Chapter 3

POPULATIONS AND SPECIES

The previous chapter emphasized the description and measurement of single individuals within a species. But no two individuals are exactly alike, and the variation between them is a central fact of biology. In this chapter, we consider the sources of variation, why it is important from an evolutionary perspective, and some of the special problems facing paleontology. Variation among individuals must be considered in light of the nature of populations, to which we turn now.

3.1 POPULATIONS IN BIOLOGY AND PALEONTOLOGY

The individual organisms whose morphology we treated in the previous chapter exist in the biological context of the **population,** which can be defined as a group of individuals of the same species that live close enough together that they have ample opportunity for interbreeding. This emphasis on breeding applies, of course, to sexually reproducing species. The population shares a single **gene pool.** The gene pools of adjacent populations of a species may be partially or completely isolated from one another. When two populations interbreed, there is said to be **gene flow** between them. Depending on the dispersal ability and behavior of organisms and on the fragmentation of the physical habitat, populations differ widely in how large a geographic area they occupy and in how isolated they are from neighboring populations.

The geographic structure of populations is commonly studied in the context of metapopulation theory. The larger population, or **metapopulation,** consists of a number of smaller subpopulations. Subpopulations that are in particularly favorable environments may produce many individuals that disperse to other areas; they are said to be *sources*. Other subpopulations, by contrast, may accumulate migrants; they are *sinks*. The factors that govern the dynamics of sources and sinks within metapopulations, as well as other fine-scale aspects of geographic structure, are quite important to ecologists. The spatial structure of populations is also important for paleontological questions, because it plays a role in the origin of new species and in the pattern of evolutionary change over time [SEE SECTIONS 3.3 AND 9.3].

Variation among Individuals within Populations

Each individual within a population has a particular **genotype**—its genetic composition encoded in its DNA sequences—and a **phenotype**—its form, structure, physiology, biochemistry, and behavior. The ultimate sources of variation within a population are **genetic mutation** and **recombination** of existing genetic material into new genotypes, through the production of sex cells and through sexual reproduction.

Importance of Variation Variation is not only a fundamental property of populations. It also underlies all

evolutionary change. **Evolution** within populations generally occurs if two simple conditions are met:

1. There is a regular relationship between genotypic and phenotypic variation, so that the phenotypic variation is heritable from parents to offspring. It is common to think of heritability in a direct sense—for example, a genetically determined trait such as eye color being inherited from the mother and father. In practice, however, heritability is studied by statistical analysis of populations, as we will see later in this section.
2. There is a relationship between heritable phenotypic variation and variation in reproductive success, reflecting both survival and fecundity. For example, suppose that bill size is heritable and that birds with larger bills tend to leave more offspring because they are able to eat larger and more nutritious seeds and therefore invest more energy into reproduction. Then the mean bill size of the population in this hypothetical case would tend to increase over time unless it were offset by other factors.

We just illustrated these simple requirements for evolution using a hypothetical case in which there was a direct, cause-and-effect relationship between the phenotype and reproductive success. For morphological traits such as body size and shape, as well as for many other traits, direct effects of this kind are generally accepted as the prevailing reason for evolutionary change. Such relationships are examples of **natural selection.** Traits can also evolve without being directly selected if they are genetically correlated with other traits that are under selection. For example, extra digits in vertebrates occur more commonly in larger individuals. Selection for increased body size could therefore lead indirectly to an increase in the number of digits.

An alternative to natural selection is that the correlation between phenotype and reproductive success is a matter of chance. This is likely in only two situations:

1. The traits are truly neutral with respect to selection, meaning that individuals are equally well adapted regardless of the trait value that they have. True neutrality is thought to be uncommon except for certain cases involving alternative forms of proteins and other biomolecules.
2. The population is so small that chance fluctuations in reproductive success are not averaged out. For example, even if it is true on average that larger-billed birds feed more effectively and leave more offspring, the occasional smaller-billed bird will be lucky, finding a cache of seeds, for instance.

In small populations, chance events can be of significance. Evolutionary change that results from such chance fluctuations is known as **genetic drift.** In a broader sense, chance fluctuations can also occur in other ways, such as by extinction of local populations that may differ in genetic composition relative to the larger metapopulation.

Heritability of Variation One often reads debates in the press on the subject of "nature versus nurture"—whether particular traits, such as aspects of human behavior, are genetic or environmental in origin. In fact, the entire phenotype arises through ontogeny from the interplay between an organism's genetic composition and its environment. For example, it is well known that growth in oysters and other animals that live on hard surfaces molds the organism to the substrate. There is clearly an environmental effect, yet the capacity to grow in such a malleable way has a genetic basis.

Many factors, including light, temperature, nutrition, water and soil chemistry, and substrate, can yield environmental variation in the phenotype. Phenotypic variation that is attributable to environmental variation is referred to as **ecophenotypic;** the tendency for a genotype to produce different phenotypes in different environments is known as **phenotypic plasticity.** The environmental effect on the phenotype may be adaptive. Many such cases of adaptive plasticity have been documented, including animals that detect the presence of predators via chemical cues and grow protective ornament in response (Figure 3.1).

Because each individual's form is both genetically and environmentally determined, evolutionary biologists study the sources of phenotypic variation among individuals rather than the phenotype of a particular individual. Figure 3.2 shows an analysis of bill size in a population of finches on an island in the Galápagos Archipelago. Each point compares the mean bill size of offspring produced by a pair of parents with the mean bill size of the parents. The open and closed symbols represent measurements that were taken in two different years. The positive relationship between parental size and offspring size, shown by the lines, indicates that the trait has a heritable component. The less scatter there is around this line, the higher the heritability.

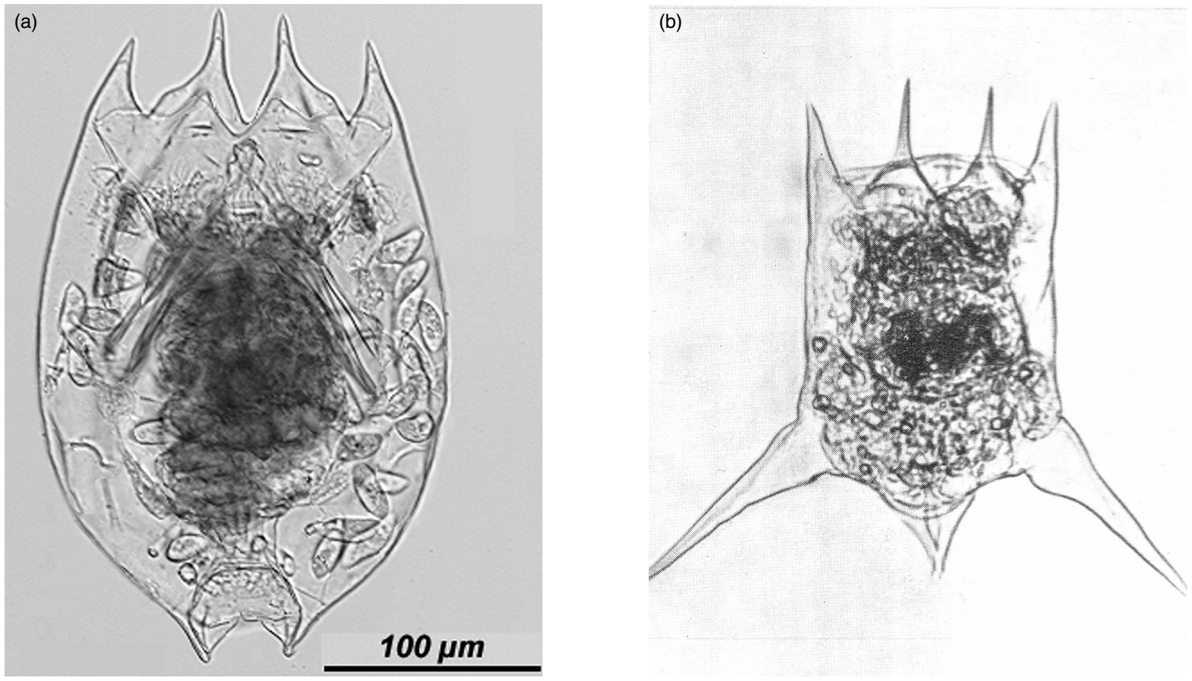

FIGURE 3.1 Two specimens of the living rotifer *Brachionus calyciflorus*. Individuals that grow in the presence of a predatory rotifer *Asplanchna*, or in chemical extracts derived from *Asplanchna,* develop elongate spines, as in the specimen on the right. The specimen in part (b) is approximately 150 microns across, excluding the spines. *(a: The Academy of Natural Sciences; b: From Gilbert, 1966, http://www.schweizerbart.de)*

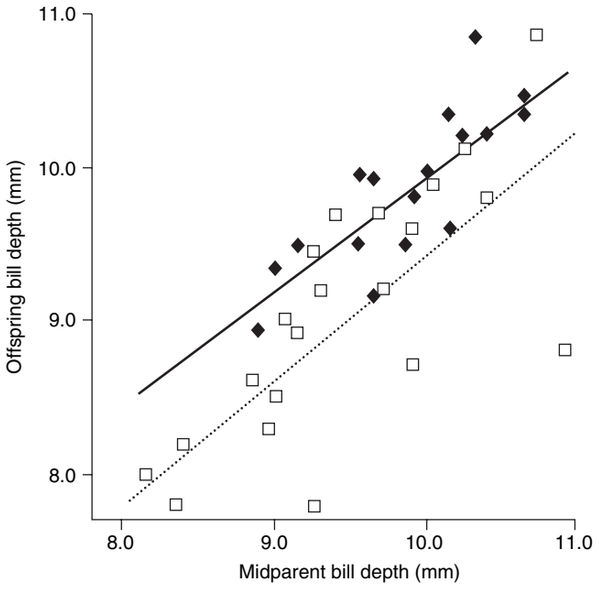

FIGURE 3.2 Heritability of bill size in the ground finch *Geospiza fortis*. Midparent value is the mean measurement of a pair of parents; offspring value is the mean of the offspring of these parents. The positive correlation between the bill size of the offspring and that of the parents indicates heritability of this trait. Open and closed symbols denote measurements taken in two different years. Lines are fitted to each set of points. *(From Boag, 1983)*

This scatter can be measured with standard statistical methods which, in this case, indicate that about 60 percent of the variation in bill size among the offspring is heritable. This is a statistical statement about the population as a whole in the given environment. It does not tell us to what extent the bill size of any individual bird is attributable to its genotype and how much to its environment.

Thus, variation among individuals of a given ontogenetic stage has both genetic and environmental components. In the population at large, change through ontogeny and differences between the sexes also contribute to the overall variation of the population. We typically try to factor out these last two sources of variation by studying the same sex at a comparable ontogenetic stage. There are still other sources of variation that affect fossil populations.

Additional Sources of Variation in Fossil Populations

The populations of paleontology, consisting of individuals collected from a given locality, differ from living populations in some important ways. They have passed through various taphonomic filters such as postmortem distortion, and they may represent a time-averaged assemblage [SEE SECTION 1.2].

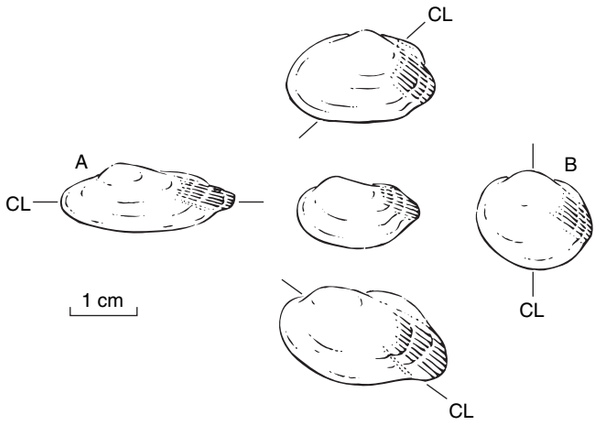

FIGURE 3.3 Effects of rock deformation on fossil morphology. The drawing in the center shows an undeformed specimen of *Arisaigia postornata*. The four drawings surrounding it show different patterns of deformation. In each case, the direction marked CL corresponds to rock cleavage and is perpendicular to the direction of maximum shortening. The specimens marked A and B were deformed in directions perpendicular to one another. *(From Bambach, 1973)*

One of the main taphonomic processes tending to increase apparent variation is distortion produced by compaction of sediments or deformation of sedimentary rocks. A particularly striking but by no means rare case is illustrated in Figure 3.3. Richard Bambach (1973) analyzed shape variation in a large sample of the infaunal bivalve *Arisaigia postornata* from Silurian rocks of Nova Scotia. For each specimen, the orientation of the rock's cleavage relative to the bivalve's morphology was noted. Because these specimens are actually two-dimensional molds of the original shells, the shapes taken by the fossils leave an accurate account of the deformational history of the rock, with the direction of maximal compression being perpendicular to the direction of cleavage.

Figure 3.3 shows drawings of four typical specimens covering the range of geometric relationships of cleavage direction to morphology. Despite the different appearance of these specimens, they can be assigned to the species in question because they possess characteristic surface ornament. The fifth drawing (center) is a reconstruction of an undeformed specimen. The reconstruction was aided by standard methods from structural geology. In essence, forms A and B are end-members that, based on orientations of rock cleavage, must have been deformed in perpendicular directions. The relative length:height ratios of A and B were used to estimate the relative degree of deformation in the two directions; this in turn was used to estimate what the original length:height ratio of undeformed specimens must have been. In Bambach's collections, we know that all the specimens were deformed because co-occurring brachiopods, which must originally have been bilaterally symmetrical, are also deformed. Thus, the form in the center of Figure 3.3 was not found.

Another potential source of added variation in fossil populations is the process of time averaging [SEE SECTION 1.2]. Figure 3.4 shows a hypothetical case in which the variation of a trait within a population is constant, as indicated by the width of the curve, while the average trait value changes over time. Typically, many successive populations will be averaged together into a single fossil sample. The resulting variation of the time-averaged sample depends on the amount of variation within the population at a moment in time and on how much the population's morphology shifts over time.

In principle, time averaging could act to the point where we could not obtain reasonable estimates of

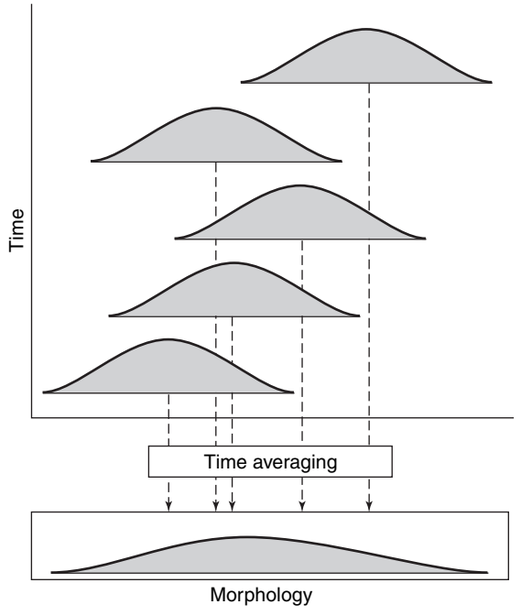

FIGURE 3.4 Hypothetical effects of time averaging on variation within fossil populations. Each curve in the upper part of the figure depicts variation within a population at a point in time, and the population average is shown by the position of the curve along the *x* axis. If the population shifts over time, the resulting time-averaged sample, depicted by the bottom curve, will be more variable than the population. *(From Hunt, 2004a)*

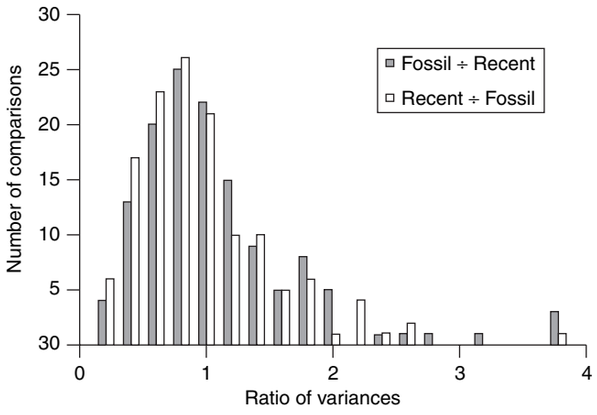

FIGURE 3.5 Comparison of variance between living and fossil populations of the same species. The solid bars show the frequency distribution of about 130 ratios between the variance of a fossil sample and the variance of a living population of the same species. The open bars depict the ratio of variances between living populations and fossil samples. The two distributions are indistinguishable, implying that these fossil samples are, in general, neither more nor less variable than their living counterparts. *(Data from Hunt, 2004b)*

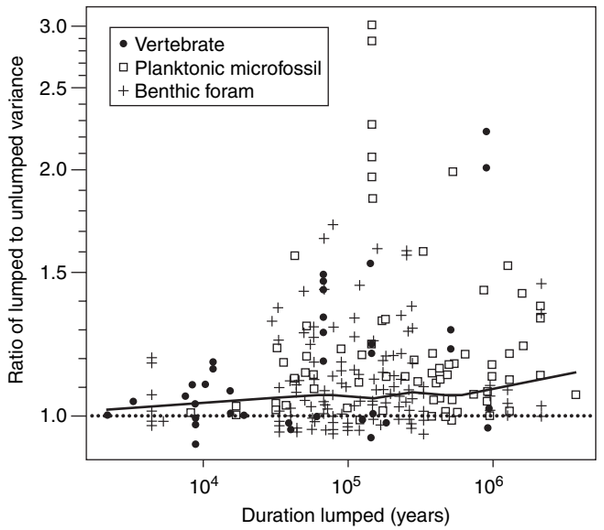

FIGURE 3.6 Inflation of population variance relative to the extent of analytical time averaging. Successive samples were lumped together, and the ratio of lumped to unlumped variance is plotted against the temporal extent of lumping. The thick line shows the median ratio. On average, there is less than a 10 percent inflation of variance due to time averaging. *(From Hunt, 2004b)*

variation within fossil populations. But how much does time averaging matter in practice? This can be assessed in two ways: (1) by comparing the variation of living populations with that of fossil samples of the same species; and (2) by comparing the variation of fossil samples with the duration of time averaging of those samples.

Figure 3.5 depicts comparisons among living populations of a number of species and fossil samples of the same species. Variation is expressed by the statistical measure known as the *variance* (see Box 3.1). For each comparision, the ratio between the variance in the fossil sample and that in the living population was calculated. These ratios are summarized by the solid bars in Figure 3.5. A ratio greater than 1 means that the fossil population has a greater variance than does the living one. Clearly, some of the fossil samples are more variable than their living counterparts and some are less so. The open bars in Figure 3.5 show the ratio of variance of living to fossil populations. Here, ratios greater than 1 mean that living populations have greater variance. There is no appreciable difference between the two sets of ratios. In other words, fossil samples are sometimes more variable than corresponding living populations and sometimes less so, but there is no predominant tendency one way or the other.

Paleontologists have no control over the amount of time represented by naturally occurring beds. The extent of time averaging can be varied artificially, however, by combining fewer or more beds together into a single sample. This practice, which results **analytical time averaging,** allows one to explore how variance changes as more time is incorporated into a sample. Figure 3.6 summarizes data on variance from a number of fossil studies. Samples of the same species were analytically time averaged to determine how the variance of the combined samples is affected by time averaging. Each point in this figure compares the variance of a time-averaged sequence of populations with the duration of time averaging. The thick line shows the average trend through the points. Although there is an overall increase in variance with time averaging, it is generally rather small—on average, less than 10 percent even when millions of years are averaged together.

Thus, available evidence suggests that the variance within fossil samples is not dominated by time averaging. Variation within samples can be meaningfully studied, provided that gross distortion of the kind shown in Figure 3.3 is ruled out. That variance barely increases with time averaging in many cases implies that the population

is morphologically relatively stable over time. We will return to this point in Chapter 7.

3.2 DESCRIBING VARIATION

In this section, we provide a brief treatment of some of the most important procedures for describing and analyzing variation. Our coverage is only introductory, and the sources listed at the end of the chapter should be consulted for additional details. The availability of high-speed computers and software makes it easy to perform a wide range of analyses, but it is essential to have a firm understanding of the goals, assumptions, and calculations underlying each analysis. We focus on variation among individuals within a population and variation among similar populations. Because of the importance of populations in evolution, these levels of analysis play a special role. Nevertheless, many of the same procedures we will describe can also be applied, often with only minor modification, to study other aspects of variation—within the growth of a single individual or among the distantly related species of a larger biologic group.

Box 3.1

DESCRIPTIVE STATISTICS

A histogram shows the number or proportion of individuals having trait values falling within specified intervals (Figure 3.7); the histogram may be smoothed into an idealized frequency curve. The graphical summary of the histogram is often accompanied by other statistics, as outlined herein.

Paleontologists and biologists are generally interested in two main aspects of univariate data within a population: the central tendency and the dispersion or variation. Which statistics are appropriate to express central tendency and dispersion depends on the nature of the variables. There are three main kinds of biologic variables: (1) **nominal** or **categorical;** (2) **ordinal** or **ranked;** and (3) **quantitative.**

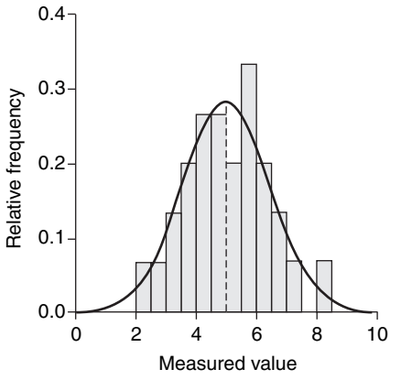

FIGURE 3.7 Hypothetical histogram and idealized, smooth frequency curve. Each bar shows the proportion of individuals having a trait value within the corresponding interval of values on the *x* axis. The dashed line shows the position of the sample mean.

Nominal data can take on only particular, distinct values, and there is no natural ordering to the values; one value is not inherently greater or smaller than another. Examples include gender (male or female); presence or absence of a specified structure; and features such as surface ornament (none, spines, tubercles, and so on).

Ordinal data also take on only distinct values, but there is a natural ordering to them. Examples include small, medium, and large; absent, rare, common, and abundant; and compressed, equidimensional, and elongate. There is often an unmeasured continuum underlying the values. We can express size as small or large even though size can take on any number of values if measured more precisely. The differences between values on an ordinal scale generally have no consistent meaning. The difference between small and large, for example, is not twice as great as the difference between small and medium.

Examples of quantitative measures include length, width, area, volume, mass, and angle. The units on the scale have a consistent meaning. For example, the difference between a length of 10.0 and 10.2 mm is twice as great as the difference between 10.0 and 10.1 mm. Quantitative measures are the most common form of data in the study of fossil

Describing Variation in One Dimension

Basic descriptive statistics are an important part of formal taxonomic work (Box 3.1). Summaries of measurements are typically univariate, involving only a single variable at a time. A graphical and tabular summary of measurements allows other workers to compare the data with similar measurements for other species [SEE SECTION 4.1]. These data are important, for example, in assessing whether two samples of specimens are likely to belong to the same species or to different species. Whether data are summarized in graphical or tabular form, it is important to make available to other workers the original measurements on which these summaries are based.

Describing Variation in Two or More Dimensions

Most univariate analyses, such as the assessment of differences between two samples of specimens (Box 3.1), populations. We therefore focus on such measures. Meristic counts [SEE SECTION 2.2] are, strictly speaking, ordinal. For many purposes, however, they are treated as quantitative, especially if the counts vary over a wide range of values. The statistics of nominal and ordinal data are discussed in the sources listed at the end of this chapter.

The **arithmetic mean** or **average** can be used to measure central tendency for quantitative data. If the measured values are denoted x, the mean of a sample, $\bar{x}$ (read "x bar"), is simply the sum of values divided by the number of individuals measured (n): $\bar{x} = \sum x/n$.

If the distribution of data is highly asymmetric or if n is small, the mean can be unduly influenced by a small number of high or low measurements. In such cases, the **median** often provides a more reliable estimate of central tendency. (The skewed distribution of wealth is one reason economic statistics commonly report median income and assets; the mean wealth of one billionaire and 1000 paupers would be roughly $1 million per person.) In the case of an asymmetric distribution, the median also better represents what we think of as the typical form. By definition, half of the observations are at or below the median, and half are at or above the median. For example, suppose we measure total length in centimeters in a small sample of seven specimens and obtain the following values, placed in increasing order: 12, 13, 13, 13, 14, 15, and 32. The median would be the fourth value, or 13 cm. This is a clear case in which one would want to report the median. The mean value of 16 cm is dominated by one large observation and is outside the range of the six remaining measurements. It does not adequately represent the typical size one is likely to encounter in the species from which this sample was drawn.

Dispersion for quantitative traits is typically measured by the average squared difference between observed values and the mean. The **variance** s^2 of a sample is defined as $\sum(x - \bar{x})^2/(n - 1)$. Variance has units that are the square of the original unit of measurement. To express dispersion in the same units as the original measurements, it is common to use the **standard deviation** s, equal to the positive square root of the variance. In the example of the previous paragraph, the variance is 51 cm^2 and the standard deviation is 7.1 cm.

One of the most common uses of univariate data is to interpret observed differences between two populations. Several samples, even if drawn from the very same population, will inevitably differ somewhat because of chance variation. Every sample statistic, such as the sample mean $\bar{x}$, has an associated **standard error,** which is a measure of the uncertainty in the statistic. If we took a very large number of samples from the same population and calculated $\bar{x}$ for each one, then the standard deviation of the values of $\bar{x}$ would be the standard error of the mean. The smaller the standard error relative

continued on next page

have bivariate and multivariate analogs. In addition, there are certain bivariate methods designed to analyze data when the focus is on the relationships between two measured variables. Such methods are useful in studying, for example, growth [SEE SECTION 2.3], function [SECTION 5.3], and heredity (Figure 3.2).

Two of the principal goals of bivariate analysis are to measure the strength of correlation between two variables and to describe the form of the relationship between them. For the first goal, a number of **correlation coefficients** are commonly employed (see Box 1.1). These typically vary between −1 and +1, with values closer to these extremes indicating stronger correlations. Negative values indicate that an increase in one variable tends to correspond with a decrease in the other, while positive values indicate that the two variables tend to increase together. Values near zero indicate that there is little relationship between the variables. Referring back to the example of bill size in Figure 3.2, the correlations between parent and offspring for the two sets of points

Box 3.1 (continued)

to the difference between two sample statistics, the smaller the probability that the observed difference reflects chance variation, and therefore the greater the chance that the two samples come from truly different underlying distributions. In general, the less variable the population and the more individuals in the sample, the smaller will be the standard error of a sample statistic. It is therefore necessary to compare any observed difference with the intrinsic variance of the population (Figure 3.8).

The samples of Figure 3.9 are quite different relative to their intrinsic variability. If the two histograms were superimposed on the same graph, they would barely overlap. Thus, it is unlikely that the difference between the samples is due to chance. The samples of Figure 3.10, by contrast, could easily have been drawn from the same statistical distribution. They have essentially the same mean and would fully overlap if plotted together. Whether a difference between samples is due to chance can be very unlikely (Figure 3.9), very likely (Figure 3.10), or anywhere in between—but one can never know with absolute certainty. Nonetheless, a low probability that the observed difference is due to sampling error provides reasonable, operational grounds for considering the difference to be meaningful unless proven otherwise. This probability is assessed with formal statistical tests described in several of the sources listed at the end of this chapter.

The pair of samples depicted in Figure 3.9 shows a large difference in mean values, whereas the pair in

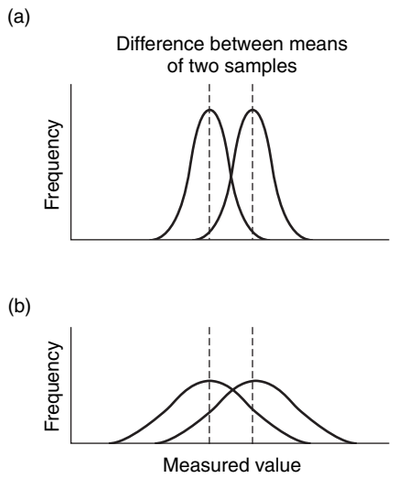

FIGURE 3.8 Pairs of hypothetical frequency distributions illustrating the importance of variation in assessing the significance of an observed difference between mean values. It is much more likely that the observed difference occurred by chance in part (b) than in part (a).

Figure 3.10 shows a small difference. In fact, each of these two comparisons involves the same pair of subspecies; the contrasting results reflect different traits that were analyzed. Because different traits can show different patterns of variation between populations, it may be necessary to measure and analyze several traits—hence, the need for multivariate analysis.

are between 0.75 and 0.80. These are relatively high values, in agreement with our earlier statement that there is clear evidence for heritability.

The form of the relationship between two variables expresses *how much* of a change in one is seen with respect to a change in the other. This is typically studied with a linear model of the form $Y = aX + b$. The slope a estimates how much of a change in Y there is for a given change in X. The intercept b is the value that the variable Y takes on when the variable X has a value of zero [SEE SECTION 2.3]. Referring again to Figure 3.2, the slopes of the lines fitted to the data are approximately 0.8. This means that for every millimeter difference in parental bill size, there is, on average, a difference of 0.8 mm between the corresponding offspring.

There are two main reasons to fit a line to data. The first is to describe a mutual relationship between two variables without giving primacy to one or the other. This use is common in the study of allometry, as in Figure 2.29. Second, a fitted line can be used for predictive

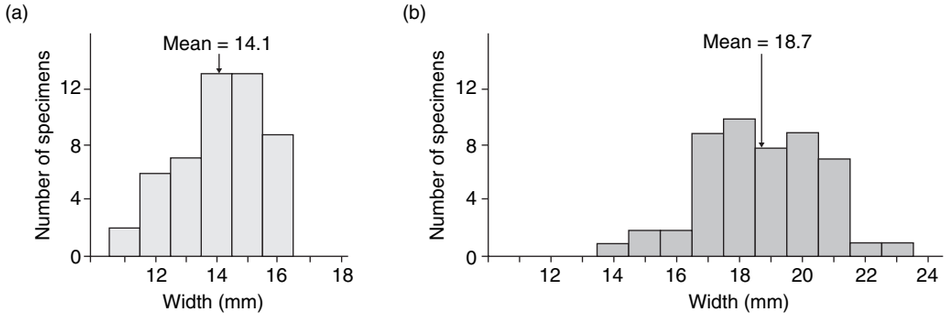

FIGURE 3.9 Frequency distributions of shell width in two subspecies of the Devonian brachiopod *Pholidostrophia*. (a) *Pholidostrophia gracilis nanus*. (b) *Pholidostrophia gracilis gracilis*. The mean values are different enough, relative to variation about the mean, that the difference is not likely to be due to chance in sampling alone. *(Data from Imbrie, 1956)*

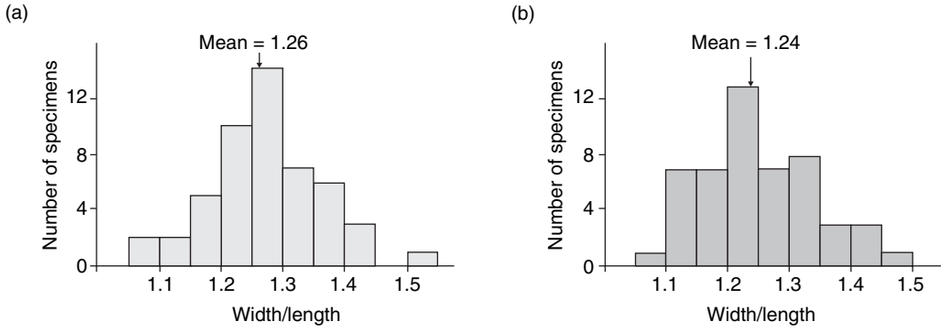

FIGURE 3.10 Frequency distributions of width-to-length ratio in two subspecies of the Devonian brachiopod *Pholidostrophia*. (a) *Pholidostrophia gracilis nanus*. (b) *Pholidostrophia gracilis gracilis*. The slight difference between the mean values could easily be due to chance errors in sampling. *(Data from Imbrie, 1956)*

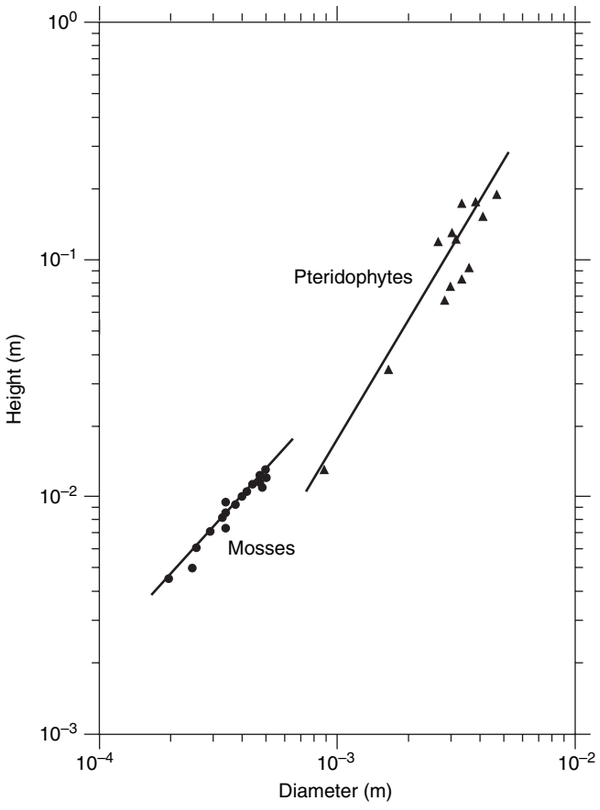

FIGURE 3.11 Predictive regression lines that can be used to estimate stem height from stem diameter in two groups of land plants. Both variables are measured on a logarithmic scale. Pteridophytes are an informal grouping of primitive vascular plants. *(From Niklas, 1994a)*

purposes. For example, Figure 3.11 shows the relationship between stem diameter and height in a number of species of living mosses and primitive vascular plants. The strength of the correlations means that stem diameter could reasonably be used to predict stem height—or vice versa. The prediction of stem height from stem diameter is actually much more useful in paleontology. Because material is often fragmentary, it is unlikely that the entire height of the stem will be preserved. Of course, this approach depends on having living or completely preserved representatives with which to establish the predictive relationship. Descriptive and predictive line fitting involve slightly different assumptions and procedures, which are covered in any elementary text on statistics.

It is important to bear in mind that our treatment of bivariate data with the equation $Y = aX + b$ assumes that the relationship between the variables is linear. Two traits may in fact be nonlinearly correlated, as with the brachiopods of Figure 2.23. In such a case, the correlation coefficient can greatly underestimate the strength of association between the traits, and a straight line fitted to the data is all but meaningless. A nonlinear relationship such as that of Figure 2.23 can sometimes be made linear by measuring the variables on a logarithmic scale. This, of course, is what the allometric equation [SEE SECTION 2.3] accomplishes. Other transformations can often be used to linearize the data.

Rarely is organic form sufficiently well represented by one or two features. It is often necessary to take measurements on many traits to gain a more complete picture of form. Doing so leads to problems, however, for the human mind cannot so easily visualize all the mutual relationships among numerous variables the way it can grasp a simple bivariate relationship. Therefore, a large class of approaches has been developed, collectively referred to as **multivariate analysis.** These approaches share the common goal of data reduction, in other words, summarizing, in a small number of dimensions, data that represent a large number of variables. The dimensions used are often synthetic in the sense that they are combinations of the original variables.

Any such reduction in dimensionality in effect represents a projection of the original data, just as a map is a projection of the globe into two dimensions. A projection generally produces distortion, and most methods have associated with them some means for assessing this distortion. Figure 3.12 shows hypothetical cases in which there are two original variables. Although there is variation in both dimensions, the strong correlations among the variables in Figure 3.12a imply that most of the variation can be summarized by the major axis running through the points from the lower left to the upper right. That is to say, if we were to treat this axis as a single, synthetic variable, and represent each point by a single number—its projected position along this axis—there would be little distortion and we would lose relatively little information. We may have measured two traits originally, but the number of meaningful variables is closer to one.

A contrasting case is shown in Figure 3.12b. Here, the variables are more weakly correlated, so there is more dispersion around the major axis. This means that the number of meaningful variables is much closer to two than to one, and we would lose a great deal of information by considering only the position of the points along

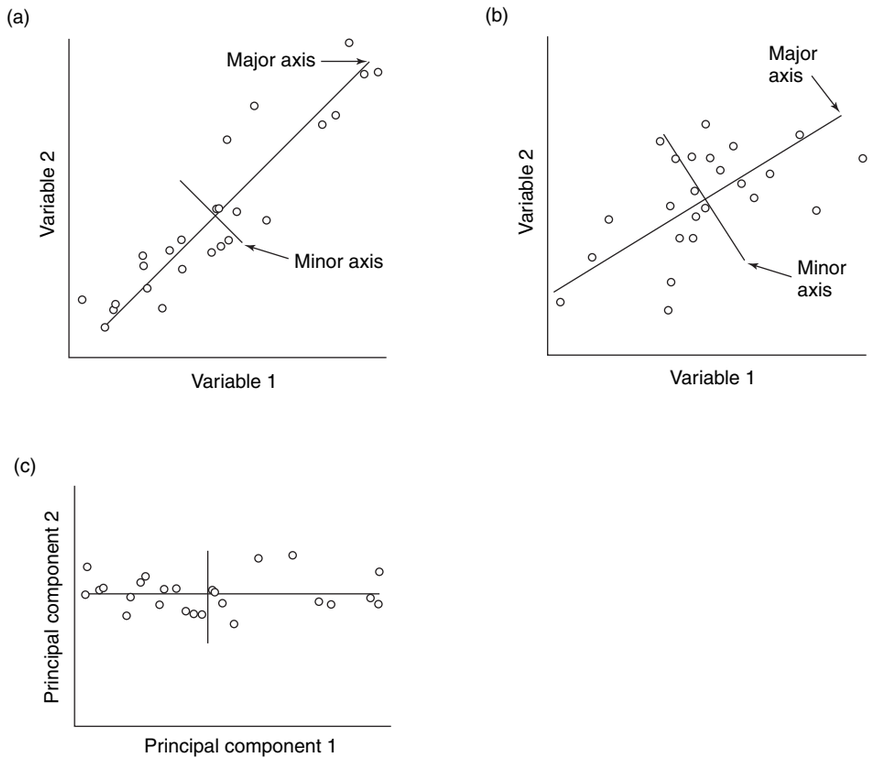

FIGURE 3.12 Illustration of the rationale behind multivariate analysis. (a) Here the two variables are highly correlated. In this case, projecting all points onto the major axis would result in little loss of information. (b) Here the variables are more weakly correlated, and therefore much information would be lost by reducing the two-dimensional data to a single dimension. (c) Shown here are the same data as in part (a). The major and minor axes are the first and second principal components, and the projections of the data onto these axes are the principal-component scores.

the major axis. Because of mutual intercorrelations among anatomical traits, biometric data are typically closer to the graph in Figure 3.12a. For example, if we measured the lengths of two limb bones in a sample of TETRAPOD vertebrates, we would find that the larger species or individuals tend to have greater lengths with respect to both measures.

In this section, we illustrate multivariate analysis with several different methods. The goal of data reduction pervades all of them, but each one focuses on a different kind of question. There are numerous other methods that are similar in spirit to the ones we present here, while differing in the particulars. Some of these will be covered later in this book.

Ordination of Specimens One of the main uses of multivariate analysis is to facilitate visual inspection of data. In a bivariate plot, it is easy to see which specimens are most similar, how specimens differ, how the data trend, and so on. To do the same with multivariate data requires an **ordination**—a representation of the positions of the specimens relative to one another. One of the most widely employed methods to achieve this goal is **principal-component analysis.** Figure 3.12c shows the same hypothetical data as Figure 3.12a. The points have simply been rotated so that the major and minor axes running through the data in Figure 3.12a are now in the same direction as the new the x and y axes of Figure 3.12c. The direction of the major axis is the direction of maximal dispersion in the data and defines the first principal component. There is still residual variation around this axis, indicated by the minor axis that is perpendicular to the first axis. This minor axis defines the second principal component.

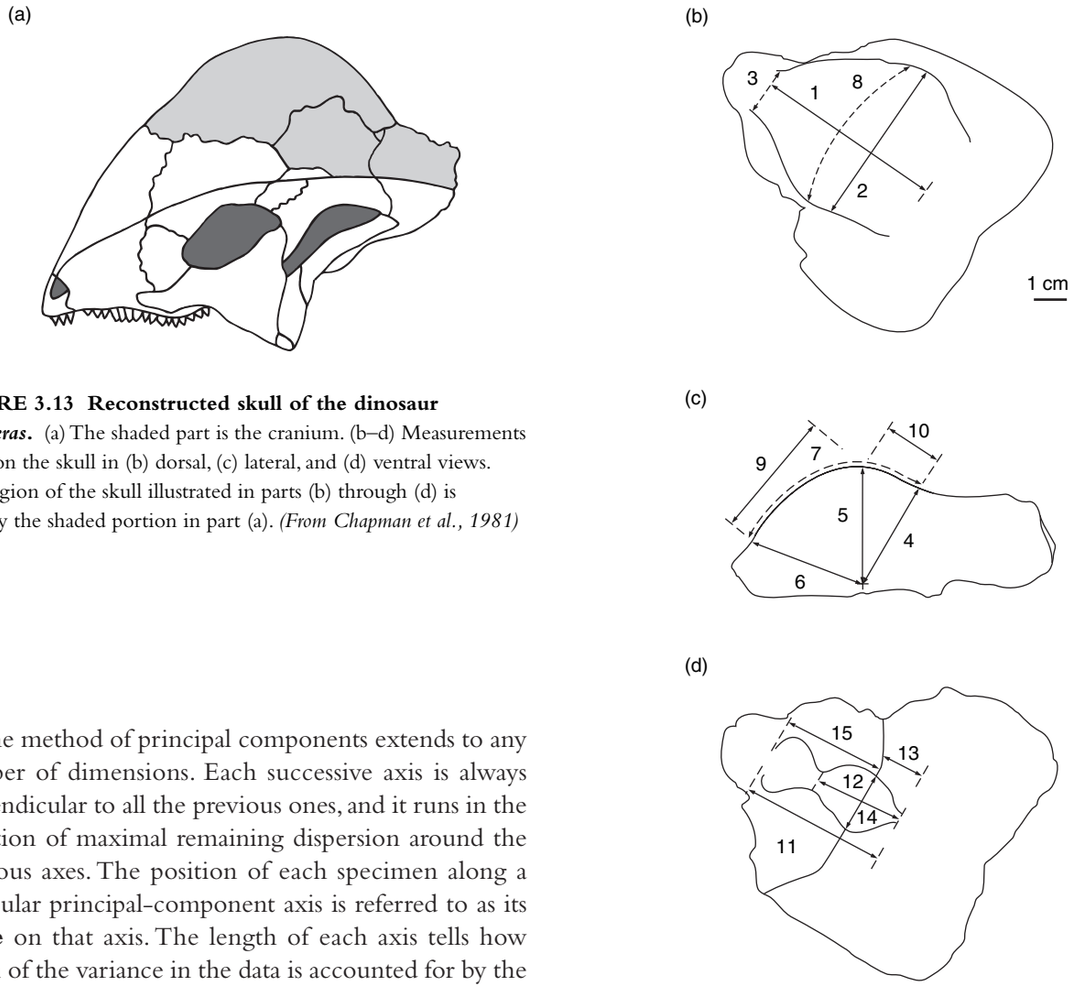

FIGURE 3.13 Reconstructed skull of the dinosaur *Stegoceras*. (a) The shaded part is the cranium. (b–d) Measurements taken on the skull in (b) dorsal, (c) lateral, and (d) ventral views. The region of the skull illustrated in parts (b) through (d) is roughly the shaded portion in part (a). *(From Chapman et al., 1981)*

The method of principal components extends to any number of dimensions. Each successive axis is always perpendicular to all the previous ones, and it runs in the direction of maximal remaining dispersion around the previous axes. The position of each specimen along a particular principal-component axis is referred to as its **score** on that axis. The length of each axis tells how much of the variance in the data is accounted for by the corresponding principal component; it is expressed by a number called the **eigenvalue** (see Table 3.1).

Let us consider a paleontological example of principal-component analysis. Figure 3.13a shows a reconstruction of the skull of the dinosaur *Stegoceras*. The shaded portion is the cranium, which includes the braincase and a prominent dome. A number of crania were measured on specimens from the Upper Cretaceous of western North America; the measurements are shown schematically in Figures 3.13b–d. There are ten measurements of the dome and five of the braincase. Figure 3.14 portrays the scores of about 30 specimens on the first two axes that result from a principal-component analysis. The specimens of the species *Stegoceras validus* appear to sort rather naturally into two groups, indicated by the closed and open circles.

Structural Relationships among Variables If our only goal were to ordinate specimens to determine how much they differ from one another or whether they seem to sort into different groups, the analysis shown in Figure 3.14 would be sufficient. If we want to understand the nature of the differences, however, then it is essential that we know something about how the original variables are combined to produce the synthetic principal components. Groups of variables that are mutually correlated will tend to be represented in similar ways in the synthetic variables. Thus, in a general way, principal-component analysis allows us to explore structural relationships among variables.

This aspect of principal-component analysis is illustrated for the *Stegoceras* data in Table 3.1, which lists the correlations between the original variables (arrayed in rows) and the principal components (arrayed in columns). Each entry in this table is termed a **loading.**

TABLE 3.1

Summary of Principal-Component Analysis of *Stegoceras* Skull Measurements

	Principal Components			
Variables	**1**	**2**	**3**	**4**
1. Dome length	0.960	−0.167	−0.050	−0.005
2. Dome width	0.954	−0.187	−0.098	0.049
3. Anterior dome width	0.918	−0.040	0.090	0.187
4. Posterior dome thickness	0.909	−0.214	−0.184	−0.087
5. Dome thickness	0.837	−0.351	0.167	−0.095
6. Anterior dome thickness	0.947	−0.086	−0.097	0.051
7. Dome length (on curvature)	0.945	−0.166	−0.071	−0.115
8. Dome width (on curvature)	0.946	−0.110	−0.138	0.028
9. Anterior dome length	0.916	−0.084	−0.045	0.034
10. Posterior dome length	0.918	−0.108	−0.143	−0.170
11. Braincase length	0.714	0.629	0.254	−0.060
12. Length to braincase constriction	0.595	0.695	0.248	−0.206
13. Posterior braincase length	0.220	0.824	−0.311	−0.363
14. Braincase width	0.277	0.636	−0.435	0.556
15. Anterior braincase length	0.685	0.238	0.597	0.270
Eigenvalue	10.0	2.32	0.91	0.66

SOURCE: Chapman et al. (1981)

NOTE: The table shows the loadings of variables (rows) on principal components (columns). Refer to Figure 3.13 for the definition of the variables. Each eigenvalue is equal to the sum of the squared loadings for the corresponding principal component. The larger the eigenvalue, the greater the proportion of information summarized by the principal component. The first two eigenvalues are much larger than the remaining ones, indicating that most of the variation in the data is summarized by the two principal components.

Relatively high loadings mean that the variable makes a substantial contribution to the principal component.

With biometric data of the kind represented here, it is common for most or all variables to have mutually high loadings on the first principal component. This component can then be interpreted, albeit only roughly, as a general measure of size. The second and higher principal components may have substantial loadings for just

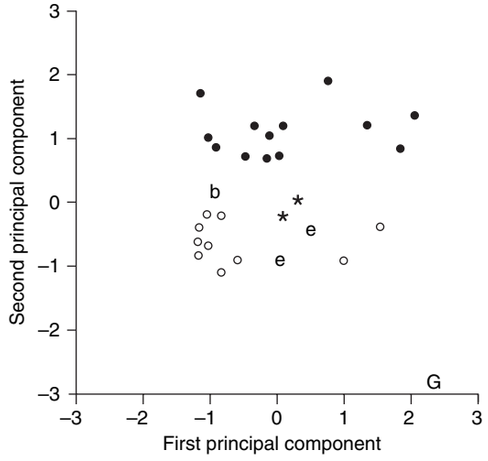

FIGURE 3.14 Principal-component scores of *Stegoceras* specimens and the related species *Gravolithus albertae*. Open and closed circles are specimens of *S. validus*. *S. browni* and *S. edmontonensis* are labeled *b* and *e*, the asterisks denote *Stegoceras* specimens of uncertain species affinity, and *Gravolithus* is labeled *G*. In contrast to the hypothetical case of Figure 3.12, this analysis has adopted a fairly common practice of standardizing the scores to have a mean of 0 and a variance of 1 on each principal component. This standardization is often done to portray each component as a biological factor of equal importance. It has little effect on the interpretation of results in this example. *(From Chapman et al., 1981)*

a few variables. The second principal component in this case has high loadings for braincase measurements only, suggesting that this component reflects the relative development of the braincase versus the dome.

The interpretation of the second principal component as a contrast between braincase and dome allows us to make some sense of the separation of individuals along this axis (Figure 3.14). Because the loadings for the dome characters are low and those for the braincase characters are high on this axis, a specimen with a high score will tend to have relatively low values of the dome variables and high values of the braincase variables. Thus, the upper group of specimens should have relatively larger braincases and relatively smaller domes.

This suggestion can be tested directly with bivariate analyses that compare dome and braincase measurements. A plot of dome length against braincase length (Figure 3.15) shows that the two sets of specimens do in fact differ in the relative development of the dome and braincase. A further question, one which

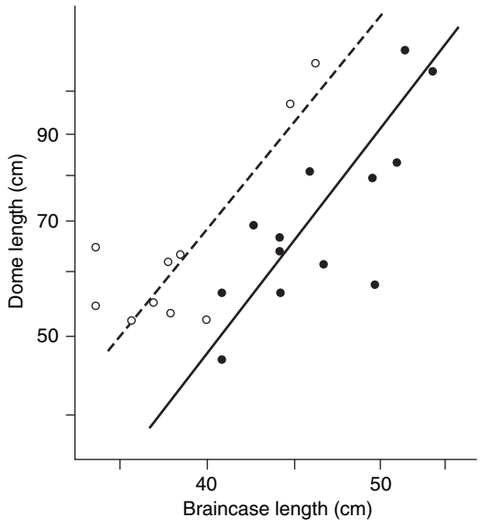

FIGURE 3.15 Bivariate comparison of braincase length and dome length in specimens of *Stegoceras validus*. Open and closed circles correspond to the two groups of specimens in Figure 3.14. Lines are fitted to each set of points. *(From Chapman et al., 1981)*

Box 3.2

CLUSTER ANALYSIS

Cluster analysis begins with a matrix of similarities or dissimilarities between specimens. In this example, we focus on dissimilarity, which can be measured in numerous ways; here, it is simply calculated as the straight-line distance between two specimens in the complete, multivariate space representing the measurement data. Thus, if the number of variables measured is m and if x_{ik} represents the value of variable k on specimen i, then the distance between any two specimens a and b is given by

$$d_{a,b} = \sqrt{\sum_{k=1}^{m} (x_{ak} - x_{bk})^2}$$

Table 3.2 shows the dissimilarities between a subset of the *Stegoceras* specimens. The **phenogram** or **dendrogram** (Figure 3.16) is constructed by finding those pairs of specimens that share mutually smallest dissimilarities; each is more similar to the other than to the rest of the specimens. In Table 3.2, these pairs are specimens 2 and 3, and specimens 6 and 7. Once these mutually most similar pairs have been found, the remaining specimens are joined with the existing clusters, and clusters are then joined together at several nested levels until all the clusters have been joined.

TABLE 3.2

Distances between a Subset of Specimens Used to Construct the Dendrogram of Figure 3.16

	1	2	3	4	5	6	7
1	—						
2	1.9	—					
3	2.2	1.1*	—				
4	2.7	2.2	2.2	—			
5	2.1	2.1	2.1	2.1	—		
6	1.7	1.8	1.6	1.3	1.0	—	
7	1.8	2.2	1.8	1.6	1.5	0.8*	—

NOTE: These specimens are identified as cluster D in the dendrogram. Mutually most similar pairs are indicated by an asterisk.

cannot be answered by multivariate analysis alone, is *why* individuals differ in the degree of braincase development. One explanation that has been offered is that this difference reflects sexual dimorphism (Chapman et al., 1981).

We have seen that principal-component analysis can provide an ordination of specimens in a reduced number of dimensions and can facilitate the study of relationships among variables. In the example of *Stegoceras*, the ordination revealed what seem to be two distinct groups, which could be understood, by studying loadings, in terms of the original variables. Finding groups that are not known in advance is indeed one of the other major uses of multivariate analysis, to which we now turn.

Classification of Specimens Paleontologists often start out studying a suite of specimens without knowing precisely how many natural groups are present. Determining the number of groups and the composition of each is the goal of **clustering** or **classification** techniques. One family of methods, collectively known as *cluster analysis*, is illustrated in Box 3.2, with the same *Stegoceras* measurements used in the principal-component analysis. The objective of cluster analysis is to summarize the morphological similarities and dissimilarities among specimens in the form of a **dendrogram.** This is a branching diagram that links similar specimens together into groups and separates them from other groups (Figure 3.16).

The *Stegoceras* specimens sort into about five clusters, labeled A through E in Figure 3.16. Comparing this dendrogram with the principal-component plot of Figure 3.14, we can see that specimens within each of the two groups identified on that plot tend to belong to the same clusters.

In contrast to the situation for which cluster analysis is used, we may want to determine whether two or more groups, designated in advance, differ appreciably in their measured traits. This is the problem of **discrimination,** as opposed to classification, and is discussed later in this

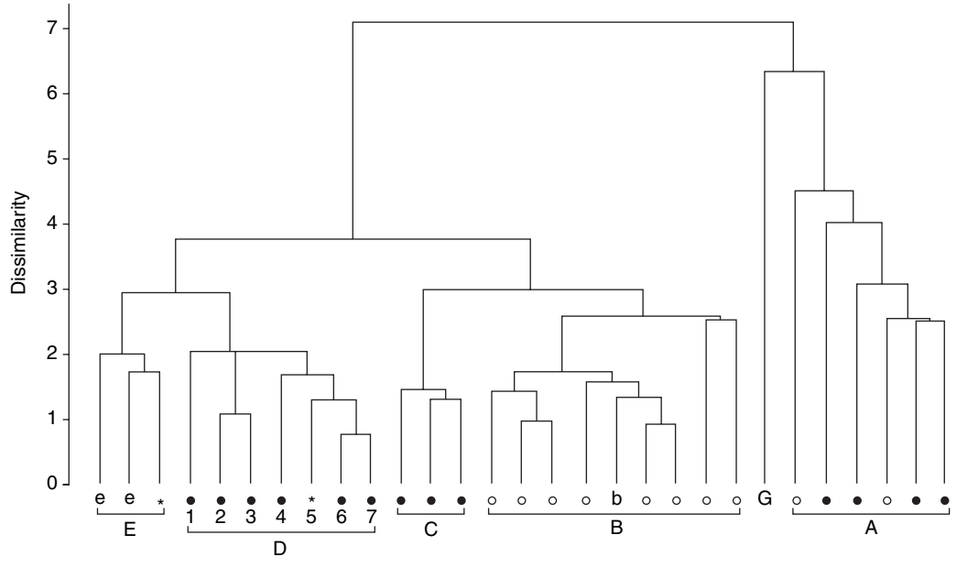

FIGURE 3.16 Dendrogram depicting results of cluster analysis of *Stegoceras* specimens. This analysis is intended to find groups based on overall morphological similarity. Symbols correspond to those in Figure 3.14. The numbered specimens in cluster D are discussed further in the text and in Tables 3.2, 3.3, and 3.4.

continued on next page

chapter. Discrimination often involves multivariate analogs of the analysis depicted in Figures 3.8 to 3.10.

3.3 THE BIOLOGICAL NATURE OF SPECIES

One of the most conspicuous and ancient observations, cutting across human cultures, is that organisms tend to sort out morphologically into relatively discrete clusters, the species of biology. Each species may have not only a distinctive form, but also physiology, behavior, trophic requirements, habitat, and so on. Species are, to varying degrees, ecologically as well as morphologically distinct. A major factor that maintains the distinctiveness of species is reproductive isolation, the evolution of which is at the core of the origin of new species from existing species. Likewise, the maintenance of reproductive isolation, once attained, is central to the maintenance of species distinctiveness.

Box 3.2 (continued)

There is a large family of methods for determining how specimens join clusters and how clusters join one another. In this example, clusters of one or more specimens are linked if they have mutually smallest dissimilarity as measured by the average of all the pairwise dissimilarities between specimens in one cluster and specimens in the other.

Table 3.3 shows the dissimilarities that are relevant to the second round of clustering. Here, specimens 2 and 3 have been replaced by the cluster 2 + 3, and likewise for specimens 6 and 7. The dissimilarity between specimen 4 and cluster 2 + 3 is equal to the mean of $d_{2,4}$ and $d_{3,4}$ from Table 3.2. The remaining dissimilarities between specimens and clusters are calculated in the same way. The dissimilarity between clusters 2 + 3 and 6 + 7 is calculated as the mean of $d_{2,6}$, $d_{2,7}$, $d_{3,6}$, and $d_{3,7}$. There is a single mutually smallest distance in Table 3.2, namely, that between specimen 5 and cluster 6 + 7. Specimen 5 therefore joins this cluster. This procedure of recalculating the distance matrix and finding mutually closest pairs is repeated until all the specimens are joined.

A potential shortcoming of cluster analysis is that a nested structure is superimposed; all specimens eventually cluster together no matter how little they have in common. Moreover, multivariate data are compressed into a single dimension of overall morphological distance. It is therefore inevitable that there will be some distortion in the representation of dissimilarities. A simple and effective way to assess this distortion is to compare the dissimilarities implied by the dendrogram with the true, original dissimilarities based on all the variables.

When two clusters join in a dendrogram, every specimen in one cluster is represented as having the same dissimilarity vis-à-vis every specimen in the other cluster, even though the original pairwise

TABLE 3.3

Average Distances between Specimens and/or Clusters after One Round of Clustering of the Specimens in Table 3.2

	1	2 + 3	4	5	6 + 7
1	—				
2+3	2.05	—			
4	2.7	2.2	—		
5	2.1	2.25	2.1	—	
6+7	1.75	1.85	1.45	1.25*	—

NOTE: The one mutually most similar pair is indicated by an asterisk.

TABLE 3.4

Implied Distances between Specimens of Cluster D, Based on the Dendrogram of Figure 3.16

	1	2	3	4	5	6	7
1	—						
2	2.0	—					
3	2.1	1.1	—				
4	2.0	2.0	2.0	—			
5	2.1	2.0	2.0	1.7	—		
6	2.1	2.0	2.0	1.7	1.3	—	
7	2.1	2.0	2.0	1.7	1.3	0.8	—

The Biologic Species Concept

The most widely accepted biologic definition of the species was formulated by Ernst Mayr (1942): "Species are groups of actually or potentially interbreeding natural populations, which are reproductively isolated from other such groups" (p.120). The species is referred to as a group of populations, emphasizing the fact that most species are divided geographically into subunits or breeding populations. It is explicit in the definition that such breeding populations are actually or potentially interbreeding with one another. Two populations are said to be reproductively isolated only if interbreeding would not occur if they both lived in the same area. Thus, "potentially" in the species definition is particularly critical. An important part of the species definition is that populations of different species are reproductively isolated from one another under natural conditions. There are many examples of species hybridizing readily in captivity or under domestication. This stems from the fact that

dissimilarities may vary quite a bit. This implied dissimilarity is equal to the height on the dendrogram at which the clusters join. Table 3.4 shows the implied dissimilarities for the specimens listed in Table 3.2, and Figure 3.17 compares the implied and original dissimilarities for all pairs of specimens used to construct the dendrogram. The correlation coefficient between original and implied dissimilarities measures how well the original data are represented by the dendrogram. In this case, it is equal to 0.83, a relatively high value, which suggests that the original dissimilarities are represented reasonably well by the dendrogram.

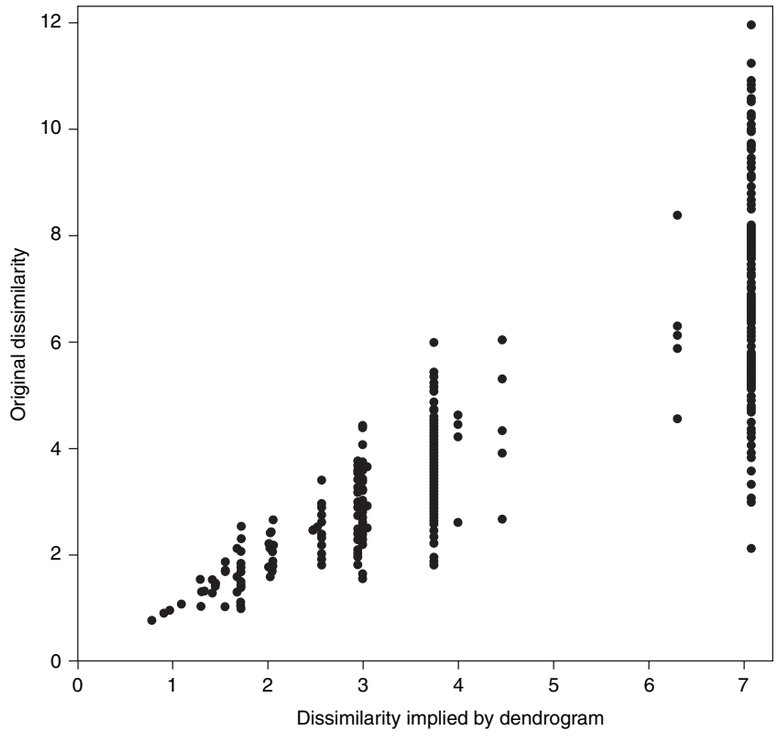

FIGURE 3.17 Comparison of original dissimilarities between *Stegoceras* specimens and dissimilarities implied by the dendrogram of Figure 3.16. Points line up vertically because all between-cluster specimen pairs have the same implied distance when two clusters join.

reproductive isolation often depends on ecologic or behavioral barriers that tend to break down in captivity.

The biologic species concept has some shortcomings. Chief among them are the occasional existence of evolutionary intermediates between species and the difficulty of applying the concept when reproduction is asexual. Biologists who study groups in which asexual reproduction is the rule sometimes adopt a species concept based on phenotypic attributes, such as biochemical properties in bacteria. We focus on the biologic concept because it is thought to apply reasonably well for many of the paleontologically important groups of organisms.

The Origin of Species

If two or more populations of a species diverge to a sufficient extent genetically, they may become reproductively isolated and thus come to be distinct species. One of the principal questions in the study of the origin of species, or **speciation,** concerns the geographic relationships of the diverging populations. Do they have overlapping geographic ranges, in which case they are referred to as **sympatric,** or do they have disjunct ranges—that is, are they **allopatric?** Because gene flow can reduce distinctions between populations, and because populations living in the same broad area may be subject to largely the same forces of natural selection, it seems reasonable to presume that speciation should occur mainly between allopatric populations. In fact, this is the prevailing view among biologists, although there are many theoretical and empirical arguments in favor of sympatric speciation as well.

For allopatric speciation to take place, a population must first become geographically isolated from other populations of the species; then it must persist for some time; and finally it must attain reproductive isolation. Geographic isolates are forming all the time, as organisms disperse and found new populations geographically separated from parental populations, and as newly created geographic barriers, such as mountains, rivers, and emergent land, split populations. The resulting populations represent potential new species, but their fate is not at all assured. Many isolates become extinct, either because they start out with relatively few individuals and therefore are susceptible to fluctuations in population size, or because the environments they colonize may be unfavorable or ephemeral.

If a geographically isolated population does become established, even occasional migration of individuals between populations can lead to sufficient gene flow to prevent reproductive isolation from developing. Gene flow on a large scale is facilitated by the spatial shift of environments over time, which promotes migration as populations track the local conditions to which they are adapted. The probability that a geographically isolated population will actually become a new species is therefore generally quite low.

Our understanding of speciation comes mainly from biology rather than paleontology. Nonetheless, how species originate—that is to say, how populations become reproductively isolated and how evolutionary change is associated with this process—has important paleontological implications that we will pursue further in Chapter 7.

Discrimination of Species

It is important to distinguish between how species are defined in principle and how they are recognized in practice. Biologists rarely perform breeding experiments to determine whether two populations are part of the same species, and of course paleontologists cannot do so with fossil populations. Except for the availability of behavioral data and the widespread analysis of genetic data in biology, the approaches of biologists and paleontologists are often rather similar: One typically starts by determining whether the phenotypic difference between two populations is large relative to the variation within the populations (see Figure 3.9).

Figure 3.18 shows an example of this approach with corals from the Silurian of Arctic Canada. Here there are three clear groups that do not overlap: *Heliolites* aff. *H. luxarboreus*, *H. diligens*, and *H. tchernyshevi*. These are accepted as distinct species on morphological grounds. A fourth form, *H.* sp., is rather similar to *H. tchernyshevi* with respect to the characters portrayed here, but it is not known from enough material to assess its variation in these characters. It is nevertheless accepted as a distinct species because it differs from the remaining species in other characters, such as the nature of the septa, or vertical plates within the corallites.

Genetic data, either in the direct form of DNA sequences or in the indirect form of proteins, have also proven invaluable in discriminating living species, and genetic analysis is now part of the standard toolkit of biologists. (See Box 3.3.) If two populations differ from each

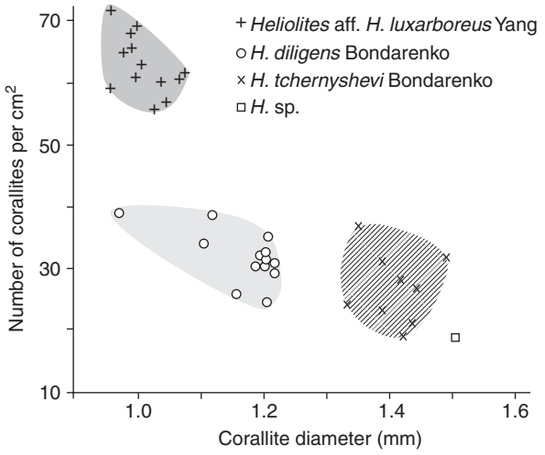

FIGURE 3.18 Morphological discrimination of species of the coral *Heliolites* from the Silurian of Arctic Canada. *(From Dixon, 1989)*

other by as much as two closely related species typically do, they are often regarded as belonging to distinct species. Genetic data can be used to great advantage when morphological differences are negligible or difficult to observe. As is true with morphological data, however, there is no formula that says how much genetic difference characterizes distinct species.

Morphologic and Biologic Species

In practice, both biologists and paleontologists usually apply a morphologic species concept. There are several important problems that stem from this approach.

Failing to take variation into consideration can lead to biologically unrealistic results. Figure 3.19 shows an example involving the Triassic ammonoid genus *Paranannites* from the Great Basin of the western United States. This graph plots two separate characters, the whorl width (*W*) and the umbilical width (*U*), against the shell diameter. Each point is a single specimen and each field in the graph represents a separate bivariate comparison. Within each bivariate comparison, the points form a continuous distribution. There are no obvious divisions or clusters that would serve as evidence for multiple species. Partly on these grounds, Bernhard Kummel and Grant Steele (1962) concluded that the material represents a single species, *Paranannites aspenensis*.

Thirty years before Kummel and Steele performed this analysis, J. P. Smith (1932) studied a subset of this material. In addition to *P. aspenensis*, Smith erected three other species, based mainly on differences relative to *P. aspenensis* in overall size, whorl width, and umbilical diameter, as well as on details of sculpture. Given that Smith studied the same traits as Kummel and Steele, how

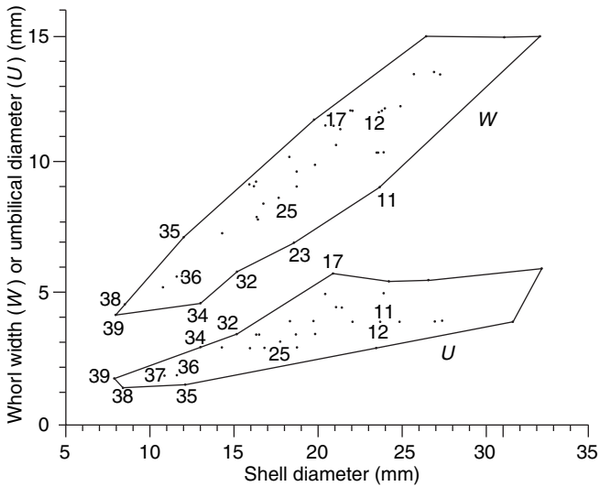

FIGURE 3.19 Biometric analysis of the ammonite species *Paranannites aspenensis* from the Triassic of the Great Basin. Two separate bivariate comparisons are shown here: whorl width (*W*) against shell diameter, and umbilical diameter (*U*) against shell diameter. Each point represents one specimen. The numbered points are type specimens that had previously been used to describe this species and three additional species. Because they show continuous variation, all the specimens are now considered to belong to a single species. The type specimens tend to fall near the extremes of the continuous distribution of form. *(From Kummel & Steele, 1962)*

can we account for the different numbers of species recognized by these authors? The numbered points on Figure 3.19 are Smith's type specimens—the exemplars he chose as representative of the species he described [SEE SECTION 4.1]. Most of these lie at the periphery of the scatter of points. Smith evidently focused on extreme forms and considered them to be representatives of separate species, rather than recognizing them as simply end-members of a continuum.

There are potential problems with the use of morphologic species, in both biology and paleontology, that cannot easily be overcome with more detailed assessment of morphological variation. First is the existence of **cryptic species,** also known as **sibling species.** Closely related species may be genetically and behaviorally distinct but may lack clear morphological differences. Second, species may contain numerous distinct morphological types, or **polymorphs.** The different forms within a polymorphic species are under genetic control, but they are not reproductively isolated and the genetic differences involved are generally small. Nonetheless, polymorphs are sometimes sufficiently different in form that they might be mistaken for distinct species on the basis of morphology alone. Finally, as we discussed earlier, some of the variation within species is ecophenotypic rather than heritable. Thus, two populations that belong to the same species could be mistaken for different species if they lived in environments that induced substantially different phenotypes.

There is no question that these problems exist in principle, but it is important to determine how common they are in reality. One study that explores this question involves living species of the cheilostome bryozoan genera *Steginoporella*, *Stylopoma*, and *Parasmittina* from the Caribbean Sea.

Using multivariate morphometric techniques similar to those we discussed earlier, Jeremy Jackson and Alan Cheetham (1990, 1994) analyzed a variety of skeletal measurements and found morphological clusters of specimens that were defined operationally as **morphospecies.** Once the morphospecies were established, Jackson and Cheetham sought to assess the importance of ecophenotypic variation. Embryos of known parentage were raised in environments different from those in which their parents had been raised. After rearing, the offspring were measured and assigned to prospective parents on the basis of morphological similarity. That is, each of the offspring was assigned to the parental colony with which it was morphologically most similar. For all seven species studied, these assignments were found to be correct—matching true parentage—99 to 100 percent of the time, despite the fact that parents and offspring did not share the same environment. On the whole, morphological variation was much more strongly affected by heritability than by variation in the environment in which the embryos grew.

Jackson and Cheetham then tested for polymorphism by asking whether morphologically distinct species have consistent genetic differences. To identify genetic differences, they used the standard technique of electrophoresis, which identifies alternative forms of proteins having different mass and electrical properties. Because proteins are coded by DNA, the alternative forms of protein are used as evidence for differences in DNA sequence. In general, different forms of the same gene are referred to as **alleles.** Here the different proteins are inferred to represent different alleles. For a given kind of gene, each individual inherits one allele from its mother and one from its father. For that gene, the combination of two alleles is the individual's genotype.

Box 3.3 gives one example of how the genetic results are interpreted to test for differences between populations. When this approach was applied to the bryozoans, every pair of distinct morphospecies within a genus was found to have at least one diagnostic genetic difference. Thus, these morphospecies are likely to be true biological species rather than polymorphs within a single species. Moreover, if genetic and morphological dissimilarity between populations are compared, it is found that the magnitudes of morphological and genetic difference are well correlated (Figure 3.20). Pairs of populations that are more dissimilar morphologically also tend to be more dissimilar genetically.

Finally, Jackson and Cheetham tested for the existence of cryptic species by determining whether different populations of the same morphospecies have diagnostic genetic differences. The analysis found no cases in which two populations of the same morphospecies could be genetically distinguished with confidence. In other words, populations that could not be distinguished morphologically could not be distinguished genetically, either. Thus, there was no compelling evidence for the existence of cryptic species in these genera.

Box 3.3

TESTING FOR DIAGNOSTIC GENETIC DIFFERENCES BETWEEN POPULATIONS

In the genus *Stylopoma*, there are four alternative forms of the protein GPI. By genetically assaying many individuals (about 40 on average) within each morphospecies, it was found that these four alleles, denoted *a* through *d*, are present in different frequencies in the two morphospecies. Given the standard assumption of random mating between individuals within a species, the allele frequencies allow the genotype frequencies to be estimated. For example, the frequencies of the *b*- and *c*-alleles in *S.* sp. 1 are $f_b = 0.139$ and $f_c = 0.583$. The frequency of the *bc* genotype is therefore inferred to be equal to $2f_bf_c$, or 0.162. (The multiplication by 2 reflects the fact that an individual can inherit the *b*-allele from either its mother or its father, and likewise for the *c*-allele.)

Once the genotype frequencies are determined, we see that most genotypes are unique to one species or the other. If an individual has the *aa* or *ab* genotype, it belongs to *S.* sp. 2. If it has the *bc*, *bd*, *cc*, *cd*, or *dd* genotype, it belongs to *S.* sp. 1. The only ambiguous genotype is *bb*. Because the vast majority of *bb* individuals are in *S.* sp. 2, our best guess would be to assign any such individual to that species. If we assume that the two species are represented by the same number of individuals, then the probability that a randomly sampled individual will belong to *S.* sp. 1 *and* will have the *bb* genotype is equal to $f_b^2 \div 2$, which in this case is only $0.019 \div 2$, or less than 1 percent. In other words, if we use the GPI genotypes to assign individuals to morphospecies, we will be wrong less than 1 percent of the time. Operationally, genetic markers with which the expected probability of misclassifying an individual is less than 1 percent are considered to be diagnostic.

TABLE 3.5

Allele and Genotype Frequencies of the Protein GPI in Two Species of the Cheilostome Bryozoan *Stylopoma*

	Symbol for Frequency	Frequency *Stylopoma* sp. 1	Frequency *Stylopoma* sp. 2
Allele			
a	f_a	—	0.188
b	f_b	0.139	0.812
c	f_c	0.583	—
d	f_d	0.278	—
	Formula for Frequency	**Frequency *Stylopoma* sp. 1**	**Frequency *Stylopoma* sp. 2**
Genotype			
aa	f_a^2	—	0.035
ab	$2f_af_b$	—	0.305
bb	f_b^2	0.019	0.659
bc	$2f_bf_c$	0.162	—
bd	$2f_bf_d$	0.077	—
cc	f_c^2	0.340	—
cd	$2f_cf_d$	0.324	—
dd	f_d^2	0.077	—

SOURCE: Jackson & Cheetham (1990)

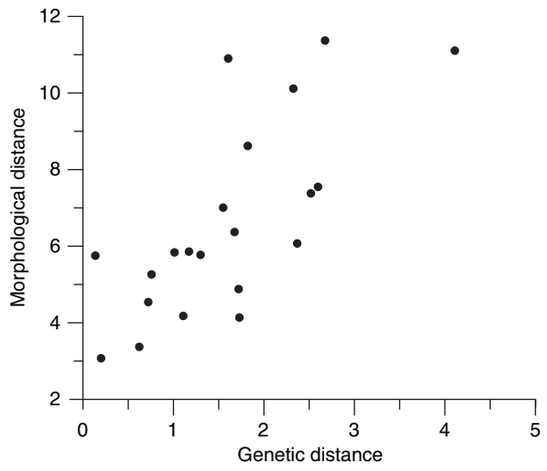

FIGURE 3.20 Comparison of morphological and genetic dissimilarity between populations of the bryozoan *Stylopoma*. Each point represents a comparison between two populations. Morphological distance is measured by a variant of the straight-line distance described in the discussion of cluster analysis in Box 3.2. Distance is measured on the basis of differences in gene frequencies. (See Table 3.5 for examples of gene frequencies.) Morphological and genetic differences are positively correlated. *(From Jackson & Cheetham, 1994)*

Taken together, these results suggest that there is an excellent concordance between biological and morphological species in this sample of cheilostome bryozoans.

The question of concordance between morphological and biological species applies to biology as much as to paleontology. There is a special problem, however, that paleontologists must face because of the temporal dimension of the history of life. Our earlier discussion of speciation was restricted to the situation in which an evolving lineage splits into two distinct lineages. It sometimes happens that a single lineage may evolve over time to the point where it becomes morphologically quite distinct from earlier populations in the lineage, even though there has been no splitting (Figure 3.21). In cases like this, some paleontologists will divide the lineage into two or more named species. Because of the added time dimension, species such as A and B in Figure 3.21 may be referred to as **chronospecies.** Many workers today prefer, if possible, to place species boundaries at branching points and at true lineage terminations. It may be difficult to avoid erecting chronospecies, however, if the intermediate forms between A and B are not sampled.

3.4 CONCLUDING REMARKS

That there is a close correspondence between morphological and genetic species in a sample of bryozoans does not imply that the same is true for other groups of organisms, or even for other bryozoans. If these results prove to be general, however, then biologists and paleontologists are in a strong position to discriminate species on the basis of morphology. It is still too soon to assess fully the correspondence between morphospecies and biological species. Nonetheless, studies on many other groups of organisms have shown that, as in the bryozoans, morphologically defined species tend to be genetically distinct. At the same time, cryptic species are known to be common in some groups.

There is thus an asymmetry in the relationships between morphological and genetic species. If two populations are morphologically distinct, there is often a good chance that they belong to different species. But if they are morphologically indistinguishable, this need not imply that they belong to the same species. This asymmetry will be relevant when we consider the relationship between speciation and morphological evolution in Chapter 7.

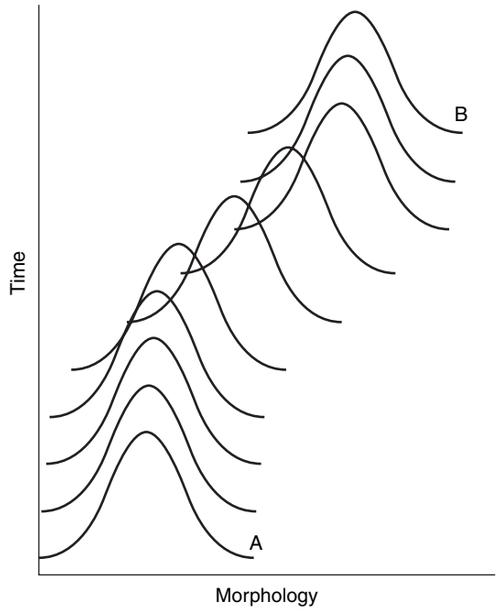

FIGURE 3.21 The problem of species delimitation in an evolving lineage. Each curve represents a frequency distribution for a trait. The lineage is a single, continuous succession of populations, yet populations at different points in time—for example, points A and B—may be so different from each other that they would be taken for different species if found together.

SUPPLEMENTARY READING

Coyne, J. A., and Orr, H. A. (2004) *Speciation*. Sunderland, Mass., Sinauer Associates, 545 pp. [An in-depth review of ecological, genetic, and geographic aspects of speciation.]

Davis, J. C. (2002) *Statistics and Data Analysis in Geology*, 3rd ed. New York, John Wiley and Sons, 638 pp. [Includes intuitive and graphical explanations of many multivariate techniques.]

Futuyma, D. J. (1998) *Evolutionary Biology*, 3rd ed. Sunderland, Mass., Sinauer Associates, 763 pp. [Comprehensive text on genetics, evolution, and speciation.]

Hammer, Ø., Harper, D. A. T., and Ryan, P. D. (2001) PAST: Paleontological statistics software package for education and data analysis. *Palaeontologia Electronica,* volume 4, issue 1, article 4, 9 pp. [Noteworthy for worked examples of many methods of data analysis. Available for free download at http://folk.uio.no/ohammer/past/.]

Hanski, I., and Gilpin, M. E. (eds.) (1996) *Metapopulation Biology: Ecology, Genetics, and Evolution*. San Diego, Academic Press, 512 pp. [A collection of theoretical papers and case studies.]

Knowlton, N. (1993) Sibling species in the sea. *Annual Review of Ecology and Systematics* **24:**189–216. [A broad survey of the problem of cryptic species.]

Reyment, R. A. (1991) *Multidimensional Palaeobiology*. Oxford, U.K., Pergamon Press, 377 pp. [Techniques of multivariate analysis applied to a range of paleontological problems.]

Siegel, S., and Castellan, N. J., Jr. (1988) *Nonparametric Statistics for the Behavioral Sciences*, 2nd ed. New York, McGraw-Hill, 399 pp. [Statistical analysis of univariate and bivariate data, with special emphasis on categorical and ordinal variables.]

Sneath, P. H. A., and Sokal, R. R. (1973) *Numerical Taxonomy*. New York, W. H. Freeman, 583 pp. [Includes important discussions of biological dissimilarity, ordination, clustering, and discrimination.]

Sokal, R. R., and Rohlf, F. J. (1995) *Biometry: The Principles and Practice of Statistics in Biological Research*, 3rd ed. New York, W. H. Freeman, 887 pp. [Comprehensive text on statistical analysis of biological data.]

SOFTWARE

McCune, B., and Mefford, M. J. (2002) *PC-ORD: Multivariate Analysis of Ecological Data, Version 4.* Glenden Beach, Oreg., MjM Software Design. [Software package with a variety of options for ordination and clustering of data. Available for purchase at www.ptinet.net/mjm.]

R Development Core Team (2004). *R: A Language and Environment for Statistical Computing.* Vienna, Austria, R Foundation for Statistical Computing. [Useful software for univariate, bivariate, and multivariate analysis, with powerful graphical capabilities. Available for free download at http://cran.r-project.org.]