ABSTRACT.—Focusing on geometric morphometrics (GMM), we review methods for acquiring morphometric data from 3-D objects (including fossils), algorithms for producing shape variables and morphospaces, the mathematical properties of shape space, especially how they relate to morphogenetic and evolutionary factors, and issues posed by working with fossil objects. We use the Raupian shell-coiling equations to illustrate the complexity of the relationship between such factors and GMM morphospaces. The complexity of these issues re-emphasize what are arguably the two most important recommendations for GMM studies: 1) always use multivariate methods and all of the morphospace axes in an analysis; and 2) always anticipate the possibility that the factors of interest can have complex, nonlinear relationships with shape.

INTRODUCTION

Morphometrics literally means ‘the measurement of morphology,’ and is a broad topic with tendrils that creep into almost every branch of paleontology, biology, statistics, and informatics. In this paper, we review the acquisition and analysis of morphometric shape data extracted from three-dimensional (3-D) objects and discuss issues related to the analysis of high-dimensional data by comparing theoretical and empirical morphospaces by analyzing 3-D brachiopod valves generated using the shell-coiling equations of Raup (1961, 1966).

Our focus is on geometric morphometric methods (GMM; Table 1), a class of methods in which the analytical variables are the Cartesian coordinates (XYZ values) of geometric points, usually referred to as landmarks or semilandmarks, which represent the locations of structures on objects such as fossils (Bookstein, 1991; Dryden and Mardia, 1998). GMM thus differs from other multivariate morphometric methods that analyze lengths, areas, volumes, angles, and other kinds of measurement data (e.g., Reyment et al., 1984). One of the primary differences is that GMM uses Procrustes superimposition (Table 1) to rescale objects to unit size, translate them to their geometric centroids, and rotate them to minimize the sum of squared differences between corresponding points, thereby standardizing the objects in a common coordinate system, removing the mathematical component of size, and minimizing shape differences between objects to facilitate hypothesis testing of patterns of shape variation (Gower, 1975; Rohlf and Slice, 1990; Rohlf, 1990). GMM is undoubtedly the most widely used type of morphometrics today. Its strengths include: 1) the ability to effectively remove size and thus focus the analysis on pure shape; 2) the comparative simplicity of collecting coordinate data (although data collection is easier from two-dimensional [2-D] photographs than from 3-D objects), and 3) the ease of visualizing results as transformations of the morphology itself rather than as tables of numbers (Bookstein, 1989; Klingenberg, 2013).

Procrustes superimposition imposes important limits on GMM. Analyses are usually restricted to
Geometric morphometrics (GMM) is a class of shape analyses that use Cartesian coordinates as variables, usually based on Procrustes superimposition to remove variation due to size, translation, and rotation (Bookstein, 1991; Dryden and Mardia, 1998). Ordination is any method for transforming variables to a new set of axes. In GMM, the most common ordination is Principal Components Analysis using the covariance method, which rigidly rotates landmark coordinates to a set of orthogonal axes. Procrustes superimposition is a least-squares algorithm that rescales, translates, and rotates landmark constellations to minimize the sum of squared distances between corresponding landmarks, a criterion that facilitates hypothesis testing that shape differences are statistically significant (Gower, 1975; Rohlf and Slice, 1990).

Shape models = landmark configurations constructed from PC eigenvectors for a particular point in shape morphospace (see Equation 3). Shape models can be constructed for any point in morphospace, regardless of whether it is occupied by a real shape (Rohlf, 1993; Dryden and Mardia, 1998).

Thin-plate splines = a two- or three-dimensional spline algorithm that extrapolates shape deformation to areas between landmarks. TPS is frequently used to illustrate shape differences as a d’Arcy Thompson-like grid and is sometimes used to decompose shape variation by spatial scale (Bookstein, 1989, 1991).

In this paper, we begin by pointing readers to seminal works and other overviews of GMM methods and applications. We then describe how GMM analysis works, with a focus on the mechanics of constructing shape variables and morphospaces. We inventory currently available software for extracting morphometric data from 3-D digital objects. For completeness, we also briefly review common statistical questions that can be answered using morphometric data and their analytical solutions. We also discuss issues associated with analyzing fossils, including breakage and deformation. However, our primary focus is on issues related to identifying biological factors from analysis of multivariate shapes. Our discussion includes an exploration of the relationship between theoretical parameter spaces and empirical shape morphospaces using the Raupian shell-coiling equations, an examination of how different strategies for quantifying the shape of 3-D objects can influence the construction of morphospace, and how phylogenetic covariances interact with other biological factors to determine the orientation of morphospace axes.

All analyses, modeling, and simulations in this paper were performed in Mathematica (Wolfram, 2015), often with assistance of the functions in the Geometric Morphometrics, Phylogenetics, and Quantitative Paleontology packages (Polly, 2014, 2016a, b), unless otherwise noted.

### PREVIOUS REVIEWS OF MORPHOMETRICS

Several comprehensive treatments and reviews of GMM methods already exist, to which we refer readers interested in topics that are not covered adequately in this review. Two books from the 1990s are the core references for GMM. Bookstein (1991), colloquially known as the ‘Orange Book,’ made the seminal leap into geometric morphometrics, combining direct analysis of Cartesian coordinates instead of interlandmark distances, and thin-plate spline deformations (Table 1) as both an analytical algorithm for decomposing shape variation and an illustrative, d’Arcy Thompson-esque method for showing the transformation of one shape into another. Dryden and Mardia (1998) presented GMM from a mathematical and statistical perspective.

Many GMM methods were first presented in an unofficial series of edited volumes that, similar to the Orange Book, are known by the colors of their covers. The ‘Red Book’ (Bookstein et al., 1985) was the antecedent of GMM, remembered especially for the presentation of truss analyses, an approach that focused on landmarks but used standard
morphometric algorithms to analyze a ‘truss’ of interlandmark distances. The highlight of the ‘Blue Book’ (Rohlf and Bookstein, 1990) was a series of papers on Procrustes- and thin-plate spline-based methods. The power and popularity of GMM methods quickly took root and were developed by papers in the ‘Black Book’ (Marcus et al., 1993), including the debut of so-called relative warps analysis, a GMM variant of principal component analysis (PCA) that is used to produce shape scores (Rohlf, 1993), and the ‘White Book’ (Marcus et al., 1996), which contains a variety of methodological and applied papers. A second ‘Blue Book’ (MacLeod and Forey, 2002) focuses specifically on issues related to phylogenetics and morphometrics. A second ‘Black Book’ (Slice, 2005) focuses on anthropological applications, but includes several important methodological papers including a full treatment of 3-D sliding semilandmarks (Gunz et al., 2005). Finally, the ‘Yellow Book’ was published as a special issue of the journal *Hystrix* (Cardini and Loy, 2013), and contains reviews of GMM, including analyses involving phenotypic trajectories, phylogenetic comparative methods, and semilandmarks. Collectively, these volumes contain much of the GMM canon.

Three important textbooks provide synthetic introductions to GMM for beginners. *Geometric Morphometrics for Biologists* (Zelditch et al., 2012) covers the basic theory and methods of analysis (including statistical analysis), and provides applied examples in several areas of evolutionary biology. *Paleontological Data Analysis* (Hammer and Harper, 2006) is a general reference for analytical methods frequently used by paleontologists, including GMM. It contains concise essays on individual methods that include succinct descriptions of their purpose and theory as well as clear equations and algorithms for implementing them. *Morphometrics with R* (Claude, 2008) also covers the theory and methods of GMM, but with a focus on implementing them in the R statistical language.

Finally, three journal article reviews of GMM are especially noteworthy in the present context. Adams et al. (2004, 2013) have twice critically reviewed the state of the field, providing readers with succinct and insightful introductions to methods and prospects for future developments. Viscosi and Cardini (2011) reviewed GMM methods with examples of quantifying and analyzing leaf shapes, and touched on important issues of measurement error, statistical analysis, sample sizes, and visualization. Their paper provides readers with the knowledge and skills needed to competently perform publishable quality GMM analyses with little other introduction. Finally, Webster and Sheets (2010) introduced GMM to paleontologists in the 2010 Paleontology Society quantitative methods short course.

**QUICK-START TECHNICAL GUIDE TO GMM**

As explained in more detail below, GMM begins with the Cartesian coordinates of landmark points placed on the objects of interest, which are then Procrustes superimposed and converted to shape variables in the form of principal components scores (Dryden and Mardia, 1998). Statistical tests, phylogenetic comparative methods, or other analyses are usually performed on the shape variables using the same techniques applied to any multivariate dataset. Many of the software packages described below carry out the computations presented in this section; however, readers can find the following algorithmic details useful for building their own custom analyses in R (R Core Team, 2013), Mathematica (Wolfram, 2015), MATLAB (MATLAB, 2015), or in another programming environment.

Selecting landmarks is the most important step of any analysis. Typical GMM analyses are based on a handful of landmarks placed on homologous structures in 2-D digital photographs, but increasingly, it is possible to acquire landmarks from 3-D objects (e.g., meshes or voxel datasets; Fig. 1). Such data are quick, easy, and inexpensive to collect. The term ‘landmark,’ although often used generally to refer to any Cartesian point used in GMM, strictly refers to a single point placed on biologically homologous structures. The term ‘semilandmark’ refers to points placed algorithmically on a curve (including an outline) or surface (Bookstein, 1991; Gunz and Mitteroecker, 2013). Usually, the curve or surface itself is biologically homologous, but the positions of the individual semilandmarks are arbitrary. Once placed, semilandmarks are analyzed exactly like landmarks, but their biological interpretation can differ. Semilandmarks, like landmarks, can be 2-D or 3-D.
The simplest strategy for applying semilandmarks to curves or outlines is to space a constant number of points equidistant from one another in the same direction from a homologous start point, a procedure that is also used in elliptical Fourier shape and eigenshape analysis (Lohmann, 1983; Rohlf, 1986; MacLeod and Rose, 1993; Polly, 2008b). The curve can also be broken into segments anchored by ordinary landmarks if it has homologous waypoints, which tends to improve correspondence between similar shapes (MacLeod, 1999). Because Procrustes superimposition removes degrees of freedom from the landmarks, which has important implications for statistical analysis. One degree of freedom is removed for scaling, two or three degrees are removed for translation of 2-D or 3-D data, and one or three degrees are removed for rotation of 2-D or 3-D data. For 2-D data, the total degrees of freedom will be \(2k - 4\) (where \(k\) is the number of landmarks) or \(n - 1\) (where \(n\) is the number of objects), whichever is smaller. For 3-D data, the total degrees of freedom will be \(3k - 7\) or \(n - 1\), whichever is smaller. Note that there will often appear to be variance in at least one more dimension, which is an artifact of the curvature of shape space (Dryden and Mardia, 1998; Rohlf, 1999). To correct for this, shape variables can be projected to a Euclidean tangent space (a noncurved approximation of shape space in which analyses can be performed using standard algorithms), although this is seldom necessary for biological data, including fossils, because shape variation is usually so small that the curvature of the space is negligible (Rohlf, 1999). The curvature of shape space can be an issue for objects that have

FIGURE 1.—Collecting landmarks from a 3-D scan of a Devonian brachiopod: (A) photograph of the Devonian brachiopod, *Spirifer murchisoni* Hall, from the Indiana University Paleontology Collection (IUPC #2400); (B) a 3-D mesh, obtained from a Creaform GoScan20 structured light scanner, with four landmarks (loosely following Collins, 2014) denoted on the specimen by solid black circles, as follows: (i) posteriormost point of pedicle valve; (ii) lowermost point of medial sulcus along commissure; (iii) lateralmost point along the flat portion of the commissure on the hinge; and (iv) the ventralmost point of the parasulcus along the commissure.
shapes that are more variable and random, such as stone tools (e.g., Costa, 2010), sedimentary particles (e.g., MacLeod, 2002), or randomly generated models.

Once the coordinates have been superimposed, their covariance matrix is calculated. First, the consensus shape is subtracted from each shape to produce residuals. This step centers the principal component space on the mean shape. The covariance matrix is then calculated from the residuals. Note that for 3-D objects, especially ones that use semilandmarks, the covariance matrix can be large and can present computational problems. Presuming that the number of objects is smaller than the number of coordinates, it could be more efficient to perform the principal components ordination in Q-mode (based on a matrix of the objects) instead of the normal R-mode (based on a matrix of the variables) (Legendre and Legendre, 1998).

Singular value decomposition (SVD) of the covariance matrix is used to compute the eigenvectors that define the axes of the principal component space:

\[
\text{SVD} [P] = UVW^T
\]

where \( P \) is the covariance matrix of the coordinate residuals (the symbol \( P \) is used in evolutionary genetics for the phenotypic covariance matrix); \( U \) is the matrix of eigenvectors with one column per axis of the principal component space; \( V \) is a diagonal matrix of singular values (which are the same as eigenvalues when the matrix being decomposed is square, symmetric, and positive definite, as it is with covariance matrices); \( W \) is identical to \( U \) for square, symmetric, positive definite matrices and can be ignored for our purposes; and \( T \) indicates a transposed matrix. The eigenvalues report the variance of the data along each eigenvector, which means along each axis of the principal components morphospace. By convention, these axes are sorted in descending order by their variances. Note that only the first \( 3k - 7 \) (or \( 2k - 4 \)) or \( n - 1 \) axes will have nonzero variances. Higher axes can be discarded.

Principal components (PC) scores, which are the shape variables used for further analysis (see below), are obtained by multiplying the residuals by the eigenvector matrix:

\[
S = RU
\]

where \( S \) is the matrix of scores, \( R \) is the matrix of residuals, and \( U \) is the eigenvector matrix obtained from SVD. Again, axes that exceed the expected dimensionality of the Procrustes superimposed data can be discarded.

**LANDMARK DATA ACQUISITION FROM 3-D OBJECTS**

Digitization of fossil specimens is now progressing at a rapid pace, in large part due to major technological advances in digital photography, CT scanning, structured light and laser-based surface-scanning systems, and photogrammetry and structure-from-motion programs. However, the acquisition of data from digitized specimens that are suitable for geometric morphometrics provides an increasingly complex set of challenges, with an ever-growing number of software programs and analytical tools to meet these challenges. This section focuses on the acquisition of data from 3-D digitized surface scans of fossil specimens. Such datasets can be obtained from any number and combination of the virtualization methods outlined in the various papers from this short course publication. For 2-D objects and voxel-based 3-D data, unsuperimposed coordinates are usually in pixel (or voxel) units unless special effort is made to scale them to real-world units (e.g., millimeters). Each unit of a digital image, termed a pixel, has color-value information and a specific location. As an extension of this format, the unit of many three dimensional images is called a voxel. Each voxel, similar to a pixel in a 2-D array, is a piece of data that records either color, or grayscale, and location information (in three dimensions, \( X \), \( Y \), and \( Z \)). Volumetric data (e.g., raw tomographic datasets or MRI scans; Sutton et al., 2012), can be analyzed in terms of the Cartesian coordinates of the voxels within that dataset. However, surface scans (e.g., those generated by laser-based scanning systems, photogrammetry, or structured light scanners) are generally presented as meshes. Meshes are arrays of nodes as 3-D points that are connected by internodes. Typical representations of meshes are in polygonal faces, usually triangles, in which each vertex of the polygon gives the specific 3-D coordinate information required for geometric morphometric analyses, and the faces are the component that can record color information.

Geometric morphometric methods, therefore, rely upon the careful selection of precise Cartesian coordinates for landmark-based analyses. Most GMM
programs use a standardized text-file format, including: 1) .TPS, which is the de facto standard for GMM as popularized by Rohlf (2015) in his TPS suite of GMM programs, 2) .NTS files (NTSYSpc format of Rohlf, 1988), 3) .CSV (comma-separated values file), or 4) .TXT (standard ASCII or binary text file). The specific format of each of these file types is one that can cause great consternation in preparing shape data for GMM, and each program treats landmark data acquisition differently. The most widely utilized and standardized format is that of the .TPS file. The ‘thin plate spline’ format of Rohlf (2015) consists of architecture that specifies a number of properties for each specimen in sequence before repeating the pattern for the subsequent specimen in the text file. The first line of text specifies the number of landmarks (‘LM = p’ or ‘LM3 = p’ for 3-D objects, where p is the number of landmarks), and the second line begins a per-row listing of the Cartesian coordinates for each landmark. The following lines can contain identifiers, image names, outline specifications, etc., ending with the scale factor. A sample .TPS file has the following basic format:

```
LM3 = 4
9.124494 10.345037 4.594074
−1.783271 −10.989462 −13.257661
10.716694 −3.154946 14.635059
−4.783275 −8.905925 4.594074
IMAGE = 2400_Spirifer_murchisoni.JPG
ID = 1
SCALE = 0.06664
```

Although the Windows program tpsUtil is quite helpful in structuring these .TPS files, there are many other programs that can produce structured files that can be imported for morphometric analysis. The following section reviews features of a selected set of programs that can be used for acquiring landmark coordinates from images or meshes for their analysis.

**PROGRAMS FOR GMM DATA ACQUISITION AND ANALYSIS**

This section contains a brief review of software for acquiring landmark coordinates from 3-D digital objects and for conducting GMM analyses (Table 2). We focus primarily on open-source packages.

![Table 2](https://www.paleontologicalsocietypapers.org/76/76.pdf)
Landmark 3.6 is a Windows-only free software package developed for evolutionary morphing (Wiley et al., 2005) that allows users to place semilandmarks on 3-D mesh surfaces and export them for analysis in other programs, but it does not perform geometric morphometric analyses per se. Landmark does provide a GMM context for ‘morphing’ one mesh (.PLY, .PTS, and .STL inputs) into another mesh (e.g. along ontogenetic trajectories, evolutionary sequences, etc.). Landmark is available from http://graphics.idav.ucdavis.edu/research/EvoMorph.

MeshLab is an open-source program (MacOS, Windows, Linux) for the processing and editing of surface meshes (Cignoni et al., 2008), and is an essential tool for the inspection, editing, cleaning, and repairing of meshes, in addition to the conversion of myriad 3-D file formats. MeshLab also contains a PointPicker tool that will allow for a .PP text file to be saved and then formatted and imported into the GMM program of your choice. MeshLab is available from http://meshlab.sourceforge.net/.

ISE-MeshTools, from InteractiveSoftwarE, is an open-source, 3-D interactive fossil reconstruction freeware (MacOS, Windows, and Linux; Lebrun, 2014), and is a good mesh manipulation software for preparing data for GMM or acquiring landmarks and exporting for use in other programs. MeshTools is designed to maximize the extensibility of generally rigid constraints of GMM by approximating missing landmarks, merging complementary datasets (e.g., landmarks from both dorsal and ventral orientations of a specimen), and augmenting the general power of statistical inference possible for suboptimal datasets. ISE-MeshTools is freely available at http://morpho.museum.com/downloadMeshtools.

ImageJ and FIJI (FIJI Is Just ImageJ) are Java-based open-source programs (MacOS, Windows, Linux) that are powerhouses for image analysis and processing of digital images. Many specialized packages are available for ImageJ (Schneider et al., 2012) and a number of the ‘standard’ packages for ImageJ are bundled and distributed as FIJI (Schindelin et al., 2012). There are a number of plugins for ImageJ that can be used to acquire 2-D coordinates from images or 3-D coordinates from image stacks (e.g., PointPicker and 3D Viewer). These plugins are available with ImageJ and FIJI from http://imagej.net/Downloads.

The TPS series of software (Windows only; open-source) arguably set the standard for GMM (Rohlf, 2015), in which .TPS file creation and landmark placement occurs in tpsUtil and tpsDig. Superimposition, image unwarping, and averaging are executed in tpsSuper; fitting and visualization of thin-plate splines on trees are performed in tpsSmall; fitting and visualization of thin-plate splines on trees are performed in tpsTree; fitting and visualization of thin-plate splines on trees are performed in tpsTree; and the general power of statistical inference possible for suboptimal datasets. The various IMP programs are available at http://www3.canisius.edu/~sheets/IMP%208.htm.

A package for the R statistical environment (open source for MacOS, Windows, Linux), geomorph (Adams and Otarola-Castillo, 2013) is a versatile and easy-to-use suite of landmark acquisition and GMM data manipulation functions, with the ability to capture landmark coordinates, import landmark data from other programs in 2-D and 3-D, and perform numerous statistical analyses of shape variation and covariation. Additionally, Adams and Otarola-Castillo (2013) provided a thorough walkthrough of data acquisition, 3-D landmark placement, and various analyses performed on sample datasets included within the R package. Although the package is restricted to .PLY file imports that are strictly of the ASCII format, additional 3-D file formats can be utilized, if converted using the Morpho package function ‘file2mesh’ (see below). Analyses include Procrustes ANOVA and pairwise tests, comparison of rates of shape evolution on phylogenies, fit of...
sliding semilandmarks to surfaces and curves, generalized Procrustes analysis (GPA), quantification of morphological integration between modules, estimates of morphological disparity, two-block partial least-squares analysis, and more. R-package geomorph provides excellent graphical depictions of shape evolution and patterns of shape variation and is available by using ‘install.packages (‘geomorph’)’ in the R console or from https://cran.r-project.org/web/packages/geomorph/index.html.

MorphoJ is an open-source package for the R statistical environment (MacOS, Windows, Linux; Schläger, 2016) is a toolbox for numerous geometric morphometric methods including sliding operations for semilandmarks, importing, exporting and manipulating of 3-D-surface meshes, and semiautomated placement of surface landmarks. Morpho can perform two-block partial least-squares regression, Riemannian distance calculations, principal components analysis (PCA), canonical variance analysis (CVA), .TPS grid interpolation, superimposition, and visualization; it also contains functions for the easy exporting of arrays from R to MorphoJ or the EVAN Toolkit (formerly Morphologika; Phillips et al., 2010). The Morpho package can batch-process landmark placement, and automatically superimposes semilandmarks on a folder of meshes using a template ‘atlas’ and is available by using ‘install.packages (‘Morpho’)’ in the R console or from https://cran.r-project.org/web/packages/Morpho/index.html.

MorphoJ is an open-source, cross-platform program (MacOS, Windows, Linux; Klingenberg, 2011) developed for the analysis of geometric morphometric datasets, and will input .CSV, .NTS, .TPS, and .TXT files that record 3-D landmark coordinates (captured in other utilities) and output standard text files (for use in other GMM utilities) as well as .MORPHOJ project files for easy return to analyses in MorphoJ. Analyses that can be performed in this program include principal components analysis (PCA), matrix correlation, covariance of shape variables, two-block partial least squares, regression, evaluation of modularity, canonical variate analysis (CVA), discriminant analysis, morphology mapping onto phylogenies, and some quantitative genetic capabilities to factor in heritability and additive genetic variation. MorphoJ is freely available at http://www.flywings.org.uk/MorphoJ_page.htm.

Morpheus et al. is an open-source cross-platform interface (MacOS, Windows, and Linux; Slice, 2013) for GMM visualization and basic preparation for analyses to be carried out in other programs, and is essentially a command line interface with an attached graphics display pane for data visualization. Although .PLY files can be input, great utility is derived from the numerous export options available. Morpheus et al. will export .LND, .NTS, .RAW, .RDATA, .TPS, and .TXT files so that it plays well with other applications and is available from http://morphlab.sc.fsu.edu/software.html.

Geometric Morphometrics for Mathematica (Polly, 2016a) is a free add-on package for the commercial Mathematica symbolic computation program (Wolfram, 2015) that performs GMM analyses, including Procrustes, two-block partial least squares, thin-plate splines, principal components of shape, multivariate regression and ANOVA, Euclidean Distance Matrix Analysis, and projections of phylogenetic trees into morphospace. It is complemented by the Phyllogenetics for Mathematica (Polly, 2014) and Quantitative Paleontology for Mathematica (Polly, 2016b) packages. It is available at http://mypage.iu.edu/~pdpolly/Software.html.

PAST, or Paleontological Statistics (MacOS and Windows; Hammer et al., 2001) is an open-source, broad-based GUI program that includes a number of data transformation tools, graphical plotting features, numerous univariate and multivariate statistical operations, modeling algorithms, geometric morphometric operations (including 2-D and 3-D principal component analysis, thin-plate spline for 2-D landmarks, linear regression of 2-D landmarks, and elliptic Fourier shape analysis of outline data), and analyses for diversity, time series, stratigraphy, cladistics, and scripting of the PAST program itself. PAST is a workhorse of paleontological statistics that is available, with excellent documentation, at http://folk.uio.no/ohammer/past/.

TINA (This Is No Acronym) is a robust machine-vision environment available for the KNOPPIX operating system, a Linux distribution (Schunke et al., 2012). The KNOPPIX OS is available as a bootable disc image that is burned onto a CD and loads when the boot sequence in the BIOS of a machine is modified to load from the CD drive. TINA is complicated, but is a very powerful and high-level analytical environment for estimating anisotropic
measurement errors on landmarks, and likelihood-based pseudolandmark placement with measurement covariance estimates (as a favored method to traditional Procrustes superimposition and PCA). TINA also performs shape alignment by linear modeling and iterative optimization, landmark covariance estimation, landmark acquisition reproducibility studies using anisotropic measurement covariance to assess statistical equivalence between replicated studies, and Monte Carlo tests that use model parameters with covariance estimates to evaluate expected/obtained measurement errors in shape analyses for both surface and volumetric datasets. TINA and KNOPPIX are available from http://www.tina-vision.net/tina-knoppix/software.html.

ANATOMY OF A PC MORPHOSPACE

Figure 2 illustrates the properties of a typical PC morphospace. Each point within the shape space represents a unique configuration of the 500 semilandmarks used to construct it. The markers A–F show the location of six brachiopod valves after Procrustes superimposition. Only two axes are shown here, but this morphospace has a total of five meaningful axes. The axes are sample limited at \( n = 1 \), but larger samples could have as many as \( 1493 \) axes \((3k − 7, \) where \( k \) is the number of landmarks\). The PC scores record the position of each shell along each axis (illustrated by the two arrows pointing to valve A). There are five scores for each brachiopod in this example, one for each of the five axes. The scores are uncorrelated (the product-moment correlation between the six scores on any two PCs is 0.0).

Each axis represents a literal spectrum of shape, illustrated here by a series of shape models (see Table 1 for definitions) constructed for the positions indicated by the scores under each one. Shape models can be constructed for any point in morphospace by reversing the operation in Equation 2 and adding back in the mean shape:

\[
M = S \cdot U^{-1} + C
\]

(3)

where \( M \) is a matrix containing the XYZ coordinates of the model shape, \( S \) is the set of scores for the point of interest in morphospace, \( U^{-1} \) is the inverse of the eigenvector matrix, and \( C \) contains the XYZ coordinates of the mean (consensus) shape. Note that for eigenvectors, the matrix transpose of \( U \) gives the same result as the inverse and is computationally more efficient to calculate. The origin of the axes (score \( 0,0 \)) represents the shape of the sample mean (consensus shape) and is therefore identical on all individual PC axes for score \( = 0 \).

The units of the axes are ‘Procrustes units,’ where 1 unit equals the centroid size of each shell after superimposition, the magnitude of which is a function of the number of landmarks. The space is scaled so that the Euclidean distance between two objects (e.g., D and E) is equal to their Procrustes distance (the sum of the distances between corresponding landmarks, also known as the partial Procrustes distance). Note that the distance between objects equals the Procrustes distance only if all the axes are used in its calculations: only lengths \( a \) and \( b \) are shown in Figure 2, but the Euclidean distance along \( c \) was calculated using all five dimensions.

Because each point in morphospace represents a unique configuration of landmarks, and because the distances between shapes is preserved, the PC scores can be used as shape variables for most other analyses, including statistical testing (Rohlf, 1993; Dryden and Mardia, 1998). The scores have the convenient property of being statistically independent from one another (the correlation between scores on different axes is 0). Note, however, that subsets of scores (e.g., scores from one of several groups in the larger sample) will be correlated. Each axis (eigenvector) consists of a linear combination of the original variables and thus represents a ‘package’ of variation among landmarks that is internally correlated but uncorrelated with other axes. The axes themselves are sample dependent, and the fact that they are orthogonal (uncorrelated) is imposed by design, not by biology. As with Procrustes distances, however, all of the scores are needed to fully capture each shape. Thus it is seldom justifiable to discard higher axes when performing shape analyses.

As illustrated below, the axes of PC space are sample dependent because the first axis lies along the major axis of variation in the sample. The orientation of the first PC will change (sometimes radically) as objects are added or removed from the analysis. The scores and the eigenvectors will therefore change as the sample changes. However, the changes are always rigid rotations and translations of the axes, and these transformations preserve interobject distances. Thus, the Procrustes distance (Euclidean distance)
between two objects remains the same if calculated using all axes, no matter which other objects were used to construct the morphospace, and no matter how much the relationship between the objects appears to change in a 2-D scatter plot.

THEORETICAL AND EMPIRICAL MORPHOSPACES

The GMM morphospace described above is empirical, derived from Cartesian coordinate sampling of shapes, not a theoretical morphospace as in the Raup (1961, 1966) shell-coiling space. The axes of GMM morphospaces are linear combinations of the Cartesian coordinate variables representing shapes of objects, whereas the axes of theoretical morphospaces are parameters of the processes that generate morphology (Thompson, 1917; Raup, 1966; McGhee, 1999). Those processes can have nonlinear relationships to shape (and thus with shape space) and nonorthogonal relationships to one another (Mitteroecker et al., 2004; Polly, 2008a, Adams and Collyer, 2009; Viscosi and Cardini, 2011; Bookstein, 2013, 2016; Gerber, 2014). Nevertheless, as long as the same configuration of landmarks is used to quantify shape in both kinds of spaces, there must exist a one-to-one mapping between them, although it could be complex and nonlinear.

The complex relationship between theoretical and empirical morphospaces can be illustrated using the Raupian (1961, 1966; Raup and Michelson, 1965) shell-coiling equations (Fig. 3). These equations, which were derived from earlier work on logarithmic spiraling in shell growth (e.g., Thompson, 1917), describe the general shapes of mollusks and brachiopods as functions of rotation of the shell aperture around a coiling axis. The parameters in the equations represent the distance of the aperture from the axis (D), the rate of whorl expansion (W), and the translation of the aperture along the axis (T). Note that T is always 0 for nonspiraled shells such as the brachiopod valves in Figure 3. Along with the shape of the aperture itself (S), these parameters can be used to generate most of the forms present in bivalve, gastropod, cephalopod, and brachiopod valves.

Here, we used the Raupian equations to create virtual brachiopod valves, the surfaces of which were prepopulated with semilandmarks by placing semilandmarks around the aperture curve (we used equally spaced points) and rotating it in increments around the coiling axis. Brachiopods are bilaterally symmetrical, therefore, we held the Raupian translation parameter constant at 0. To more efficiently
model brachiopods, we rescaled the Raupian original parameters so that our W parameter is ~1/100th of the magnitude of the Raupian parameter and our D parameter is scaled with respect to the hinge line instead of the aperture center. By experimentation, a range of values consistent with brachiopod shape was identified for W (1.4 to 2.0) and D (−2.2 to −1.2). Because brachiopod apertures are semicircular, the centers must be shifted negatively toward the coiling axis to produce a realistic shape. We added our own parameters to vary aperture shape (S). A Gabor wavelet function was used to deform an elliptical semicircle into a bilaterally symmetrical pattern of waves to define a 3-D curve representing the aperture, onto which 20 equally spaced semilandmarks were placed. Shape variation in the aperture is controlled by three parameters: S₁ is the number of peaks in the wavelet (which varied from 2 to 6 in this example); S₂ is the amplitude scaling parameter (0.3–0.5); and S₃ is the arc length of the aperture with respect to half of an ellipse (±0.0–0.3 Pi). Samples of brachiopod valves were then created by simulating tip values for each of these five parameters using Brownian motion on a phylogenetic tree (Martins and Hansen, 1997). We used the virtual clade of 28 brachiopods from Figure 4 to illustrate how theoretical morphospace as defined by the generating parameters in the Raupian equations maps onto empirical morphospace as defined by shape variation in the resulting brachiopod valves. The valves (numbered dots) are graphed in two dimensions of the theoretical parameter space in Figure 4B. The axes of this morphospace are the parameters used to generate the valves (W, D, S₁, S₂), two of which are shown here. Their phylogenetic tree was projected into the parameter space by reconstructing the node values of each parameter.
(Martins and Hansen, 1997; Revell and Collar, 2009) and connecting the branches of the tree (Rohlf, 2002). An empirical morphospace was constructed for the same valves using Procrustes superimposition and PCA on the valve shapes as represented by the semilandmarks generated by the Raupian equations. The valves are shown in the first two dimensions of GMM empirical morphospace (Fig. 4C) with their phylogeny.
projected using the same method, but based on the PC
shape scores instead of the Raupian parameters.

The theoretical and empirical morphospaces thus
represent two different ordinations of the same
objects, one based on generating parameters and the
other based on shape analysis. They have many broad
similarities. The relative positions of the valves are
similar in both spaces; the W axis of parameter space
roughly corresponds to PC 1 of shape space, and the D
axis roughly corresponds to PC 2 (if the latter is
flipped). However, many small discrepancies can be
observed that are due to the nonlinear mapping of
parameters into shapes that arises from the non-
linearities of the Raupian logarithmic equations. As
discussed in detail below, this mapping between
theoretical parameter space and empirical shape
morphospace is very complex.

One of the most intuitive mathematical differ-
ences between the two morphospaces is that their
axes have different units and scaling. The GMM
morphospace is a truly Euclidean morphospace
(sensu Mitteroecker and Hutteger, 2009) in which
distances between objects can be calculated using
Euclidean geometry (e.g., the Pythagoras theorem for
calculating the distance between two points based on
the lengths of sides of a right triangle). This is
possible in the empirical shape morphospace because
the axes all have the same units (Procrustes distance
units), thus creating a natural scaling between them.
Furthermore, the two PC axes that define the space
are orthogonal by design. Distances between objects
in the empirical space are therefore also in Procrustes
units, and they equal the square root of the sum
of squared differences between corresponding land-
marks on the valves. The theoretical parameter space
does not have all of these properties, in large part
because its axes have different units. The units of W
are rates equal to the proportional change in aperture
size per revolution, and the units of D are in distances
of whatever units the aperture size is measured (often
a proportion of its circumference or area). Euclidean
distances between points in the theoretical space,
although mathematically calculable, cannot be
expressed in common units. Although the distances
and directions in the empirical shape morphospace
can be interpreted in terms of shape differences and
transformations, which in turn allow processes such
as growth to be studied in the shape morphospace,
the distances and directions in theoretical morphospace
are more difficult to interpret. The theoretical
parameter space is a type of affine space (Mitter-
oecker and Hutteger, 2009) that preserves some
aspects of spatial geometry, including parallel
trajectories, but not others (e.g., direction, angles
between trajectories, and distances between points).
All of the latter features will vary depending on
arbitrary decisions about how to scale one axis to
another (i.e., how to equate a change in W with a
change in D). One of the consequences of these
mathematical differences between the two morpho-
spaces is that the ancestral shapes reconstructed in
them cannot be expected to match because the algo-
rithms used to reconstruct ancestral node values relies
on Euclidean mathematical assumptions that are not
met by the theoretical parameter space.

This is not to say that the equations of theoretical
morphology and their associated parameter spaces
are not useful—indeed, theoretical morphospaces are
arguably invaluable to the quantitative study of form.
Many kinds of operations in parameter space can
be imagined, including likelihood and Bayesian
estimations of the parameters that underlie the evo-
lution of form in a species or clade. The parameters of
theoretical morphology equations are (as an ideal)
linked to the biological processes (evolutionary or
developmental) that produce morphology.

**STATISTICAL ANALYSES**

Statistical analysis of shape, by which we mean test-
ing a hypothesis for statistical significance against a
random null model, is a topic too varied to cover
adequately in a review such as this (the literature
reviewed above includes many papers on statistical
analysis). Briefly, however, a large number of statis-
tical GMM questions in paleontology fall into one of
two general types: testing two or more groups for
difference in shape (e.g., samples from different
localities or stratigraphic intervals), and testing shape
variation for association with another variable, either
continuous or categorical (e.g., body size or dietary
type). Testing shape for association with a continuous
variable is a type of regression question, whereas
testing groups for shape differences and testing for
association with a categorical variable are both
types of analysis of variance (ANOVA) questions.
Because shape is multivariate, both kinds of tests
will be multivariate (Dryden and Mardia, 1998).
Frequently, paleontological data points are from different species, in which case phylogeny must also be taken into account (Martins and Hansen, 1997). Because shape is a complex variable that is unlikely to be multivariate normal, permutation or other non-parametric resampling statistical tests are usually preferred over parametric tests (Manly, 2004; Kowalewski and Novack-Gottshall, 2010). We will present examples of multivariate regression and MANOVA, both with and without phylogenetic correction in the phylogenetic section of this paper.

DEALING WITH BREAKAGE AND DEFORMATION IN FOSSILS

All else being equal, analyzing the shape of fossils is no different than analyzing the morphology of extant organisms (Goswami and Polly, 2010a). As long as the same morphological structures are preserved intact, extinct and extant organisms can be combined into the same GMM analysis with no special consideration. Fossils frequently present special challenges, however, because of breakage and distortion.

Broken fossil specimens can result in missing landmarks, a problem that is difficult or impossible to overcome for GMM analysis because Procrustes analysis requires all landmarks to be present to optimize the superimposition. In ordinary morphometrics, various strategies exist for bypassing or imputing missing data, such as substituting the sample mean value for missing data points in PCA, which is possible because only deviations from the mean contribute to the ordination (e.g., Reyment et al., 1984; Strauss et al., 2003). These tactics do not work with GMM because the mean value of each landmark coordinate is dependent on the superimposition of the entire sample, which is optimized based on the fit of all landmarks. Imputed landmarks thus have a greater potential to affect the outcome of GMM analyses than imputed data points do in ordinary distance-based statistics. Nevertheless, several strategies have been proposed to estimate the positions of missing landmarks. Sometimes they can be estimated by eye based on the topography of the morphology around the missing point (e.g., the position of a broken tooth cusp can often be estimated accurately based on the slope of the surrounding area). In symmetrical structures, missing landmarks can be estimated from their symmetrical homologs by mirroring (Mardia et al., 2000; Webster and Sheets, 2010). If the broken fossil is part of a larger sample with complete sets of landmarks, the position of missing landmarks can be estimated based on their covariance with other landmarks in the unbroken sample (Gunz, 2005; Gunz et al., 2009; Mitteroecker and Gunz, 2009). This strategy depends on the covariance structure being similar in the fossil with the missing landmark as in the more complete sample. All else being equal, the more nonmissing landmarks there are and the more closely related the sample, the more accurate the reconstruction will be.

Sediment compaction, tectonic stress, and other processes frequently deform fossils. Undeformed fossils should always be given preference for GMM studies that depend on accurate comparisons of living morphology. Techniques exist, however, for retrodeforming fossils (Hughes and Jell, 1992; Zollikofer and de León, 2005; Angielczyk and Sheets, 2007; Gunz et al., 2009). The simplest methods are used on symmetrical specimens that are intact but plastically deformed (plastic deformation changes the shape of the object without breaking it), such as equalizing the distances of symmetric landmarks across a reconstructed sagittal plane (Ogihara et al., 2006). Other methods find the minimum amount of stretch required to make the two halves symmetrical (Zollikofer and de León, 2005), or find the best fit of a distorted specimen to an undeformed one using algorithms similar to those used to place sliding semilandmarks (Gunz et al., 2009). Brittle deformation, which can include both breakage and plastic deformation that has independently affected the broken pieces, requires both piecing the object back together and finding optimal parameters for plastically retrodeforming it (Zollikofer and de León, 2005). This procedure has so many independent parameters that many solutions can be found that differ in shape (Boyd and Motani, 2008).

THE COMPLEX RELATIONSHIP OF MORPHSPACE AXES TO BIOLOGICAL FACTORS

A common goal of morphometric analysis is to identify trajectories in morphospace that are produced by biological factors such as developmental
processes, functional relationships to the environment, or phylogenetic history (e.g., Monteiro, 1999; Kim et al., 2002; Mitteroecker et al., 2004; Bastir et al., 2006; Hunt, 2007a, b; Figueirido et al., 2009; Adams and Collyer, 2009; Pierce et al., 2009; Polly et al., 2013b; Schreiber et al., 2014). GMM is a powerful tool for this kind of study, but it must be remembered that the PC axes of morphospace cannot be expected to naturally align with the shape variation produced by individual biological factors; therefore, univariate analysis of a single PC axis is unlikely to fully recover the relationship between it and an independent variable.

The paths of trajectories through morphospace of Raupian shell-coiling parameters illustrate why this is so. Because the Raupian equations were used to generate the brachiopod valves in Figure 4, they can serve as proxies for morphogenetic processes and their relationship to shape space. The Raupian parameters are analogous to genotypes, and his equations are analogous to the developmental interactions that produce phenotypes. Just as a developmental biologist might systematically alter genetic backgrounds to understand developmental processes from the phenotypes that are produced, so, too, can we systematically change the Raupian parameters to trace the corresponding phenotypes in shape space. Knowing the Raupian equations, we know the true relationship between parameters, processes, and shape, and can evaluate the extent to which we could correctly recover them if we only knew the parameters and phenotypes.

Figure 5 shows the trajectories of four of the Raupian parameters (W, D, S₁, and S₂) projected into the first three dimensions of the valve morphospace (compare with Fig. 4C). Valves 3, 23, 35, 40, 46, and 49 are labeled in Figure 5A to aid in orienting it with respect to Figure 4C (note that Fig. 5B is based on a different set of random valves). Each trajectory represents shapes generated when one parameter is varied and the others are held constant at the value of the mean shape (which causes the four trajectories to intersect at the morphospace origin). In this example, W roughly parallels PC 1, in part because varying the rate of whorl expansion affects the positions of nearly every semilandmark on a logarithmic scale, thus generating a lot of shape variance. W is at a slight angle to the PC axes and is gently curved and thus is not fully congruent with PC 1. D roughly parallels PC 2. S₁ is more complex in that it is not linear with respect to shape space, but rather forms an undulating wave traveling along PC 2 with peaks and troughs in PC 3. S₁ is the parameter controlling number of waves in the valve margin. The path of its trajectory through shape space can be understood by imagining the shapes of the valve margins with three and five waves respectively: because of bilateral symmetry, three of the five sulci would be in the same position and thus not contribute to shape difference. Now, imagine margins with three and four waves respectively: none of the sulci would line up, which would make valves with four waves more different from one with three waves than a valve with three sulci is from one with five. The oscillating trajectory of S₁ is therefore caused by alternation between even and odd numbers of waves. S₂, the circumference, has an even more complex trajectory, tracing a stair-step pattern at an angle along both PC 2 and PC 3. This erratic path is caused by ranks of semilandmarks along the radial axis slipping on and off of sulci as wave spacing is changed by changes in the circumference.

Although the correlation between shape W and D can be captured by univariate regression of PC 1 or PC 2 onto them, respectively, the two aperture-shape parameters produce trajectories that are both multivariate in their orientations (traveling along more than one axis) and nonlinear in their paths. These relationships are evident in the univariate summary statistics in Table 3, which reports the amount of variance accounted for by each of the first ten PC axes (including as percent of the total variance) and the coefficient of determination ($R^2$) of the regression of the scores from each axis onto each of the Raupian parameters. $R^2$ can be interpreted as the proportion of shape variation along each axis that is explained by the parameter. As expected, W explains 97% of variance along PC 1 and very little on any other axis. Similarly, D explains a high percentage of PC 2 variation (79%), although it also explains a substantial proportion of PC 3 (16%). The univariate results for the two aperture-shape parameters are more complicated, as one might expect. S₁ has the highest explanatory value for PC 3, but has some variance scattered across other PCs. Note that for all three of these parameters, the sums of the $R^2$ values across the first 10 PCs is nearly 1.0 (100%) because all of the shape variance associated with the

**Figure 5**

The paths of trajectories through morphospace of Raupian shell-coiling parameters illustrate why this is so. Because the Raupian equations were used to generate the brachiopod valves in Figure 4, they can serve as proxies for morphogenetic processes and their relationship to shape space. The Raupian parameters are analogous to genotypes, and his equations are analogous to the developmental interactions that produce phenotypes. Just as a developmental biologist might systematically alter genetic backgrounds to understand developmental processes from the phenotypes that are produced, so, too, can we systematically change the Raupian parameters to trace the corresponding phenotypes in shape space. Knowing the Raupian equations, we know the true relationship between parameters, processes, and shape, and can evaluate the extent to which we could correctly recover them if we only knew the parameters and phenotypes.

Figure 5 shows the trajectories of four of the Raupian parameters (W, D, S₁, and S₂) projected into the first three dimensions of the valve morphospace (compare with Fig. 4C). Valves 3, 23, 35, 40, 46, and 49 are labeled in Figure 5A to aid in orienting it with respect to Figure 4C (note that Fig. 5B is based on a different set of random valves). Each trajectory represents shapes generated when one parameter is varied and the others are held constant at the value of the mean shape (which causes the four trajectories to intersect at the morphospace origin). In this example, W roughly parallels PC 1, in part because varying the rate of whorl expansion affects the positions of nearly every semilandmark on a logarithmic scale, thus generating a lot of shape variance. W is at a slight angle to the PC axes and is gently curved and thus is not fully congruent with PC 1. D roughly parallels PC 2. S₁ is more complex in that it is not linear with respect to shape space, but rather forms an undulating wave traveling along PC 2 with peaks and troughs in PC 3. S₁ is the parameter controlling number of waves in the valve margin. The path of its trajectory through shape space can be understood by imagining the shapes of the valve margins with three and five waves respectively: because of bilateral symmetry, three of the five sulci would be in the same position and thus not contribute to shape difference. Now, imagine margins with three and four waves respectively: none of the sulci would line up, which would make valves with four waves more different from one with three waves than a valve with three sulci is from one with five. The oscillating trajectory of S₁ is therefore caused by alternation between even and odd numbers of waves. S₂, the circumference, has an even more complex trajectory, tracing a stair-step pattern at an angle along both PC 2 and PC 3. This erratic path is caused by ranks of semilandmarks along the radial axis slipping on and off of sulci as wave spacing is changed by changes in the circumference.

Although the correlation between shape W and D can be captured by univariate regression of PC 1 or PC 2 onto them, respectively, the two aperture-shape parameters produce trajectories that are both multivariate in their orientations (traveling along more than one axis) and nonlinear in their paths. These relationships are evident in the univariate summary statistics in Table 3, which reports the amount of variance accounted for by each of the first ten PC axes (including as percent of the total variance) and the coefficient of determination ($R^2$) of the regression of the scores from each axis onto each of the Raupian parameters. $R^2$ can be interpreted as the proportion of shape variation along each axis that is explained by the parameter. As expected, W explains 97% of variance along PC 1 and very little on any other axis. Similarly, D explains a high percentage of PC 2 variation (79%), although it also explains a substantial proportion of PC 3 (16%). The univariate results for the two aperture-shape parameters are more complicated, as one might expect. S₁ has the highest explanatory value for PC 3, but has some variance scattered across other PCs. Note that for all three of these parameters, the sums of the $R^2$ values across the first 10 PCs is nearly 1.0 (100%) because all of the shape variance associated with the
parameter is distributed somewhere in the morphospace, most of it in the first ten axes. $S_2$ has some association with PC 2 (22%) and PC 3 (17%), but its variance is distributed more erratically across many PCs. In fact, only 67% of $S_2$ shape variance is accounted for by the first 10 PCs—43% of its variance is on PCs 11–27, even though they account for <0.1% of the total shape variance in the dataset. The reason is that circumference changes create only small differences in the positions of semilandmarks compared to $W$, $D$, or $S_1$, thus appearing insignificant in the total pool of variation in this dataset, and they produce shape variation that is almost randomly distributed.

**THE SAMPLE DEPENDENCE OF MORPHSPACE ORIENTATION**

Added complexity comes from the fact that GMM morphospace axes are sample dependent, but the trajectories produced by underlying biological processes are not. PC axes are constructed to satisfy a set of linear mathematical rules for orthogonality (each axis is at right angles to all others) and direction (the first axis is aligned to the major axis of variance in the sample, the second is aligned at right angles along the major axis of residual variance, and so on until all variance is accounted for). The changed orientation of the PC morphospace is nothing more than a rigid rotation of the original data: it can be considered as a new directional viewpoint of the original Cartesian landmark coordinates in multivariate space. The orientations of GMM morphospace axes are therefore driven by the sample on which they are based, unlike theoretical morphospaces in which axes are firmly tied to parameters. Add or subtract objects from a sample, and the orientation of morphospace axes will change, even though the relative positions of the objects within it will not (the change is a rigid rotation, which preserves the distances between them, ensuring they remain equal to the Procrustes distance).

Sample-dependent changes in morphospace orientation can have a startling effect on the relationship between individual PC axes and biological factors. Add an outgroup, for example, and what seemed to be a strongly significant statistical association between PC 1 and body mass based on univariate tests can suddenly become insignificant. Figure 5B shows a morphospace constructed from a different random simulation of brachiopod valves based on the same phylogeny with the same parameters. The Raupian equations that relate shape to the underlying parameters are identical, but the sample is
The Raupian parameter trajectories appear to be flipped into a new orientation and transformed. W now parallels PC 2 (although not truly parallel), D parallels PC 1 (also oblique), S1 appears to be more hook-shaped than a sine wave, and S2 appears to have a more irregular trajectory, the major direction of which is parallel to PC 3. The switch of W and D between PC 1 and 2 is straightforward in the univariate variance explained by them (Table 3), but not so the change in S1 and S2, both of which have complex trajectories through shape space so that their variance is still distributed across several PC axes, but differently than with the original sample. In reality, the trajectories of the four parameters have precisely the same relationship to each other and to shape; what has changed is simply the viewpoint in multivariate space, which is positioned relative to the major axes of variation in our sample (i.e., the orientation of the morphospace relative to the trajectories).

Sample dependency of principal components ordination means that individual PC axes should not be expected to be associated exclusively with any particular developmental, functional, ecological, temporal, or phylogenetic factors (Monteiro, 1999; Bookstein, 2013, 2016). For a factor to be reliably associated with only one PC axis, it would either have to be responsible for always producing the bulk of shape variance in every sample and therefore would always dominate variance enough to align with PC 1 no matter what the sample, or the factor would have to be consistently uncorrelated with whatever combination of factors determines the orientation of PC 1, which is the only way in which one of the higher PCs could come to lie parallel to it.

Sample-dependent changes in morphospace orientation will be more pronounced when the sample consists of multiple taxa rather than individuals from the same biological population. Adding and subtracting taxa changes the phylogenetic covariance structure of the shapes, which is part of the covariance structure that drives the principal components ordination along with the morphogenetic processes.

<table>
<thead>
<tr>
<th>Sample 1</th>
<th>Variance % Explained</th>
<th>W (R²)</th>
<th>D (R²)</th>
<th>S1 (R²)</th>
<th>S2 (R²)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PC 1</td>
<td>0.01148</td>
<td>69.4%</td>
<td>0.97</td>
<td>0.02</td>
<td>0.01</td>
</tr>
<tr>
<td>PC 2</td>
<td>0.00286</td>
<td>17.3%</td>
<td>0.00</td>
<td>0.79</td>
<td>0.02</td>
</tr>
<tr>
<td>PC 3</td>
<td>0.00146</td>
<td>8.8%</td>
<td>0.01</td>
<td>0.16</td>
<td>0.82</td>
</tr>
<tr>
<td>PC 4</td>
<td>0.00030</td>
<td>1.8%</td>
<td>0.00</td>
<td>0.00</td>
<td>0.02</td>
</tr>
<tr>
<td>PC 5</td>
<td>0.00024</td>
<td>1.4%</td>
<td>0.00</td>
<td>0.00</td>
<td>0.05</td>
</tr>
<tr>
<td>PC 6</td>
<td>0.00009</td>
<td>0.5%</td>
<td>0.00</td>
<td>0.00</td>
<td>0.06</td>
</tr>
<tr>
<td>PC 7</td>
<td>0.00006</td>
<td>0.4%</td>
<td>0.00</td>
<td>0.00</td>
<td>0.01</td>
</tr>
<tr>
<td>PC 8</td>
<td>0.00002</td>
<td>0.1%</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>PC 9</td>
<td>0.00001</td>
<td>0.1%</td>
<td>0.00</td>
<td>0.00</td>
<td>0.01</td>
</tr>
<tr>
<td>PC 10</td>
<td>0.00001</td>
<td>0.0%</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Total</td>
<td>0.01653</td>
<td>99.9%</td>
<td>0.99</td>
<td>0.97</td>
<td>1.00</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Sample 2</th>
<th>Variance % Explained</th>
<th>W (R²)</th>
<th>D (R²)</th>
<th>S1 (R²)</th>
<th>S2 (R²)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PC 1</td>
<td>0.00412</td>
<td>56.1%</td>
<td>0.05</td>
<td>0.97</td>
<td>0.16</td>
</tr>
<tr>
<td>PC 2</td>
<td>0.00270</td>
<td>36.7%</td>
<td>0.95</td>
<td>0.01</td>
<td>0.02</td>
</tr>
<tr>
<td>PC 3</td>
<td>0.00036</td>
<td>4.9%</td>
<td>0.00</td>
<td>0.01</td>
<td>0.61</td>
</tr>
<tr>
<td>PC 4</td>
<td>0.00012</td>
<td>1.6%</td>
<td>0.00</td>
<td>0.00</td>
<td>0.19</td>
</tr>
<tr>
<td>PC 5</td>
<td>0.00002</td>
<td>0.3%</td>
<td>0.00</td>
<td>0.00</td>
<td>0.01</td>
</tr>
<tr>
<td>PC 6</td>
<td>0.00001</td>
<td>0.2%</td>
<td>0.00</td>
<td>0.00</td>
<td>0.01</td>
</tr>
<tr>
<td>PC 7</td>
<td>0.00001</td>
<td>0.1%</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>PC 8</td>
<td>0.00001</td>
<td>0.1%</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>PC 9</td>
<td>0.00000</td>
<td>0.0%</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>PC 10</td>
<td>0.00000</td>
<td>0.0%</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Total</td>
<td>0.00735</td>
<td>100.0%</td>
<td>1.00</td>
<td>0.99</td>
<td>1.00</td>
</tr>
</tbody>
</table>
(note that these two factors are not the only ones contributing to covariance structure among taxa). Adding or subtracting new individuals to an intraspecific population-level sample should not affect the orientation of the PC axes much, as long as there is a sufficiently large number to estimate the covariance structure reliably (e.g., Cheverud, 1982; Ackermann and Cheverud, 2000; Arnold et al., 2001; Goswami and Polly, 2010a, b; Marroig et al., 2012). Changing the number of intraspecific populations represented in a sample might also alter the orientation of PC axes because phylogeographic and environmental differences in phenotypic covariances occur frequently (e.g., Via and Lande, 1985; Phillips and Arnold, 1999; Polly, 2005). Therefore, multivariate statistics on the entire set of shape variables should always be used when testing for statistical associations between factors such as body size, diet, geographic location, or taxonomic identity (e.g., Monteiro, 1999; Viscosi and Cardini, 2011; Zelditch et al., 2012).

THE DEPENDENCE OF MORPHOSPACE ORIENTATION ON LANDMARKING SCHEMES

As discussed above, many strategies for placing landmarks and semilandmarks are available. For 3-D surfaces, semilandmarks will usually be the obvious solution, but the choice of placement method will seldom be straightforward and will often be based on what aspect of shape a researcher wants to capture and what tools are available. These choices will always have an effect because shape distances between objects and the total shape variance in the sample are directly related to the interaction between point placements and the underlying shape. However, whether that effect is large or small depends on the context.

We illustrate the effects of landmarking in Figure 6 by applying different strategies to the six brachiopod valves from Figure 3. The morphospace ordination in Figure 6A is based on the semilandmarks produced by the Raupian equations: semilandmarks around the aperture are equally spaced and the ranks from the umbo to the margin are logarithmically spaced (closer together on the umbo than close to the valve margins). One could argue that this placement strategy is biologically homologous because each rank is analogous to a growth line. The morphospace in Figure 6B is based simply on an analysis of aperture shape, with semilandmarks equally spaced around it. The third morphospace in Figure 6C is based on equal spacing of semilandmarks around the aperture and along arcs running from the semilandmarks back to the umbo. This scheme is similar to the one in Figure 6A, except that the density of landmarks is roughly the same everywhere.

The differences between these three semilandmarking schemes are small in this example. Allowing for the reversal of PC 2, the first two schemes produce almost identical morphospace ordinations. The similarity might seem surprising given the radical difference in the number of points and the complete absence of points that would measure the dorsal curvature of the valve. However, the similarity in this case is expected because each rank of points in Figure 6A is itself identical to the aperture shape except in scale and orientation. It is unlikely that ordinations of real brachiopods based on surface and apertural shape would be as similar as these hypothetical examples. Although broadly similar, the ordination in morphospace Figure 6C is more different than the other two (e.g., taxon F lies between C and D in this ordination instead of between A and D). This difference might also seem surprising because the placement of semilandmarks is visually similar to that in Figure 6A, but in Figure 6C, the ranks from umbo to margin are equally spaced, which means that only the final rank has the same cross-sectional shape as the aperture.

The similarity among these three ordinations should not be taken for granted. Other published examples demonstrate that the strategy used to place points can have a profound effect on the outcome. Salient examples have been provided by Gunz and Mitteroecker (2013) that demonstrate how sliding semilandmark strategies can substantially improve the fit between shapes (see also Bookstein, 1997; Gunz et al., 2005; and Perez et al., 2006). Polly (2008b) illustrated an example in which the same semilandmarking strategy (eigensurface, which is similar to the equal-spaced point scheme in Fig. 6C) works well for one bone (the calcaneum in carnivores), but exaggerates shape differences in another (the astragalus) because the long neck of the latter produces a sharp boundary that causes substantial changes in the positions of many semilandmarks in response to changes in the proportional length of the
next, which is only a small part of shape variation in the astragalus.

**LINEAR AND NONLINEAR TRAJECTORIES**

Nonlinear relationships between biological factors of interest and shape space are a different matter. Although standard multivariate tests might be adequate to assess the overall association between shape and the factor of interest, estimating the factor trajectory through morphospace requires more specialized techniques. The trajectories can be traced in Figure 5 only because we have access to the equations that were used to generate the shapes. In real
datasets, the goal could be to estimate an analogous morphogenetic process when we know only one underlying parameter (e.g., body size, ontogenetic age, whorl expansion rate, or genetic mutation). In this case, we need a nonlinear model that can be fit to the data. The issue of nonlinearity is important because many developmental and ecological processes are known to have nonlinear relationships in shape space, e.g., human cranial development (Mitteroecker et al., 2004; Bastir et al., 2006), ammonite growth (Gerber et al., 2007), trilobite ontogeny (Gerber and Hopkins, 2011), and sexual dimorphism of fish in different environments (Collyer and Adams, 2007).

Trajectories can be traced by connecting points in shape space or by fitting curvilinear (nonlinear) regressions (Mitteroecker et al., 2004). Once trajectories have been identified, an approach for comparing them is phenotypic trajectory analysis (PTA; Collyer and Adams, 2007, 2013; Adams and Collyer, 2009). Any given trajectory can be described by parameters of its own (i.e., size, direction, shape and orientation in morphospace) that allow it to be analytically compared to others. Trajectories have a position, direction, and length (size) in shape space that correspond to the starting and ending shapes, the change in shape along the trajectory, and the amount of change. Because GMM morphospace is truly Euclidean, these PTA parameters describe the differences between the kind and magnitude of shape change associated with the biological factor of interest. Adams and Collyer (2009) measured trajectory size using the path length along the segments of the trajectory, which is equivalent to the total shape change in Procrustes units, overall trajectory orientation using the major axis (PC 1) of the trajectory points (which can be imagined as the straight path that parallels the trajectory as closely as possible), and trajectory shape using Procrustes analysis of the trajectory points, which works as comparison of the landmark points themselves. Statistical tests (e.g., MANOVA) can then be used to determine whether differences in these parameters are statistically significant.

SHAPE SPACE AND PHYLOGENETIC PATTERNS

Paleontological shape data frequently have a phylogenetic component in the form of lineages of a single species sampled through a series of stratigraphic units or a collection of tip species on a phylogenetic tree. Shared ancestry introduces its own pattern of covariances that not only influence the similarities and differences between taxon shapes (i.e., Procrustes distances), but also affect the orientation of morphospace (i.e., the shape variables and the alignment of the principal components relative to other biological factors). These covariances also cause nonindependence in the data points, which must be taken into account for certain kinds of statistical analyses (but not for all questions). They can also produce sometimes-unexpected distributions of taxa in shape space.

Methods for phylogenetic comparative analysis of tip taxa are well described elsewhere (e.g., Felsenstein, 1985, 2012; Graffen, 1989; Harvey and Pagel, 1991; Garland et al., 1992; Martins and Hansen, 1997; Butler and King, 2004; Revell and Collar, 2009; Revell, 2010), and all involve procedures for correcting for phylogenetic covariances as part of performing statistical analyses such as regressions or ANOVAs. The two most common procedures are: 1) phylogenetic independent contrasts (PIC), which removes the effects of shared common ancestry by reducing the data to distances between pairs of taxa and/or nodes (contrasts) so that the change that took place along each branch of the phylogeny is represented only once (Felsenstein, 1985); and 2) phylogenetic generalized linear models (PGLM), which adjust regression parameters for the expected covariance associated with the phylogenetic topology (Martins and Hansen, 1997). Both approaches normally require the assumption that the traits have evolved by Brownian motion (random walk), but variants of the PGLM method can accommodate overdispersion (e.g., disruptive selection) or underdispersion (e.g., stasis, constraint, Ornstein-Uhlenbeck processes). The PGLM and related likelihood methods can also be used to estimate the most likely ancestral states for traits, also assuming a Brownian motion mode of evolution (e.g., Martins and Hansen, 1997; Mooers and Schluter, 1999; Pagel, 1999; Polly, 2001; Pagel et al., 2004). Several authors have specifically discussed adaptations of these methods for geometric morphometrics, which generally involves applying multivariate versions of these standard methods to shape scores (Rohlf, 2001; Caumul and Polly, 2005; Klingenberg and Gidaszewski, 2010;
and there is no nonphylogenetic correlation between the two. Real world examples are likely to be more complicated, if for no other reason than evolution cannot be expected to follow a Brownian motion model.

Species lineages sampled through stratigraphic intervals are a special case of phylogenetically structured data, one that is simpler to deal with in many respects. A species lineage is equivalent to a single branch on a phylogenetic tree—a series of putatively ancestor-descendant population samples with no branching events—such that later samples share phylogenetic history with earlier ones. Such data constitute a classic time series, from which spurious correlations can arise between the biological data of interest and independent factors (e.g., Raup, 1977; Roopnarine et al., 1999; Gingerich, 2009; Hunt, 1977; Rabosky, 2010). The correlation is caused by the fact that most time-series data, including Brownian motion random walks, have general trends over time. Two such trends, even though unrelated, will produce a serial correlation, either positive or negative, simply because they are measured over the same interval. Arguably the simplest approach to correcting for this phenomenon is the method of first differencing (McKinney, 1990). Instead of analyzing the values of the dependent and independent variables, the differences (changes) between points in the time series are measured. If the variables are truly correlated, they should, on average, change in concert, so an underlying connection between the variables should produce a significant correlation between their first differences. This approach is almost identical to phylogenetic independent contrasts in that the time series is decomposed into its constituent independent changes.

Evolutionary time series can produce surprising patterns in GMM morphospace (Fig. 7C–D) (Mitteroecker et al., 2004; Bookstein, 2013). All things being equal, a sequence of taxa in an evolutionary time series forms a shape gradation in which shapes from adjacent intervals are more similar than shapes at the series beginning and end. When data have high dimensionality, which GMM semilandmarks on 3-D objects usually have, the first PC of a densely sampled lineage will usually contrast its beginning and end, much as the trajectory orientation parameter of Adams and Collyer (2009) is estimated with the first PC of its points. The second PC, which is always at right angles to the first,
will usually contrast the middle of the series with its ends, thus producing a ‘horseshoe-shaped’ pattern of points in morphospace. Furthermore, the third PC will contrast the centers of the regions lying between the middle and the endpoints. The last two panels in Figure 7 illustrate this phenomenon. Ten Brownian motion random walks with 150 variables (50 landmarks with three dimensions each) were generated for 1000 steps, shown with their first two variables in Figure 7C. Each of these was submitted to principal components, all of which are shown together in a single graph of the first three PC axes in Figure 7D. The same kind of curved pattern can be expected from any gradational series, such as analysis of segmental structures such as vertebrae (e.g., Head and Polly, 2015). Bookstein (2013) described the mathematics behind this phenomenon, the predictability of which increases along with the number of variables. This phenomenon is not a problem for analysis per se (other than the issue of serial correlation if shape is being regressed onto another variable of interest), because the original paths of each time series are preserved in the full multidimensional shape space. Corrections such as detrended correspondence analysis should never be applied to shape data because they distort shape space in unpredictable ways that can invalidate statistical conclusions and compromise the ability to map trajectories in the detrended space back to shape. Instead, all dimensions of the shape variables should

FIGURE 7.—Phylogenetic effects on morphospace ordinations: (A) multivariate regression of shape scores onto a randomly evolved variable; shape and the random variable are not correlated, except as an artifact of phylogenetic covariance; the univariate relationship of PC 1 is shown here; (B) multivariate regression of independent contrasts of shape onto independent contrasts of the same random variable; phylogenetic covariances no longer influence the regression and the apparent significance of the relationship has disappeared; (C) ten high-dimensional, Brownian-motion, random-walk lineages illustrated with the first two of 150 variables; (D) the same ten random walks, each ordinated with principal components and plotted together. See text for discussion.
be used so that the analysis captures the true pattern in the data rather than the partial pattern on the first few axes. For example, evolutionary model fitting would likely find a directional trend in these data if only the first two or three PCs were considered, but Brownian motion would be the obvious pattern based on the entire dataset.

The curvature of gradational shapes in the first few axes of shape space is usually present in any sample with a phylogenetic component, such as is visible in the W trajectory in Figure 5A, but the pattern is often subtle because sampling of the tree is sparse (usually tips only), and because the branching pattern adds noise to the gradation. However, principal component plots of tip taxa often have their own regular patterns: the first PC axis often separates the most distantly related taxa because they are, all other things being equal, the most different in shape. The second axis separates the next closest clade, the third axis the next closest, and so on. Differences between closely related tips are often found on the highest axes. With the caveats already mentioned, this is not a problem as long as all dimensions of shape space are included in analyses, and phylogenetic comparative methods are applied when necessary. Revell (2009) described an ordination method called phylogenetic principal components analysis (pPCA) that removes the expected covariance due to phylogeny from the covariance matrix used to calculate the eigenvectors (P in Equation 1). This procedure aligns objects with an axis that is not affected by phylogenetic covariances, but it does not remove phylogenetic correlation from the PC scores (Revell, 2009; Polly et al., 2013a). The properties of pPCA complicate many of the ordinary interpretations of morphospace (e.g., the first PC of pPCA might account for less variance than higher PCs) and should be used with caution. Fully multivariate statistical results should be identical regardless of whether PC or pPC shape scores are used.

CONCLUSIONS

Geometric morphometrics provides a powerful set of methods for exploring, analyzing, and modeling 3-D shape. Practical hurdles for acquiring GMM data from 3-D objects are substantial, but are becoming easier as new software tools are developed. Theoretical hurdles for how to represent the complexity of 3-D objects with landmarks and semilandmarks are only beginning to be addressed, partly because some of the issues of homology, weighting, and artifacts were not apparent until recent years when 3-D analyses began to be feasible. The complexity of 3-D shape analysis re-emphasizes what are arguably the two most important recommendations for GMM studies: 1) always use multivariate methods and all of the morphospace axes in an analysis; and 2) anticipate the possibility that the factors of interest have a complex, nonlinear relationship with shape.

ACKNOWLEDGMENTS

L. Tapanila and I. Rahman generously invited us to participate in this short course and were patient with our schedules. A. Reese scanned the brachiopods used in Figure 1. P. Novack-Gottshall provided insightful comments on some of our analyses. Special thanks to S. Marcus for fielding questions, handling special formatting requests, and seeing both this paper and the entire volume through to publication. Scanners used to acquire data were supported by the Indiana University Center for Biological Research Collections. Supercomputing was supported by Lilly Endowment through the Indiana University Pervasive Technology Institute and the Indiana METACyt Initiative. This research was partially supported by National Science Foundation grant NSF EAR-1338298 to PDP.

REFERENCES


Adams, D.C., 2013, Quantifying and comparing phylogenetic evolutionary rates for shape and other high-dimensional phenotypic data: Systematic Biology, v. 63, p. 166–177.


evolution are rare in comparative data: Evolution, v. 64, p. 2385–2396.


Polly, P.D., 2008a, Developmental dynamics and G-matrices: Can morphometric spaces be used to model phenotypic evolution?: Evolutionary Biology, v. 35, p. 83–96, DOI: 10.1007/s11692-008-9020-0.


Polly, P.D., 2014, Phylogenetics for Mathematica, ver. 3.0: Bloomington, Indiana, Department of Geological Sciences, Indiana University.

Polly, P.D., 2016a, Geometric Morphometrics for Mathematica, ver. 12.0: Bloomington, Indiana, Department of Geological Sciences, Indiana University.

Polly, P.D., 2016b, Quantitative Paleontology for Mathematica, ver. 4.3: Bloomington, Indiana, Department of Geological Sciences, Indiana University.


