CHAPTER 4

Landmark Morphometrics and the Analysis of Variation

JOAN T. RICHTSMEIER*, SUBHASH R. LELE†, AND THEODORE M. COLE III‡

*Department of Anthropology, The Pennsylvania State University, University Park, Pennsylvania, USA; Center for Craniofacial Development and Disorders, The Johns Hopkins University, Baltimore, Maryland, USA

†Department of Mathematical and Statistical Science, University of Alberta, Edmonton, Alberta, Canada

‡Department of Basic Medical Science, School of Medicine, University of Missouri—Kansas City, Kansas City, Missouri, USA

INTRODUCTION

But the problems of variability, though they are intimately related to the general problem of growth, carry us very soon beyond our limitations.

Sir D'Arcy Thompson (1917)

Variation

Landmark-based morphometrics consists of a series of approaches to the analysis of form and form difference in two or three dimensions. Landmarks represent the location of biologically relevant features that can be recorded from a form with an acceptable degree of accuracy and precision in two or three dimensions (Richtsmeier *et al*., 1998). Morphometric methods provide measures of form and shape differences based on the relative location of landmarks. Modern morphometric approaches were inspired by the early work of Sir D'Arcy Thompson (1992) and matured with the belief that geometry of the whole, rather than analysis of disjointed measurements, might provide more complete answers to questions pertaining to morphology, phylogeny, functional morphology, and development. There are certain limitations associated with using the locations of landmarks to study morphology, however. For example, no information about surface features between the landmark locations is included in landmark data. But, other considerations (e.g., ease of data collection; developmental, evolutionary, or biomechanical correspondence of landmarks within, and potentially across, species; degree of precision; and the fact that coordinate data can be used to calculate linear distances or angles among landmarks) make them an appropriate choice for analysis in many biological research designs (see Lele and Richtsmeier, 2001).

Morphometrics has grown from a novel tool adopted by a few authors of morphological studies of the 1980s to a fairly standard methodology in biological inquiry. Still, problems persist in the estimation of certain parameters from landmark data, especially the parameters associated with the mean and variance–covariance structure (see Richtsmeier *et al*., 1992; Lele, 1993; Lele and McCulloch, 2002). Measures of phenotypic variation estimated from linear measures have proven incredibly useful for understanding the relative roles of genes and environment in the production of the phenotype, in part because the genetic properties of a population can be characterized by partitioning phenotypic variation into components ascribed to various causative sources using methods from quantitative genetics (Falconer, 1981; Falconer and Mackay, 1996). These more traditional measures of variation have proven useful in statistical analysis of the phenotype and in quantitative genetics studies, but the ability to properly measure variation local to biologically significant landmarks could provide additional advantages. First, the ability to localize measured variation to a landmark, rather than to linear distances that traverse developmental and anatomical boundaries, could improve the delineation of the sources of variation within an organism. Given that landmark coordinate-based morphometric techniques enable the researcher to localize observations of morphological differences to particular points in space, landmark-based measures of variation should enable a fine-tuning of local measures of variation without the unavoidable overlap among linear distances that share a common endpoint. Second, using morphometric methods that enable comparison of growth patterns, estimation of anatomically localized measures of change in form from one generation to the next (Richtsmeier and Lele, 1993) should enable subsequent partitioning of local

variance components, each attributable to a different cause, including localized selection differentials. Third, and more practically, if we can provide valid estimates of the variation around landmarks, rather than local to all unique distances between landmarks, there will be fewer parameters to estimate and to study. These advantages suggest that estimation of variation from landmark data may be useful. However, obtaining valid estimates of variation from landmark data has proven to be more difficult than originally thought.

In this chapter, we discuss the standard approaches that have been used to estimate variation in landmark data and explain why these methods do not properly estimate variation in biological forms. Though some of these approaches have become rather mathematically involved (Dryden and Mardia, 1998), their suitability to the realities of biological data has not improved in parallel. We present a generalized model for variation in landmark data. It has been shown that only certain features of this model can be consistently estimated (Lele, 1993; Lele and Richtsmeier, 2001; Lele and McCulloch, 2002). This model and the estimators are used in morphometric analysis as the basis for parametric bootstrapping procedures to test for differences in form using landmark data (Lele and Richtsmeier, 2001). Recognizing the limitation of the ability to estimate only certain features of this general model of variation, we discuss the need for the development of less general models that may reasonably characterize variation in landmark data. We show how biological knowledge pertaining to the organisms under study can be used to impose certain constraints on the models of variance–covariance structure. We suggest ways to integrate such constraints into the proposal of several less generalized models, some statistically convenient but biologically improbable, others less streamlined statistically but more biologically reasonable.

I. COORDINATE DATA AND THE COORDINATE SYSTEM

Most investigators credit Sir D'Arcy Thompson (1992) with the stimulus that initiated the field of morphometrics. Early geometry-based methodologies were proposed by Boas (1905) and Sneath (1967) (see Cole [1996] for the specifics of Boas's contribution), but were largely ignored until much later. The modern field of *geometric morphometrics*, defined as the fusion of geometry and biology (Bookstein, 1982), flourished from the 1980s onward and originated from a desire to analyze biological forms in ways that preserve the physical integrity of form in two or three dimensions. The nature of landmark data was particularly appealing to biologists. The expression of the relative location of landmarks in a coordinate system that displays geometric relationships among parts of the whole stood in stark contrast to the usual expression of a form as disarticulated series of linear and angular measures. Although the use of landmark data in the estimation of an average and variation around that average appears straightforward, the innate nature of

landmark data, in particular the presence of the nuisance parameters of translation and rotation, makes estimation of the true variation complicated.

The researcher starts with coordinate data collected from a group of organisms. Biological organisms that constitute a group resemble each other to such a degree that we have an intuitive idea of the appearance of the typical or "average" form that is representative of all members of the group. Using statistics, an average or mean form can be estimated from the observations. When coordinate data are used, this mean form provides the average relative locations of landmarks. Genetic and environmental influences combine to affect structures so that all forms differ from each other in diverse ways. This variation in the sample is estimated as divergences in the relative location of landmarks from their configuration in the mean form.

When the researcher accumulates landmark data from a sample of forms, both the mean of the sample and the relationship of individual observations to the mean (the perturbation structure) are unknown. To estimate the mean and variance, it is necessary to specify a mathematical construct or *model*, that attempts to characterize certain aspects of the properties of the population. A *model* is used here as a mathematical construct that attempts to characterize certain aspects of the underlying phenomena (e.g., dimensions, dynamics, properties, interactions). Models include quantities called parameters, and these parameters are estimated according to the specified model using the sample data. Models and methods used to estimate mean and variance of a sample of coordinate data sets require that attention be paid to the unique coordinate system of each data set. Because of the need to address the varied coordinate systems within a sample of data sets, models for landmark data typically define a group of parameters that are unique to the properties of coordinate data. These parameters include rotation, translation, and in some cases, reflection and are required to place coordinate data sets into a common coordinate system so that the mean and variance can be estimated. This seems simple enough, and intuitively it makes sense: If we are going to build a mean form and determine phenotypic variation for a sample, the forms must all be expressed within the same coordinate system. The question is, of the infinite coordinate systems available, which one should be chosen for analysis?

From a biological perspective, one might argue that an orientation scheme should be adopted with reference to the primitive state of the organism or with reference to whatever is defined as standard anatomical position of the adult for a given species. Compelling arguments for the adoption of a particular coordinate system on mathematical grounds (i.e., ease of computation) could also be made. If adoption of varying coordinate systems gave the same sample estimates, this choice would be trivial. Unfortunately it has been shown that estimates of the mean and variance change with the coordinate system adopted (Richtsmeier *et al.*, 1992; Lele, 1993; Lele and McCulloch, 2002).

Because estimation of the mean and variance–covariance structure from landmark data requires that all forms be placed into a common coordinate system and the

estimates of these parameters change according to the chosen coordinate system, that choice must be critically evaluated. How can a scientist know that the parameters estimated for a sample are valid if there is no information regarding the relative validity of any particular coordinate system? This question is the source of the complexity of estimating sample parameters for coordinate data sets. We detail the problems involved in estimation of mean and variance–covariance structure of coordinate data sets in the following text.

II. THE GENERAL PERTURBATION MODEL FOR LANDMARK VARIATION

Suppose we are interested in a sample of n organisms and we measure them using a series of K landmarks in D (=2 or 3) dimensions. A $K \times D$ matrix that we designate M describes the mean for the population, where each row represents the D dimensional coordinates of a single landmark. Although M is a mathematical construct and is not directly observable from a sample of specimens (even an infinite one), we can imagine the mean configuration of landmarks in a coordinate system that we will refer to as "natural space," where within-sample variation arises. No single individual is likely to be identical in form to the mean, and no two individuals are likely to be identical.

In the natural space, phenotypic variation is manifested as perturbations around the mean landmark configuration, M (Figure 4-1). Note that the dispersion patterns of these perturbations can vary in size and shape from landmark to landmark. Some landmarks are more variable than others, as indicated by the relative sizes of their dispersions. The perturbation scatters for landmarks can also vary in shape, with some being round and others being elliptical. Finally, there may be covariances between the landmarks because of developmental, biomechanical, or physical constraints, so that the relative positions of the observations at one landmark may be correlated with the positions of other landmarks. If we want to understand the biological processes and mechanisms that produce such a pattern of variation, we need first an objective way of describing it quantitatively. In statistical analysis, a model must be specified before data are analyzed. Parameters are then estimated from each sample using the specified model. Knowledge of particular properties of the phenomena that the data represent can and should be included in proposing a model. Whatever characteristics the biologists deem important or explanatory should be included as assumptions of the model. If a model appears appropriate and correct, methods for checking the assumptions can be developed (if they do not already exist) and applied.

In morphometrics, phenotypic variation as described by landmark data has been modeled routinely using a *general perturbation model*, sometimes called the *Gaussian* perturbation model (see Goodall, 1991; Lele, 1993). Goodall (1991)

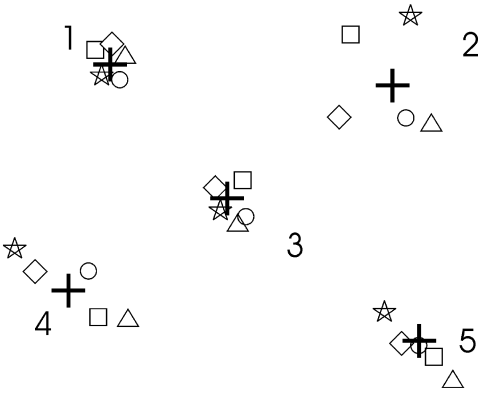

FIGURE 4-1. The "natural space" where individual differences in form originate. Plus signs (+) represent the parametric mean configuration for a hypothetical organism with five landmarks. The symbols represent the landmark locations of different specimens (where like symbols belong to the same specimen). These locations are phenotypic perturbations of the mean, which reflect underlying genetic and environmental variability. Note that the dispersion patterns differ from one landmark to the next. Some landmarks have roughly circular distributions (1, 2, and 3), while others are elliptical. Landmarks 1 and 3 have relatively small dispersions, while the dispersion around landmark 2 is large. In addition, some of the perturbations may be correlated—note the similarity in the rank order of perturbations (from upper left to lower right) for landmarks 4 and 5. As detailed in the text, the positions of the perturbations in natural space cannot be reconstructed; however, some descriptors of the dispersion patterns can be estimated. Adapted from Cole *et al.,* 2002.

used this model in the development of morphometric methods based on Procrustes superimposition, and Lele and colleagues (e.g., Lele and Richtsmeier, 1990, 1991; Lele and McCulloch, 2002) used it in the development of Euclidean distance matrix analysis, or EDMA (pronounced ĕd·mă). Using the general perturbation model, the set of landmark data for each observation in a sample is described by a $K \times D$ matrix called X_i. Each X_i is related to the population mean, M, as follows:

$$X_i = (M + E_i)\Gamma_i + t_i$$

where E_i is a $K \times D$ matrix of perturbations that describe how X_i differs from M in the natural space (the *true* difference between X_i and M). For the population, these perturbations are assumed to have a multivariate normal distribution with a $K \times D$ mean matrix $\mathbf{0}$ and a covariance structure $\mathbf{\Sigma}_K \otimes \mathbf{\Sigma}_D$, where $\otimes$ denotes a Kronecker product. $\mathbf{\Sigma}_K$ is a $K \times K$ matrix that describes the variances and covariances of the landmarks, while $\mathbf{\Sigma}_D$ is a $D \times D$ matrix that describes the variances and covariances of the perturbations with respect to the natural space's coordinate-system axes (i.e., $\mathbf{\Sigma}_K$ and $\mathbf{\Sigma}_D$ describe the sizes, shapes, and orientations of the perturbation scatters). Γ_i is an orthogonal $K \times K$ matrix that describes the rotation of X_i (in the

coordinate system where the data have been collected) relative to M (as it lies in the natural space), while t_i is a $K \times D$ matrix that describes the translation of X_i relative to the position of M in the natural space. Realize that both the position and orientation of M are arbitrary.

The mean, M, and the variance–covariance matrix of the perturbations, $\mathbf{\Sigma}_K \otimes \mathbf{\Sigma}_D$, are obviously of great biological interest. Unfortunately, they are not estimable given the general perturbation model because of the presence of the other terms in the equation. These entirely arbitrary parameters, Γ_i and t_i, are "nuisance parameters" (*sensu* Neyman and Scott, 1948). Although they have no biological meaning, these nuisance parameters are essential in describing the statistical nature of the observed data. Consequently, Γ_i and t_i are nuisance parameters from the scientific, but not the statistical perspective. This is because identification of the orientation of an object with respect to the mean is not needed to characterize the biological nature of phenotypic variation, but is essential in describing the statistical nature of the observed landmark data. Unfortunately, Γ_i and t_i are unobservable and cannot be estimated, for reasons explained in the following text. This means that reconstruction of the natural space from empirical data is impossible (Lele and Richtsmeier, 1990; Lele, 1993; Lele and McCulloch, 2002; Richtsmeier *et al*., 2002). Lele and colleagues provide mathematical proof of the inestimability of M, $\mathbf{\Sigma}_K$, and $\mathbf{\Sigma}_D$ (Lele, 1993; Lele and McCulloch, 2002).

First, let us consider the number of unknowns in the general perturbation model, $X_i = (M + E_i)\Gamma_i + t_i$. Only X_i is known from the data. All parameters, including the mean form, M, the variance–covariance structure of the errors, E_i, and the rotation and translation parameters, Γ_i and t_i, are unknown. The parameters M, $\mathbf{\Sigma}_K$, and $\mathbf{\Sigma}_D$ are fixed and common to all the specimens. However, the parameters Γ_i and t_i are different for *every specimen* because each individual has a unique location and orientation with reference to its position in the natural space. This means that we have a total of $(2 + 2n)$ unknowns for the given equation, the first term referring to M and the variance–covariance structure of E_i for the sample, the second term referring to Γ_i and t_i for *each individual* in the sample of size n. The number of unknowns $(2 + 2n)$ is therefore larger than the sample size (n). A basic tenet of inferential statistics is that one cannot estimate more parameters than the number of observations. Therefore the parameters cannot be estimated and must remain unknown. Although, M, Γ_i, t_i, etc., are matrices, for the sake of exposition, we consider them as single entities. Strictly speaking and in mathematical terms, the number of parameters is of $\mathbf{O}(n)$, which is the same order of magnitude as the sample size (Richtsmeier *et al*., 2002; see Lele and Richtsmeier, 2001; pp. 63–66).

It is the nature of landmark data—that they require specification of a coordinate system for their expression—that makes them so attractive to biological inquiry, but also makes them difficult to analyze. No information is available from landmark data *per se* regarding how the coordinate systems for each specimen relate to the natural space. In morphometrics, values for translation and rotation are often chosen based upon a "rule" or goal. For example, in Generalized Procrustes

Analysis (GPA), rotation and translation are estimated so that the sums of the squared distances between corresponding landmarks among individuals in a sample are minimized. An alternate rule is used in the robust or resistant fitting Procrustes analysis (Siegel and Benson, 1982; Chapman, 1990; Rohlf and Slice, 1990), and the application of this rule results in a different superimposition and thus a different mean and variance–covariance structure. In fact, because no information is available from the landmark data about how the coordinate system of each individual and the natural space fit to one another, there are infinite ways to place all individuals from a sample into a common coordinate system, and therefore there are infinite choices for the possible values of Γ_i and t_i. If all these choices were equivalent, the issue of choice would be irrelevant. However, as proven elsewhere (Lele, 1993; Lele and McCulloch, 2002) and shown graphically later in this chapter, this choice directly affects the estimation of the mean and especially the variance. If the estimate of the variance–covariance structure is invalid, it certainly cannot be used to partition variance components.

It is often stated in analyses using Procrustes methods that the nuisance parameters of rotation and translation are "effectively removed" from the analysis once the specimens have been oriented into a consensus coordinate system. In reality, the nuisance parameters are not removed, effectively or otherwise, and this statement is a misrepresentation of what occurs when data sets are analyzed using Procrustes analyses. Even after the data have undergone rotation, translation, and in some cases reflection, according to whatever rule the Procrustes method specifies, those parameters are still embedded (although further obfuscated) within the data. As a consequence, they affect the estimates of the mean and variance–covariance structure.

In short, nuisance parameters are not just irritants that can be eliminated by distributing their effects among individuals. Nuisance parameters must be dealt with in ways that do not cause them to affect the estimates of the very parameters that we are trying to estimate. Fortunately, the nuisance parameters, Γ_i and t_i, are of no real interest (except for some specific research questions) because knowledge of the orientation of an object with respect to its original position or to the mean is unimportant. We are really only interested in estimating the mean form, M, and the variance–covariance structure of the errors, E_i. In the following section, we show how these nuisance parameters can be properly eliminated.

III. PROPER ELIMINATION OF NUISANCE PARAMETERS USING A COORDINATE SYSTEM INVARIANT METHOD OF ESTIMATION

An alternative general method for estimating the mean and variance–covariance structure does not require an attempt to estimate Γ_i and t_i (Lele and Richtsmeier, 1991; Lele, 1993). If we limit the focus to the mean, M, and the variance–covariance

of E_i, then following the logic presented in the preceding text, the *number* of unknowns is fixed and does not change with sample size.

These biologically interesting components of the model that are identifiable can be estimated in several ways, one of which is by using method-of-moments techniques developed by Lele (1993) and Lele and McCulloch (2002). Although the coordinate locations of the landmarks of the population mean form, M, cannot be observed directly, we can compute the coordinates of a consistent sample estimate of the mean, called $\hat{M}$, up to translation, rotation, and reflection. The sample estimate, $\hat{M}$, consists of all linear distances measured between unique pairs of landmarks and is equivalent to a coordinate data set, except that it does not require a coordinate system for its expression. Similarly, although one cannot estimate the among-landmarks variance–covariance matrix ($\mathbf{\Sigma}_K$), one can obtain a consistent estimate of a singular version of it, called $\mathbf{\Sigma}_K^*$ (see Lele, 1993). Although neither the among-axes variance–covariance matrix ($\mathbf{\Sigma}_D$) nor its eigenvectors is estimable, the eigenvalues of this matrix can be estimated (Lele and McCulloch, 2002), describing the overall eccentricity of the perturbation scatters at each landmark. In summary, the following features of the mean and the variance parameters can be identified and estimated:

1. All pairwise distances in the mean form M
2. A singular version of $\mathbf{\Sigma}_K$
3. Eigenvalues of $\mathbf{\Sigma}_D$

We emphasize that these results hold true irrespective of the particular method of estimation, whether method of moments or maximum likelihood. Moreover, these quantities and their estimators are *coordinate-system invariant*, meaning that they are not affected by the positions and orientations of the observations in any arbitrary coordinate system. However, the centering that is required in the calculation of $\mathbf{\Sigma}_K^*$ is by its nature not coordinate system invariant and, therefore, $\mathbf{\Sigma}_K^*$ can only be used as a suitable measure of the variation in landmark coordinate data for specific purposes as described in the following text. Details of the calculations of $\hat{M}$ and $\mathbf{\Sigma}_K^*$ are provided by Lele (1993), Lele and Cole (1996), and Lele and Richtsmeier (2001).

At this point, it is critical to understand the relationship between $\mathbf{\Sigma}_K^*$, which we can estimate, and $\mathbf{\Sigma}_K$, which we cannot. $\mathbf{\Sigma}_K^*$ is a singular version of $\mathbf{\Sigma}_K$, and the two are related as follows:

$$\mathbf{\Sigma}_K^* = \mathbf{L}\mathbf{\Sigma}_K\mathbf{L}^T$$

where $\mathbf{L}$ is a *centering matrix*. In both superimposition methods and in EDMA, the centering of all observations into a common coordinate system (that is, the translation of all observations to the origin) is a necessary first step. However, the purpose of the centering is different under the two methods. For superimposition methods, the centering is an initial step in attempting to reconstruct the natural space and,

as a result, the true variance–covariance structure $\mathbf{\Sigma}_K$. However, we know that the natural space is actually impossible to reconstruct and that $\mathbf{\Sigma}_K$ is not identifiable, so whatever is reconstructed using superimposition methods, it cannot be the true variance–covariance structure. Using EDMA, the purpose of centering is different: The observations are centered to get a *version* of $\mathbf{\Sigma}_K$, called $\mathbf{\Sigma}_K^*$, that is identifiable and consistently estimable. We emphasize again that in estimating $\mathbf{\Sigma}_K^*$ in EDMA, we do *not* intend to reconstruct the natural space. The preceding equation may lead the reader to ask why we cannot obtain $\mathbf{\Sigma}_K$, given both $\mathbf{\Sigma}_K^*$, which we can estimate from data, and the known matrix, **L**. The answer is that the system of equations $\mathbf{\Sigma}_K^* = \mathbf{L}\mathbf{\Sigma}_K\mathbf{L}^T$ does not possess a unique solution. There are more unknowns than the number of equations in the system. So even if **L** and $\mathbf{\Sigma}_K^*$ are known completely, one cannot solve for $\mathbf{\Sigma}_K$ in a unique fashion. In estimating $\mathbf{\Sigma}_K^*$ from real data, we must decide on the form of **L**, and we must realize that this decision is ultimately arbitrary. We may choose to center all of the observations on a particular landmark, or we may choose to center them on some function of two or more landmarks, such as the centroid. If we experiment with different centering criteria, we will see that different decisions will lead to different pictures of variability.

To demonstrate this, we have randomly generated a sample of 100 observations (each with three landmarks in two dimensions) in the natural space, using the mean form (**M**) and variance–covariance matrix ($\mathbf{\Sigma}_K$) shown in Table 4-1. The data are illustrated in the upper left-hand corner of Figure 4-2. Because these are artificial data that we generated using known parameters, we know that this is a true picture of the sample variation. To illustrate the effect of different choices of centering, we centered $\mathbf{\Sigma}_K$ several different ways to get $\mathbf{\Sigma}_K^*$, using the relationship $\mathbf{\Sigma}_K^* = \mathbf{L}\mathbf{\Sigma}_K\mathbf{L}^T$, where **L** is the centering matrix (Table 4-1). If we center, in turn, on the centroid (Figure 4-2A), the first landmark (Figure 4-2B), and the second landmark (Figure 4-2C), we see that we get $\mathbf{\Sigma}_K^*$ matrices with structures that look very different (Table 4-1). When we generate random data using these matrices (see Cole *et al*., 2002), we also get very different visual impressions of the variability in the samples. Which, if any of these is the most accurate depiction of $\mathbf{\Sigma}_K$? We simply have no objective way of knowing.

Because the centering criterion is arbitrary, we must resist the temptation to examine a particular estimate of $\mathbf{\Sigma}_K^*$ as if it were a picture of the natural space. Had we chosen a different centering criterion, we may have obtained a very different picture. Given that fact, we are led to an obvious question about $\mathbf{\Sigma}_K^*$: If $\mathbf{\Sigma}_K^*$ does not give us an accurate picture of the true variation in landmarks ($\mathbf{\Sigma}_K$), how is it useful? In practical applications, estimates of $\mathbf{\Sigma}_K^*$ should not be seen as ends in themselves but as aids in asking other interesting questions about variation in form, shape, or growth. For example, we have used $\mathbf{\Sigma}_K^*$ to generate random data sets under the general perturbation model (given sample estimates of

TABLE 4-1. Example Illustrating the Effects of Different Centering Criteria on $\mathbf{\Sigma}_K^*$ Matrices, where $\mathbf{\Sigma}_K^* = \mathbf{\Sigma}_K L^{T}$.[a]

True parameters:

$$\mathbf{M} = \begin{bmatrix} 0 & 0 \\ 10 & 0 \\ 0 & 10 \end{bmatrix} \qquad \mathbf{\Sigma}_K = \begin{bmatrix} 1.0 & 0.3 & 0.6 \\ 0.3 & 3.0 & 0.8 \\ 0.6 & 0.8 & 2.0 \end{bmatrix}$$

Centering on the centroid:

$$\mathbf{\Sigma}_K = \begin{bmatrix} 0.778 & -0.656 & -0.122 \\ -0.656 & 1.311 & -0.656 \\ -0.112 & -0.656 & 0.778 \end{bmatrix}$$

Centering on landmark 1:

$$\mathbf{\Sigma}_K = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 3.4 & 0.9 \\ 0 & 0.9 & 1.8 \end{bmatrix}$$

Centering on landmark 2:

$$\mathbf{\Sigma}_K = \begin{bmatrix} 3.4 & 0 & 2.5 \\ 0 & 0 & 0 \\ 2.5 & 0 & 3.4 \end{bmatrix}$$

[a]The mean form **M** is used to generate the random data shown in Figure 4-2, using both $\mathbf{\Sigma}_K$ and the different versions of $\mathbf{\Sigma}_K^*$.

M and $\mathbf{\Sigma}_K$* and assuming multivariate normality) (Richtsmeier *et al.*, 2000). We have used these random data as part of a parametric bootstrapping procedure in testing for differences between the mean shapes of two samples (Lele and Cole, 1996). We used the same random data to generate confidence intervals for sample differences in the relative sizes of specific linear distances. In another type of application, we used $\mathbf{\Sigma}_K^*$ estimates as part of tests for phylogenetic signals in landmark data (Cole *et al.*, 2002). Importantly, Lele and McCulloch (2002: Theorem 2) have demonstrated that the choice of centering does not have any effect on the results of these applications. In addition, as noted by Lele and Cole (1996: Appendix),

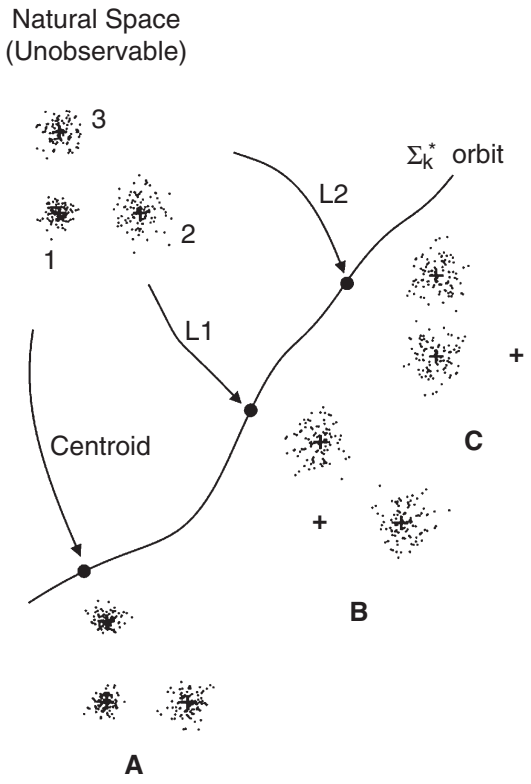

FIGURE 4-2. Sample of 100 randomly generated observations (three landmarks in two dimensions) in the natural space, based on the mean from *M* and variance–covariance matrix $\mathbf{\Sigma}_K$ shown in Table 4-1. The mean locations of the landmarks are indicated by the plus signs (+). The plot shows the true dispersion pattern of the sample in the natural space (top left). Also shown are random samples generated under different singular versions of the variance–covariance structure, calculated using different centering criteria, **L**, that used (A) the centroid; (B) landmark 1 (L1); and (C) landmark 2 (L2). $\mathbf{\Sigma}_K^*$ for each of these centering patterns are given in Table 4-1. The mean locations and the scales of the plots are the same in every case.

confidence intervals and the distributions of test statistics will be the same when calculated with any $\mathbf{\Sigma}_K^*$ as they would be if we could use $\mathbf{\Sigma}_K$ itself.

IV. ADDING ASSUMPTIONS TO THE PERTURBATION MODEL

Estimation of a $\mathbf{\Sigma}_K^*$ matrix, which is a nonunique, singular version of $\mathbf{\Sigma}_K$, is the best that we can do with the general perturbation model that we have specified. However, if we can add some additional assumptions to the model, and we are

confident that these assumptions have a reasonable theoretical base, we can move another step closer to seeing a picture of the natural space, which is a picture of variation that we can interpret in biological terms. In essence, by imposing a more restricted structure, we make it possible for the system of equations, $\mathbf{\Sigma}_K^* = \mathbf{L}\mathbf{\Sigma}_K\mathbf{L}^T$, to have a unique solution. We impose the constraints in such a fashion that the number of unknowns is *at most* equal, and hopefully smaller, than the number of equations.

Remember that we are trying to estimate the mean form, M, and the variance–covariance structure of the perturbations, $\mathbf{\Sigma}_K \otimes \mathbf{\Sigma}_D$. E_i is a $K \times D$ matrix of perturbations that describe how the coordinates for each individual within a sample, X_i, differs from M in a space that will be specified by the assumptions of our model. The perturbations are assumed to have a multivariate normal distribution with a covariance structure $\mathbf{\Sigma}_K \otimes \mathbf{\Sigma}_D$. $\mathbf{\Sigma}_K$ is a $K \times K$ matrix that describes the variances and covariances of the landmarks, while $\mathbf{\Sigma}_D$ is a $D \times D$ matrix that describes the variances and covariances of the perturbations with respect to the coordinate-system axes.

A. Model 0: Isotropic Error Model

A simplifying assumption that often has been added to the general perturbation model to make it estimable is that $\mathbf{\Sigma}_K = \sigma^2\mathrm{I}$ and $\mathbf{\Sigma}_D = \mathrm{I}$, I being the identity matrix (Bookstein, 1986). In fact, this assumption is *necessarily* made by generalized Procrustes analysis in order for the estimate of the mean to be consistent (Lele, 1993; Kent and Mardia, 1997; Lele and Richtsmeier, 2001). What this assumption means is that the variances of all landmarks (that is, the amount of dispersion) are expected to be the same. It also means that the patterns of dispersion across landmarks are expected to be uncorrelated, so that different landmarks vary independently. Biologically, the idea of equal variances at each landmark seems unlikely. Lele and Richtsmeier (1990) demonstrated strong rejection of the hypothesis of equality of variances local to landmarks in an analysis of three biological data sets. Sometimes, especially in Procrustes analysis (Dryden and Mardia, 1998), it is assumed that the variation around the landmarks is small enough that the equal variation model is a reasonable approximation of reality. Assuming small variance local to landmarks seems biologically ill-advised as well, especially when this assumption cannot be validated using the real data. Moreover, an assumption of landmark independence seems extremely improbable, and it logically precludes the use of these models for studies of morphological integration or modularity, which focus on and interpret the covariances and correlations among landmarks. If we assume the covariances are all zero, there is, by definition, nothing for us to study. Biology has presented us with a number of constraints on the description of form by landmark data based solely on anatomy and morphology. It seems unwise to place additional constraints on analyses solely based on statistical

considerations (e.g., the assumptions of Procrustes analyses) that may not reflect biological reality.

We now propose some models for the variance–covariance matrix that are more restricted than the general perturbation model, but are not as restricted as the isotropic error model. To demonstrate our models, we use the mouse dentary as the mean form on which we have located 11 landmarks, each landmark associated with one of six morphogenetic units, as defined by Atchley and Hall (1991) (Figure 4-3). As we progress through the development of models, we add increasingly complex assumptions in an effort to increase their biological validity and realism.

B. Model 1: Σ_K = Independent Local Variation

In this model, we introduce heterogeneity of phenotypic variation at landmarks by allowing different magnitudes of variation local to each landmark. Solely, for exposition of our models, we adopt the convention that variation can take any value

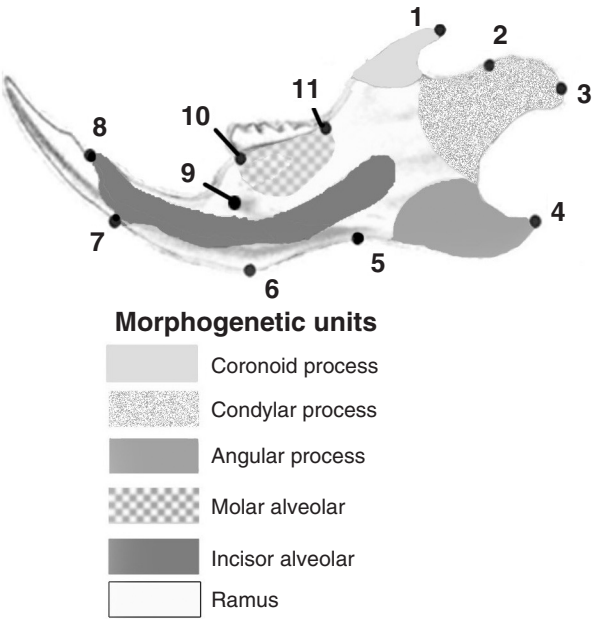

FIGURE 4-3. The adult mouse hemimandible (dentary) subdivided into the mandibular morphogenetic units (modules) proposed by Atchley and Hall (1991).

from 0 to 6, 0 representing a lack of variation (or covariation), and 6 representing the maximum phenotypic variation possible. In this particular model, correlations among landmarks are not permissible, so off-diagonal elements remain constrained to zero. This is an extremely simple model that is relatively easy to estimate. However, it is biologically unrealistic since studies of morphological integration and allometry have demonstrated that local measures of variability are correlated within an organism because of relationships that have a developmental or functional origin. Because this model does not allow the most basic relationships among neighboring structures, further assumptions are required to build a more realistic model.

C. Model 2: Σ_K = Independent Modules

In this model, landmarks are grouped into modules that exhibit common properties of phenotypic variation based on developmental or functional considerations. The developmental organization of the mammalian dentary provided by Atchley and Hall (1991) divides the mandible into six morphogenetic units, each being derived from a separate cell condensation (Figure 4-3). Studies of modularity and morphological integration suggest that landmarks within the same module may share aspects of the perturbation structure.

In the independent modules model, variation local to landmarks (values of diagonal elements) that coexist within a module are similar, as are the values of the off–diagonal elements within a module (Figure 4-4). Correlation patterns exist for the perturbation pattern of landmarks within a module, but there is no correlation in patterns between modules. For example, landmarks 2 and 3 are both located within the condylar process morphogenetic unit. We set our expected value of variation to be similar local to landmarks 2 and 3, and we set an expected level of covariation within this module that reflects our knowledge (or hypotheses pertaining to identification) of the relative potential variation within this part of the mandible. Landmarks that represent the ramus (5, 6, 9) are expected to show a lesser magnitude of variation and covariation because formation of the ramus has proven to be relatively stable even in the case of knock-out experiments (see codependent module model, in the following text).

In our example, modules are defined and values are assigned using developmental information, but can be based on other relevant biological considerations. There are many other sources of information for modeling the variance–covariance structure (e.g., phylogenetic considerations, trajectories of disease processes, fate maps, teratogen targets, functional anatomy). Depending on the factors that define the modules, landmarks within a module may or may not be physically close to one another. Importantly, in our example, we are assigning values to demonstrate

LANDMARKS												
LANDMARKS		1	2	3	4	5	6	7	8	9	10	11
	1	**6**	*0*	*0*	*0*	*0*	*0*	*0*	*0*	*0*	*0*	*0*
	2	*0*	**3**	*4.2*	*0*	*0*	*0*	*0*	*0*	*0*	*0*	*0*
	3	*0*	*4.2*	**3**	*0*	*0*	*0*	*0*	*0*	*0*	*0*	*0*
	4	*0*	*0*	*0*	**4**	*0*	*0*	*0*	*0*	*0*	*0*	*0*
	5	*0*	*0*	*0*	*0*	**2**	*2.1*	*0*	*0*	*2.1*	*0*	*0*
	6	*0*	*0*	*0*	*0*	*2.1*	**2**	*0*	*0*	*2.1*	*0*	*0*
	7	*0*	*0*	*0*	*0*	*0*	*0*	**3**	*2.9*	*0*	*0*	*0*
	8	*0*	*0*	*0*	*0*	*0*	*0*	*2.9*	**3**	*0*	*0*	*0*
	9	*0*	*0*	*0*	*0*	*2.1*	*2.1*	*0*	*0*	**2**	*0*	*0*
	10	*0*	*0*	*0*	*0*	*0*	*0*	*0*	*0*	*0*	**3.5**	*2.6*
	11	*0*	*0*	*0*	*0*	*0*	*0*	*0*	*0*	*0*	*2.6*	**3.5**

FIGURE 4-4. Variance–covariance structure for independent modules. In this model, we constrain measures of variation at a landmark (diagonal elements of the matrix) and covariation between landmarks (off-diagonal elements of the matrix) to values between 0 and 6, 0 indicating a lack of variation or covariation and 6 indicating the maximum amount of variation or covariation possible. In this model, variation local to landmarks within a morphogenetic unit or module are set to similar values and covariation between landmarks within a module is allowed. No covariation is allowed between landmarks that occupy different morphogenetic units.

the expected patterns based on our knowledge of developmental relationships. When estimating the variances and covariances from data, we only specify the pattern (constraints within and between modules). The specified constraints comprise the pattern. Variances and covariances are estimated using the data.

D. Model 3: Σ_K = Co-dependent Modules

In this model, correlation in the perturbation pattern for landmarks within and between modules is allowed. One potential way to model such relationships is to predict a correlation structure that reflects what is known about the effects of particular genes, functions, selection differentials, or teratogens on different modules. An alternate way to model the relationships is to predict a correlation structure that is a function of the dissimilarity between landmarks within modules. If a dissimilarity metric is used, we might predict that all landmarks will be correlated but that

the *degree* of correlation will be a function of the dissimilarity of the landmarks. The measure of dissimilarity does not need to be based on the physical proximity of landmarks or modules but can be based on an alternate measure of "adjacency" (*sensu* Chernoff and Magwene, 1999) that indicates whether two traits can be considered "neighbors" with respect to a specific criterion. The criterion for specifying adjacency may be functional, developmental, spatial, or based on any sensible biological data (Cheverud, 1982). For example, if development proves to be a driving force in patterns of variability, we might construct a dissimilarity measure based on the embryonic source of particular structures. With the appropriate data sets, we might expect that landmarks representing features derived from a specific population of neural crest cells might share a dissimilarity metric of lower magnitude (indicating a stronger relationship) even though the structures are anatomically remote.

As a simplified example of the codependent modules model using the mouse dentary, we assign expected values of variation and covariation based on information gained from development and from knock-out experiments in mice as summarized by Hall (2003). Normal dentary bones develop in mice in which Msx-1 has been knocked out, but the teeth and the associated alveolar bone fails to form (Satokata and Maas, 1994). When working with a developmental system in which we know that some aspect of the molecular pathway involving Msx-1 is affected, we might propose that the expressed phenotypic variation in the incisor and molar alveolar morphogenetic units would be similar and amplified, compared with phenotypic variation expressed local to the other morphogenetic units (ramus, coronoid process, condylar process, and angular process). Because the Msx-1 pathway has been manipulated in some way, it is expected that the proposed experiment will yield mice that display increased variation local to the molar and incisive alveolar modules. Moreover, since the same genetic pathway disrupts development in these two modules, we expect covariation between the incisive and molar alveolar modules. To accommodate these observations, we adjust our model by maintaining the expectations of the perturbation pattern for landmarks within those modules, increase the expectation of variation within the molar and incisive alveolar modules, and introduce covariation between modules (Figure 4-5). Similar magnitudes of perturbations are expected for the incisive and molar alveolar modules, while a lower magnitude of variation is proposed between the two alveolar modules and the ramus, coronoid, condylar, and angular morphogenetic units.

How can estimation of $\mathbf{\Sigma}_K$ be achieved under any given model? What we seek is the model of $\mathbf{\Sigma}_K$ that is the best "fit" to $\mathbf{\Sigma}_K^*$, which we compute from real data. To measure the fit of $\mathbf{\Sigma}_K^*$ (which we observe) to the $\mathbf{\Sigma}_K$ model (which we propose based on theory), we consider the following statistic:

$$\gamma = trace\,[(\mathbf{\Sigma}_K^* - \mathrm{L}\mathbf{\Sigma}_K\mathrm{L}^T)(\mathbf{\Sigma}_K^* - \mathrm{L}\mathbf{\Sigma}_K\mathrm{L}^T)^T]$$

LANDMARKS												
LANDMARKS		1	2	3	4	5	6	7	8	9	10	11
	1	**6**	*0*	*0*	*0*	*0*	*0*	*0*	*0*	*0*	*0*	*0*
	2	*0*	**3**	*4.2*	*0*	*0*	*0*	*0*	*0*	*0*	*0*	*0*
	3	*0*	*4.2*	**3**	*0*	*0*	*0*	*0*	*0*	*0*	*0*	*0*
	4	*0*	*0*	*0*	**4**	*0*	*0*	*0*	*0*	*0*	*0*	*0*
	5	*0*	*0*	*0*	*0*	**2**	*2.1*	*0*	*0*	*2.1*	*0*	*0*
	6	*0*	*0*	*0*	*0*	*2.1*	**2**	*0*	*0*	*2.1*	*0*	*0*
	7	*0*	*0*	*0*	*0*	*0*	*0*	**5**	*4.2*	*0*	*3.7*	*3.7*
	8	*0*	*0*	*0*	*0*	*0*	*0*	*4.2*	**5**	*0*	*3.7*	*3.7*
	9	*0*	*0*	*0*	*0*	*2.1*	*2.1*	*0*	*0*	**2**	*0*	*0*
	10	*0*	*0*	*0*	*0*	*0*	*0*	*3.7*	*3.7*	*0*	**5**	*4.2*
	11	*0*	*0*	*0*	*0*	*0*	*0*	*3.7*	*3.7*	*0*	*4.2*	**5**

FIGURE 4-5. Variance–covariance structure for codependent modules. In this model, we constrain measures of variation at a landmark (diagonal elements of the matrix) and covariation between landmarks (off-diagonal elements of the matrix) to values between 0 and 6, 0 indicating a lack of variation or covariation possible. In this model, variation local to landmarks within a morphogenetic unit or module are set to similar values and covariation between landmarks within a module is allowed. In addition, covariation is allowed between landmarks that occupy different morphogenetic units when the scientist has knowledge of an influence that affects more than one module simultaneously. In our example, we propose a murine model in which the Msx-1 pathway has been manipulated. Because knock-out experiments have demonstrated Msx-1 to be necessary to the development of the incisive and molar alveolar morphogenetic units, we expect phenotypic variation to be similarly elevated in these two modules and coviaration to exist among landmarks in these two modules.

Given $\mathbf{\Sigma}_K^*$, the best-fitting estimate of $\mathbf{\Sigma}_K$ (subject to the constraints of the model) is the one that minimizes γ. The estimate must be found by solving a system of nonlinear, simultaneous equations, which can be a formidable computational task when the matrices are large. If $\mathbf{\Sigma}_K^*$ were simply a centered version of our theoretical model, then γ would equal zero. While this exact relationship is unlikely with real data, we recognize values of γ that are closer to zero as indicative of better-fitting models.

In comparing the fits of two different models, we could use their respective γ values to develop a test statistic that would determine which model was the better fit to the data. The bootstrap could be then used to measure uncertainty in the test statistic, allowing us to decide whether one model was significantly better

supported than another. We should reemphasize that estimation of $\mathbf{\Sigma}_K$ using this method requires us to specify a constrained model of variation *a priori;* this specification is far from being an automated process and can be very labor intensive. In theory, there are infinite numbers of potential models, and our most imposing challenge is to use our biological knowledge and experience to specify models that will be as realistic as possible.

V. CONCLUSIONS

The translation of genetic information into phenotype is central to biology. We are, however, a long way from understanding the connection between the underlying genetic architecture and expression of phenotypic variation. Because the entire range of genetic effects varies relative to the genetic variants present in any given genome within any given environment (Rutherford, 2000), evaluation of phenotypic variation and measures of the effects of various factors on the production of the phenotype remain critical to our understanding of many fundamental processes, among them development and evolution. To be of any use to science, however, these quantitative observations need to be precise, and sample parameters estimated from observations must be valid.

Our discussion has focused on the nature of landmark coordinate data and how their dependence upon a coordinate system limits what we can know about phenotypic variation within a sample. The position of a set of landmarks on a form provides a geometric representation of the relative locations of a set of salient features. When landmarks are collected from a sample of forms, each form is captured in its own coordinate system. We have shown that the true correspondence among the various coordinate systems of individuals within a sample can never be known. Models proposed to characterize the perturbation structure of landmark data sets contain nuisance parameters (Neyman and Scott, 1948) that include rotation, translation, and sometimes, reflection. These nuisance parameters make the determination of a valid common coordinate system impossible and prevent the estimation of the variance–covariance structure from empirical data (Lele and McCulloch, 2002; Lele and Richtsmeier, 1990). This limitation should be regarded as a challenge, forcing us to be clever in using what we *can* know to probe the data further.

We have presented alternative, coordinate-system-invariant features of the mean and the variance parameters that can be identified and estimated, and we have suggested ways that they might be used in biological contexts. Importantly, we have proposed how the addition of assumptions based on scientific knowledge might move our model closer to a characterization of variation that can be interpreted in biological terms. As always, the investigator's knowledge of the biology of the forms under study and proper incorporation of this knowledge into the model are key to its correctness and suitability.

ACKNOWLEDGMENTS

We thank Benedikt Hallgrímsson and Brian Hall for inviting us to contribute to this fine volume on the critical issue of variation. The ideas presented in this chapter are the result of years of collaborative effort to which all three authors contributed equally, but with varying points of emphasis. This work was supported in part by Public Health Service awards, 1F33DE/HD05706, 1 P60 DE13078 (Project VI), HD24605, and an NSERC grant.

REFERENCES

Atchley, W., and Hall, B. (1991). A model for development and evolution of complex morphological structures. *Biological Reviews* **66**, 101–157.

Boas, F. (1905). The horizontal plane of the skull and the general problem of the comparison of variable forms. *Science* **21**, 862–863.

Bookstein, F. L. (1982). Foundations of morphmetrics. *Annual Review of Ecology Systematics* **13**, 451–470.

Bookstein, F. (1986). Size and shape spaces for landmark data in two dimensions. *Statistical Science* **1**, 181–242.

Chapman, R. (1990). Conventional Procrustes approaches. In *Proceedings of the Michigan Morphometrics Workshop* (F. Rohlf and F. Bookstein, eds.), pp. 251–267. Ann Arbor, MI: University of Michigan Museum of Zoology.

Chernoff, B., and Magwene, P. (1999). Afterword. In *Morphological Integration* (E. Olson and R. Miller, eds.), pp. 319–348. Chicago: University of Chicago Press.

Cheverud, J. (1982). Phenotypic, genetic, and environmental morphological integration in the cranium. *Evolution* **36**, 1737–1747.

Cole, T., III (1996). Historical note: Early anthropological contributions to "geometric morphometrics." *American Journal of Physical Anthropology* **101**, 291–296.

Cole, T. M., III, Lele, S., and Richtsmeier, J. T. (2002). A parametric bootstrap approach to the detection of phylogenetic signals in landmark data. In *Morphology, Shape, and Phylogeny* (N. MacLeod and P. Forey, eds.), pp. 194–219. London: Taylor and Francis.

Dryden, I., and Mardia, K. (1998). *Statistical Shape Analysis*. Chichester, England: Wiley.

Falconer, D. (1981). *Introduction to Quantitative Genetics*. London: Longman.

Falconer, D., and Mackay, T. (1996). *Introduction to Quantitative Genetics*. Essex, England: Longman.

Goodall, C. (1991). Procrustes methods in the statistical analysis of shape. *Journal of the Royal Statistics Society Series B* **53**, 285–339.

Hall, B. (2003). Unlocking the black box between genotype and phenotype: Cell condensations as morphogenetic (modular) units. *Biology and Philosophy* **18**, 219–247.

Kent, J. T., and Mardia, K. (1997). Consistency of Procrustes estimators. J. R. *Statist. Soc. Ser. B* **59**, 281–290.

Lele, S. (1991). Some comments on coordinate free and scale invariant methods in morphometrics. *American Journal of Physical Anthropology* **85**, 407–418.

Lele, S. (1993). Euclidean distance matrix analysis (EDMA) of landmarks data: Estimation of mean form and mean form difference. *Mathematical Geology* **25**, 573–602.

Lele, S., and Cole, T. M. III. (1996). A new test for shape differences when variance–covariance matrices are unequal. *Journal of Human Evolution* **31**, 193–212.

Lele, S., and McCulloch, C. (2002). Invariance, identifiability, and morphometrics. *Journal of American Statistical Association* **971**, 796–806.

Lele, S., and Richtsmeier, J. T. (1990). Statistical models in morphometrics: Are they realistic? *Systematic Zoology* **39**, 60–69.

Lele, S., and Richtsmeier, J. T. (1991). Euclidean distance matrix analysis: A coordinate-free approach for comparing biological shapes using landmark data. *American Journal of Physical Anthropology* **86**, 415–427.

Lele, S., and Richtsmeier, J. (2001). *An Invariant Approach to the Statistical Analysis of Shapes*. London: Chapman and Hall/CRC Press.

Neyman, J., and Scott, E. (1948). Consistent estimates based on partially consistent observations. *Econometrika* **16**, 1–32.

Richtsmeier, J. T., and Lele, S. (1993). A coordinate-free approach to the analysis of growth patterns: Models and theoretical considerations. *Biological Reviews of the Cambridge Philosophical Society* **68**, 381–411.

Richtsmeier, J., Cheverud, J., and Lele, S. (1992). Advances in anthropological morphometrics. *Annual Review of Anthropology* **21**, 231–253.

Richtsmeier, J. T., Cole, T. M., 3rd, Krovitz, G., Valeri, C. J., and Lele, S. (1998). Preoperative morphology and development in sagittal synostosis. *Journal of Craniofacial Genetics and Developmental Biology* **18**, 64–78.

Richtsmeier, J. T., Baxter, L. L., and Reeves, R. H. (2000). Parallels of craniofacial maldevelopment in Down syndrome and Ts65Dn mice. *Developmental Dynamics* **217**, 137–145.

Richtsmeier, J., DeLeon, V., and Lele, S. (2002). The promise of geometric morphometrics. *Yearbook of Physical Anthropology* **45**, 63–91.

Rohlf, F., and Slice, D. (1990). Extensions of the Procrustes method for the optimal superimposition of landmarks. *Systematic Zoology* **39**, 40–59.

Rutherford, S. L. (2000). From genotype to phenotype: Buffering mechanisms and the storage of genetic information. *Bioessays* **22**, 1095–1105.

Satokata, I., and Maas, R. (1994). Msx1 deficient mice exhibit cleft palate and abnormalities of craniofacial and tooth development. *Nature Genetics* **6**, 348–356.

Siegel, A. F., and Benson, R. H. (1982). A robust comparison of biological shapes. *Biometrics* **38**, 341–350.

Sneath, P. H. A. (1967). Trend-surface analysis of transformation grids. *Journal of Zoology London* **151**, 65–122.

Thompson, D. (1917). *On Growth and Form*. Cambridge, England: Cambridge University Press.

Thompson, D. (1992). *On Growth and Form*. The complete revised edition. New York: Dover.