MULTIVARIATE ALLOMETRY

Christian Peter Klingenberg

Department of Biological Sciences
University of Alberta
Edmonton, Alberta T6G 2E9, Canada

ABSTRACT

The subject of allometry is variation in morphometric variables or other features of organisms associated with variation in size. Such variation can be produced by several biological phenomena, and three different levels of allometry are therefore distinguished: static allometry reflects individual variation within a population and age class, ontogenetic allometry is due to growth processes, and evolutionary allometry is the result of phylogenetic variation among taxa. Most multivariate studies of allometry have used principal component analysis. I review the traditional technique, which can be interpreted as a least-squares fit of a straight line to the scatter of data points in a multidimensional space spanned by the morphometric variables. I also summarize some recent developments extending principal component analysis to multiple groups. "Size correction" for comparisons between groups of organisms is an important application of allometry in morphometrics. I recommend use of Burnaby's technique for "size correction" and compare it with some similar approaches. The procedures described herein are applied to a data set on geographic variation in the waterstrider *Gerris costae* (Insecta: Heteroptera: Gerridae). In this example, I use the bootstrap technique to compute standard errors and perform statistical tests. Finally, I contrast this approach to the study of allometry with some alternatives, such as factor analytic and geometric approaches, and briefly analyze the different notions of allometry upon which these approaches are based.

INTRODUCTION

Variation in size of organisms usually is associated with variation in shape, and most metric characters are highly correlated among one another. These associations are the subject of allometry (Huxley, 1932; Cock, 1966; Gould, 1966, 1975). Although allometry is often used to examine the consequences of size for ecological or physiological variables (Günther, 1975; LaBarbera, 1989; Reiss, 1989), this review deals only with measurements of traits used to characterize the morphological form of organisms.

Advances in Morphometrics, Edited by L. F. Marcus et al
Plenum Press, New York, 1996

Unlike other approaches in morphometrics, which are built on geometric theory, allometry has a largely empirical basis. Huxley (1932) realized that scatter plots of two trait measurements in growing organisms often closely follow a curved line, and that this relationship usually becomes linear if both measurements are transformed to logarithms. From this, he derived his formula of simple allometry

$$y = bx^{\alpha}$$

or, in log-transformed notation,

$$\log y = \log b + \alpha \log x,$$

where x and y are trait measurements, and b and α are constants. The constant a, the slope in log-log plots of x and y, is often called the allometric coefficient (terminology is not uniform; some authors call b coefficient). The special case when $\alpha = 1$ is called isometry, and indicates direct proportionality between x and y. If $\alpha > 1$, there is positive allometry, whereas for negative allometry, $\alpha < 1$ (Huxley and Teissier, 1936). In humans, for example, the long bones of the limbs show positive allometric growth relative to overall stature, and the height of the head shows negative allometry.

In most morphometric data sets, measurements are positively correlated, i.e., x and y increase or decrease simultaneously. Even if there is negative allometry, a still is positive; negative allometry implies only that the relative variation in y is smaller than that in x, e.g., y grows by 10% for every 20% growth increment in x. If a is negative, however, there is an absolute reduction in y associated with an increase in x. This case is called enantiometry (Huxley and Teissier, 1936). Reduction of the absolute size of organs during growth is a real phenomenon, although it is not found commonly in morphometric studies. The most striking example is the shrinking of larval structures during metamorphosis, e.g. the gills and tail of anuran tadpoles; but in a subtler way, enantiometry even occurs in cranial growth of primates (Corner and Richtsmeier, 1991).

Huxley's approach is not restricted to pairs of measurements. In many multivariate data sets, log-log plots of all pairwise combinations of morphometric variables show approximately linear relationships. Therefore, Huxley's bivariate allometry can be generalized to multiple dimensions. Moreover, it is not confined to growth data, as straight-line relationships are also found in log-log plots of intra- and interspecific variation within one particular ontogenetic stage (most often adults). From this perspective of allometry, some major questions follow: How much variation is there? Are the data points clustered around a straight line and, if so, how closely? What is the direction of that line in multidimensional space? Do different groups of organisms share the same allometric relationship?

In this paper I review concepts and techniques used in studies of multivariate allometry. First, I introduce the main levels of variation that have been the subject of allometric studies within and between species,. Then, I present the multivariate generalization of allometry using principal component analysis and some more recent developments, such as the bootstrap and techniques dealing with multiple groups. Finally, I briefly contrast some alternative approaches to allometry.

LEVELS OF ALLOMETRY

Huxley (1932) devised allometry mainly as a tool to study the relative growth of parts in various organisms. Growth, however, is not the only origin of variation in overall size and

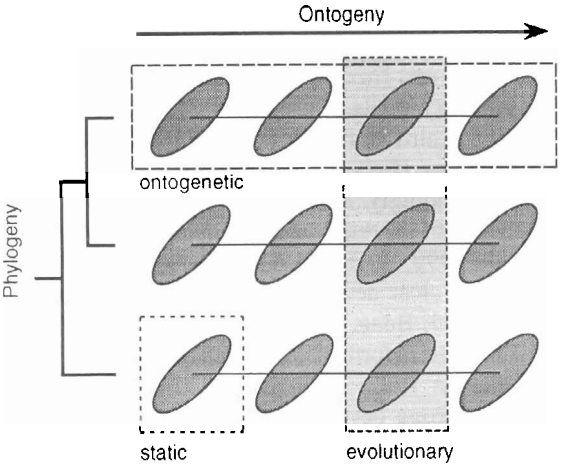

Figure 1. The three levels of allometry. The diagram shows three species, each with four different ontogenetic stages, that are considered to be homologous among species. Rectangles enclose the species and stage groups included in an analysis of allometry at each of the three levels. Ontogenetic allometry can be separately analyzed for all three species, evolutionary allometry for each of the four stages, and static allometry for each of the 12 species and stage groups.

associated variation in shape, because evolutionary changes and individual variation also can generate allometric relations. These levels or types of allometry have been included in elaborate classification schemes (see Cock, 1966; Gould, 1966, 1975). Because of its simplicity, I prefer the terminology proposed by Cock (1966) who distinguished static, ontogenetic and evolutionary allometry (Fig. 1). This classification has also been used in most empirical comparisons between levels of allometry (Cheverud, 1982; Leamy and Bradley, 1982; Boag, 1984; Gibson et al., 1984; Leamy and Atchley, 1984a; Shea, 1985; Klingenberg and Zimmermann, 1992a).

Static Allometry. Static allometry, which is also referred to as size allometry, results from variation among individuals of the same population and age group (intraspecific scaling, Gould, 1975). Static allometry is particularly easy to study in organisms with discrete growth stages, such as insects (Cuzin-Roudy, 1975; Klingenberg and Zimmermann, 1992a), or in adults of organisms with determinate growth, such as birds (Boag, 1984; Gibson et al., 1984). These studies, among others, found that the largest proportion of multivariate variation is contained in one dimension, and that the model of simple static allometry therefore is appropriate. This phenomenon has been termed morphometric or phenotypic integration (e.g., Leamy and Atchley, 1984b; Zelditch, 1987, 1988; Zelditch and Carmichael, 1989). Although there is an extensive literature describing static allometry and morphometric integration, relatively little is known about their developmental basis (but see Cowley and Atchley, 1990; Atchley and Hall, 1991; Shea, 1992; Paulsen and Nijhout, 1993). In theoretical models, Riska (1986) explored how developmental processes can affect static correlations among the traits they produce (see also Cowley and Atchley, 1992). Patterns of static allometry have sometimes been used to deduce underlying developmental processes (e.g., Zelditch, 1987; Wheeler, 1991). Such inferences, however, should be substantiated by direct observations.

Ontogenetic Allometry. Ontogenetic allometry or growth allometry deals with covariation among characters during growth. Simple allometry occurs if the ratio between the specific growth rates (percentage increment per time unit) of two different characters is constant (Huxley, 1932; Reeve and Huxley, 1945; Shea, 1985; Blackstone, 1987). Theoretical studies showed that various models of growth as a function of time can result in simple allometry (e.g., Laird, 1965; Laird et al., 1968; Katz, 1980). The rule of simple allometry holds often, but not always, as Huxley (1932) demonstrated with an impressive list of bivariate examples. Correspondingly, multivariate studies often find that one dimension contains the largest proportion of the total variation, sometimes more than 99% (e.g., Jungers et al., 1988; Solignac et al., 1990; Strauss, 1990b; Klingenberg and Zimmermann, 1992a). Some studies, however, show clear deviations from simple allometry in certain structures (Cuzin-Roudy and Laval, 1975; Boag, 1984; Cane, 1993) or subtle curvatures of growth trajectories in the space of log-transformed characters (Bookstein, 1991: Figs. 4.2.2 to 4.2.4; Klingenberg and Zimmermann 1992a; Klingenberg and Spence, 1993). Studies of plant growth showed particularly strong deviations from simple allometry (e.g. Kampny et al., 1993; McLellan, 1993). Three types of data are used to study ontogenetic allometry: longitudinal data based on measurements of the same individuals at several developmental stages, cross-sectional data with different specimens in several known stages, and mixed cross-sectional data collected without information on ontogenetic stage (Cock, 1966).

Evolutionary Allometry. Evolutionary allometry reflects covariation among changes in different traits along the branches of a phylogeny. It is concerned with character covariation among contemporaneous species sharing a common ancestor (Fig. 2a) or among fossil members of an evolutionary lineage (Fig. 2b). I do not distinguish separate levels for analyses that use these two types of data (for different terminology, see Gould, 1966). Evolutionary processes leading to the associations between trait changes presumably do not differ depending on whether these changes occur within one lineage successively or in different lineages giving rise to sister groups. It is important to use specimens in comparable ontogenetic stages to avoid confounding evolutionary and ontogenetic variation. This is straightforward in organisms with determinate growth, such as birds or insects (e.g., Livezey, 1989; Strauss, 1990a; Klingenberg and Zimmermann, 1992a), but it is more difficult in organisms with indeterminate growth, for which some studies assume that specimens are "typical" for the respective species (e.g., Strauss, 1985). In these studies, among others, the model of simple allometry fits the data fairly well, indicating that evolutionary variation is constrained in its dimensionality (Maynard Smith et al., 1985; Gould, 1989; Arnold, 1992). Some of this covariation among traits may be determined by developmental processes, as Riska (1989) showed with a simulation study.

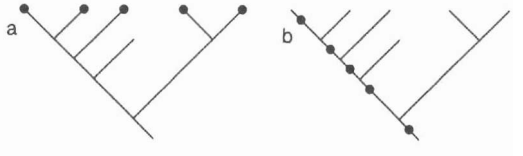

Figure 2. Types of data for evolutionary allometry. (a) In neontological studies, the data always are measurements taken on recent species from several lineages in a clade, that are related as sister groups rather than ancestors and descendants. (b) In paleontological studies, evolutionary allometry often refers to character covariation among members of a single evolutionary lineage. Because it is difficult to distinguish with certainty whether two fossil species are related as ancestor and descendant or as sister groups, many of these studies may in fact use designs that are mixtures of (a) and (b).

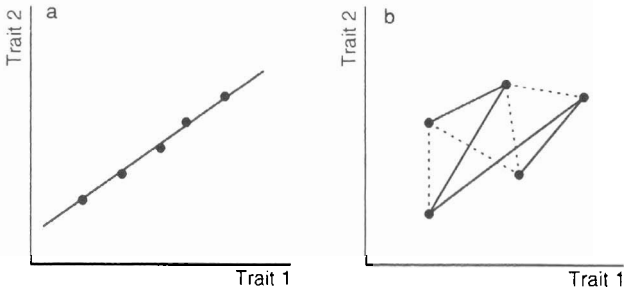

Figure 3. Influence of the correlation between characters on the robustness of evolutionary allometry. For simplicity, I assume that the species are members of a single evolutionary lineage (as in Fig. 2b). (a) If the traits are highly correlated, most morphometric variation is in a single direction. Estimates of evolutionary allometry, the "average" direction of evolutionary changes, will yield almost the same result regardless of the phylogenetic relations among species (i.e., how the points are connected). (b) If the correlation between traits is low, however, the "average" direction of changes depends strongly on the phylogenetic scenario, i.e., on how the points are connected. The solid and dashed lines represent two alternative hypotheses of ancestor–descendant relationships, which have drastically different directions of evolutionary changes.

Evolutionary allometry, like all interspecific comparisons, presents some statistical problems because the species are not independent of one another but are parts of a hierarchically structured phylogeny (e.g., Felsenstein, 1985; Pagel and Harvey, 1988). This interdependence is most evident for species presumed to be members of a single, unbranched lineage (Fig. 2b), but it also applies to comparisons among terminal taxa in a clade. A possible solution is the method of phylogenetically independent contrasts (Felsenstein, 1985; Martins and Garland, 1991; Garland et al., 1992), which analyzes character changes along the branches in the phylogeny instead of the measurements (character states) of terminal taxa. Changes are either directly measured, if actual ancestor–descendant series of fossils are available, or inferred by indirect methods such as parsimony or maximum likelihood. This approach requires knowledge of the phylogeny of the group under consideration and a multivariate theory for reconstructing state vectors of quantitative characters for internal nodes, which still needs to be developed (for univariate approaches, see Maddison and Maddison, 1992). For morphometric data, however, the problem of phylogenetic dependence may not be as severe as for other data types, because often most of the variation is in a single dimension. These high correlations make estimates of the direction of evolutionary changes relatively robust against errors in phylogenetic reconstruction (Fig. 3).

Further levels of allometry exist in organisms with modular organization, such as colonial animals and most vascular plants. In addition to the ontogeny of individual zooids, colonial animals have an additional level of colony wide development, which is called astogeny (e.g., Pandolfi, 1988). Buss and Blackstone (1991) showed that colony growth in a marine hydroid follows well-determined trajectories and that colonies react to experimental perturbations in an integrated manner (see also Anstey, 1987). Similarly, the structure of plant parts changes with the age of the entire plant (heteroblasty; e.g., McLellan, 1993). Jones (1992, 1993) studied correspondences between the development of individual leaves and the succession of leaves during whole-plant ontogeny. New methodological approaches, such as "process morphology" (Sattler, 1992; Jeune and Sattler, 1992), reflect the morphological flexibility of modular organization in plants but are only semiquantitative and cannot be directly related to allometry (a similar approach in zoology is the "skeleton space," Thomas and Reif, 1993).

The causes of allometric variation at different levels are mutually interrelated. Static variation, which is caused by variation in ontogenetic processes that produce the structures of interest, is the raw material upon which natural selection can act. Response to selection, in turn, generates evolutionary changes affecting the developmental processes. One way to study these interactions is to compare empirically the patterns of variation between different levels of allometry. Various such comparisons have been made (Cheverud, 1982; Leamy and Bradley, 1982; Boag, 1984; Gibson et al., 1984; Leamy and Atchley, 1984a; Shea, 1985; Klingenberg and Zimmermann, 1992a). Most of these studies found that patterns of allometry at different levels were similar but not equal. It is not possible, however, to make further generalizations of the results because the studies differ widely in the kinds of data and methods used. As an alternative to this observational approach, the mechanisms that are the basis of allometric variation can be investigated by experimental techniques, e.g., using genetically engineered organisms (Shea et al., 1990) or directly manipulating the size of eggs or embryos (Sinervo, 1993).

PRINCIPAL COMPONENTS AND ALLOMETRY

Under the model of simple allometry, bivariate plots of pairs of log-transformed morphometric variables are straight lines. If there are three variables, and all pairwise combinations satisfy this condition, then the data points follow a line in the three-dimensional space defined by the variables. This argument can be extended to more than three variables: data points still are arranged along a straight line under simple allometry (e.g., Teissier, 1955), but this line now is in the multidimensional space defined by all the variables. Therefore, dimensionality of morphometric variation is a prime concern of allometry (e.g., Hopkins, 1966; Sprent, 1972). As in bivariate allometry, points may be scattered around the line, rather than exactly lying on it, and one has to find a line that "optimally" fits the scatter of data points (Pearson, 1901). Jolicoeur (1963) proposed the first principal component, estimated from the covariance matrix of log-transformed measurements, as a multivariate generalization of simple allometry.

Many texts of multivariate statistics introduce principal component analysis (PCA) as a technique for summarizing most of the variation in a multivariate data set in fewer dimensions (e.g., Pimentel, 1979; Jolliffe, 1986; Flury, 1988; Flury and Riedwyl, 1988; Johnson and Wichern, 1988; Jackson, 1990; Jobson, 1992). The first principal component (PC1) is the linear combination that accounts for the maximum variance. Geometrically, it corresponds to the direction of the longest axis through the scatter of data points. Subsequent principal components take up maximal variance, subject to being orthogonal to all preceding component axes.

Fig. 4 shows a contour ellipse of the bivariate distribution of two variables X_1 and X_2 with its centroid (mean vector) at the point labeled 0. For simplicity, data are centered by subtracting the means of X_1 and X_2, which shifts the coordinate system to the new axes x_1 and x_2 each with a sample mean of zero. Therefore, the x_1 and x_2 values themselves are the deviations from their mean; their sample variances can be calculated as the sum of squared x_1 and x_2 values divided by $(n - 1)$. In Fig. 4, one data point is labeled P, and its projection onto the x_1 axis is Q. The sum of squared x_1 values is the sum of the squared distances between 0 and Q for all data points. By the same argument, the sum of squared x_2 values corresponds to the sum of squared distances between P and Q. According to the Pythagorean theorem, the squared distance between 0 and P is the sum of the squared distances between 0 and Q and between P and Q. It follows that the sum of the squared distances of all data points from the sample centroid, divided by $(n - 1)$, is the sum of the variances of x_1 and x_2, or total variance. Now, consider the same data set after rotating the coordinate system to the

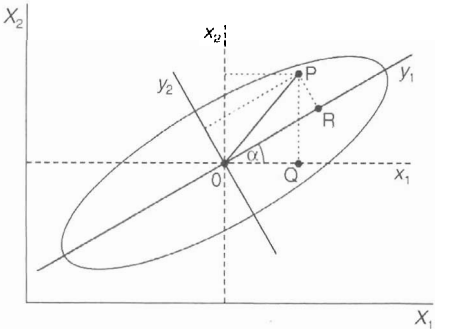

Figure 4. Principal component analysis. The diagram shows the contour ellipse of a bivariate distribution with its centroid at the point labeled 0. The centered coordinates x_1 and x_2 are derived from the original variables X_1 and X_2 by subtracting their mean values. The principal components y_1 and y_2 are the directions of maximal and minimal variance, respectively. See text for details.

directions of the principal component axes y_1 and y_2. Because the PC1 axis, y_1, is defined as the direction that has maximal variance, the sum of the squared distances between 0 and R, the projection of P onto the PC1 axis, is maximal. Because the rotation of the coordinate system does not change the distances between the data points and the centroid, maximizing the sum of squared distances between 0 and R also results in minimizing the sum of squared distances between P and R. The sum of squared distances between P and R, divided by $(n - 1)$, is the part of the total variance not accounted for by the PC1, i.e., the residual variance. This argument also holds for more than two dimensions: because the PC1 is the direction that has maximal variance, all other principal components taken together have minimal variance. Hence, the PC1 axis can be seen as a least-squares fit of a straight line to the scatter of data points in the space of log-transformed, bivariate or multivariate data. This justifies Jolicoeur's (1963) multivariate generalization of allometry (see also Hopkins, 1966).

PCA decomposes a covariance matrix **S** into eigenvectors and eigenvalues, so that **S** = **BLB**'. The matrix **B** of eigenvectors is used to transform the original data **X** into a set of new variables **Y** = **XB**, the principal components (PCs). The matrix **L** is the covariance matrix of the PCs, and as the PCs are uncorrelated among each other, all off-diagonal elements of **L** are zero. The diagonal elements of **L**, the eigenvalues, are the variances for which the associated eigenvectors account. They are difficult to interpret by themselves, because they depend on the measurement units and the base of the logarithm used to transform the data. However, the proportion of the total variance for which the PC1 accounts is important to assess how well the model of simple allometry fits the data.

Principal component analysis can be interpreted geometrically as a rotation of the coordinate system. The PC axes are aligned with the directions of the axes of the multidimensional "scatter ellipsoid" (in two dimensions, this is an ellipse, Fig. 4). The PC coefficients of the original variables can be interpreted as "direction cosines," i.e., the cosine of the angle between the PC axis and the coordinate axis of the respective variable (α for PC1 and x_1 in Fig. 4). Principal component axes are mutually orthogonal, and the vectors of PC coefficients are usually normalized to have unit length so that the squares of the coefficients sum up to unity (**b'b** = 1, where **b** is an eigenvector). As a result, the coefficient values depend on the number of variables. Nevertheless, translating PC coefficients to bivariate allometric coefficients (Huxley's α) is quite easy (Jolicoeur, 1963; Shea, 1985). The ratio of PC1 coefficients for two variables corresponds to the variables' bivariate

allometric coefficient. For example, in a study of two species of voles, Airoldi and Flury (1988) found that the PC1 coefficients of skull length, width and height were approximately 2/3, 2/3, and 1/3, respectively. Thus, in allometric plots with skull length as the independent variable, skull width would be isometric with a slope of about 1, whereas skull height, with a slope of about 0.5, would show strong negative allometry. With p variables, isometry in all pairwise combinations of variables results in a PC1 in which all coefficients are equal, and have the value 1 divided by the square root of p, i.e., $p^{-0.5}$. Isometry can be assessed with Anderson's (1963) test, which is based on normal theory (Pimentel, 1979: 70; Flury, 1988: 34), or by comparison with jackknifed or bootstrapped confidence intervals (see below; e.g., Diaconis and Efron, 1983; Gibson et al., 1984; Klingenberg and Zimmermann, 1992a). Multiplying PC1 coefficients by the square root of p yields values that can be interpreted as bivariate allometric coefficients for each of the variables against a measure of overall size (a weighted geometric mean of all variables). These allometric coefficients or the principal component coefficients can be graphed as Huxley's (1932) growth gradients (e.g., Boitard et al., 1982; Solignac et al., 1990), or they can be displayed on diagrams of the measurements, such as the truss network (Strauss and Bookstein, 1982; Bookstein et al., 1985). Another type of graphical display for PC coefficients is the biplot (e.g., Marcus, 1993).

Empirical comparisons of bivariate and multivariate approaches found that both estimated corresponding patterns of allometry (Davies and Brown, 1972; Shea, 1985). Jungers and German (1981) criticized the multivariate approach because allometric coefficients derived from principal components of skeletal measurements did not match those from bivariate regressions on a known variable for size that was not included in the analysis, either body weight or length. Hills (1982) showed that these discrepancies disappear if one considers allometry between the traits and the size measure that is taken as a reference, e.g., between skeletal size and body weight.

It is possible to perform PCA by use of a correlation matrix instead of a covariance matrix (see also Pimentel, 1979; Bookstein et al., 1985; Johnson and Wichern, 1988). This corresponds to an analysis of standardized variables. Geometrically, it means that all the variables are adjusted to have standard deviations of 1 by stretching or shrinking their coordinate axes before the analysis. This maneuver can be used to remove scaling effects if variables are measured in different units. Standardizing may also be useful if one is interested in only ordination of specimens, as in some applications in systematics, where giving equal weight to all variables may be more important than scaling. For allometry, however, scaling is essential. After removing scale by standardization one still can determine whether the data points lie along a straight line, but it is impossible to estimate allometric coefficients because standardization changes the direction of the allometric axis. A simple hypothetical example, constructed from purely allometric variation without any residual scatter, can show this. Let the multivariate allometric coefficients (eigenvector) be 2/3, 2/3 and 1/3; these coefficients correspond to strong deviations from isometry. Then, the covariance matrix is a multiple of

$$\begin{bmatrix} 4 & 4 & 2 \\ 4 & 4 & 2 \\ 2 & 2 & 1 \end{bmatrix}$$

and the correlation matrix is

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

The PC1 of this correlation matrix has coefficients that all take the value $3^{-0.5}$, and thus falsely indicate isometry, with no residual variation at all (see also Johnson and Wichern, 1988: 350). Likewise, in allometric interpretations of PCAs based on correlation matrices of real data (e.g., Teissier, 1955; Somers, 1986), all "non-isometric" variation that may be inferred from PC1 coefficients merely results from the residual scatter about an allometric relationship, but does not reflect allometry. Therefore, it is crucial to use the covariance matrix for allometry. Routines for PCA in many statistical software packages use the correlation matrix as the default option. Users of these programs should make sure to specify the option for PCA based on the covariance matrix.

A related point is the transformation of data to logarithms. There are numerous practical and theoretical reasons why it is often useful to transform data to logarithms (Pimentel, 1979; Reyment et al., 1984; Bookstein et al., 1985; Bookstein, 1991; Reyment, 1991). For ontogenetic allometry, Huxley (1932) justified the use of power functions, and therefore also of logarithms, by his rule of constant ratios among specific growth rates of different organs (see also Reeve and Huxley, 1945; Günther, 1975; Katz, 1980; Bookstein et al., 1985; Shea, 1985; Blackstone, 1987). Such a theoretical justification, however, is more difficult to find for static and evolutionary allometry. The multiplicative nature of growth processes also may be important for these levels because all variation in morphological structures is due to variation in the developmental processes that generate them. Mosimann (1970) and Mosimann and James (1979) pointed out statistical advantages of the log-normal distribution (but see Smith, 1993, for biases in predicted values from allometric regression). In practice, log transformation often renders relations among variables more linear and also can make variances more homogeneous. Finally, log-transformed data are independent of measurement units (e.g., millimeters or the units of an eyepiece micrometer) but retain the information about scale (e.g., lengths versus surfaces). It does not matter which base for the logarithms is chosen, as long as the same base is used consistently throughout a given analysis.

Results of PCAs are estimates of allometric patterns in the populations from which the study samples are drawn. To assess how reliable these estimates are, standard errors or confidence intervals should be calculated. Formulas for these statistics (e.g., Flury and Riedwyl, 1988) are based on the assumption of multivariate normal distribution and on large sample sizes. In most allometric studies, however, the distribution of data cannot be assumed to be multivariate normal. In ontogenetic allometry, for example, the distribution of measurements depends on the age composition in the sample, as well as on the growth dynamics of the structures investigated. In these cases, the bootstrap and jackknife procedures are helpful (an excellent introduction is Efron and Tibshirani, 1993; other useful references are Diaconis and Efron, 1983; Efron and Gong, 1983; Efron and Tibshirani, 1986; Manly, 1991). The bootstrap is a computer-intensive procedure that substitutes repeated sampling from the sample distribution for a theoretical model of that distribution. The only assumption that must be made is that the specimens have been sampled randomly. Applications of the bootstrap to PCA include Diaconis and Efron (1983), Stauffer et al. (1985), Daudin et al. (1988), and Efron and Tibshirani (1993). In multivariate allometry, Gibson et al. (1984) and McGillivray (1985) used the jackknife, whereas Klingenberg and Froese (1991) and Klingenberg and Zimmermann (1992a, b) used the bootstrap. Marcus (1990) compared the jackknife and the bootstrap with each other and with results from large-sample theory.

The fundamental idea of the bootstrap, and the procedures to apply it, follow immediately from the definitions of standard errors and confidence intervals for a statistic q (e.g., mean value or PC coefficients) estimated from a sample of n specimens. Both standard errors and confidence intervals provide answers to the same question: If the same study were repeated numerous times, estimating q from a sample with n specimens each time, how variable would the estimates be? The standard error of a statistic is the standard deviation of

these estimates and a confidence interval is the interval containing a certain percentage of the estimates (e.g., the 95% confidence interval is delimited by the 2.5% and 97.5% quantiles). This is exactly what the bootstrap does, assuming that the sample distribution is representative of the totality of organisms about which statements are made (e.g., all members of a local population or all females of a species). Repeatedly, a "bootstrap sample" of n specimens is drawn randomly, with replacement, from the original sample, and q is estimated for each bootstrap sample. The standard deviation of these estimates is the bootstrapped standard error, and confidence intervals can be derived from the distribution of bootstrap estimates (for details, see Efron and Tibshirani, 1993). About 100 bootstrap samples usually are sufficient to establish standard errors, but at least about 1000 are necessary for confidence intervals (Efron and Tibshirani, 1993). The bootstrap can even be used for hypothesis tests by generating bootstrap replicates of a test statistic under a particular null hypothesis and then comparing the resulting distribution to the test statistic calculated for the observed data (Efron and Tibshirani, 1993; for a related topic, permutation tests, see also Manly, 1991). An advantage of the bootstrap is that it can be adapted to the particular design of a study. For instance, if there are discrete growth stages that can be identified unambiguously (see Fig. 1), there may be no sampling error in the stage composition of the data set (e.g., a design with equal numbers from each stage). The bootstrap procedure for ontogenetic allometry can be adapted by drawing "bootstrap subsamples" from these stages separately. These are then pooled into one bootstrap sample and principal components are calculated (Klingenberg and Zimmermann, 1992a).

ANALYSES OF MULTIPLE GROUPS

Many morphometric studies deal with several groups of specimens, e.g., different sexes, species or ecomorphs. In all these cases, variation within and between groups has to be separated. Otherwise, levels of allometry may be confounded, or within-group variability may invalidate discrimination between groups. Separation of size-related variation within groups from between-group differences has been a traditional topic in morphometrics (Burnaby, 1966; Gower, 1976; Reyment and Banfield, 1976; Pimentel, 1979; Humphries et al., 1981; Thorpe, 1983; Reyment et al., 1984; Bookstein et al., 1985; Rohlf and Bookstein, 1987; Marcus, 1990; Reyment, 1991).

Multivariate comparisons of allometric patterns often focus on the directions of the major axes of scatter ellipsoids in several groups. A straightforward measure for differences between two groups is the angle between their first principal components. For normalized principal components (i.e., squared coefficients sum up to unity), the angle a between components **b** and **c** in two groups is the arc cosine of the inner product of the two vectors,

$$\alpha = \arccos(\Sigma b_i c_i) = \arccos(\mathbf{b}'\mathbf{c})$$

($\mathbf{b}'\mathbf{c}$ is sometimes called vector correlation; note that this is not the correlation between corresponding elements of the two vectors; see also Pimentel, 1979; Bryant, 1984). Angles can even be calculated from published tables of PC coefficients. Applications of angular comparisons include Boitard et al. (1982) and Gibson et al. (1984). Cheverud (1982) and Klingenberg and Zimmermann (1992a) used Monte Carlo simulations of angles between random vectors to assess statistical significance. For the example in this paper, I used the bootstrap to test the more appropriate null hypothesis of equal PC vectors (0° angles).

Another method for comparison among multiple groups of organisms is based on a multivariate ordination of the directions of allometric axes (Solignac et al., 1990; Klingenberg and Froese, 1991; Klingenberg and Spence, 1993). The first principal component of

each group is considered as a data point in the space spanned by the coefficients of the original variables. The vectors of PC1 coefficients are entered as observations in an ordination by a second PCA. The results of this analysis can then be displayed as plots of the "meta-PC" scores. Bootstrap estimates of allometric coefficients can be used to draw confidence ellipses (e.g., Owen and Chmielewski, 1985; Johnson and Wichern, 1988) as a visual indication of statistical accuracy (Klingenberg and Froese, 1991; Klingenberg and Spence, 1993).

Comparisons of allometry within several groups often, but not always, show that the coefficients of the PC1s differ only minimally. In these cases, it may be feasible to use one of the models of common covariance structure (Fig. 5; Airoldi and Flury, 1988; Flury, 1988). These models are based on the assumption that the groups share a common allometric pattern, i.e., that the major axes of their scatter ellipsoids are parallel. Therefore, the differences between the observed PCs of the samples are regarded as effects of sampling error. Principal component analysis, however, is a procedure for analyzing variation in a single sample, and therefore the method needs to be generalized for the context of multiple groups. Usually, the PC1 of the pooled within-group covariance matrix has been used to characterize this common allometric pattern, e.g., in multigroup PCA (Pimentel, 1979; Thorpe, 1983), Burnaby's procedure (Burnaby, 1966; Rohlf and Bookstein, 1987) and in the shearing procedure (Humphries et al., 1981; Bookstein et al., 1985; Rohlf and Bookstein, 1987). Airoldi and Flury (1988) criticized the use of the pooled within-group covariance matrix, because it implicitly assumes that the covariance matrices of all groups are identical (Fig. 5, right panel). They proposed an alternative procedure, common principal component analysis (CPCA), which only assumes that the PCs are common to all groups (see also Flury, 1988; Flury and Riedwyl, 1988). Whereas the directions of the principal axes are assumed to be the same, the amount of variation associated with each PC can vary between groups (Fig. 5, center). Common principal component analysis is available in the NTSYS software package, FORTRAN routines are contained in the IMSL/STAT library (routines KPRIN and DKPRIN), a MATLAB program was written by L. F. Marcus and a SAS/IML version is available from the author. In applications of CPCA to multivariate allometry, the first common principal component (CPC1) is interpreted as an allometric pattern shared by all groups (Airoldi and Flury, 1988; Klingenberg and Zimmermann, 1992a, b; Klingenberg and Spence, 1993).

Discrimination between groups is often difficult because of allometric variation within groups. Especially in organisms with indeterminate growth, the amount of within-group variation may far exceed between-group differences. For instance, fish can increase in size by several orders of magnitude during their life cycle. Other causes, such as nutrition, also contribute to variability within groups (Bernays, 1986; Patton and Brylski, 1987; Meyer, 1990; Smith and Palmer, 1994). Depending on the particular organisms of interest, within-group variation is mostly ontogenetic or mostly static allometry, or a mixture of both.

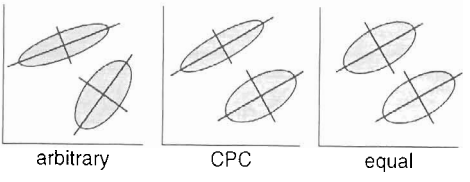

Figure 5. Three levels of similarity between covariance structures. Groups can have arbitrarily different covariance matrices; scatter ellipses then differ both in the directions and lengths of their principal axes. Under the common principal component model (CPC), groups share the same directions of principal axes but may differ in the amount of variation associated with each axis. Groups with equal covariance matrices have the same directions and lengths of principal axes. More details, including additional levels of similarity, are given by Airoldi and Flury (1988) and Flury (1988).

Because most of this variation is often confined to a single dimension, along the allometric axis, it can be removed from an analysis by eliminating the variation in that direction. This approach uses allometry as a criterion of subtraction (Gould, 1975), as it has been done traditionally in bivariate studies. The central assumption of all methods for "size correction," is that the groups share the same allometric vector. If the groups have different allometric patterns, size correction is not possible because all corrections that are suitable for one group will not work in other groups. Several methods have been developed following this principle (Burnaby, 1966; Gower, 1976; Humphries et al., 1981; Thorpe, 1983; Bookstein et al., 1985; Rohlf and Bookstein, 1987). None of these methods, however, can be a substitute for careful examination of the specimens: hidden heterogeneity within the groups, e.g., undetected sex dimorphism or cryptic species, may invalidate the entire analysis.

The procedure proposed by Burnaby (1966) eliminates the effects of growth from multivariate data by projecting data points onto a subspace that is orthogonal to the growth vector (see also Gower, 1976; Rohlf and Bookstein, 1987; Reyment, 1991). This growth-invariant subspace has one dimension fewer than the original space. With three variables, for example, the growth-invariant space is a plane, and for two variables, it is a line (Fig. 6). The growth-adjusted data are coordinates of the projected points, expressed in the coordinate system of the original variables. The growth-invariant data for an n x p data matrix $\mathbf{X}$ and a p x 1 growth vector $\mathbf{b}$ can be obtained as

$$\mathbf{X(I - b(b'b)^{-1}b')},$$

where I is an identity matrix of rank p. With a normalized vector, such as a principal component, the formula simplifies to

$$\mathbf{X(I - bb')}.$$

Usually, the PC1 of the pooled within-groups covariance matrix has been used as the growth vector in Burnaby's procedure (e.g., Reyment and Banfield, 1976; Riska, 1981; Rohlf and Bookstein, 1987). Because CPCA is based on less-stringent assumptions (see above),

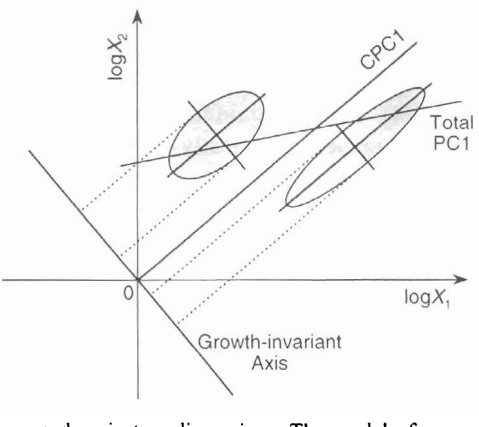

Figure 6. Burnaby's (1966) procedure in two dimensions. The model of common principal components is appropriate, because the principal axes of the scatter ellipses in the two groups are parallel. Therefore, the groups share a common axis of allometric growth, the CPC1. All the variation is projected onto an axis perpendicular to the CPC1 by setting the CPC1 scores to zero. Subsequent analyses focus on between-group differences in this "growth-invariant" axis. The PC1 of the combined samples ("Total PC1") confounds variation within and between groups.

the CPC1 seems preferable as an estimate of a common allometric pattern. This version of Burnaby's technique is equivalent to a procedure involving three consecutive steps: (1) a rigid rotation to the common principal components, (2) setting the CPC1 scores of each data point to zero, and (3) rigid rotation back to the original coordinate system.

For analyses of growth-adjusted data, it may be more convenient to omit step (3). First, common principal components are computed as an estimate of within-group variation, and the CPC scores, except for the CPC1, are used as variables in subsequent analyses. The results of these analyses, e.g., discriminant analysis or MANOVA, are identical to the results based on data adjusted by Burnaby's original procedure, but within-group covariance matrices are of full rank. This technique is almost identical to the one proposed by Thorpe (1983; the PC1 of the pooled within-groups covariance matrix as an estimate of the allometric axis), which has been used in numerous studies (e.g., Wiig, 1985; Thorpe and Baez, 1987; Corti et al., 1988; Lessa and Patton, 1989).

Interpreting the values resulting from Burnaby's procedure is somewhat difficult. With some caution, they can be seen as measurements adjusted to "unit size." This is possible for most morphometric data sets because the CPC1 has high positive correlations with all variables, and therefore can be interpreted as an overall size variable. Setting the CPC1 score for log-transformed measurements to zero corresponds to setting the untransformed value of overall size to one. The adjusted variables therefore allow one to compare organisms of different sizes by rescaling all measurements allometrically to unit size. For example, in Fig. 6 the group to the left has higher X_2 and lower X_1 values at corresponding sizes than does the group to the right. The difference between this approach and comparisons of ratios or geometric shape is that Burnaby's procedure takes into account the shape changes caused by allometric growth.

Another way to understand Burnaby's procedure is by analogy to the most familiar method of size correction in bivariate allometry, regression residuals. Regression residuals are an appropriate way of correcting for size if one is interested in the relationship of a dependent variable y (e.g., an ecological or physiological parameter) to a size variable x known a priori and measured independently (e.g. body weight). Then, the residuals are the deviations of actual measurements from the value expected for an "average" specimen of that particular size. These deviations are computed in the direction of the y axis by subtracting the y value estimated by regression from the observed value. Hence, the interpretation of allometry as a "criterion of subtraction" (Gould, 1975) can be taken literally. In multivariate allometry, however, there is no a priori size variable that can be measured independently from the other variables. All measurements are affected simultaneously by overall size, which only can be estimated from them. As I explained above, the PC1 is a good choice for such an estimate. The "residuals" are subsequent PCs, which are perpendicular to the PC1 (step [2] above). In bivariate regression, comparison of intercepts of several groups makes sense only when the slopes in all groups are equal. Analogously, Burnaby's procedure only works if all the groups share a common allometric pattern, as it is estimated by the CPC1. (For groups that differ in their growth vectors, Burnaby [1966] suggested removing all growth vectors from the data; after this, however, there may not be much meaningful variation left to study [e.g., Humphries et al., 1981].)

Because it includes only rigid rotations, Burnaby's procedure conserves the spatial relationships among data points in all directions except that of the growth vector. Therefore, data adjusted by this technique can be used to quantify variation perpendicular to the growth trajectories. For instance, lateral transpositions of growth trajectories (also called vertical transpositions) can be distinguished from group differences produced by shifts along the growth axis (ontogenetic scaling, Gould, 1975; Shea 1985, 1992). These two types of group differences, as well as within-group variation, are confounded in PCA of combined samples (Fig. 6; for discussion, see Voss et al., 1990; Voss and Marcus, 1992; Klingenberg and Spence,

1993). Burnaby-adjusted data also can be used to examine within-group deviations from simple ontogenetic allometry, such as curvatures of the growth trajectories. Klingenberg and Spence (1993) used a MANOVA of Burnaby-adjusted data to separate lateral transposition from nonallometric growth in a data set containing samples of six discrete ontogenetic stages from each of six waterstrider species (i.e., a data structure corresponding to Fig. 1). The PC1 scores of the between-species matrix can be used to display lateral transposition, and likewise the between-stage matrix for nonallometric growth.

Procedures for size correction in several groups all assume that the groups share the same allometric pattern. But what if that is not true? In this case, the formal procedures fail, and one has to find a visual and at best "semistatistical" way to assess group differences. A very useful technique of this kind is the "tomographic representation" introduced by Boitard et al. (1982). First, they calculated the PCs for the covariance matrix of the pooled data. Then, they plotted the second versus third PC scores separately for data points grouped by the data points PC1 scores. Finally, they combined these layers into a plot showing the scatter ellipsoids "suspended" by their within-group PC1 axes in a box representing the space of the first three PCs of the pooled data.

EXAMPLE: GEOGRAPHIC VARIATION

Some of the techniques described above are applied in a simple example, a morphometric data set taken from a larger study on geographic variation in the waterstrider *Gerris costae* (Klingenberg, 1992). Only adult specimens were measured, and because these bugs do not grow after they reach the adult stage, I made no attempt to correct for size in the original study. Here, using allometric techniques, I reanalyze a part of the data.

Three samples, each representing a different subspecies, are included here: the nominate subspecies *G. c. costae* is represented by a sample from the Swiss Alps ($n = 32$), *G. c. fieberi* is represented by a sample from northern Greece ($n = 33$), and *G. c. poissoni* by a sample from the eastern Pyrenees in France ($n = 28$). In this example, I consider only adult males. Allometric variation within the three samples is therefore purely static allometry.

Four measurements were chosen for the example: (1) total thorax length, (2) the length of the first antennal segment, (3) the middle femur length and (4) the hind femur length. The raw data are presented in the Appendix. Data were transformed to natural logarithms. For the convenience of presentation, variances and covariances are multiplied by 10^4 (this also applies to eigenvalues). The covariance matrices are

$$\begin{bmatrix} 12.67 & 8.39 & 8.75 & 11.09 \\ 8.39 & 14.34 & 9.04 & 11.61 \\ 8.75 & 9.04 & 12.05 & 13.63 \\ 11.09 & 11.61 & 13.63 & 17.83 \end{bmatrix}$$ for the sample from the Alps,

$$\begin{bmatrix} 5.55 & 2.07 & 2.53 & 4.06 \\ 2.07 & 9.66 & 4.03 & 4.67 \\ 2.53 & 4.03 & 7.82 & 9.02 \\ 4.06 & 4.67 & 9.02 & 12.98 \end{bmatrix}$$ for the sample from Greece, and

$$\begin{bmatrix} 4.93 & 5.97 & 3.90 & 5.16 \\ 5.97 & & 6.78 & \\ 3.90 & & 6.94 & \\ 5.16 & 8.35 & 7.56 & 10.70 \end{bmatrix}$$ for the sample from the Pyrenees.

As in many morphometric data sets, all covariances are positive and most of them are relatively high. There are, however, some differences between groups and among variables. The sample from the Alps is more variable than the other two samples. Whereas

the two femur lengths are highly correlated in all groups, correlations involving thorax length and the first antennal segment tend to be lower and more variable (e.g., the correlation between these two measurements is only 0.28 in the Greek sample).

Principal components were computed for each sample, and parametric standard errors for the estimates of PC coefficients and eigenvalues were calculated using the formulas given by Flury (1988). Moreover, standard errors were also determined with a bootstrap procedure. For each group, 1000 bootstrap samples were randomly drawn (with replacement), and principal components were computed for each. (This number of replications is more than actually needed for standard errors; Efron and Tibshirani, 1993.) The standard deviations of these 1000 estimates of the PC coefficients or eigenvalues are their bootstrap standard errors.

The amount of variance (± parametric and bootstrapped standard errors) for which PC1 accounts is 46.1 (± 11.7, 12.5) or 81% of total variance in the sample from the Alps; for 24.1 (± 6.0, 4.5), or 67% in the Greek sample; and for 31.4 (± 8.5, 7.2), or 76% in the sample from the Pyrenees. Standard errors determined by the two approaches are similar in magnitude, although the bootstrap standard error for the Greek sample differs from the parametric estimate by about 25%. The percentages of total variance taken up by the PC1 are quite typical for static allometry (e.g. Cuzin-Roudy, 1975; Klingenberg and Zimmermann, 1992a). The estimated coefficients of the PC1 and their parametric and bootstrapped standard errors are

$$\begin{bmatrix}0.441\\0.471\\0.477\\0.597\end{bmatrix}, \begin{bmatrix}0.057\\0.061\\0.036\\0.034\end{bmatrix}, \text{and } \begin{bmatrix}0.060\\0.055\\0.032\\0.030\end{bmatrix} \text{ for the sample from the Alps,}$$

$$\begin{bmatrix}0.269\\0.409\\0.527\\0.695\end{bmatrix}, \begin{bmatrix}0.085\\0.119\\0.045\\0.054\end{bmatrix}, \text{and } \begin{bmatrix}0.093\\0.113\\0.045\\0.060\end{bmatrix} \text{ for the Greek sample, and}$$

$$\begin{bmatrix}0.315\\0.695\\0.400\\0.507\end{bmatrix}, \begin{bmatrix}0.050\\0.082\\0.057\\0.070\end{bmatrix}, \text{and } \begin{bmatrix}0.037\\0.057\\0.078\end{bmatrix} \text{ for the sample from the Pyrenees.}$$

The estimates of PC coefficients are fairly stable, as indicated by the relatively small standard errors. Bootstrap standard errors agree well with parametric estimates. The coefficient values clearly do not conform to the pattern for overall isometry, where all coefficients would be equal to $p^{-0.5}$, i.e., the vector [0.5, 0.5, 0.5, 0.5]'. This is confirmed by Anderson's test (Anderson, 1963; Pimentel, 1979:70; Flury, 1988:34), which was significant for all three groups (Alps: $\chi^2 = 11.52$, $P < 0.01$; Greece: $\chi^2 = 20.37$, $p < 0.001$; Pyrenees $\chi^2 = 23.82$, $p < 0.0001$; df = 3 in each case).

Coefficients vary considerably between samples. In the sample from the Alps, the middle femur (third measurement) is almost isometric relative to thorax length (first measurement), as the ratio of their coefficients is 0.477 / 0.441 = 1.08. In the Greek sample, however, the corresponding ratio is 0.527 / 0.269 = 1.96, which indicates strong positive allometry of the middle femur relative to the thorax. Whereas the lengths of the first antennal segment and of the middle femur (second and third measurements) are almost isometric in the sample from the Alps, the middle femur shows positive allometry relative to the first antennal segment in the Greek sample, and negative allometry in the Pyrenees. As an overall measure of these differences, I computed the angles between PC1 axes of different groups;

they are 12.22° between Alps and Greece, 16.29° between Alps and Pyrenees, and 21.19° between Greece and the Pyrenees.

If we expect a common allometric pattern, these angles seem quite large. But are the differences from zero statistically significant? Because there is no parametric test for the angles between PCs, I used the bootstrap approach to test the null hypothesis that the common principal component model holds. The bootstrap test procedure is straightforward: use the data to generate a modified data set that conforms to the null hypothesis, then repeatedly draw bootstrap samples and compute the test statistic, and finally compare the empirical distribution of the test statistic to the value calculated for the original data (Efron and Tibshirani, 1993). To produce a data set conforming to the CPC model, I rotated the data in each sample so that the within-group PC axes are aligned exactly with the CPC axes. This is easy to do because the matrix of PC coefficients **B** can be used to rotate the data points from the original coordinate system to the PC coordinates $\mathbf{Y} = \mathbf{XB}$, whereas the transpose of **B** performs the reverse rotation $\mathbf{YB'} = \mathbf{X}$. A data set conforming to the CPC model can be obtained by using each group's own PC coefficients for the first rotation, but the coefficient matrix from a CPCA for the reverse rotation. Thus, the modified data in the i-th group are $\mathbf{X^*}_i = \mathbf{X}_i\mathbf{B}_i\mathbf{B'}_{CPC}$, where $\mathbf{B}_i$ is the matrix of within-group PC coefficients, and $\mathbf{B}_{CPC}$ is the matrix of CPC coefficients (using the PC scores for each group, $\mathbf{Y}_i = \mathbf{X}_i\mathbf{B}_i$ is equivalent). For this test, 5000 bootstrap replications were performed. For each replication, bootstrap samples were drawn from the modified data of all three groups, and the angles between the PC1 axes of the three groups were computed. The 95% quantiles of the angles are 18.56° between the samples from Alps and from Greece, 16.32° between the Alps and the Pyrenees, and 20.59° between the Pyrenean and the Greek samples. Therefore, the two angles that involve the sample from the Pyrenees seem to indicate a significant difference from the Greek sample and a borderline significance for the difference from the Alps. Because there are three comparisons, however, the chance of rejecting a true null hypothesis by chance at an α of 0.05 is $1 - (1 - \alpha)^3 = 1 - 0.86 = 0.14$. The Bonferroni technique can be used to adjust the significance level for m individual comparisons to a new significance threshold at $(1 - \alpha/m) =$ 98.33%. None of the comparisons was significant at this adjusted level (98.33% quantiles are 22.23, 20.63, and 25.48, respectively). Therefore, the data seem to be consistent with a model of an allometric pattern that all three groups have in common.

This preliminary conclusion justifies using common principal components to estimate an allometric pattern for all three groups simultaneously. Again, standard errors were estimated both by using formulas based on large-sample theory (Flury, 1988) and by bootstrapping with 1000 bootstrap iterations. The estimates of the CPC1 coefficients with their parametric and bootstrapped standard errors are

$$\begin{bmatrix} 0.361 \\ 0.510 \\ 0.482 \\ 0.614 \end{bmatrix}, \begin{bmatrix} 0.034 \\ 0.047 \\ 0.024 \\ 0.028 \end{bmatrix}, \text{ and } \begin{bmatrix} 0.039 \\ 0.053 \\ 0.024 \\ 0.031 \end{bmatrix}, \text{ respectively.}$$

Although in a CPC model there is only one set of eigenvectors, which is shared by the groups, each group has its own eigenvalues. The variance taken up by the CPC1 (± parametric and bootstrap standard errors) in each sample is 45.8 (± 11.4, 12.2) for the Alps; 23.6 (± 5.8, 4.7) for Greece, and 30.0 (± 8.0, 7.4) for the Pyrenees. This corresponds to 80.4%, 65.7% and 72.8% of the total variance in the respective samples. These values are only a little lower than corresponding values from the separate one-group PCAs of each sample; one common allometric pattern can account for almost as much of the variation as the PC1 of each group separately. Further support for a common model comes from the log-likelihood ratio test of the CPC model (Flury, 1988; Airoldi and Flury, 1988), which does

not show a significant difference from the unrestricted model in which every group has its own PCs ($\chi^2 = 10.55$, df = 12, $p = 0.57$). For the same χ^2 statistic, the bootstrap test with 5000 iterations yielded a 95% quantile of 20.90, which agrees nicely with the corresponding value of 21.03 in statistical tables (the same bootstrap replicates as for the angles between one-group PCs, above). The angles between the CPC1 axis and the one-group PC1 axes indicate that the CPC estimate is a "compromise" between the one-group PCs (5.2° for Alps; 9.4° for Greece; 13.5° for Pyrenees). In the bootstrap test, none of these angles exceeded the 98.33% quantiles (Bonferroni adjustment for $\alpha = 0.05$; quantiles are 9.85°, 18.25°, and 17.75°, respectively). Therefore, it is reasonable to assume that the three groups share the same allometric pattern and that the differences between the PC1 estimates of individual groups are due to sampling error.

To examine whether the differences between groups are simply extensions of within-group variation, I used Burnaby's approach. "Size-invariant" variation between groups was analyzed by a MANOVA of CPC scores, omitting the CPC1. The first eigenvector of the resulting between-groups matrix indicates the axis that contains the most variation among groups, subject to being perpendicular to the within-group allometric axis. Because there are three groups, the matrix of between-groups sums of squares (mean squares would yield equivalent results) has only two non-zero eigenvalues. The first eigenvector of this matrix accounted for 82% of the total between-group variation, and it therefore summarizes most of the differences between samples after adjusting for size.

Fig. 7 is a plot of this axis of group differences versus the CPC1 scores. The Greek specimens differ from the Alpine ones mainly by their higher CPC1 scores, which indicate greater overall size. Therefore, the bugs of the Greek sample can be seen as "scaled-up" versions of their counterparts from the Alps, corresponding to intraspecific scaling along the axis of static allometry (Gould, 1975; although these data deal with subspecies, and not with different species, I do not think it would be helpful to coin a new term). The differences between the samples from the Alps and Pyrenees are largely unrelated to within-group variation. These conclusions are also supported in a more quantitative way by the Mahalanobis distances between the three groups. The Mahalanobis D^2 value, computed from the second to fourth CPC scores, is clearly smaller between the samples from the Alps and from Greece

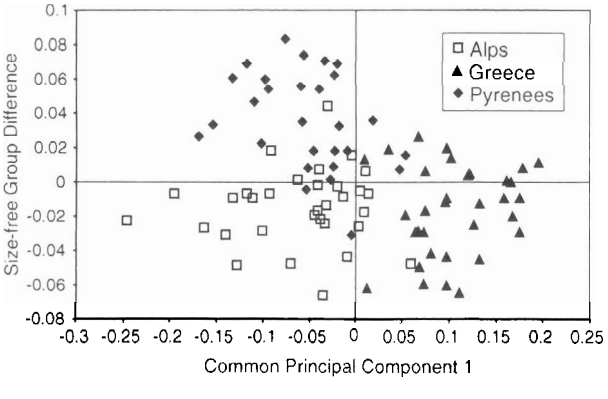

Figure 7. Geographic variation in the waterstrider *Gerris costae*. The first common principal component of the three samples (abscissa) can be interpreted as a measure of overall size. It indicates that Greek specimens are considerably larger than those from the Alps or the Pyrenees. The vertical axis is the PC1 of the between-group matrix from a MANOVA of "size-free" data (the scores of CPC2–CPC4) and summarizes differences between groups independent of within-group allometry. It shows that the samples from the Alps and Pyrenees are separated fairly well, although they are of similar overall size.

($D^2 = 2.38$) than between either of these and the sample from the Pyrenees (Alps versus Pyrenees $D^2 = 6.55$; Greece versus Pyrenees $D^2 = 4.92$). These values, however, also show that scaling does not account for all the difference between the specimens from the Alps and from Greece.

The bootstrap technique has several advantages for assessing statistical accuracy and even for hypothesis testing in morphometric analyses. First, it does not require that the data conform to any particular probability distribution as other techniques do. Nevertheless, assumptions about distributions can be incorporated in the simulations (parametric bootstrap; Efron and Tibshirani, 1993). For the present example, the parametric bootstrap for a multivariate normal CPC model gave results similar to the nonparametric results presented above. Second, the bootstrap can be used for any test statistic, even if its statistical properties are unknown. In the example, I extensively used the angles between PC axes, because angles are particularly intuitive as a measure of overall similarity for allometric vectors. Moreover, the bootstrap can be adapted easily to a variety of experimental designs or hypothesis tests. For the bootstrap test of the CPC model, rotation of the original data was sufficient to generate a data set conforming to the null model. With tools such as the singular value decomposition (e.g., Marcus, 1993), a variety of other null models could be simulated with real data.

These advantages have a cost, however, as the bootstrap technique is based on massive amounts of numerical calculation. As computers become faster and cheaper, this may not be a serious problem except for extremely large or complex data sets. For this example, all bootstrap analyses were done by a personal computer with a 486/50MHz processor, and using SAS/IML software (version 6.08). Although the number of bootstrap replications was substantial, the computation time was moderate. For standard errors in one-group PCA, with 3000 analyses (1000 bootstrap replications for three groups), the entire bootstrap procedure took less than 1.5 minutes. The 1000 bootstrap replicates for CPCA took about 45 minutes, much longer than for ordinary PCA, because the computational procedure for CPCA is more complex. The most effort was required for the bootstrap test: the 5000 iterations, each with three one-group PCAs and CPCA, took a little more than 3 hours. The high number of bootstrap replications for the test was necessary because one is interested in the tails of the test statistic's empirical distribution (Efron and Tibshirani, 1993). The computational effort used for this example shows that the bootstrap is a reasonable option, even with a personal computer and for relatively complex problems.

ALTERNATIVE APPROACHES IN ALLOMETRY

In the preceding sections, I presented allometry in a pragmatic way, extending the familiar logic of two-dimensional scatter plots to a multivariate context. This approach focuses on the patterns of variation by determining amount, dimensionality and direction of morphometric variation in the space of log-transformed variables.

Some readers may have noticed that I only used the words "size" and "shape" in a rather informal way, although they are the central concepts for other approaches in morphometrics (Bookstein et al., 1985; Bookstein, 1989, 1991, 1993; Rohlf, 1990; Rohlf and Marcus, 1993). In studies based on the approach described above, "size" and "shape" may appear in interpretations of the results, but they are not parts of the analyses themselves. The analyses are exploratory or they test simple hypotheses about the structure of variation, such as whether or not the scatter ellipsoids of several groups have major axes that are parallel. Principal components, used in many allometric studies, can "account for" or "take up" variation, but do not "cause" or "explain" it. Interpretation and explanation are extrinsic to the analyses, and they consist of arguments about biological processes producing the observed patterns of variation, e.g., growth dynamics or evolutionary constraints.

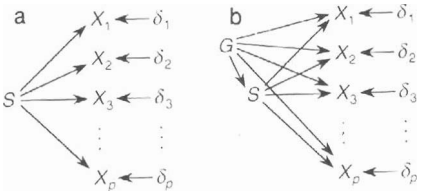

Figure 8. Path models in allometry. (a) Path diagram for simple allometry in a single group. All covariation among variables is caused by a general size factor *S*. Residual variation (δ) is uncorrelated among variables. (b) Allometric variation in two groups. Both groups share the same general size factor and, therefore, also the same within-group covariance structure. Group differences are determined by the group factor *G* in two different ways: directly as group shape differences (arrows from *G* to the variables), or as group size difference via the general size factor (arrow from *G* to *S*).

A very different framework underlies the factor analytic approach, which starts with an explicit model of the origin of variation in measurement variables (see Bookstein et al., 1985; Bookstein, 1989, 1991). Factors are formally included in the analysis as causes of covariation among several morphometric variables; explanation is therefore an intrinsic part of the analysis. A path model is constructed according to biological knowledge and specifies a hypothesis of relationships between factors and observed variables (Wright, 1968; Bookstein et al., 1985; Loehlin, 1987; Zelditch, 1987; Marcus, 1990; Bookstein, 1991). In a path model of allometry, general size is a factor, or latent variable, simultaneously affecting all morphometric measurements and causing the covariance among them (Fig. 8a). As a latent variable, size cannot be measured directly, but it can be estimated. This is usually done using the within-group PC1 (Bookstein et al., 1985; Rohlf and Bookstein, 1987).

Hopkins (1966) proposed a similar factor model of allometry. The observed covariance matrix **S** of log-transformed characters is composed of two parts, $\mathbf{S} = \mathbf{T} + \mathbf{D}$. The matrix **T**, which reflects systematic covariance, is of rank one. Therefore, it has only one principal component, which corresponds to the allometric axis, or to the general size factor in Fig. 8a. The PC coefficients of **T** (which cannot be observed) are proportional to the factor loadings of general size. The matrix **D** stands for the residual variation, which is assumed to be uncorrelated among variables; therefore, **D** is a diagonal matrix. The structure of **D** is crucial for the choice of methods to estimate the parameters of the model. The PC1 of **S** is only appropriate to estimate the size factor if the diagonal elements of **D** are equal (Hopkins, 1966); otherwise, the size factor should be estimated from the off-diagonal elements (see Bookstein, 1991). If all variables are highly correlated to each other and **D** therefore makes only a minor contribution to the total covariance matrix **S**, as in many morphometric data sets, the PC1 is a reasonable estimator.

If there is more than one group of specimens, one or more additional factors for group differences (*G* in Fig. 8b) explain differences between two or more groups (Bookstein et al., 1985; Rohlf and Bookstein, 1987; Bookstein, 1991). Group factors can affect the measurements through general size or in a size-invariant manner. Group size differences (arrow from *G* to *S* in Fig. 8b) cause shifts along the growth axis and correspond to ontogenetic scaling, whereas size-invariant differences (arrows from *G* to the variables in Fig. 8b) correspond to lateral shifts of trajectories (Shea, 1985, 1992; Klingenberg and Spence, 1993). This is the path model for the shearing procedure, which was originally introduced by Humphries et al. (1981) and later reformulated by Bookstein et al. (1985), Rohlf and Bookstein (1987) and Bookstein (1991). The main purpose of this procedure is to obtain factor loadings interpretable as path coefficients (Rohlf and Bookstein, 1987) rather than to obtain an ordination. The geometric basis for the shear is more complex than for Burnaby's procedure because it

is not just a rigid rotation (see Humphries et al., 1981; Bookstein et al., 1985; Rohlf and Bookstein, 1987). As a consequence, it does not conserve the spatial relationships among data points and is difficult to use, e.g., to quantify lateral transposition of growth trajectories. Applications of the shear include Bookstein et al. (1985), Strauss (1985), Voss et al. (1990), and Voss and Marcus (1992).

Several studies have used factor analysis to investigate more complex models of correlation or covariance structure among morphometric variables (Bookstein et al., 1985; Zelditch, 1987, 1988; Zelditch and Carmichael, 1989; Marcus, 1990). In addition to general size, these models include factors explaining joint variation in groups of variables that are developmentally or functionally related. An alternative procedure (Cowley and Atchley, 1990; Paulsen and Nijhout, 1993) uses hypotheses about relations among characters to predict the pattern of a correlation matrix, and then compares these to observed correlation matrices by means of randomization tests (Cheverud et al., 1989; Manly, 1991). These models, however, are beyond the scope of allometry.

Yet another, more general method of characterizing size was introduced by Mosimann (1970), when he defined standard size variables. Any positive, real-valued function $G(\mathbf{x})$ of a vector of measurements $\mathbf{x}$ is a standard size variable, if multiplication of each measurement by a constant a results in an a-fold value of the size function, i.e., $G(a\mathbf{x}) = aG(\mathbf{x})$. This condition ensures that the variable scales as a linear dimension. A standard size variable transformed to logarithms is called a log-size variable. A class of log-size variables important for multivariate allometry is defined as linear combinations of log-transformed measurements, i.e., $\log G(\mathbf{x}) = \Sigma b_i (\log x_i)$, with $\Sigma b_i = 1$ (Mosimann and James, 1979; Darroch and Mosimann, 1985). Rescaling the PC coefficients for ontogenetic allometry so that they sum up to unity yields a log-size variable that indicates each specimen's position along the growth trajectory (Klingenberg and Zimmermann, 1992b; Klingenberg and Spence, 1993). Similar measures of size, but without rescaling, were used by Creighton and Strauss (1986), Strauss (1990b) and Voss and Marcus (1992), among others.

If size alone is of interest, the choice of a size measure often does not matter very much. In many morphometric data sets, the variables and the size measures derived from them are highly correlated among one another. Hence, different size measures may produce different scaling factors, due to allometry, but they will yield basically the same ordering from small to large specimens and similar size differences between them. In studies of allometry based on either PCA or factor analysis, "shape" often does not appear explicitly at all, or if it does, it is used in a sense very different from everyday language (e.g., Bookstein, 1989). In those cases, "shape" is usually (though rarely explicitly) defined as "everything that is not size." It is through this notion of "shape" that the choice of a size measure matters for morphometric studies: because "size" often takes up a large fraction of the total variation in a data set, relatively small changes in the size measure produce proportionally large changes in what remains after "size" is removed.

In multivariate allometry based on PCA, the second and subsequent PCs often are interpreted as "shape scores." Nevertheless, they do not reflect a geometric concept of shape. If two specimens, differing in size, have the same "shape scores," they can be interpreted as geometrically similar only if the corresponding size vector is isometric; otherwise, there are allometric changes in shape (see also Bookstein, 1989). To separate size and geometric shape, Somers (1986, 1989) proposed a size-constrained version of PCA in which variation in the direction of an isometric vector is removed first. Unfortunately, Somers used the correlation instead of the covariance matrix, thereby removing not only isometric size from the data but also all allometric variation (see above). As an alternative procedure to achieve Somers's original objective, Burnaby's procedure can be used to eliminate a vector representing isometric variation, which is mathematically equivalent to performing a PCA on the covariance matrix of doubly centered data (Somers, 1989) or on the "principal components of

shape" proposed by Darroch and Mosimann (1985; see also Jungers et al., 1988). Notice, however, that removing isometric size adjusts for only variability in overall size itself, but not for any size- or age-related shape variation. For example, although Barbie dolls are smaller than are many other dolls clearly representing infants, it is easily recognizable that Barbie dolls are modeled after human adults.

Mosimann (1970) presented a definition of a shape space based on geometric similarity. Each vector of measurements $\mathbf{x}$, divided by a standard size variable $G(\mathbf{x})$, constitutes a measure of shape. Darroch and Mosimann (1985) developed principal components and canonical variates for the space of these shape measures and applied them to two examples. Further applications are found in Mosimann and James (1979) and Jungers et al. (1988). In this framework, allometry exists if variations in size and shape are associated; isometry means that variation in size and shape are statistically independent. Mosimann's theory of size and shape links some aspects of multivariate allometry to the landmark-based methods of geometric morphometrics (Bookstein, 1989, 1991, 1993). The underlying concept of allometry, however, differs fundamentally from the other frameworks presented here, as it abandons the straight-line relationship among log-transformed variables, which is the basis of allometry as devised by Huxley (1932). In this point, Mosimann's concept is closer to the much broader notion of allometry adopted by Gould (1966), who characterized it as "the study of size and its consequences."

Whereas Mosimann's (1970) approach, although based on considerations of geometric similarity, still uses vectors of length measurements, geometric morphometrics goes one step further and analyzes shape as geometric configurations of morphological landmarks (e.g., Rohlf, 1990; Bookstein, 1991, 1993; Rohlf and Marcus, 1993; other chapters in this volume). The strong emphasis on shape in geometric morphometrics is reflected in two recent definitions of morphometrics, characterizing it as "the quantitative description, analysis and interpretation of shape and shape variation in biology" (Rohlf 1990) and as "the geometrically reified description of effects on geometric shape" (Bookstein, 1993) without even mentioning size. Clearly, geometric morphometrics presents a dramatically different framework for allometry. Variation in size is removed from the data (by the two-point registration for Bookstein's shape coordinates or by standardizing for centroid size; Bookstein, 1991; Rohlf, 1993), and shape changes alone are included in the analysis. Allometry can be assessed by combining the results from shape analysis with additional information, either directly by nonlinear regression of "shape scores" (relative warps, shape coordinates, or Procrustes residuals) on a measure of size (Bookstein, 1991; Walker, 1993), or by subdividing specimens into size classes (MacLeod and Kitchell, 1990) or age groups (Reilly, 1990; Bookstein, 1991; Zelditch et al., 1992) and comparing their mean shapes. Allometric variation over an extended size range, as it occurs in many growth studies, often leads to highly nonlinear trajectories (Bookstein, 1991; Zelditch et al., 1992; Walker, 1993).

The choice of methods for a particular study depends on what questions a study is supposed to answer. The results obtained from analyses using the geometric methods can be interpreted directly in terms of shape. Here, "shape" is used in its intuitive sense, meaning a geometric configuration. The disadvantage of these methods, however, is the complexity of allometric relations. For example, a procedure analogous to Burnaby's technique to adjust for shape differences due to allometric growth would have to use nonlinear regression of shape measures on overall size. On the other hand, the results of analyses that use distance data are more difficult to describe in everyday language; graphical displays, like Fig. 7 are abstractions rather than pictures of real organisms. If the notion of "shape" is used at all, it denotes the relative sizes of parts of the organism. The advantage of methods using log-transformed distances is that these data often fit linear models due to their relationship to growth dynamics, which was used by Huxley (1932) to justify his formula for simple allometry.

From an extreme point of view, the configuration of morphological landmarks of an organism could be considered as merely an epiphenomenon of the growth processes affecting the tissues between the landmarks. Ideally, therefore, morphometric methods should be based on models of biological processes rather than geometrical or statistical considerations (e.g., Sattler, 1992). Although this view is correct in principle, our knowledge of the mechanisms involved in developmental processes is incomplete even for simple and well-studied experimental systems (e.g., Atchley and Hall, 1991). For less well-known organisms and for more complex problems, such as evolutionary comparisons, landmark configurations or length measurements must be used as the basis for our understanding of organismic form.

ACKNOWLEDGMENTS

I thank Michel Baylac, Marco Corti, Bernard Flury, Les Marcus, Rich Palmer, John Spence, Manfred Zimmermann and various participants of the Advanced Study Institute for discussions about multivariate allometry and constructive comments on earlier versions of this paper. Financial support was provided by NATO, a University of Alberta Ph.D. Scholarship, an Izaak Walton Killam Scholarship, and a Mary Louise Imrie Graduate Award (Faculty of Graduate Studies and Research, and the Vice-President [Research] of the University of Alberta) to the author and by grants from the Natural Sciences and Engineering Research Council of Canada to John R. Spence.

REFERENCES

Airoldi, J.-P., and B. K. Flury. 1988. An application of common principal component analysis to cranial morphometry of *Microtus californicus* and *M. ochrogaster* (Mammalia, Rodentia). Journal of Zoology (London) 216:21–36.

Anderson, T. W. 1963. Asymptotic theory for principal component analysis. Annals of Mathematical Statistics 34:122–148.

Anstey, R. L. 1987. Astogeny and phylogeny: Evolutionary heterochrony in Paleozoic bryozoans. Paleobiology 13:20–43.

Arnold, S. J. 1992. Constraints on phenotypic evolution. American Naturalist 140:S85–S107.

Atchley, W. R., and B. K. Hall. 1991. A model for development and evolution of complex morphological structures. Biological Reviews 66:101–157.

Bernays, E. A. 1986. Diet-induced head allometry among foliage-chewing insects and its importance for graminivores. Science 231:495–497.

Blackstone, N. W. 1987. Specific growth rates of parts in a hermit crab *Pagurus longicarpus*: A reductionist approach to the study of allometry. Journal of Zoology (London) 211:531–545.

Boag, P. T. 1984. Growth and allometry of external morphology in Darwin's finches (*Geospiza*) on Isla Daphne Major, Galapagos. Journal of Zoology (London) 204:413–441.

Boitard, M., J. Lefebvre, and M. Solignac. 1982. Analyse en composantes principales de la variabilité de taille, de croissance et de conformation des espèces du complexe *Jaera albifrons* (Crustacés Isopodes). Cahiers de Biologie Marine 23:115–142.

Bookstein, F. L. 1989. "Size and shape": A comment on semantics. Systematic Zoology 38:173–180.

Bookstein, F. L. 1991. Morphometric tools for landmark data: Geometry and biology. Cambridge University Press: Cambridge.

Bookstein, F. L. 1993. A brief history of the morphometric synthesis. Pages 15–40 *in* L. F. Marcus, E. Bello, and A. Garcia-Valdecasas, (eds.), Contributions to morphometrics. Monografías del Museo Nacional de Ciencias Naturales 8, Madrid.

Bookstein, F. L., B. Chernoff, R. L. Elder, J. M. Humphries, Jr., G. R. Smith, and R. E. Strauss. 1985. Morphometrics in Evolutionary Biology: The geometry of size and shape change, with examples from fishes. Academy of Natural Sciences of Philadelphia Special Publication 15.

Bryant, P. 1984. Geometry, statistics, probability: Variations on a common theme. American Statistician 38:38–48.

Burnaby, T. P. 1966. Growth-invariant discriminant functions and generalized distances. Biometrics 22:96–110.

Buss, L. W., and N. W. Blackstone. 1991. An experimental exploration of Waddington's epigenetic landscape. Philosophical Transactions of the Royal Society of London B 332:49–58.

Cane, W. P. 1993. The ontogeny of postcranial integration in the common tern, *Sterna hirundo*. Evolution 47:1138–1151.

Cheverud, J. M. 1982. Relationships among ontogenetic, static, and evolutionary allometry. American Journal of Physical Anthropology 59:139–149.

Cheverud, J. M., G. P. Wagner, and M. M. Dow. 1989. Methods for the comparative analysis of variation patterns. Systematic Zoology 38:201–213.

Cock, A. G. 1966. Genetical aspects of metrical growth and form in animals. Quarterly Review of Biology 41:131–190.

Corner, B. D., and J. T. Richtsmeier. 1991. Morphometric analysis of craniofacial growth in *Cebus apella*. American Journal of Physical Anthropology 84:323–342.

Corti, M., R. S. Thorpe, L. Sola, V. Sbordoni, and S. Cataudella. 1988. Multivariate morphometrics in aquaculture: A case study of six stocks of the common carp (*Cyprinus carpio*) from Italy. Canadian Journal of Fisheries and Aquatic Sciences 45:1548–1554.

Cowley, D. E., and W. R. Atchley. 1990. Development and quantitative genetics of correlation structure among body parts of *Drosophila melanogaster*. American Naturalist 135:242–268.

Cowley, D. E., and W. R. Atchley. 1992. Quantitative genetic models for development, epigenetic selection, and phenotypic evolution. Evolution 46:495–518.

Creighton, G. K., and R. E. Strauss. 1986. Comparative patterns of growth and development in cricetine rodents and the evolution of ontogeny. Evolution 40:94–106.

Cuzin-Roudy, J. 1975. Étude de la variabilité et de l'allométrie de taille chez *Notonecta maculata* Fabricius (Insectes, hétéroptères), par les méthodes classiques et par la méthode des composantes principales. Archives de Zoologie Expérimentale et Générale 116:173–227.

Cuzin-Roudy, J., and Ph. Laval. 1975. A canonical discriminant analysis of post-embryonic development in *Notonecta maculata* Fabricius (Insecta: Heteroptera). Growth 39:251–280.

Darroch, J. N., and J. E. Mosimann. 1985. Canonical and principal components of shape. Biometrika 72:241–252.

Daudin, J. J., C. Duby, and P. Trecourt. 1988. Stability of principal component analysis studied by the bootstrap method. Statistics 19:241–258.

Davies, R. G., and V. Brown. 1972. A multivariate analysis of postembryonic growth in two species of *Ectobius* (Dictyoptera: Blattidae). Journal of Zoology (London) 168:51–79.

Diaconis, P., and B. Efron. 1983. Computer-intensive methods in statistics. Scientific American 248 (May):116–130.

Efron, B., and G. Gong. 1983. A leisurely look at the bootstrap, the jackknife, and cross-validation. American Statistician 37:36–48.

Efron, B., and R. Tibshirani. 1986. Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy. Statistical Science 1:54–77.

Efron, B., and R. J. Tibshirani. 1993. An introduction to the bootstrap. Chapman and Hall: New York.

Felsenstein, J. 1985. Phylogenies and the comparative method. American Naturalist 125:1–15.

Flury, B. 1988. Common principal components and related multivariate models. Wiley: New York.

Flury, B., and H. Riedwyl. 1988. Multivariate statistics: A practical approach. Chapman and Hall: London.

Garland, T., Jr., P. H. Harvey, and A. R. Ives. 1992. Procedures for the analysis of comparative data using phylogenetically independent contrasts. Systematic Biology 41:18–32.

Gibson, A. R., A. J. Baker, and A. Moeed. 1984. Morphometric variation in introduced populations of the common myna (*Acridotheres tristis*): An application of the jackknife to principal component analysis. Systematic Zoology 33:408–421.

Gould, S. J. 1966. Allometry and size in ontogeny and phylogeny. Biological Reviews 41:587–640.

Gould, S. J. 1975. Allometry in primates, with emphasis on scaling and the evolution of the brain. Contributions to Primatology 5:244–292.

Gould, S. J. 1989. A developmental constraint in *Cerion*, with comments on the definition and interpretation of constraint in evolution. Evolution 43:516–539.

Gower, J. C. 1976. Growth-free canonical variates and generalized inverses. Bulletin of the Geological Institutions of the University of Uppsala, N. S. 7:1–10.

Günther, B. 1975. Dimensional analysis and theory of biological similarity. Physiological Reviews 55:659–699.

Hills, M. 1982. Bivariate versus multivariate allometry: A note on a paper by Jungers and German. American Journal of Physical Anthropology 59:321–322.

Hopkins, J. W. 1966. Some considerations in multivariate allometry. Biometrics 22:747–760.

Humphries, J. M., F. L. Bookstein, B. Chernoff, G. R. Smith, R. L. Elder, and S. G. Poss. 1981. Multivariate discrimination by shape in relation to size. Systematic Zoology 30:291–308.

Huxley, J. S. 1932. Problems of relative growth. Methuen: London. Reprinted 1972, Dover Publications: New York.

Huxley, J. S., and G. Teissier. 1936. Terminology of relative growth. Nature 137:780–781.

Jackson, J. E. 1990. A user's guide to principal components. Wiley: New York.

Jeune, B., and R. Sattler. 1992. Multivariate analysis in process morphology of plants. Journal of Theoretical Biology 156:147–167.

Jobson, J. D. 1992. Applied multivariate data analysis. Volume II: Categorical and multivariate methods. Springer-Verlag: New York.

Johnson, R. A., and D. W. Wichern. 1988. Applied multivariate statistical analysis. 2nd edition. Prentice-Hall: Englewood Cliffs, New Jersey.

Jolicoeur, P. 1963. The multivariate generalization of the allometry equation. Biometrics 19:497–499.

Jolliffe, I. T. 1986. Principal component analysis. Springer-Verlag: New York.

Jones, C. S. 1992. Comparative ontogeny of a wild cucurbit and its derived cultivar. Evolution 46:1827–1847.

Jones, C. S. 1993. Heterochrony and heteroblastic leaf development in two subspecies of *Cucurbita argyrosperma* (Cucurbitaceae). American Journal of Botany 80:778–795.

Jungers, W. L., and R. Z. German. 1981. Ontogenetic and interspecific skeletal allometry in nonhuman primates: Bivariate versus multivariate analysis. American Journal of Physical Anthropology 55:195–202.

Jungers, W. L., T. M. Cole, III, and D. W. Owsley. 1988. Multivariate analysis of relative growth in the limb bones of Arikara Indians. Growth, Development and Aging 52:103–107.

Kampny, C. M., T. A. Dickinson, and N. G. Dengler. 1993. Quantitative comparison of floral development in *Veronica chamaedrys* and *Veronicastrum virginicum* (Scrophulariaceae). American Journal of Botany 80:449–460.

Katz, M. J. 1980. Allometry formula: A cellular model. Growth 44:89–96.

Klingenberg, C. P. 1992. Multivariate morphometrics of geographic variation of *Gerris costae* (Heteroptera: Gerridae) in Europe. Revue Suisse de Zoologie 99:11–30.

Klingenberg, C. P., and R. Froese. 1991. A multivariate comparison of allometric growth patterns. Systematic Zoology 40:410–419.

Klingenberg, C. P., and J. R. Spence. 1993. Heterochrony and allometry: Lessons from the waterstrider genus *Limnoporus*. Evolution 47:1834-1853.

Klingenberg, C. P., and M. Zimmermann. 1992a. Static, ontogenetic, and evolutionary allometry: A multivariate comparison in nine species of water striders. American Naturalist 140:601–620.

Klingenberg, C. P., and M. Zimmermann. 1992b. Dyar's rule and multivariate allometric growth in nine species of waterstriders (Heteroptera: Gerridae). Journal of Zoology (London) 227:453–464.

LaBarbera, M. 1989. Analyzing body size as a factor in ecology and evolution. Annual Review of Ecology and Systematics 20:97–117.

Laird, A. K. 1965. Dynamics of relative growth. Growth 29:249–263.

Laird, A. K., A. D. Barton, and S. A. Tyler. 1968. Growth and time: An interpretation of allometry. Growth 32:347–354.

Leamy, L. and W. Atchley. 1984a. Static and evolutionary allometry of osteometric traits in selected lines of rats. Evolution 38:47–54.

Leamy, L., and W. R. Atchley. 1984b. Morphometric integration in the rat (*Rattus* sp.) scapula. Journal of Zoology (London) 202:43–56.

Leamy, L., and D. Bradley. 1982. Static and growth allometry of morphometric traits in randombred house mice. Evolution 36:1200–1212.

Lessa, E. P., and J. L. Patton. 1989. Structural constraints, recurrent shapes, and allometry in pocket gophers (genus *Thomomys*). Biological Journal of the Linnean Society 36:349–363.

Livezey, B. C. 1989. Morphometric patterns in Recent and fossil penguins (Aves, Sphenisciformes). Journal of Zoology (London) 219:269–307.

Loehlin, J. C. 1987. Latent variable models: An introduction to factor, path, and structural Analysis. Lawrence Erlbaum Associates: Hillsdale, New Jersey.

MacLeod, N., and J. A. Kitchell. 1990. Morphometrics and evolutionary inference: A case study involving ontogenetic and developmental aspects of foraminiferal evolution. Pages 283–299 *in* F. J. Rohlf and

F. L. Bookstein, (eds.), Proceedings of the Michigan morphometrics workshop. University of Michigan Museum of Zoology Special Publication No. 2: Ann Arbor.

Maddison, W. P., and D. R. Maddison. 1992. MacClade: Analysis of phylogeny and character evolution. Sinauer Associates: Sunderland, Massachusetts.

Manly, B. F. J. 1991. Randomization and Monte Carlo methods in biology. Chapman and Hall: London.

Marcus, L. F. 1990. Traditional morphometrics. Pages 77–122 *in* F. J. Rohlf and F. L. Bookstein, (eds.), Proceedings of the Michigan morphometrics workshop. University of Michigan Museum of Zoology Special Publication No. 2: Ann Arbor.

Marcus, L. F. 1993. Some aspects of multivariate statistics for morphometrics. Pages 95–130 *in* L. F. Marcus, E. Bello, and A. García-Valdecasas, (eds.), Contributions to morphometrics. Monografías del Museo Nacional de Ciencias Naturales 8, Madrid.

Martins, E. P., and T. Garland, Jr. 1991. Phylogenetic analyses of the correlated evolution of continuous characters: A simulation study. Evolution 45:534–557.

Maynard Smith, J., R. Burian, S. Kauffman, P. Alberch, J. Campbell, B. Goodwin, R. Lande, and L. Wolpert. 1985. Developmental constraints and evolution. Quarterly Review of Biology 60:265–287.

McGillivray, W. B. 1985. Size, sexual size dimorphism, and their measurement in Great Horned Owls in Alberta. Canadian Journal of Zoology 63:2364–2372.

McLellan, T. 1993. The roles of heterochrony and heteroblasty in the diversification of leaf shapes in *Begonia dregei* (Begoniaceae). American Journal of Botany 80:796–804.

Meyer, A. 1990. Morphometrics and allometry in the trophically polymorphic cichlid fish, *Cichlastoma citrinellum*: Alternative adaptations and ontogenetic changes in shape. Journal of Zoology (London) 221:237–260.

Mosimann, J. E. 1970. Size allometry: Size and shape variables with characterizations of the lognormal and generalized gamma distributions. Journal of the American Statistical Association 65:930–945.

Mosimann, J. E., and F. C. James. 1979. New statistical methods for allometry with application to Florida red-winged blackbirds. Evolution 33:444–459.

Owen, J. G., and M. A. Chmielewski. 1985. On canonical variates analysis and the construction of confidence ellipses in systematic studies. Systematic Zoology 34:366–374.

Pagel, M. D., and P. H. Harvey. 1988. Recent developments in the analysis of comparative data. Quarterly Review of Biology 63:413–440.

Pandolfi, J. M. 1988. Heterochrony in colonial marine animals. Pages 135–158 *in* M. L. McKinney, (ed.), Heterochrony in evolution: A multidisciplinary approach. Plenum Press: New York.

Patton, J. L., and P. V. Brylski. 1987. Pocket gophers in alfalfa fields: Causes and consequences of habitat-related body size variation. American Naturalist 130:493–506.

Paulsen, S. M., and H. F. Nijhout. 1993. Phenotypic correlation structure among elements of the color pattern in *Precis coenia* (Lepidoptera: Nymphalidae). Evolution 47:593–618.

Pearson, K. 1901. On lines and planes of closest fit to systems of points in space. Philosophical Magazine and Journal of Science, 6th series 2:559–572.

Pimentel, R. A. 1979. Morphometrics: The multivariate analysis of biological data. Kendall/Hunt: Dubuque, Iowa.

Reeve, E. C. R., and J. S. Huxley. 1945. Some problems in the study of allometric growth. Pages 121–156 *in* W. E. Le Gros Clark and P. B. Medawar (eds.), Essays on growth and form presented to D'Arcy Wentworth Thompson. Clarendon Press: Oxford.

Reilly, S. M. 1990. Comparative ontogeny of cranial shape in salamanders using resistant fit theta rho analysis. Pages 311–321 *in*. F. J. Rohlf and F. L. Bookstein (eds.), Proceedings of the Michigan morphometrics workshop. University of Michigan Museum of Zoology, Special Publication 2.

Reiss, M. J. 1989. The allometry of growth and reproduction. Cambridge University Press: Cambridge, U.K.

Reyment, R. A. 1991. Multidimensional palaeobiology. Pergamon Press: Oxford.

Reyment, R. A., and C. F. Banfield. 1976. Growth-free canonical variates applied to fossil foraminifers. Bulletin of the Geological Institutions of the University of Uppsala, N. S. 7:11–21.

Reyment, R. A., R. E. Blackith, and N. A. Campbell. 1984. Multivariate Morphometrics. 2nd edition. Academic Press, London.

Riska, B. 1981. Morphological variation in the horseshoe crab *Limulus polyphemus*. Evolution 35:647–658.

Riska, B. 1986. Some models for development, growth, and morphometric correlation. Evolution 40:1303–1311.

Riska, B. 1989. Composite traits, selection response, and evolution. Evolution 43:1172–1191.

Rohlf, F. J. 1990. Morphometrics. Annual Review of Ecology and Systematics 21:299–316.

Rohlf, F. J. 1993. Relative warp analysis and an example of its application to mosquito wings. Pages 131–159 *in* L. F. Marcus, E. Bello, and A. García-Valdecasas (eds.), Contributions to morphometrics. Monografías del Museo Nacional de Ciencias Naturales 8, Madrid.

Rohlf, F. J., and F. L. Bookstein. 1987. A comment on shearing as a method for "size correction." Systematic Zoology 36:356–367.

Rohlf, F. J., and L. F. Marcus. 1993. A revolution in morphometrics. Trends in Ecology and Evolution 8:129–132.

Sattler, R. 1992. Process morphology: Structural dynamics in development and evolution. Canadian Journal of Botany 70:708–714.

Shea, B. T. 1985. Bivariate and multivariate growth allometry: Statistical and biological considerations. Journal of Zoology (London) 206:367–390.

Shea, B. T. 1992. Developmental perspective on size change and allometry in evolution. Evolutionary Anthropology 1:125–134.

Shea, B. T., R. E. Hammer, R. L. Brinster, and M. R. Ravosa. 1990. Relative growth of the skull and postcranium in giant transgenic mice. Genetical Research, Cambridge 56:21–34.

Sinervo, B. 1993. The effect of offspring size on physiology and life history. BioScience 43:210–218.

Smith, L. D., and A. R. Palmer. 1994. Effects of manipulated diet on size and performance of brachyuran crab claws. Science 264:710–712.

Smith, R. J. 1993. Logarithmic transformation bias in allometry. American Journal of Physical Anthropology 90:215–228.

Solignac, M., M.-L. Cariou, and M. Wimitzki. 1990. Variability, specificity and evolution of growth gradients in the species complex *Jaera albifrons* (Isopoda, Asellota). Crustaceana 59:121–145.

Somers, K. M. 1986. Multivariate allometry and removal of size with principal components analysis. Systematic Zoology 35:359–368.

Somers, K. M. 1989. Allometry, isometry and shape in principal component analysis. Systematic Zoology 38:169–173.

Sprent, P. 1972. The mathematics of size and shape. Biometrics 28:23–37.

Stauffer, D. F., E. O. Garton, and R. K. Steinhorst. 1985. A comparison of principal components from real and random data. Ecology 66:1693–1698.

Strauss, R. E. 1985. Evolutionary allometry and variation in body form in the South American catfish genus *Corydoras* (Callichthyidae). Systematic Zoology 34:381–396.

Strauss, R. E. 1990a. Patterns of quantitative variation in lepidopteran wing morphology: The convergent groups Heliconiinae and Ithomiinae (Papilionoidea: Nymphalidae). Evolution 44:86–103.

Strauss, R. E. 1990b. Heterochronic variation in the developmental timing of cranial ossifications in poecilid fishes (Cyprinodontiformes). Evolution 44:1558–1567.

Strauss, R. E., and F. L. Bookstein. 1982. The truss: Body form reconstructions in morphometrics. Systematic Zoology 31:113–135.

Teissier, G. 1955. Allométrie de taille et variabilité chez *Maïa squinado*. Archives de Zoologie Expérimentale et Générale 92:221–264.

Thomas, R. D. K., and W.-E. Reif. 1993. The skeleton space: A finite set of organic designs. Evolution 47:341–360.

Thorpe, R. S. 1983. A review of the numerical methods for recognizing and analyzing racial differentiation. Pages 404–423 *in* J. Felsenstein, (ed.), Numerical taxonomy. Springer-Verlag: Berlin.

Thorpe, R. S., and M. Baez. 1987. Geographic variation within an island: Univariate and multivariate contouring of scalation, size, and shape of the lizard *Gallotia galloti*. Evolution 41:256–268.

Voss, R. S., L. F. Marcus, and P. Escalante. 1990. Morphological evolution in muroid rodents I. Conservative patterns of craniometric covariance and their ontogenetic basis in the neotropical genus *Zygodontomys*. Evolution 44:1568–1587.

Voss, R. S., and L. F. Marcus. 1992. Morphological evolution in muroid rodents II. Craniometric factor divergence in seven neotropical genera, with experimental results from *Zygodontomys*. Evolution 46:1918–1934.

Walker, J. A. 1993. Ontogenetic allometry of threespine stickleback body form using landmark-based morphometrics. Pages 193–214 *in* L. F. Marcus, E. Bello, and A. García-Valdecasas, (eds.), Contributions to morphometrics. Monografías del Museo Nacional de Ciencias Naturales 8, Madrid.

Wheeler, D. E. 1991. The developmental basis of worker caste polymorphism in ants. American Naturalist 138:1218–1238.

Wiig, Ø. 1985. Multivariate variation in feral American mink (*Mustela vison*) from Southern Norway. Journal of Zoology (London) 206:441–452.

Wright, S. 1968. Evolution and the genetics of populations. Volume 1. Genetic and biometric foundations. University of Chicago Press: Chicago.

Zelditch, M. L. 1987. Evaluating models of developmental integration in the laboratory rat using confirmatory factor analysis. Systematic Zoology 36:368–380.

Zelditch, M. L. 1988. Ontogenetic variation in patterns of phenotypic integration in the laboratory rat. Evolution 42:28–41.

Zelditch, M. L., F. L. Bookstein, and B. L. Lundrigan. 1992. Ontogeny of integrated skull growth in the cotton rat *Sigmodon fulviventer*. Evolution 46:1164–1180.

Zelditch, M. L., and A. C. Carmichael. 1989. Ontogenetic variation in patterns of developmental and functional integration in skulls of *Sigmodon fulviventer*. Evolution 43:814–824.

APPENDIX

Table 1. Morphometric data (raw values, in millimeters) for the example in the text. Four measurements were made on male *Gerris costae* from three different locations in Europe (Klingenberg, 1992). The measurements are the lengths of the thorax (Tho), the first antennal segment (Ant), and the femora of the middle and hind legs (MF and HF), corresponding to numbers 35, 43, 49, and 53 of Klingenberg (1992)

Alps				Greece				Pyrenees			
Tho	Ant	MF	HF	Tho	Ant	MF	HF	Tho	Ant	MF	HF
5.11	1.51	6.07	6.13	5.74	1.72	6.65	7.05	5.15	1.45	6.05	6.36
5.10	1.53	6.15	6.25	5.40	1.59	6.39	6.76	5.05	1.46	5.86	6.11
5.08	1.60	6.28	6.41	5.34	1.67	6.47	6.79	4.97	1.52	6.12	6.35
5.16	1.55	6.18	6.3	5.51	1.72	6.65	7.05	5.03	1.43	6.01	6.21
4.78	1.54	5.79	5.86	5.44	1.66	6.46	6.76	5.34	1.53	6.13	6.40
4.64	1.41	5.55	5.46	5.49	1.69	6.27	6.48	4.99	1.41	5.81	5.89
5.27	1.60	6.20	6.29	5.64	1.75	6.75	6.95	5.31	1.59	6.39	6.67
4.71	1.43	5.74	5.67	5.45	1.71	6.70	7.12	5.29	1.54	6.23	6.62
5.42	1.55	6.28	6.35	5.36	1.71	6.84	7.12	5.33	1.59	6.35	6.60
4.91	1.48	5.70	5.71	5.44	1.72	6.31	6.71	5.18	1.46	6.25	6.54
5.03	1.48	6.07	5.82	5.51	1.69	6.34	6.55	5.22	1.52	6.08	6.49
5.11	1.48	6.29	6.38	5.71	1.69	6.44	6.88	5.23	1.47	6.02	6.46
5.01	1.56	6.10	6.23	5.42	1.65	6.44	6.59	5.16	1.51	6.06	6.20
5.25	1.56	6.06	6.13	5.35	1.67	6.36	6.61	5.07	1.42	5.89	6.27
5.15	1.57	6.03	6.21	5.45	1.57	6.18	6.62	5.13	1.54	5.91	6.25
5.18	1.57	6.20	6.3	5.28	1.56	6.20	6.49	5.14	1.55	6.06	6.35
4.96	1.53	5.92	5.93	5.43	1.66	6.28	6.58	5.26	1.47	6.18	6.49
5.02	1.48	5.87	6.17	5.44	1.63	6.28	6.52	4.96	1.40	5.81	6.15
4.82	1.51	6.02	6.08	5.48	1.75	6.44	6.61	5.10	1.42	5.97	6.07
5.22	1.62	6.13	6.21	5.51	1.71	6.17	6.52	5.17	1.54	6.11	6.35
5.16	1.63	5.98	6.07	5.32	1.67	6.52	7.02	5.06	1.48	6.02	6.35
5.34	1.58	6.19	6.41	5.58	1.62	6.51	6.80	5.08	1.53	6.22	6.38
5.18	1.55	6.04	6.14	5.46	1.61	6.45	6.66	5.08	1.39	5.92	6.15
4.94	1.59	6.36	6.46	5.44	1.66	6.59	6.92	5.25	1.43	6.08	6.39
5.05	1.48	5.83	5.92	5.63	1.61	6.46	6.80	5.06	1.46	6.18	6.57
4.87	1.48	5.79	5.93	5.61	1.70	6.75	7.16	5.21	1.61	6.05	6.34
5.10	1.55	6.28	6.44	5.42	1.74	6.30	6.65	5.05	1.40	6.19	6.33
5.16	1.69	6.42	6.50	5.67	1.71	6.75	7.27	4.96	1.40	5.75	5.82
5.13	1.53	6.15	6.20	5.15	1.66	6.24	6.24				
4.82	1.50	5.85	5.79	5.36	1.75	6.52	6.86				
4.97	1.58	5.90	6.06	5.57	1.65	6.29	6.48				
5.02	1.56	6.17	6.25	5.62	1.74	6.57	7.06				
				5.53	1.71	6.45	6.86				