

Evolutionary Psychologists claim that their account of human nature follows from applying the principles of evolutionary biology to the study of the human mind. Consequently, to truly understand Evolutionary Psychology, and to be in a position to critically evaluate it, it is essential to have a basic understanding of evolutionary theory. This chapter provides the necessary introduction to the fundamentals of evolutionary biology. For the initiate, this may be a slow go. But theoretical principles and concepts explained in this chapter will repeatedly turn up later in our examination of Evolutionary Psychology, so understanding them is a necessary first step toward understanding Evolutionary Psychology.

In developing their account of human nature, Evolutionary Psychologists build on (their interpretation of) the reigning orthodoxy in evolutionary biology. Aspects of this reigning orthodoxy are currently being challenged by a number of researchers in developmental biology. As a result, one could endorse one of these recent challenges and criticize Evolutionary Psychology for erecting itself on a mistaken biological foundation. Although some have taken this approach, I will not. For I think it is far too early to tell whether any of these challenges will fundamentally change the way we think about evolution. Instead, throughout this book I will take for granted the reigning orthodoxy in evolutionary biology, just as Evolutionary Psychologists do. Here, then, is a brief introduction to orthodox neo-Darwinian evolutionary biology.

### The Nature of Evolution

For Darwin, and for several generations of biologists after him, evolution was conceived of as *descent with modification*. Each component of this definition, *descent* and *modification*, requires some comment. Consider first descent.

According to this conception of evolution, evolution occurs only in *lineages*, which are populations of organisms that are related by descent. A population, in the biological sense, is a group of reproductively interacting organisms. As organisms in a population reproduce, they create a new generation, which itself reproductively interacts to spawn yet another generation of reproductively interacting organisms. This process creates a temporally extended sequence of populations, the later of which are descended from the earlier by reproduction, and such a temporal sequence of populations is a lineage. In a lineage, offspring tend to inherit their characteristics from their parents, so that offspring resemble their parents more than they resemble unrelated organisms in their lineage. “Descent,” then, indicates a lineage of organisms that are characterized by hereditary similarity between parents and their offspring.

“Modification” refers to change across generations in the distribution of characteristics, or traits, in a lineage. A trait can be any one of an organism’s observable properties, from an organ or bit of morphology to a form of behavior. As the organisms in a population reproduce to create a new generation, there may or may not be changes in the frequencies of traits from one generation to the next. If one generation of a human population is 65 percent brown-eyed, 25 percent green-eyed, and 10 percent blue-eyed, for example, and if the percentages of these eye colors are different in the next generation, then there has been “modification” of that lineage. Thus, for Darwin and several generations of biologists after him, evolution was change in the frequencies of hereditary characteristics across generations in a lineage. It is important to note that, according to this definition, evolution does not concern changes that individual organisms undergo during their lifetimes. Rather, evolution consists only in the changes *across generations* within a lineage in the frequencies of characteristics of organisms.

There were two important holes in this conception of evolution, which Darwin and his early successors did not adequately fill. First, descent with modification requires some mechanism of inheritance, which is causally responsible for the resemblance between parents and their offspring. But the process by which offspring inherit their characteristics from their parents was not successfully explained by Darwin or his early successors. Second, descent with modification clearly requires variation in populations, since the frequencies of hereditary characteristics cannot change from one generation to another unless those characteristics occur in more than one form. Further, the variation in a population occasionally includes evolutionary novelties, characteristics that didn’t appear in a parent gen-

eration but that make their appearance in some members of the offspring generation, who can then transmit that characteristic to their progeny. The source of these evolutionary novelties was also not successfully explained by Darwin's theory.

As it turned out, the development of genetics in the twentieth century illuminated both of these issues. It was discovered, first of all, that offspring inherit their parents' characteristics because parents transmit their *genes* to their offspring in the process of reproduction, and genes causally influence the *phenotypes*—the anatomical structures, physiological states, or behavioral forms—that organisms exhibit. Thus, offspring resemble their parents because the genes that causally influenced parental phenotypes are directly transmitted to offspring, in whom those same genes causally influence the development of the same phenotypes. Genes were consequently recognized as playing a dual role in evolution: They are the units of heredity, which get directly transmitted from parents to offspring in reproduction, and they guide the development of organisms in ways that influence the phenotypes they possess. It was also discovered, however, that an organism's phenotype does not affect the genes it can transmit to its offspring. As a result, no modifications to an organism's phenotype during the course of its life affect the genes its offspring possess. So, genes were seen as the locus of two causal arrows, one running from the genes of an organism to its phenotypes and the other running from the genes of an organism to the genes of its offspring. But there is no causal arrow running from the phenotypes of an organism to the genes of its offspring. No matter how much body-building you do in your life, your babies won't be any stronger than they would be if you were a couch potato. Similarly, breaking your arm will not affect the bones of your offspring.

Developments in genetics also led to the discovery that, in the process of reproduction, genes sometimes *mutate* into new forms. Consequently, the evolutionarily novel phenotypes that occasionally appear in a lineage are the result of mutated genes, which produce novel phenotypes in the individuals with those genes. Once a mutated gene appears in a population, it can be transmitted to the offspring of organisms with that gene, and the novel phenotype it produces can be transmitted along with it.

Since genes are the key to both inheritance and the appearance of evolutionary novelties, they came to be seen as central to the process of descent with modification. Indeed, since phenotypes are produced by genes, and phenotypes have no effect on the genes available to be transmitted across generations, genes came to be seen as the very locus at which evolution occurs. The discoveries of modern genetics thus gave rise to a

wholly new definition of evolution. According to this new definition, which is now standard within contemporary evolutionary biology, *evolution is change in gene or genotype frequencies (at a particular locus) across generations in a lineage*. Thus, by this genetic definition of evolution, transgenerational changes in the frequencies of phenotypes do not constitute evolution unless they reflect changes in gene or genotype frequencies.

There are a number of concepts in this last paragraph, however, that are so far undefined, and the modern definition of evolution will consequently make little sense to the initiate. In order to understand the modern genetic theory of evolution, it is necessary to take a brief excursion into elementary genetics. Since this book is about human psychology, I will focus on human genetics. But the initiate should be aware that there is far more in heaven and on (and under) earth—much of it incredibly bizarre—than can be captured by a brief introduction to human genetics. The initiate should also be aware that, while some of the following may not be especially titillating compared with the evolutionary psychology of human mating, for example, concepts to be introduced here will appear again later. (It should also be noted that the definition of “evolution” explained here is a definition of *microevolution*, evolutionary change *within* species. Macroevolution concerns the birth and extinction of species, and the mechanics of macroevolution are irrelevant to the topics discussed in this book.)

First, then, human bodies contain cells, the nuclei of which contain chromosomes, which are long strings of deoxyribonucleic acid, or DNA. DNA itself is a long string composed of the four nucleic acid bases adenine, guanine, cytosine, and thymine (known as A, G, C, and T, respectively). For heuristic purposes, a chromosome can be thought of as containing a sequence of slots, called *loci*, each of which is occupied by a *gene*, which is a short, replicable segment of DNA, or nucleic acid bases. The different forms of a gene that can occupy a locus are called *alleles*, which are alternative sequences of A, G, C, or T at a particular locus. Alleles can be thought of as “rivals” for occupying that locus.

The nuclei of the cells that make up a human body—the cells that form the liver, brain, skin, and so on—contain 23 pairs of chromosomes and are called *diploid* cells. In diploid cells, the pairs of chromosomes are aligned so that we can think of the opposing loci on paired chromosomes as a single (diploid) locus that is occupied by a pair of alleles, where the pair of alleles an organism has at a locus is called its *genotype*. If different genes occur at a locus in a population, then an organism can be either *homozygous* or *heterozygous* at that locus. For example, consider a simple case in

which there are two different alleles, designated  $A$  and  $a$ , that can occur at some locus in a population. Then a pair of these alleles can be a pair of identical alleles (the pair  $AA$  or the pair  $aa$ ) or a pair of different alleles (the pair  $Aa$ ). If an organism has the same allele in each opposing slot, if it has the  $AA$  or  $aa$  genotype, it is a *homozygote*; and if it has different alleles in the opposing slots, if it has the  $Aa$  genotype, it is a *heterozygote*.

In addition to diploid cells human bodies contain some *haploid* cells, the nuclei of which contain 23 single, unpaired chromosomes. These cells, called *gametes*, are formed by a process called *meiosis*. In meiosis a diploid cell first undergoes a process of DNA replication, which generates another copy of each chromosome contained in the nucleus. This is followed by two rounds of cell division, in which the chromosomes separate from one another and divide into four haploid cells. To make this less abstract, consider the process of meiosis with respect to a single locus containing the  $Aa$  genotype. Meiosis is a process whereby that single diploid  $Aa$  cell replicates and divides to produce two haploid cells containing  $A$  and two haploid cells containing  $a$ . Consequently, the result of meiosis is that an organism's DNA gets split in half: Half of the genes in a diploid cell take up residence in one haploid cell and the other half take up residence in a different haploid cell. Thus, while all an organism's diploid cells are genetically identical (with the exception of cells in which there has been a mutation), its gametes are routinely genetically different from one another.

The gametes produced in meiosis are important in the process of reproduction, since they form the egg cells in females and the sperm cells in males. During fertilization an egg cell and sperm cell fuse to form a new diploid cell, called a *zygote*, from which a new organism develops. Reproduction is thus a process whereby each of two parents contributes a gamete, which contains half of the parent's genes, to the formation of a diploid cell that will develop into an organism of the next generation. Half of the genes in the diploid cells that form that newly developing organism's body will thus have come from its mother's egg (which contains half of the mother's genes) and the other half from its father's sperm (which contains half of the father's genes).

Now consider how zygote genotypes are determined. For simplicity, consider again a single locus at which the three genotypes  $AA$ ,  $aa$ , and  $Aa$  occur in a population. And suppose that mating in this population is *random*—that is, there is no overall statistical tendency for like genotypes to mate with one another. If two  $AA$  organisms reproduce, each will contribute only  $A$  gametes, which will fuse to form  $AA$  zygotes; so all the offspring of two  $AA$  organisms will also be  $AA$ . Similarly, all offspring of two  $aa$  organisms

will also be *aa*. If an *AA* organism reproduces with an *aa* organism, on the other hand, all their offspring will be heterozygotes with the *Aa* genotype.

Things are more complicated, however, when reproduction involves heterozygotes. Recall that half of an *Aa* organism's gametes will be *A* and the other half *a*. (There are exceptions to this, but they need not concern us here.) If an *Aa* organism reproduces with an *AA* organism, 50 percent of all possible zygotes created through their matings will be *AA* and the other 50 percent will be *Aa*. This is because the *AA* organism will contribute only *A* gametes to their union, which will fuse with the 50 percent *A* gametes contributed by the heterozygote to form the *AA* zygotes and with the other 50 percent *a* gametes from the heterozygote to form the *Aa* zygotes. Similarly, if a heterozygote reproduces with an *aa* organism, 50 percent of all possible zygotes created through their matings will be *Aa* and 50 percent will be *aa*. If two heterozygotes mate with one another, on the other hand, half the female's eggs will be *A* and half *a*, and half the male's sperm will be *A* and half *a*. Of the *A* eggs, half will thus be fertilized by *A* sperm and half by *a* sperm, so 25 percent of the fertilized eggs will be *AA* and 25 percent *Aa*. Similarly, of the *a* eggs, half will be fertilized by *A* sperm and half by *a* sperm, producing an additional 25 percent of the fertilized eggs that are *Aa* and 25 percent that are *aa*. In total, then, 25 percent of the zygotes will be *AA*, 50 percent will be *Aa*, and 25 percent will be *aa*.

I've spoken as though all an organism's gametes go to form zygotes. This, of course, is false; many parents have only one child, for example. If heterozygote parents have one child, it will be just one of the three possible genotypes. The way the above principles apply to such cases is in terms of probabilities. That is, there is a 25 percent chance that a child of two heterozygotes will be *AA*, a 50 percent chance that it will be a heterozygote like its parents, and a 25 percent chance that it will be *aa*. This use of probabilities assumes that the genotypes of zygotes in an indefinitely large population of heterozygotes would occur in the 25/50/25 percent frequencies mentioned above, even if many heterozygote pairs in that population produce only one child.

You will have noticed that, while matings between two *AA* organisms and between two *aa* organisms produce only *AA* and *aa* offspring respectively, matings between heterozygotes can produce both heterozygous and homozygous offspring. This has implications for the understanding of evolution as change in gene or genotype frequencies across generations. For suppose that there is a very small population of heterozygotes that reproduces in replacement numbers—that is, each couple produces only two offspring. Since we are supposing that each organism in this population has

the  $Aa$  genotype, there are two alleles,  $A$  and  $a$ , that occur at the locus that interests us. Further, since every organism is  $Aa$ , half the alleles that occur at that locus are  $A$  and half are  $a$ ; in other words, the frequency of  $A$  is 50 percent and the frequency of  $a$  is 50 percent. Now, although unlikely, it is possible for each couple to produce one  $AA$  and one  $Aa$  offspring. In that case, in the offspring generation the frequency of the  $A$  allele will increase to 75 percent (since three out of every four slots in the diploid locus are occupied by  $A$ ) and the frequency of the  $a$  allele will decrease to 25 percent. Under the modern genetic definition of evolution, this constitutes significant evolution.

While this example of a change in gene frequencies also involves a change in genotype frequencies (since the population evolved from 100 percent  $Aa$  to 50 percent  $AA$  and 50 percent  $Aa$ ), it is possible for there to be a change in genotype frequencies across generations without a corresponding change in gene frequencies. To see how, suppose there is a population consisting of just eight organisms—three  $AA$ , three  $aa$ , and two  $Aa$  organisms. The alleles  $A$  and  $a$  each occur with a frequency of 50 percent in this generation of the population. (There are six copies of  $A$  from the three  $AA$  organisms and two copies of  $A$  from the two  $Aa$  organisms; and there are six copies of  $a$  from the three  $aa$  organisms and two copies from the two  $Aa$  organisms. There are thus eight copies each of  $A$  and  $a$ , out of a total of sixteen alleles at the locus.) Now suppose that two  $AA$  organisms mate with each other as do two  $aa$  organisms, and one  $Aa$  organism mates with the remaining  $AA$  organism while the other  $Aa$  organism mates with the remaining  $aa$  organism. Suppose further that each of these pairs produces just two offspring, so the next generation of the population also contains just eight organisms. We know that the offspring from the homozygote matings will be homozygotes of the same genotype and that there is a 50 percent chance that the offspring from the heterozygote-homozygote matings will also be homozygotes. Suppose that in fact the heterozygote-homozygote pairs produce only homozygotes. Then the next generation of the population will consist of four  $AA$  organisms (two from the  $AA-AA$  mating and two from the  $AA-Aa$  mating) and four  $aa$  organisms (two from the  $aa-aa$  mating and two from the  $Aa-aa$  mating). Although the gene frequencies have not changed, both  $A$  and  $a$  remaining at 50 percent, the genotype frequencies have. For whereas 37.5 percent of the parental generation was  $AA$ , 37.5 percent  $aa$ , and 25 percent  $Aa$ , the offspring generation is 50 percent  $AA$  and 50 percent  $aa$ . Under the modern genetic definition of evolution, this also constitutes significant evolution.

The kinds of evolution we have just considered can also produce changes across generations in the frequencies of phenotypes in a population. The reason is that genes regulate the synthesis of proteins, the stuff of which our bodies are made, and differences between bodies or between parts of the same body are a product of differences in the proteins of which they are made. By regulating protein synthesis, genes consequently guide the development of organisms, and this influences the phenotypes that organisms possess. When a gene influences a particular phenotype in this way, biologists say it is a *gene for* that phenotype. Thus, a change in gene or genotype frequencies can produce a change in the frequencies of the phenotypes influenced by those genes or genotypes.

But why do I say that genes “influence” phenotypes, rather than saying that genes “determine” phenotypes? This is because, by themselves, genes don’t determine anything. One can’t simply put some carefully selected genes in a petri dish, for example, and grow a cute little button nose. For how a gene affects the phenotype of an organism depends on precisely when (or if) it is switched on and off in the process of development, and that in turn depends on the properties of the gene’s environment. The environment of a gene includes not only the environment outside the organism (which affects the surface of the organism), but also the cells surrounding the one in which the gene resides (which can affect gene action, sometimes as a result of cascading effects from the environment outside the organism) and the other genes within the same cell (whose patterns of activity can affect when a gene is switched on or off). In short, the development of an organism is not simply a matter of gene action, but a matter of causal *interaction* between genes and their environment.

For this reason, there is no straightforward relationship of “determination” between genotypes and phenotypes. Indeed, given the interaction between genes and environment in development, even if two individuals possess the same genotype, they can differ in phenotype as a result of developing under different environmental conditions. For example, if we plant corn seeds of the same genotype in different soil conditions, and fertilize and water those plants differently, the resulting corn plants can differ significantly in phenotypes such as height of plant and sweetness of kernels. So the same genotype can produce a range of different phenotypes across a range of different developmental environments. Some genotypes tend to produce the same phenotype across a very wide range of different environments. But rarely is there a straightforward one-to-one relation between a genotype and a phenotype. Genotypes typically produce different phenotypes if developmental conditions are varied sufficiently.

But, if this is true, what sense does it make to speak of a “gene for” a particular phenotype? To say that a gene or genotype  $G$  is “for” a phenotype  $P$  means first of all that, other things being equal, an organism with  $G$  is more likely to have  $P$  than is any organism without  $G$  (that is, with a possible rival allele of  $G$ ). The clause “other things being equal” is important here, since it includes the environments in which the organisms develop. The point is to compare organisms within developmentally similar environments to see whether having  $G$  makes a difference with respect to having  $P$ . For, if we compared organisms with  $G$  in one developmental environment to organisms without  $G$  in a different developmental environment, then any difference among them with respect to their having  $P$  could be due to the differences in their developmental environments. The clause “other things being equal” enables us to focus on how a change in having  $G$  produces a change in having  $P$ , rather than on how a change in the environment produces a change in having  $P$ . But this first condition is purely correlational, requiring that  $G$  be correlated with  $P$  in relevant environments in order to be “for”  $P$ . Consequently, this condition alone fails to distinguish the case in which  $G$  actually produces  $P$  from the case in which  $G$  produces some other phenotype that is correlated with  $P$ . Thus, to say that  $G$  is “for”  $P$  means, second, that  $G$  must play a causal role in the development of  $P$  (in those organisms with  $P$ ). When these two conditions are met, it is perfectly sensible to speak of genes or genotypes as being “for” phenotypes.

The fact that phenotypes are produced by the interaction of genes and environment has a couple of implications with respect to understanding the connection between evolution at the genetic level and changes across generations in the phenotypes of organisms in a lineage. First, if a gene increases in frequency across generations, the phenotype that it is the gene for can increase in frequency only if the developmentally relevant aspects of the environment remain relatively constant. This is because there will only be a particular range of developmental environments in which that gene will produce the phenotype it is for. So, if the environment changes so as to fall outside the range in which that gene produces that phenotype, then any increase in the gene’s frequency will not be accompanied by an increase in the frequency of the phenotype it is for. Thus, in order for patterns of phenotypic change across generations to parallel patterns of evolution at the genetic level, the developmentally relevant properties of the environment must remain relatively stable across those generations.

Second, because genotypes can produce different phenotypes in different developmental environments, it is possible for there to be phenotypic

change across generations in the absence of genetic evolution. This can occur simply in virtue of changes in the environment in which genotypes develop. For example, even if we control the genes of corn plants from one generation to the next, so that there is no genetic change, it is still possible to produce taller corn plants in the later generation by altering how much the plants are fertilized and watered. Such transgenerational changes in the environment can produce what I will call *phenotypic evolution* in the absence of any underlying genetic evolution. Thus, phenotypic evolution—changes across generations in phenotype frequencies—can be strictly environmentally driven. This happened in many human populations during the twentieth century, when improved diets produced an increase in average height over the course of the century.

### The Causes of Evolution

So far we have been concerned with what evolution *is*. But what *causes* evolution? As mentioned earlier, evolution can occur only if there is *variation* in a population. For, if evolution is change in gene or genotype frequencies, there must be at least two genotypes occurring at a particular locus in a population, the frequencies of which then get altered across generations. So, if a population is composed of organisms that are genetically identical, the only way that evolution can occur is if a new genetic variant gets introduced into the population. With this in mind, the causes of evolution can be divided into two very broad types: One type of cause introduces new variants into a population and the other changes the frequencies of already existing variants. Consider these in turn.

There are two main processes that cause evolution by introducing new variants into a population, one of which is mutation. Recall that the first stage of meiosis involves the replication, or copying, of the genes on each chromosome in a diploid cell. In the process of gene replication, there are occasional copying errors, in which one of the nucleic acid bases in a sequence gets translated incorrectly—for example, an A in a sequence of bases gets copied as T. The result is a new gene—a new sequence of bases—that differs from the gene from which it was copied at that one position in the sequence of bases. A copying error of this kind is called a *mutation*. The mutation is then shuttled into one of an organism's gametes where it can be transmitted to one of its offspring, in whom the new mutant gene can then produce some novel phenotype.

The other process that causes evolution by introducing new variants into a population is *recombination*. To illustrate, consider a double heterozygote,

an organism with the  $Aa$  genotype at one locus and the  $Bb$  genotype at another locus, where  $A$  and  $B$  occur on one chromosome while  $a$  and  $b$  occur on the other. In such a double heterozygote, meiosis without recombination produces two  $AB$  gametes and two  $ab$  gametes. (Recall that  $A$  and  $B$  are alleles at different loci on the *same* chromosome, as are  $a$  and  $b$ . So  $AB$  is not a genotype; it is a chromosome type.) Sometimes, however, after chromosome replication but before the first cell division, chromosomes align themselves and exchange genes in a process called *crossing over*. For example, an  $AB$  chromosome may align with an  $ab$  chromosome and exchange its  $B$  with its partner's  $b$ , thereby transforming the  $AB$  chromosome into an  $Ab$  chromosome and its partner into an  $aB$  chromosome. The second stage of meiosis will then produce four distinct gametes:  $AB$ ,  $Ab$ ,  $aB$ , and  $ab$ . In this process, genes get recombined, and new genetic variants get introduced, specifically the  $Ab$  and  $aB$  chromosomes.

Recombination has a significant effect in reproduction. For, in the absence of recombination, if two double heterozygotes—that is, two  $AaBb$  organisms—reproduce, their offspring have a 25 percent chance of being  $AABB$ , a 50 percent chance of being  $AaBb$ , and a 25 percent chance of being  $aabb$ . But, if recombination occurs during meiosis in one of these organisms, their offspring have instead a 12.5 percent chance of being  $AABB$ , a 25 percent chance of being  $AaBb$ , a 12.5 percent chance of being  $aabb$ , and an additional 12.5 percent chance each of being  $AABb$ ,  $AaBB$ ,  $Aabb$ , and  $aABb$ . And, if recombination occurs during meiosis in both parents, there are further possibilities. Recombination can thus introduce into an offspring generation significant genetic variation that wasn't in the parent generation. The difference between this and mutation is that mutation introduces new variants by creating *new genes*, while recombination does so by creating *new combinations* of genes on a chromosome.

It is important to note that both mutation and recombination are *non-directed*, or *random*, processes. This means that the fact that a new variant might be beneficial to an organism does not increase the probability that it will be produced. Indeed, the overwhelming majority of mutations are either neutral or detrimental. Thus, the processes that generate new variation in a population operate independently of the processes that determine what is beneficial or detrimental to the organisms in that population. But, while new variants are random in origin, their frequency in a population once they have arisen may or may not be random, as we are about to see.

There are also two main processes that cause evolution by altering the frequencies of already existing variants in a population. One of these is

*natural selection*, which is a process that occurs when three conditions obtain in a population. First, there must be preexisting *phenotypic variation* in the population. Second, the variant phenotypes must be *hereditary*—that is, there must be genes for each of the variant phenotypes, which parents transmit to their offspring. Third, these hereditary phenotypic differences must be responsible for *differences in fitness*.

This third condition requires some explaining. *Fitness*, as it is most commonly characterized, is a measure of an organism's ability to survive and reproduce. Thus, if one organism is fitter, or has greater fitness, than another, the former has a greater ability to survive and reproduce than the latter. This does not mean that the fitter organism actually will survive longer and reproduce more than the less fit organism. You may be better able than I to lift 300 pounds, but your greater ability may never have the chance to show itself in actual performance, since you may never have the opportunity to attempt to lift 300 pounds. Similarly, one organism may be better able than another to survive and reproduce even though it doesn't actually outlive and outreproduce the other. It may die from a freakish accident before puberty, for example. So fitness is not a measure of an organism's *actual* survival and reproduction, but a measure of its *ability* to survive and reproduce.

In addition, an organism's ability to survive and reproduce depends not simply on its physical characteristics, but on how well adapted those characteristics are to the environment the organism inhabits, which in turn depends on the precise nature of the environmental demands, or *selection pressures*, an organism faces. For heuristic purposes (and with serious qualifications to be discussed in chapter 3), we can think of these selection pressures as posing *adaptive problems*, which an organism must solve in order to survive and reproduce. Such problems would include finding food, avoiding predators, and attracting mates. An organism's phenotypes can then be thought of as providing potential "solutions" to these problems. Some organisms may thus be endowed with a phenotype (for example, greater running speed) that provides a better solution to an adaptive problem (escaping predators) than the phenotypes with which other (slower) organisms are endowed. To say that fitness is a measure of an organism's ability to survive and reproduce, then, is to say that fitness is a measure of how well an organism's characteristics solve the adaptive problems posed by its environment. Thus, an organism's fitness is always relative to its environment; its characteristics may make it better able to survive and reproduce in one environment than in another.

As many biologists have pointed out, however, conceiving of fitness as a measure of the ability to survive and reproduce in an environment is a

little misleading. For survival, in itself, means nothing in evolutionary terms. Surely, if one fails to survive a childhood illness, one will not contribute to the gene pool of the next generation. But one's impact on future gene pools is no greater if one is a THINKER (half of a couple with Two Healthy Incomes, No Kids, and Early Retirement). What's important in evolution is whether one *reproduces*; survival matters only insofar as it enables reproduction. But the evolutionary significance of reproduction, in turn, lies in the fact that, in contributing offspring to the next generation, one is transmitting (half of) one's genes to that generation and thereby affecting the gene and genotype frequencies in that generation. Once we reconceive fitness as a measure of the ability to survive *to* reproduce, then, and recognize that reproduction is a matter of transmitting one's genes to the next generation, we can redefine "fitness" as a measure of an organism's *expected genetic contribution* to future generations. In this refined definition, the term "expected" reflects the degree to which an organism's characteristics enable it to reproduce: Its ability to reproduce is measured as the *probability* of its reproducing. And the term "genetic contribution" reflects how many copies of its genes an organism contributes to future generations via the number of its offspring. Thus, to say that the organisms in a population differ in fitness is to say that they differ in their expected genetic contributions to future generations (in the specific environment they inhabit). Given this definition, we can then define the fitness of a genotype or phenotype as the average fitness of all the organisms with that genotype or phenotype.

Returning now to the three conditions under which natural selection occurs, when (1) phenotypic variation is (2) hereditary and (3) responsible for fitness differences in a population, the phenotypic traits that enhance fitness in that population (that is, the phenotypic traits that make their possessors *fitter than* organisms possessing alternative traits) will increase in frequency across generations. This is because organisms with a fitness-enhancing trait will, *on average*, outreproduce the other organisms, thereby transmitting more of their genes to the next generation than those other organisms transmit. These genes, of course, will include the gene for the fitness-enhancing trait. And, as more copies of that gene get transmitted to the next generation, proportionately more of the population will develop the fitness-enhancing trait, and it will thereby increase in frequency in the population. This process is natural selection, and it changes the frequencies of genes in a population as a function of the phenotypic effects they produce, increasing in frequency those genes with fitness-enhancing phenotypic effects and decreasing in frequency those genes with fitness-reducing phenotypic effects. Of course, as it changes the

frequencies of the genes with these phenotypic effects, it also changes the frequencies of the phenotypes they produce. When a phenotypic trait increases in frequency as a result of natural selection in this way, biologists say that there has been *selection for* that trait—that the trait has conferred a *selective advantage*, or *reproductive advantage*, on its bearers.

It is worth noting that some biologists apply the term *natural selection* only to selection for traits that affect survival, while applying the term *sexual selection* to selection for traits that affect the ability to attract and mate with members of the opposite sex. In other words, traits that are solutions to adaptive problems posed by members of the opposite sex evolve under sexual selection, whereas traits that are solutions to adaptive problems posed by the rest of the environment evolve under natural selection. Other biologists treat sexual selection as an aspect of natural selection. But distinguishing the two can be useful when analyzing some traits, since some traits are detrimental with respect to survival, yet enhance reproductive success by appealing to members of the opposite sex. The classic example is the peacock's tail, which is detrimental to survival (since it attracts predators and impairs the ability to escape), yet appeals to peahens and, hence, increases the mating ability of well-endowed peacocks. For the most part I will simply use the term *selection*, encompassing both natural and sexual selection. But, when necessary, I will refer specifically to natural or sexual selection.

Finally, the other process that can cause evolution by altering the frequencies of already existing variants in a population is *genetic drift*, which is due to two types of chance event: random survival and random sampling of gametes.

Random survival is due to random events—for example, floods, fires, or lightning strikes—that kill a much larger number of organisms with one allele than those with the rival allele. This would have the effect of making the latter allele more frequent in the next generation, since its bearers would have survived to reproduce at a higher rate than the bearers of the unlucky allele. This would constitute evolution, but it would be due to chance rather than to differences in fitness.

We have already touched on the random sampling of gametes. Recall that every organism produces many more gametes than will go to form zygotes. We can thus think of fertilization as a process that randomly “draws” one gamete from the total pool of gametes created by each parent organism. When a population contains  $Aa$  organisms, which produce gametes that are 50 percent  $A$  and 50 percent  $a$ , the random drawing of their gametes, *in each case*, has a 50 percent chance of yielding an  $A$  and

a 50 percent chance of yielding an *a*. But there is a possibility that the total number of drawings of gametes from heterozygotes in a population will contain many more copies of one allele than of the other. This is analogous to flipping a coin a number of times. Each coin toss has a 50 percent chance of landing heads and a 50 percent chance of landing tails. Nonetheless, it is possible that a string of twelve coin flips will yield nine heads and three tails. When the random sampling of gametes draws a greater number of one allele than of its rival in this way, there is a change in allele frequencies across generations, but it is due strictly to a randomness built into the process of fertilization rather than to selection for one allele over the others.

Drift is a causal force in evolution in every generation, since random survival and random sampling of gametes occurs in every generation. But frequently the effects of drift are offset by selection. In order for drift to be the cause of a *long-term evolutionary trend*, the rival alleles at a locus must be *selectively neutral* (that is, no one of the alleles can confer a selective advantage on its bearers). When rival alleles are selectively neutral, the frequencies of those alleles can change greatly over many generations due strictly to genetic drift. Indeed, drift can drive an allele to fixation, or extinction, in a population just as surely as selection can, since the effects of drift can be compounded over many generations just like the effects of selection. But these effects of drift are greatest in small populations. This is because in larger populations the frequencies of alleles in the pool of actually sampled gametes more closely approximate the frequencies of those alleles in the total pool of gametes available to be sampled. This, again, is analogous to the coin-flipping case. A three-to-one ratio of heads to tails is more common in series of twelve coin tosses than in series of twelve thousand coin tosses. As the number of tosses in the series increases, there are more and more series in which the frequency of heads and tails closely approximates 50 percent. Similarly, drift is far more likely to have significant effects in small populations than in large ones, since the alleles in the actually sampled heterozygote gametes in large populations more closely approximate a 50/50 frequency. In large populations, therefore, selection tends to be the primary cause of long-term evolutionary trends.

## Adaptation

These are the nuts and bolts of evolution, but how do they fit together to build all the complex, functionally integrated organisms that we see in the world? For organisms are composed of numerous and diverse parts that are

well adapted to one another and to particular features of the world, and that appear very intricately designed for their functionally specialized roles. Darwin called such functionally specialized parts of organisms “organs of extreme perfection and complication” and, in illustration, marveled at the human eye, “with all its inimitable contrivances for adjusting the focus to different distances, for admitting different amounts of light, and for the correction of spherical and chromatic aberration.”<sup>1</sup> The eye, of course, is merely one of many examples of such “perfection and complication.” The wings of birds are very well designed for flight, the echolocation (sonar) system of bats is very well designed for detecting flying insects at night, and the coloration of many species provides excellent camouflage from predators.

Such “organs of extreme perfection and complication” appear to be designed for a *purpose*. Echolocation, for example, appears to be designed precisely so that bats can detect the flying insects that make up their diet. And to say that a part is designed for a purpose is to say that an organism possesses it *because* that part solves a particular adaptive problem. So, bats appear to possess echolocation precisely because possessing echolocation enables them to eat. But, if all apparent design in nature is the product of evolution, rather than the product of creation by some intelligent being, how can there be such design-for-a-purpose in nature? How, in other words, can the processes discussed so far account for the apparent purposeful design of functionally specialized “organs of extreme perfection and complication”?

Whether it is a matter of building a trait that appears well designed for solving some adaptive problem or building an entire organism composed of numerous such traits that are all functionally integrated, the process is the same: iterated cycles of modifying a preexisting structure and retaining the modification. This process created all of the world’s diverse organic forms out of simple replicating molecules. Of course it took a very long time. But this book is not about the origins of species. For our purposes, it is sufficient to understand how complex traits that solve adaptive problems are created by the causes of evolution just discussed—to understand how traits can develop *within* a species that make organisms well adapted to the specific demands of their environment. Consider first how such traits can evolve under selection, then consider whether they can evolve under drift.

Suppose there is a population of birds whose beaks vary slightly in size. The sole food supply for this population is seeds that are digestible only once they have been extracted from their hulls. To extract the seeds, the

birds must use their beaks. Suppose that birds with the slightly broader beaks are the most efficient at hulling the seeds, hence get the most nutrition, and consequently enjoy a slight reproductive advantage over the other birds in the population. The gene for the broad beak will thus increase in frequency in the population, as will the broad beak itself.

But suppose also that the broader beak would be even more efficient at hulling seeds if it had slightly sharper edges. And suppose that there is a gene in the population that would produce sharper-edged beaks if it mutated. Of course, since mutation is random, the fact that sharper beaks would be beneficial doesn't increase the probability that the desirable mutation will occur. Also, since mutation is random, the mutation for sharper-edged beaks is just as likely to occur in a bird without the broad beak as it is to occur in a bird with the broad beak (in whom it would be most beneficial). But, as the gene for the broad beak becomes ever more frequent in the population, there is an increased probability that, if the mutation for sharper-edged beaks occurs, it will occur in conjunction with the gene for the broader beak, and thereby provide a beak that is even better designed for hulling seeds. This increased probability of a better beak is analogous to rolling dice. Suppose you need a three to turn up on a rolled die. If you roll just one die, there is a one-sixth probability of getting a three. And, if you roll twelve dice, each die has a one-sixth probability of turning up three. But your odds of getting a three are greatly improved if you can roll twelve dice rather than one die. Similarly, as the gene for a broad beak spreads in the population, there is an increased probability that a mutation for sharper edges might occur with it and, hence, further modify the beak in a way that makes it even better designed for hulling seeds. If the beak does get further modified in this way, the improved beak will become more frequent in the population over succeeding generations. And so on.

This process—a new mutation introduces a beneficial modification that is retained by selection—can be repeated many times over a very long period of time. After a very large number of generations, the population can come to be composed of a large number of birds with beaks that are extremely well designed for hulling seeds, beaks that have a shape that conforms to the demands of the seed hulls and are powered by muscles that exert efficient force in cracking those hulls. And that design will have been produced by a process of cumulative retention of slight design improvements introduced by random mutations. (Of course, there may have been many other mutations that impaired design; but they would have been selected against and driven to extinction.) In this way, the

combination of mutation (which adds modifications to preexisting traits) and selection (which preserves the new modifications that are beneficial and subtracts those that are not) can build traits of great complexity, which make their bearers highly adapted to their environment and highly successful at solving adaptive problems related to survival and reproduction.

Note that what is essential to building complex traits is the process of cumulative retention of modifications that further elaborate the design of the trait. In principle, such cumulative retention of design elaborations could be accomplished by genetic drift. For drift can increase the frequency of a design-elaborating allele in a population and thereby increase the probability that another design-elaborating mutation could occur in conjunction with it. But, if this process is guided by drift alone, by definition each new modification must be *selectively neutral*. If a modification provides a reproductive advantage to its bearers, then by definition it is undergoing selection, not drift. In addition, if each new modification were affected by drift alone, it would be as likely to drift to extinction as to drift to near fixation. (Actually, since each new modification would be introduced by an initially rare allele, the initial rarity of the allele would make it more likely to drift to extinction than fixation.) Thus, it would be monstrously unlikely that drift alone would accumulate a *whole series* of modifications to build a trait as complex as the human eye, for example. When a modification provides a selective advantage, however, selection is a force that favors its persistence, and typically proliferation, in the population and actively works against the possibility of its extinction. Since selection preserves modifications to traits that are advantageous to their bearers, it increases the probability that organisms in a population will develop complex traits that serve a purpose.

This is the crux of the issue about whether drift can create traits that solve adaptive problems so effectively as to give the appearance of having been designed for the purpose of solving those problems. For recall that, insofar as a trait's *purpose* is to serve a particular function, it is present in organisms *because of* the beneficial function it serves. A trait that has evolved under drift, however, is present in the organisms in a population only because of chance, even if it is the result of cumulative modification. In fact, even if a trait that has evolved under drift provides some benefit to the organisms that currently possess it, they possess that trait only because the gene for it has randomly drifted to a high frequency in the population, not because of the benefit the trait provides. When a trait evolves under selection, in contrast, organisms possess that trait because it provided a benefit to their ancestors from whom they inherited the gene

for that trait—specifically, the benefit because of which the trait was selected. The benefit the trait provides is thus the reason why the trait spreads or persists in the population; that benefit *is* the purpose of the trait, since the trait's providing that benefit is the reason organisms possess the trait (via inheritance from ancestors in whom the trait was selected for).

A trait that is present in a current population because it performed a function (solved an adaptive problem) that enhanced fitness in an ancestral population, and was thus preserved or proliferated under selection for it, is called an *adaptation*. In other words, an adaptation is a trait that has a history of having been preserved, and possibly modified, by selection for the beneficial role it plays in an organism. Thus, an adaptation is a trait that contributed to its own persistence or proliferation; for, by enhancing the fitness of its bearers, an adaptation contributed to the reproductive success of its bearers, which contributed to the transmission of the genes for that adaptation, which in turn contributed to the development of that trait in other organisms. Adaptations, in short, are self-perpetuating design features of organisms. Organisms have those traits *because* they were beneficial to their ancestors.

It is important not to confuse adaptation with adaptiveness. A trait is adaptive if it enhances fitness, but it is an adaptation if it is possessed by organisms in a current population because they inherited it from ancestors in whom that trait enhanced fitness. As the philosopher of biology Elliott Sober so nicely puts it: "To say that a trait is an adaptation is to make a claim about the cause of its presence; to say that it is adaptive is to comment on its consequences for survival and reproduction."<sup>2</sup>

This distinction is important to bear in mind because of the following two implications. First, just because a trait is adaptive doesn't mean that it is an adaptation. A trait could evolve in a population under drift, but then come to enhance the fitness of its bearers if the environment of the population changes so as to make the trait useful. In such a case, the trait would be adaptive, but since it did not evolve under selection it would not be an adaptation. For adaptation is a historical concept, applying only to traits with the right sort of evolutionary history. Second, a trait could be an adaptation yet fail to be adaptive. This, too, could result from a change in a population's environment. A trait could evolve under selection, and even go to fixation in a population, yet the environmental demands to which that trait was responsive could cease, thereby rendering the trait useless. In such a case, the adaptation would no longer be adaptive. For adaptiveness is an ahistorical concept, applying only to traits that currently enhance fitness. A trait is adaptive, then, if it has *current utility*; it is an

adaptation if it had *past utility*, if it evolved and is present in a current population because it *was adaptive*.

The fact that organisms possess adaptations because of the benefits those traits provided to the organisms' ancestors means that questions about why an organism possesses a particular adaptation are always ambiguous, admitting of two very distinct types of answer. To illustrate, suppose we ask why black-headed gulls remove the eggshells from their nests after their fledglings have hatched. We could answer in terms of the functioning of the neurophysiological behavior-control mechanisms in the gull and how those mechanisms respond to stimuli in the gull's environment. This would answer in terms of the immediate causal antecedents, the *proximate causes*, of the eggshell-removal behavior. We could also trace these causes back a little further in time and answer in terms of the developmental processes by which a gull comes to have the mechanisms that control eggshell removal. While this would cite causes that are not the immediate antecedents of the phenomenon to be explained, it would nonetheless explain that phenomenon by citing causes within the lifetime of the individual gull whose eggshell-removal behavior we are explaining. In that sense, it would still be an explanation citing proximate causes.

In sharp contrast, however, we could explain the eggshell-removal behavior in terms of the history of selection that caused that behavior to become widespread in, and characteristic of, black-headed gulls. Such an explanation in terms of selection would cite the fitness-enhancing benefit provided by eggshell removal, because of which eggshell removal increased in frequency in ancestral gull populations. This explanation would consist in pointing out that eggshells are conspicuous and attract the attention of birds that prey on gull fledglings; thus, removing eggshells from the nest helps protect fledglings against predation. Consequently, gulls that removed eggshells from their nests made a greater genetic contribution to subsequent generations than gulls that didn't remove their eggshells; so eggshell removal evolved to (near) fixation in gull populations. This would be an explanation in terms of the *ultimate causes* of eggshell removal, what caused the evolution of eggshell removal in gull populations. Unlike the explanation in terms of proximate causes, the explanation in terms of ultimate causes explains a gull's eggshell-removal behavior in terms of causes that acted during the evolutionary history of the lineage leading up to that gull, not in terms of causes acting during that gull's lifetime.

It is important to note that proximate explanations (those citing proximate causes) and ultimate explanations (those citing ultimate causes) do not compete with one another. It's not the case that, if one explanation is

right, the other must be wrong. Rather, they complement one another by providing different kinds of information about the same phenomenon. Indeed, one could see ultimate explanations as explaining why particular proximate causes are operative. For example, the ultimate explanation of eggshell removal explains why gulls have neurophysiological mechanisms that respond to particular stimuli in a way that results in eggshell removal. But that doesn't mean that the ultimate explanation can *replace* a proximate explanation. Knowing the ultimate causes of eggshell removal doesn't give us any information about how eggshell removal gets accomplished by any individual gull. Similarly, a proximate explanation doesn't exclude an ultimate explanation, since knowing how a particular neurophysiological mechanism causes eggshell removal doesn't inform us about the causes of the evolution of eggshell removal. Thus, every adaptation can be explained in terms of both proximate and ultimate causes, where the former cites the immediate antecedent "mechanistic" causes and the latter cites the evolutionary causes.

### Phenotypic Variation

Up to this point, I have spoken of selection as a process in which some trait consistently enhances the fitness of its bearers over a very large number of generations. When this occurs, in each new generation the fitness-enhancing trait will increase in frequency in the population. If this process continues for enough generations, the trait will eventually go to *fixation* (become possessed by every organism) in a population, thereby wiping out all rival traits. While selection does sometimes drive traits to fixation in this way, it doesn't always act to eliminate phenotypic variation and create a uniform population. Indeed, there are several reasons why selection doesn't always eliminate phenotypic variation.

First, mutation and recombination introduce new variation into a population in every generation. Thus, even if selection reduces variation in each generation, by increasing the frequency of the fittest variant, it may never completely eliminate variation, since new variation is continually introduced.

Second, some phenotypic variation is selectively neutral, in which case selection won't favor any of the variants over the others (or won't favor any of the variants that are within a certain range over the others in that range). For example, population members may vary in height or weight in ways that don't affect their fitness. In such a case, selection won't winnow this variation, since no one of the variants is fitter than any of the others.

(Of course, it could be the case that extreme heights or weights would be selected *against*, while all of the nonextreme variation in height and weight would be selectively neutral.)

Third, even if a genotype for a fitness-enhancing phenotype goes to fixation under selection, the fact that the same genotype can produce different phenotypes under different developmental conditions means that the phenotype it's for won't necessarily go to fixation also. For a genotype to be selected, it needn't always produce the fitness-enhancing phenotype. It is only necessary that the *average fitness* of all the phenotypes it produces (under all its developmental conditions) be higher than the average fitness of all the phenotypes produced by alternative genotypes (under all their developmental conditions). So, even a genotype that is increasing in frequency under selection can sometimes produce phenotypes that provide no selective advantage or are positively maladaptive. Indeed, even if that genotype goes to fixation, it might still, in certain developmental conditions, produce a phenotype other than the fitness-enhancing phenotype it was selected for producing. Thus, variation in developmental conditions can produce phenotypic variation even when a beneficial genotype has gone to fixation.

These are cases in which phenotypic variation persists in a population in spite of selection, as it were. More interesting, however, are the ways in which selection can actively maintain phenotypic variation in a population. There are several ways in which selection can maintain phenotypic variation, but only two of these will be relevant to later discussions. So here I'll confine my discussion to those two ways: frequency-dependent selection and adaptive plasticity.

Consider first *frequency-dependent selection*. To get a really good handle on frequency-dependent selection, it is best to take a brief excursion into cost-benefit analyses of fitness.

Fitness, recall, is a measure of the ability to survive and reproduce in a particular environment. Many activities in which organisms engage enhance or diminish that ability. For example, female black-tipped hangingflies mate with males who offer them edible insects. When the male presents the insect, the female feeds on it while copulation occurs. Consequently, a male hangingfly enhances his ability to reproduce by capturing an insect that will entice a female. Capturing an insect is thus a *fitness benefit* for male hangingflies. Conversely, if a female lays eggs to be fertilized by a male who turns out to be sterile, she diminishes her ability to reproduce as a function of the lost eggs. Losing the eggs is for her a *fitness cost*.

Fitness costs and benefits need not be so drastic or so obvious. Each meal that we eat contains nutrients that sustain us and thereby enhances our ability to survive and reproduce relative to the ability we would possess in the absence of receiving those nutrients. We can thus think of very simple acts such as eating an apple as having an associated, yet small, fitness benefit, measured in terms of the nutrients the apple provides and the role those nutrients play in facilitating survival and reproduction. Similarly, the very act of engaging in some activity has metabolic costs, diminishing the energy available for engaging in other activities. Three hours spent in fruitless foraging diminishes one's energy store for fruitful copulations. So every activity has an associated, though perhaps small, fitness cost, measured in terms of the depletion of energy available for other activities essential to survival and reproduction.

The fitness costs and benefits of some activities in which organisms engage are independent of the behavior of other members of the organism's population. The energy gained from eating a particular food item, for example, is independent of what other individuals in the population are eating or doing. For male dung flies, there is an optimal amount of time spent copulating, which maximizes the rate of egg fertilization per unit of copulation time, and for any given male this optimum is independent of how long other males spend copulating. And for many animals there is an optimal amount of time spent foraging for food, which maximizes the energy intake per unit of foraging time, and this optimum is independent of how long other population members spend foraging. An activity with fitness costs and benefits that are independent of how other population members behave has *frequency-independent fitness*, since its fitness is independent of the frequency of that activity in a population—independent, that is, of how many population members engage in that activity.

But the fitness costs and benefits of many activities in which humans and other animals engage are not independent of the behavior of other population members. For example, in some species males fight with one another for territory. If most males in such a species only engage in threatening displays and retreat when attacked, a tactic of extreme aggression might accrue high fitness benefits to any male adopting it. However, if most males are extremely aggressive in conflicts, then aggression could exact the fitness costs of injury or death. So the fitness costs and benefits of any particular form of behavior in a conflict depend on the tactics adopted by other males in the population. Similarly, when members of one sex compete to mate with members of the opposite sex, the best tactic to employ to attract members of the opposite sex can depend on what other

members of your own sex are doing to attract mates. If all other members of your sex send roses, rather than competing to find the best roses, you may be better off sending orchids, which are easier to obtain since they aren't in demand. It may pay just to be different. In fact, in general, the fitness costs and benefits associated with any activity that involves competition with some other population members will be a function of how one's competitors behave. Such an activity has *frequency-dependent fitness*, since its fitness is dependent on the frequency of that activity in a population. Consequently, the fitness of an activity with frequency-dependent fitness changes as the frequency of that activity in a population changes.

When activities have frequency-dependent fitnesses, selection often maintains a particular proportion of alternative variants. To see how this can occur, consider a simple model known as the "Hawk-Dove game," which was first developed in a classic article by the evolutionary biologists John Maynard Smith and Geoffrey A. Parker. For purposes of illustrating the game, we'll represent fitness costs and benefits by whole numbers, or "fitness points," where benefits are represented by positive numbers and costs by negative numbers.

The Hawk-Dove game is a contest for a resource worth +40 points, and contestants can "play" either Hawk or Dove in competing for the resource. Hawks always attack and fight aggressively until they win or get seriously injured. Doves always exhibit a threatening display, but never attack, and retreat if attacked by their opponent. Since Hawks immediately attack and Doves retreat when attacked, Doves always immediately lose to Hawks. But we'll assume that Hawks have a 50 percent chance of defeating another Hawk and that Doves have a 50 percent chance of defeating another Dove. Finally, we'll assume that the cost of a serious injury is -60 points and that wasting time and energy in a very prolonged contest costs -10 points.

Given these assumptions, neither Hawk nor Dove can evolve to fixation and remain there. To see why, consider first a population of Doves. Since Doves never attack and only retreat when attacked, the absence of attack in every Dove-Dove contest results in a very prolonged contest of display, so each Dove accrues -10 points. The eventual winner, however, gets +40 points for acquiring the resource. Since each Dove has a 50 percent chance of winning, the average payoff for a Dove in a population of Doves is thus +10 (+40 times 50 percent, plus -10 for wasting time and energy). But suppose that a mutant Hawk arises in this population. This Hawk will win every contest, so it will enjoy an average payoff of +40 compared to the Dove average of +10. Consequently, Hawks will begin to increase in fre-

quency in subsequent generations in the population. So Dove can never evolve to and remain at fixation.

But neither can Hawk. In a population of Hawks, nature is truly red in tooth and claw, for a pair of competing Hawks will attack one another aggressively, and the contest will end only when one of them is injured. The winner scores +40 points for acquiring the resource, but the loser accrues -60 points for injury. Since each Hawk has a 50 percent chance of winning (hence of losing) the contest, the average payoff for a Hawk in a population of Hawks is -10 fitness points (+40 times 50 percent, plus -60 times 50 percent). Now suppose a mutant Dove arises in this population. The Dove never wins, but it also never pays the cost of injury. So it averages 0 compared to the Hawks' -10. Consequently, Doves will begin to increase in frequency. So Hawk can't evolve to and remain at fixation either.

Interestingly, given the fitness costs and benefits assumed in this simple model, selection will favor an evolutionarily stable mix of 75 percent Hawks and 25 percent Doves, since Hawk and Dove have equal fitnesses when coexisting in this ratio. For, given this ratio of Doves to Hawks, 75 percent of one's contests are against Hawks and 25 percent are against Doves, so the average payoff to both Hawk and Dove is +2.5 fitness points. This proportion of Hawks to Doves is thus *evolutionarily stable* because any departure from it—caused by drift or mutation—will be corrected by selection, and the three-to-one ratio will be restored. This is because the fitnesses of both Hawk and Dove are dependent on the frequencies of those two phenotypes in the population. If the proportion of Hawks drops below 75 percent, then Hawks will enjoy proportionately more contests against Doves, so Hawks will have higher fitness than Doves; and, if the proportion of Hawks rises above 75 percent, then Hawks will have proportionately more potentially costly contests with one another, so Doves will have higher fitness than Hawks. Since selection will favor the phenotype with highest fitness, which phenotype selection favors thus depends on their frequencies in the population. This is *frequency-dependent selection*. And, in this example, frequency-dependent selection will maintain a three-to-one ratio of Hawks to Doves, because that is the ratio at which both Hawks and Doves enjoy equal fitness. (Of course, if we assumed different costs and benefits the stable ratio would be different.)

Note that the evolutionarily stable ratio results in no one's enjoying the greatest possible fitness. The average payoff to a Dove in a population of Doves is +10, but a population of Doves is not evolutionarily stable. Evolutionary stability, instead, turns out to be a three-to-one ratio of Hawks

to Doves in which the average payoff is +2.5. It would clearly be best to be a Dove in a population of Doves, but the best in this case can't evolve, since it isn't evolutionarily stable.

I've illustrated how selection can maintain a balanced ratio of alternative types with reference to behavioral types. But frequency-dependent selection can maintain an evolutionarily stable ratio of alternative phenotypes of any kind. For example, variation in eye color, variation in size, variation in the age at first reproduction, variation in number of offspring produced, and the sex ratio in a population could all be maintained by frequency-dependent selection. It doesn't act only on behavior.

A consistent result of mathematical models of frequency-dependent selection is that balanced proportions of alternative phenotypes, rather than just single phenotypes, turn out to be evolutionarily stable. But such balanced proportions of phenotypes can be achieved in two very different ways. To see this, consider the balanced proportion of Hawks and Doves described above. What is essential to the evolutionarily stable three-to-one ratio of Hawks to Doves is having 75 percent of one's contests against Hawk and 25 percent against Dove. One way this can be achieved is in a *mixed population*—that is, a population in which 75 percent of individuals are dedicated Hawks and 25 percent are dedicated Doves. But it can also be achieved in a population of individuals who are identical in playing a *mixed strategy*—that is, a population of individuals who randomly play Hawk 75 percent of the time and Dove 25 percent of the time.

When a mixed population is evolutionarily stable it is called a *stable polymorphism*. In a stable polymorphism, there is a genotype for each alternative phenotype, and selection maintains a stable ratio of the alternative phenotypes by maintaining a stable ratio of the alternative genetic types in the population. Thus, a genetic polymorphism—a locus at which different genotypes occur—is essential to each stable polymorphism. When a population consists of individuals playing a mixed strategy, on the other hand, the individuals in the population are genetically *monomorphic*—they share the same genotype—for that strategy. This genotype produces some mechanism that is capable of randomly generating the alternative phenotypes, and selection just sets the frequencies at which the alternative phenotypic forms are randomly generated. In both stable polymorphisms and mixed strategies, however, the alternative phenotypes have equal fitness.

Mathematical models of frequency-dependent selection are typically neutral with respect to whether an evolutionarily stable ratio of alternative phenotypes is a polymorphism or a mixed strategy. But, in nature,

mixed strategies are probably rarer than stable polymorphisms (although there are a couple of documented examples of mixed strategies, the most well known of which is the determination of the sex of offspring). The best evidence for this claim is the simple paucity of documented cases of mixed strategies compared to the number of documented cases of stable polymorphisms. This comparative paucity is evident in the numerous studies of alternative within-sex reproductive behaviors (that is, alternative behavioral tactics for attracting mates and securing copulations). Although there is intrasexual variation in reproductive behaviors in most major taxa, there is not a single documented case of a mixed reproductive strategy. In contrast, there are a number of documented cases of stable polymorphisms of alternative within-sex reproductive behaviors.

One particularly well documented case derives from the work of the biologist Stephen Shuster on *Paracerceis sculpta*, a marine isopod crustacean. Males of this species come in small, medium, and large, and these sizes perfectly correlate with distinct mating behaviors. Large males secure and “guard” harems of females in the recesses of sponges, acquiring their copulations with the females in the harem. Small males are unable to compete with large males for the acquisition of a harem, so they acquire copulations by “sneaking” past inattentive large males and thereby gaining access to the females in the harem. Medium males morphologically resemble females, so they “mimic” the female courtship display to a large male; thinking he is acquiring another female for his harem, the large male allows the medium male to enter the harem, where the medium male then copulates with some of the females inside. These three mating strategies have equal reproductive success, and the genes underlying them have been identified. So this is a clear case of a stable polymorphism. Similar polymorphisms have been found in the swordtail, *Xiphiphorus nigrensis*, the field cricket, *Gryllus integer*, a tree lizard, *Urosaurus ornatus*, and the ruff, *Philomachus pugnax*.

A second way in which selection can maintain phenotypic variation is through *adaptive plasticity* (sometimes called a *conditional strategy*). Adaptive plasticity is the capacity of a single genotype to produce more than one phenotype—more than one anatomical form, physiological state, or behavior—in response to environmental conditions. Like a mixed strategy, then, adaptive plasticity involves a single genotype that produces multiple phenotypes. But adaptive plasticity differs from a mixed strategy in two very important ways. First, the alternative phenotypes of a mixed strategy are produced *randomly*; they are not produced in response to environmental conditions. That is, in a mixed strategy a particular phenotype gets

produced just because “its number has come up,” not because that phenotype is especially suited to the particular environmental conditions in which the organism happens to find itself. In contrast, the alternative phenotypes produced by adaptive plasticity are generated nonrandomly, in response to the conditions that obtain in the organism’s environment; the phenotypes are produced to match the environmental conditions. Second, the alternative phenotypes of a mixed strategy have equal fitness at their evolutionarily stable ratio. In contrast, the alternative phenotypes produced by adaptive plasticity need not have equal fitness; in fact, the alternative phenotypes can vary significantly in their fitnesses. It need only be the case that the fitness of each phenotype is greater, in the environment in which it occurs, than any of the alternative phenotypes would be in that same environment. This is compatible with one of the alternative phenotype’s having lower fitness than the others; it just means that, *in those circumstances*, all the other phenotypes would have even lower fitness.

In biology, the concept of adaptive plasticity is applied to a very wide range of phenomena. But I will discuss just two distinct forms of adaptive plasticity, *developmental plasticity* and *phenotypic plasticity*. These two forms of adaptive plasticity don’t exhaust the phenomena, but they are the forms that will be important in later discussions. Before elaborating this distinction, however, I should issue a caveat. I will not be using the terms *adaptive plasticity*, *developmental plasticity*, and *phenotypic plasticity* in a way that conforms with standard usage in biology. The reason is that there simply is no standard usage of these terms in biology. Indeed, discussions of plasticity in the biological literature are characterized by widespread terminological inconsistency. Some biologists use all three of the above terms interchangeably, while other biologists distinguish developmental plasticity from general phenotypic plasticity. In what follows, then, I will be defining these terms so as to serve my purposes. And, for my purposes, I will treat developmental plasticity and phenotypic plasticity as distinct forms of the more general phenomenon of adaptive plasticity.

To illustrate developmental plasticity, consider the caterpillars of the moth *Nemoria arizonaria*, the larvae of which develop in oak trees. Caterpillars hatched in spring feed on the staminate flowers of the oak and develop to strongly resemble those flowers. Caterpillars hatched in summer feed on the leaves of the oak and develop to strongly resemble twigs on the oak. A difference in diet, due to a difference in chemical composition of the flowers and the leaves, is responsible for the development of the very different “flower” and “twig” phenotypes. And each phenotype is adaptive in its circumstances, since each serves the function of camou-

flaging the caterpillars and thereby protecting them against predation. Overall, however, the “flower” phenotype has the highest fitness, so it is definitely better to be a “flower” than a “twig” in the spring. But, since developing the “flower” phenotype in the summer would be maladaptive (because it would be conspicuous to predators in an oak without staminate flowers), it is more beneficial to be a “twig” than a “flower” in the summer, even though being a “twig” is suboptimal overall. As Richard Dawkins colorfully puts it, “twigs” are simply “making the best of a bad job.” Consequently, a single genotype has evolved in this species that is capable of producing both phenotypes, and it does so by selectively matching the phenotype to the environment in response to chemical cues in the caterpillars’ diets during development.

Similarly, bryozoans, or “sea moss,” are sometimes preyed upon by sea slugs. Sea slugs are detectable by a chemical cue that is present in the water around them. So, when bryozoans develop in the presence of this chemical cue, they grow spines that deter predation by sea slugs. In the absence of this chemical cue, they do not grow spines. However, since the growth of the spines is developmentally very costly (detracting from the allocation of resources to other aspects of bryozoan life history), nonspiny bryozoans have higher overall fitness. But, it would be clearly maladaptive not to grow spines in an environment populated by predatory sea slugs. Here again, a single genotype has evolved that is capable of producing two different phenotypes, each of which has higher fitness *in its circumstances* than the other, but one of which has the highest overall fitness. A similar developmental plasticity is present in aphids. If aphids develop in very crowded populations, which are likely to run out of food, they grow wings that enable migration. If they develop in uncrowded populations, they don’t grow wings.

For developmental plasticity to evolve by selection, several conditions must be met. First, there must be variation in a population’s environment. That is, there must be at least two different environmental conditions that affect fitness and that are consistently encountered across many generations of the population. Second, this environmental variation must be predictable. If a population’s environment varied, but in unpredictable ways, no hereditary mechanism could evolve to “match” offspring phenotypes to their environmental conditions, because those conditions would likely not have been encountered by ancestral generations. Third, a mix of alternative phenotypes, each occurring in its own environmental conditions, has to have a higher average fitness than any single phenotype would have across the range of variable environmental conditions. If some particular

phenotype had the highest fitness in each of the different environmental conditions, then that single phenotype would be selected. Finally, there must be “cues” in each of the different developmental conditions that are reliable predictors of the selection pressures to be encountered in the environment and to which some mechanism of adaptive plasticity can respond. If there were no waterborne chemical cue correlated with the presence of sea slugs, for example, bryozoans would not “know” when it is appropriate to grow spines and when not. The presence of sea slugs would then be unpredictable, and no developmental mechanism could evolve to *selectively* grow spines in response to the presence of sea slugs.

If all these conditions are satisfied, a genotype that is capable of producing alternative phenotypes that match alternative environmental conditions can be favored by selection over competing genotypes. In such a case, selection could drive that genotype to fixation. But, given the developmental plasticity of that genotype, selection for it would actually maintain variation at the phenotypic level as a function of the environmental variation encountered by a population. And the phenotypic variation maintained by selection in this way would be adaptive.

In cases of developmental plasticity, then, a genotype can produce two or more phenotypes, and the genotype is responsive to particular environmental cues during development in “deciding” which of those phenotypes to produce. In cases of *phenotypic plasticity*, in contrast, the genotype produces a phenotype that is capable of phenotypic change or reorganization in response to changing conditions in the organism’s environment. An example of phenotypic plasticity in the Hawk-Dove game would be an organism that played Hawk against all smaller opponents, but played Dove against all opponents of equal or greater size. In such a case, the organism’s behavioral phenotype would vary flexibly in response to fluctuating environmental demands.

A nontheoretical example of phenotypic plasticity is provided by the African cichlid fish, *Haplochromis burtoni*. There are two sexual phenotypes among *H. burtoni* males: territorial and nonterritorial. Territorial males are brightly colored, maintain and defend visually isolated territories, have mature testes, are reproductively active, and allocate all of their energy to defending their territories and reproducing. Nonterritorial males are cryptically colored, swim in schools with females, do not have mature testes, are not reproductively active, and allocate all of their energy to somatic growth. Territorial males accrue the direct fitness benefits of reproduction, which nonterritorial males do not accrue. However, because of their bright coloration, territorial males suffer far higher rates of predation than do

nonterritorial males, so they also incur greater fitness costs than nonterritorial males. Conspicuous coloration is worth its high costs only if a male is actively reproducing.

The territories defended by territorial males typically lie within the recesses of vegetation or behind leaves. In a natural environment, such territories prove highly unstable. Leaves move, territories are exposed, and other areas become visually isolated, hence suitable candidates for defensible territory. When these changes occur, nonterritorial and displaced territorial males compete for new territories. If a nonterritorial male captures a territory, within days it becomes brightly colored and develops mature testes. If a displaced territorial male fails to secure a new territory, within days it loses its bright coloration, becoming cryptically colored, and its testes begin to atrophy. As a displaced territorial male makes the transition to nonterritorial male, it begins to once again allocate all of its energy to somatic growth in preparation for later competition for new territories. If their habitats fluctuate greatly, male *H. burtoni* can cycle several times in this way through the territorial and nonterritorial phenotypes.

In order for such phenotypic plasticity to evolve by selection, there must be variation in some aspect of a population's environment that is relevant to fitness, just as with the evolution of developmental plasticity. But, in order for phenotypic plasticity rather than developmental plasticity to evolve in response to environmental variation, the environmental variation must occur relatively rapidly and unpredictably. That is, a population must face several different environmental conditions within the course of a single generation, there must be no reliable pattern in the order in which those different environmental conditions are encountered, and each different environmental condition must be of uncertain duration. In short, there must be *fluctuation* in some aspect of a population's environment that is relevant to fitness. Finally, the ability to vary phenotype in response to these fluctuating conditions must have higher average fitness than any single phenotype would have across all of the conditions. If it weren't better to revert to the nonterritorial phenotype when not holding a territory, for example, *H. burtoni* males would be brightly colored with mature testes throughout adulthood.

The following illustration will help to make the distinction between developmental plasticity and phenotypic plasticity less abstract and more intuitive. The phenomena to which I am applying the label *developmental plasticity* tend to conform to the following model: Organisms can encounter one of two types of environment, either an environment characterized by round holes or one characterized by square holes. A genotype

then evolves to build round pegs in the round-hole environment and square pegs in the square-hole environment. The phenomena to which I am applying the label *phenotypic plasticity*, however, tend to conform to the following, rather different model: Organisms encounter both round holes and square holes in their environment, in random order, and in sometimes rapid succession. A genotype then evolves to build Silly Putty, which can take the shape of a round peg or square peg as needed. In the case of developmental plasticity the genotype exhibits a flexible response to different environmental conditions, whereas in the case of phenotypic plasticity the phenotype itself exhibits the flexible ability to remake itself in response to fluctuating environmental demands. But both types of plasticity result in alternative phenotypes that are uniquely adapted to their circumstances.

In conclusion, then, selection can maintain phenotypic variation in a population by maintaining genetic variation in the population (a stable polymorphism) or by maintaining a genotype that adaptively produces alternative phenotypes (through a mixed strategy or adaptive plasticity). But, although I have explained each of them separately, you shouldn't infer that these different mechanisms of phenotypic variation are mutually exclusive. Indeed, it is possible to have a stable polymorphism of different mixed strategies, or a stable polymorphism of two "pure" phenotypes and one mixed strategy, or a stable polymorphism of two adaptively plastic genotypes, and so on. Life can be exceedingly complex. But we have seen enough of it to be able to move on.