 in race ε) and two inversions (e.g., the X chromosome in race α). Karyotypes β and ε are most similar to the hypothetical ancestral karyotype, which was inferred by phylogenetic comparison with other species of *Perognathus*. *2N* is the diploid number of chromosomes, *FN* the number of chromosome arms. (After Patton 1969.)

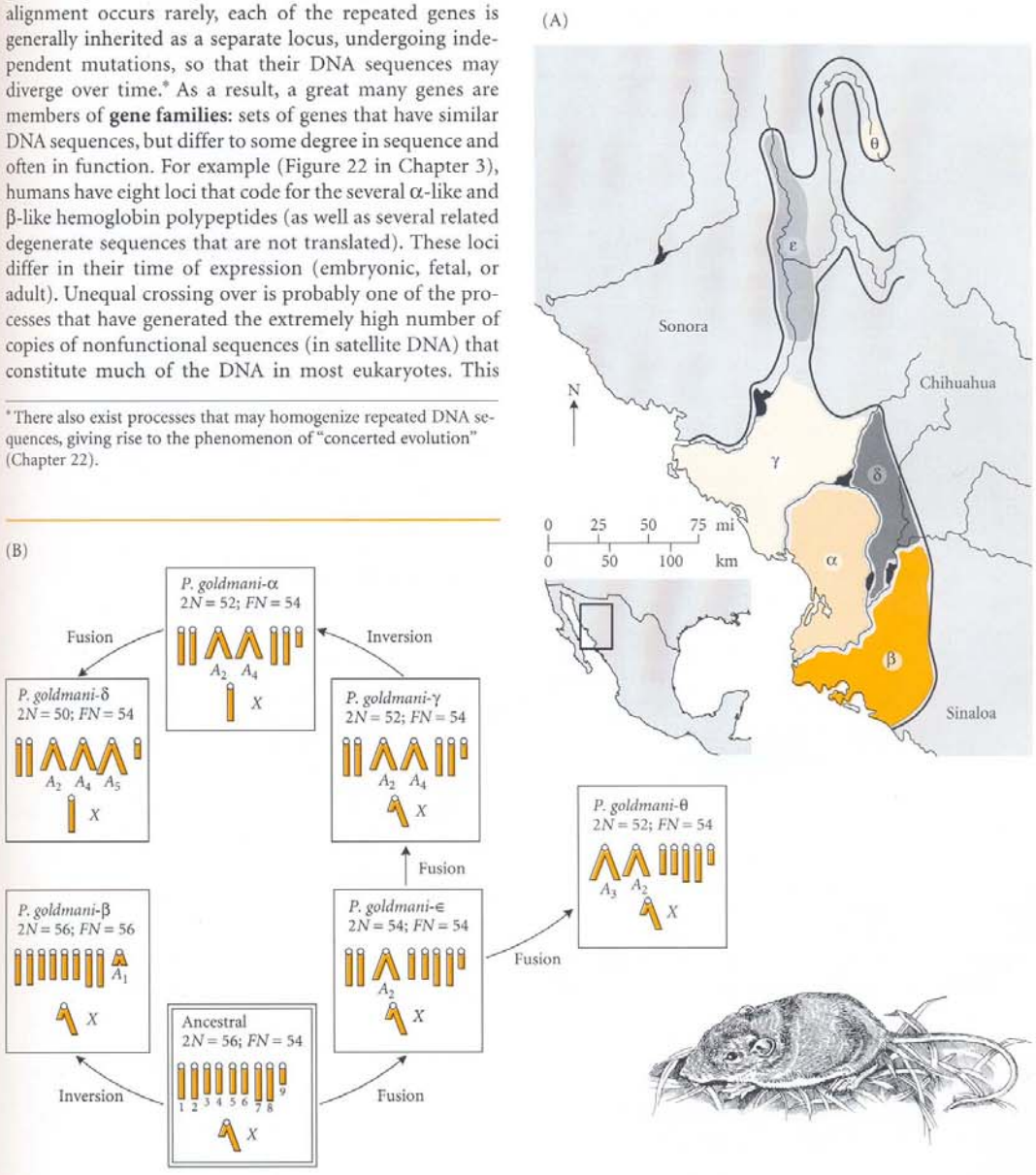
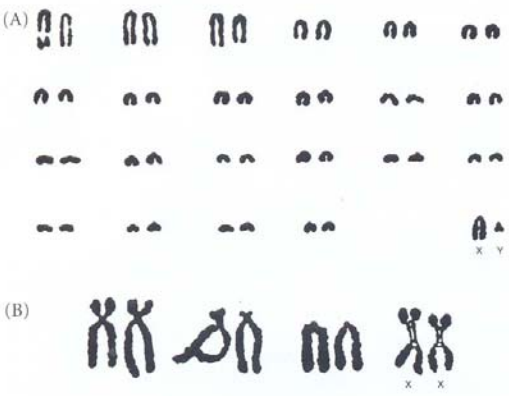


FIGURE 10.20 The diploid chromosome complements, taken from photographs, of two closely related species of muntjac deer. This is one of the most extreme differences in karyotype known among closely related species. Despite the difference in chromosomes, the species are phenotypically very similar. (A) *Muntiacus reevesi* ($2N = 46$). (B) *Muntiacus muntiacus* ($2N = 8$). (From White 1978.)



process has also been extremely important in the evolution of greater numbers of genes and of total DNA (see Chapter 22).

Unequal exchange may also occur on a very fine scale, namely, within a single gene. This gives rise to short repeats of base pair sequences, and to variation in the number of repeats. For example, in the *per* (*period*) gene of *Drosophila melanogaster*, which affects various behavioral rhythms, a six-base-pair motif, encoding threonine and glycine, is tandemly repeated 14 to 23 times, with several length variants occurring within natural populations. Related species also differ in the number of repeats (Costa et al. 1991).

External Sources of Genetic Variation

So far, we have discussed the origin of genetic variation by mutation and recombination within a single population of a species. In almost all species, an important source of genetic variation is **gene flow** from other populations of the species: the influx of alleles that have arisen or have reached high frequency in other populations, and are carried by migrants that join the local mating population. Genes may be carried by organisms (including seeds and spores) or by their gametes (e.g., pollen). We will return to this important topic in Chapter 11 and subsequent chapters. Suffice it to say that in the short term, the increase in a population's genetic variation due to gene flow is often far greater, per generation, than that provided by new mutations.

In some cases, genetic variation enters the population from *other species*, either by **hybridization** or by **horizontal gene transfer**.

Hybridization

Closely related species often can hybridize and produce at least some partly or fully fertile offspring, which, by backcrossing, introduce genes from one species into the other. Hybridization (more fully discussed in Chapter 15) may occur at a low level wherever the species meet, or may be limited to certain localities, such as disturbed habitats. Some hybrids can differ considerably from both parent forms (Figure 10.22). In some cases it may be arbitrary whether the hybridizing forms are called species or some other term, such as "semispecies." Hybridization has been documented in many groups of animals, and appears to be especially frequent in plants (Harrison 1990; Stebbins 1950; Grant 1971; Mayr 1963; but see also Mayr 1992).

The botanist Edgar Anderson (1949) applied the term **introgressive hybridization** to the incorporation into a gene pool of genes from other species that nevertheless remain fairly distinct. An example from animals is presented by two morphologically different forms of house mice, *Mus musculus* and *M. domesticus*, that are distributed throughout eastern and western Europe respectively, meeting along a narrow zone of hybridization. Robert Selander and his colleagues (1969) described patterns of enzyme polymorphism from several regions on either side of the hybrid zone in the Jutland Peninsula of Denmark (Figure 10.23). The two forms differ substantially in allele frequencies at 13 of 17 polymorphic loci; at 6 of the loci, the two species are fixed, or nearly so, for different alleles. However, some of these alleles do cross the species boundary—especially alleles typical of *M. domesticus*, such as *Es-1b*, which penetrates into *M. musculus* populations, particularly those located nearest *M. domesticus*.

FIGURE 10.21 Ordinarily, as at left, crossing over results in reciprocal exchange. But in unequal crossing over, at right, a segment of one chromosome, marked by locus B, is transferred to the other chromosome. Thus one chromosome has a deficiency, and the other bears a tandem duplication for one or more loci. Deletions are often deleterious, but duplications are sometimes advantageous.

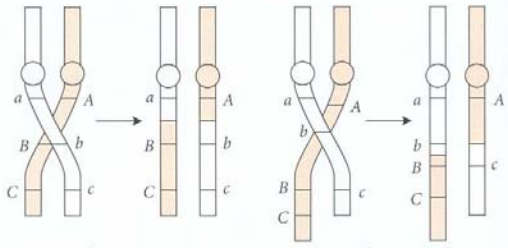
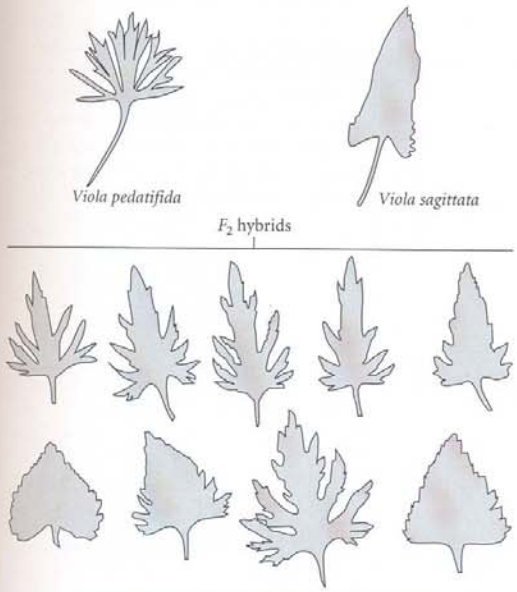


FIGURE 10.22 An example of the phenotypic variation that can arise from hybridization between species. These variations in leaf form have been found in natural hybrids between two species of violets. (After Stebbins 1950.)



Horizontal Gene Transfer

Species that have diverged so greatly that they cannot hybridize should not be able to exchange genes. Nevertheless, a few cases have emerged in which transfer of genetic material between very distantly related species is suspected. The evidence, in most cases, consists of a striking similarity of certain base pair sequences that is incongruent with phylogeny. For example, sequences called virogenes are normal constituents of vertebrate genomes, yet are similar to, and surely originally derived from, retroviruses. One virogene is found in Old World monkeys. It is apparently ancient, because the phylogeny of monkey species based on the virogene DNA sequence matches that based on other information. A similar virogene sequence is found in six closely related species of cats, but not in other Felidae or other Carnivora. Its sequence closely matches that of only certain monkeys, namely, baboons (Figure 10.24). This sequence appears, therefore, to have been transferred to the ancestor of these cat species (probably about 5–10 million years ago) from the ancestor of the baboon species. (The reader is invited to suggest why the transfer was not in the opposite direction.)

How might such transfers occur? The most probable route is cross-infection by agents such as retroviruses, which can incorporate host DNA into their genomes. Plasmids play a similar role in transferring genes among unrelated species of bacteria (Heinemann 1991). Because some similarity of DNA sequences may result from convergent evolution, only very similar sequences that are highly in-

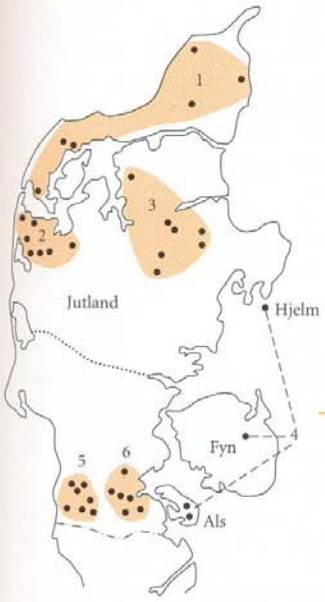


FIGURE 10.23 Introgressive hybridization between species of house mice in the Jutland Peninsula of Denmark. *Mus musculus* (regions 1–4) and *Mus domesticus* (regions 5 and 6) form a narrow zone of hybridization in the region indicated by the dotted line. The frequencies of allozyme alleles at several loci in each of the six regions show introgression of *domesticus* alleles as far north as region 1. Locus *Idh*-1 reveals introgression of *musculus* alleles into *domesticus* populations as well. (After Selander et al. 1969.)

LOCUS:		*Es*-1		*Hpd*-1			*Idh*-1	
ALLELE:		*a*	*b*	*a*	*b*	*c*	*a*	*b*
	Site							
M. musculus	1	1.00	0.00	0.00	0.00	1.00	0.17	0.83
	2	1.00	0.00	0.00	0.00	1.00	0.23	0.77
	3	1.00	0.00	0.00	0.00	1.00	0.00	1.00
	4	0.95	0.05	0.03	0.03	0.95	0.08	0.92
M. domesticus	5	0.00	1.00	0.50	0.50	0.00	0.93	0.07
	6	0.00	1.00	0.53	0.47	0.00	0.93	0.07

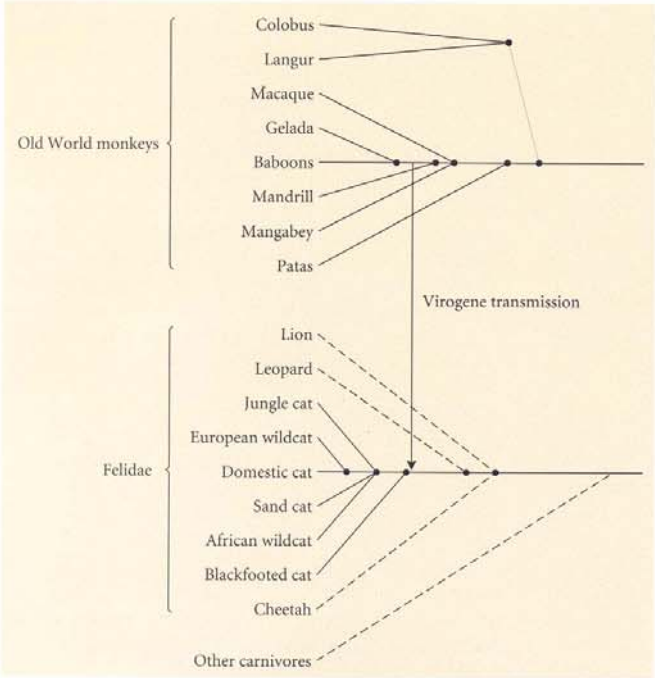


FIGURE 10.24 Horizontal gene transfer, as inferred from the phylogenies of Old World monkeys and cats (Felidae). The solid branches of the phylogenetic trees lead to species whose genomes include similar sequences of a virogene. Broken lines lead to species that lack the virogene. The sequence of the virogene in cats is most similar to that in baboons, suggesting that the gene was transferred from the ancestor of these primates to the ancestor of the small cats, after the lion, cheetah, and leopard lineages had diverged. (After Li and Graur 1991.)

congruent with the phylogeny of organisms can be considered strong evidence for horizontal gene transfer. So far, the number of well-documented cases is small (Li and Graur 1991), and it is unlikely that this phenomenon will prove to be common.

Finally, we should touch on a related subject, the evolution of extremely *intimate associations of unrelated organisms* (see Chapter 18). Lichens, formed by associations of algae with fungi, are the classic example, but many other mutualistic or parasitic associations have formed in which two species effectively become a single organism: intracellular bacteria that provide vitamins and amino acids to aphids, the algal endosymbionts essential for the growth of corals, the many viruses and plasmids that become integrated into the genomes of bacteria and other hosts. Molecular evidence suggests that transfer of genes between hosts and symbionts may occur in such instances (Heinemann 1991), but even without DNA exchange, an endosymbiont that is tightly integrated into its host's life can provide, in effect, a major addition to the host's genome, and an expansion of the host's functional capacities. Antibiotic resistance in bacteria, for example, is usually conferred by genes carried by plasmids. The ne plus ultra of such a functional integration of different genomes is the eukaryotic cell (see Chapter 7), in which mitochondria and chloroplasts, derived from two lineages of bacteria, have joined forces with a proto-eukaryote represented by the cell nucleus—a union that has proved successful in the extreme.

Summary

1. Mutations of chromosomes or genes are alterations that are subsequently replicated. They ordinarily do not constitute new species, but rather variant chromosomes or genes (alleles, haplotypes) within a species.
2. At the molecular level, mutations of genes include base pair substitutions and several other types of changes. Intragenic recombination also gives rise to new DNA sequences (haplotypes).
3. The rate at which any particular mutation arises is quite low: on average about 10^{-6} to 10^{-5} per gamete for mutations detected by their phenotypic effects, and about 10^{-9} per base pair. The mutation rate, by itself, is too low to cause substantial changes of allele frequencies. However, the total input of genetic variation by mutation, for the genome as a whole or for individual polygenic characters, is appreciable.
4. The magnitude of change in morphological or physiological features caused by a mutation can range from none to drastic. Mutations alter pre-existing biochemical or developmental pathways, so not all conceivable mutational changes are possible. Some adaptive changes may not be possible without just the right mu-

tation of just the right gene. For these reasons, the rate and direction of evolution may in some instances be affected by the availability of mutations.

5. In part because most mutations have pleiotropic effects, the *average* effect of mutations on fitness is deleterious, but there is abundant evidence, including many experiments, that some mutations are advantageous. Mutations with large effects are often deleterious, but some evolutionary biologists believe that such mutations have sometimes been important in evolution.

6. Mutations appear to be random, in the sense that their probability of occurrence is not directed by the environment in favorable directions, and in the sense that specific mutations cannot be predicted. The likelihood that a mutation will occur does not depend on whether or not it would be advantageous.

7. Recombination of alleles can potentially give rise to astronomical numbers of gene combinations, and in sexual organisms generates far more genetic variation per generation than mutation alone. However, recombination also breaks apart favorable gene combinations, and constrains the amount of variation displayed by polygenic characters.

8. Mutations of the karyotype (chromosome complement) include polyploidy (which can give rise to new species) and rearrangements that alter the chromosome number or the arrangement of genes. Many such rearrangements reduce fertility in the heterozygous condition.

9. Unequal crossing over causes deletions and duplications of genes, and is one of the processes responsible for gene families and increases in genome size and gene number.

10. The genetic variation in most populations is augmented by gene flow from conspecific populations. In some cases, genes acquired by hybridization with closely related species add genetic variety. A few examples are known of horizontal gene transfer between very distantly related organisms, but this is probably extremely rare.

Major References

Lewin, B. 1994. *Genes V.* Oxford University Press, New York. Includes molecular aspects of mutation.

Li, W.-H. 1997. *Molecular Evolution.* Sinauer Associates, Sunderland, MA. Treats evolutionary aspects of mutation at the molecular level.

White, M. J. D. 1973. *Animal cytology and evolution.* Third edition. Cambridge University Press, Cambridge. The authoritative work on chromosome evolution in animals.

Stebbins, G. L. 1971. *Chromosomal variation in higher plants.* E. Arnold, London. See also the classic by Stebbins, *Variation and evolution in plants* (Columbia University Press, New York, 1950), on chromosome evolution in plants.

Problems and Discussion Topics

1. Consider two possible studies. (*a*) In one, you capture 3000 wild male *Drosophila melanogaster*, mate each with laboratory females heterozygous for the autosomal recessive allele *vestigial*, *vg* (which causes miniature wings when homozygous), and examine each male's offspring. You find that half the offspring of each of three males have miniature wings and have genotype *vg*/*vg*. (*b*) In another study, you determine the nucleotide sequence of 1000 base pairs for 20 copies of the cytochrome *b* gene, taken from 20 wild mallard ducks. You find that at each of 20 nucleotide sites, one or another gene copy has a different base pair from all others. From these data, can you estimate the rate of mutation from the wild-type to the *vg* allele (case *a*) or from one base pair to another? Why or why not?

2. From a laboratory stock of *Drosophila* that you believe to be homozygous wild type (++) at the *vestigial* locus, you obtain 100,000 offspring, mate each of them with heterozygous +/*vg* flies, and examine a total of 1 million progeny. Two of these are *vg*/*vg*. Estimate the rate of mutation from + to *vg* per gamete. What assumptions must you make?

3. The following DNA sequence represents the beginning of the coding region of the alcohol dehydrogenase (*Adh*) gene of *Drosophila simulans* (Bodmer and Ashburner 1984), arranged into codons:

   CCC ACG ACA GAA CAG TAT TTA AGG AGC TGC GAA GGT

   (*a*) Find the corresponding mRNA sequence, and use Figure 12 in Chapter 3 to find the amino acid sequence. (*b*) Again using the figure, determine for each site how many possible mutations (nucleotide changes) would cause an amino acid change, and how many would not. For the entire sequence, what proportion of possible mutations are synonymous versus nonsynonymous? What proportion of the mutations at first, second, and third base positions within codons are synonymous? (*c*) What would be the effect on amino acid sequence of inserting a single base, G, between sites 10 and 11 in the DNA sequence? (*d*) What would be the effect of deleting nucleotide 16? (*e*) For the first 15 (or more) sites, classify each possible mutation as a transition or transversion, and determine whether or not the mutation would change the amino acid. Does the proportion of synonymous mutations differ between transitions and transversions?

4. A genus of Antarctic fishes, *Channichthys*, lacks hemoglobin. In its relatives, such as *Trematomus*, hemoglobin serves its usual functions. Assuming that the gene encoding hemoglobin in *Channichthys* has no function, and is not transcribed, how might you expect the nucleotide sequence of this gene to differ between these genera? As a challenge problem, is it possible to estimate how long the gene has been nonfunctional in the *Channichthys* lineage? How might you attempt such an estimate?

5. Many mutations with large phenotypic effects that have been found in laboratory stocks of *Drosophila* are caused by insertions of transposable elements. However, it is not yet clear that transposable elements have caused many of the mutations that have contributed to adaptive evolution in nature. Discuss research strategies that might address this gap in our knowledge, and the possible pitfalls that might be encountered.

6. Researchers have used artificial selection (see Chapter 9) to alter many traits in *Drosophila melanogaster*, such as phototactic behavior and wing length. No one has selected *Drosophila* to be as large as bumblebees (although I'm not sure if anyone has tried). Do you suppose this could be done? How would you attempt it? If your attempts were unsuccessful, what hypotheses could explain your lack of success? What role might mutation play in your experiment?

7. Ultraviolet light (UV) can induce mutations in organisms such as *Drosophila*. Because it damages DNA, and therefore essential physiological functions, it can also reduce survival. Suppose you expose a large number of *Drosophila* to UV, screen their offspring for new mutations, and discover some mutations that increase the amount of black pigment, which protects internal organs from UV. The progeny of an equal number of control flies, not exposed to UV, show fewer or no mutations that increase pigmentation. Can you conclude that the process of mutation responds to organisms' need for adaptation to the environment?