Zoological Systematics, 42(1): 4–33 (January 2017), DOI: 10.11865/zs.201702

REVIEW

Morphometrics: History, development methods and prospects

Norman MacLeod

Abstract Morphometrics has been pursued by graphical and computational means since the European Renaissance, drawing on core geometric principles first discovered in China and Classical Greece. Through the late 1800s, two distinct approaches to such analyses were pursued: a deformationist approach, epitomized by D'Arcy Thompson's graphical trans-formation grids and the statistical approach popularized by Francis Galton, Karl Pearson, and Julian Huxley in which Cartesian spaces were employed to summarize patterns of variation in size and/or shape variables. Unification of these approaches was an oft-stated goal throughout the 20th century, but proved elusive until the mid-1980s when David Kendall, Fred Bookstein, and Colin Goodall proposed a radically new way of understanding form — as the locations of configurations of landmarks on the surfaces of a nested series of hyperdimensional manifolds. Once this new mathematics of form was understood development of basic concepts, procedures, graphical tools, and statistical tests followed quickly such that the core of the long-hoped for synthesis took less than a decade to achieve. The result — geometric morphometrics — continues to develop into an ever-more extensive toolkit that can be used by researchers to describe and understand a wide range of problems involving the characterization of morphological similarities and differences in all of their many and varied contexts. In particular, the new approaches involving the direct analysis of image pixels and new tools such as machine learning and artificial intelligence are set to reinvigorate (and possibly to revolutionize) the field once again.

Key words Morphometrics, form, size, shape, biology, geometry.

1 Introduction

Morphology is a critical source of information throughout the natural sciences, materials sciences, and engineering. Across all these fields an object's form consists of two primary components: size and shape. Size is the physical scale of an object, usually determined by comparing one or more of its spatial dimensions (e.g., diameter, height, length, width, perimeter, area, volume, mass) to a reference which serves as a basis for measurement. Shape is also usually conceptualized via comparison to some reference (e.g., circle, triangle, mean distance, mean point configuration) and has been defined operationally as that component of form which remains after differences in size, position and rotation between two or more objects have been discarded (Kendall, 1977). Morphometrics, then, is the quantitative analysis of form and covariances with form (Bookstein, 1991). While this domain overlaps strongly with that of geometry, to date morphometrics has been pursued in descriptive biological contexts of far more restricted scope than those of geometry, biometry and/or spatial analysis, despite the fact that all these fields share a common origin.

The Chinese philosopher Mozi (470-390 BC) authored the earliest known description of a mathematical point which he defined as the part of a line which cannot be divided into smaller parts (Needham, 1959). Mozi, like Plato and Euclid who established geometry as the foundation of mathematics in western Europe, noted that a line segment is a one-dimensional figure joining two mathematical points in space and that the length of a line segment records the separation between its defining endpoints. These concepts of point and separation are fundamental to morphometrics where point

Special Issue: Geometric morphometrics: Current shape and future directions
Received 6 October 2016, accepted 6 January 2017
Executive editor: Fuqiang Chen; guess editor: Ming Bai

locations and the endpoints of line segments are often referred to as "landmarks" and their separations as "distances" respectively.

The first use of morphometric approaches to depict organic forms can be traced back to the origins of illustrative art in which the sizes of stylized, but recognizably human and animal, figures were represented as drawings, carvings or sculpture, usually at sizes that illustrated their importance to the artist rather than their true physical scale (Gombrich, 2007). Accuracy in the depictions of bodies and the component parts thereof became increasingly important — especially in terms of their three-dimensional representation as bas reliefs, carvings and sculpture — from the Iron Age onward. Noteworthy centres of geometrically advanced art were established in China, Mesopotamia, Egypt, Greece and Rome through ancient and classical times.

Following regression to a geometrically primitive artistic style during the European Middle Ages, interest in increasing the accuracy of depictions of humans, animals, plants and landscapes was revived throughout Europe in the Renaissance, especially through the (re)discovery and later mathematization of the principles of linear perspective by Ambrogio Lorenzetti, Filippo Brunelleschi and Leon Battista Alberti, the latter two of which were architects as well as artists. Perspective was a graphical procedure originally whereby the angles and lengths of lines on a two-dimensional surface could be adjusted to give the viewer the illusion of three-dimensional space (Fig. 1, see Andersen, 2007). This procedure was later employed by artists such as Leonardo da Vinci and Albrecht Düer to give their paintings a realism that artists who did not utilize perspective could not match. Da Vinci and Düer were particularly noteworthy for their systematic geometric studies of the human body, undertaken in an effort to identify ideal (= average) proportions and to express deviations from this ideal via a unified system of geometric transformations that included both isotropic (uniform in all directions) and anisotropic (directionally non-uniform) components (Fig. 2). Bookstein (1993) referred to this school of morphometrics as "distortionist" though others also refer to it as "transformationist" (Bookstein, 1977) and "deformationist" (Gould, 1971, 2002; Reyment, 2005).

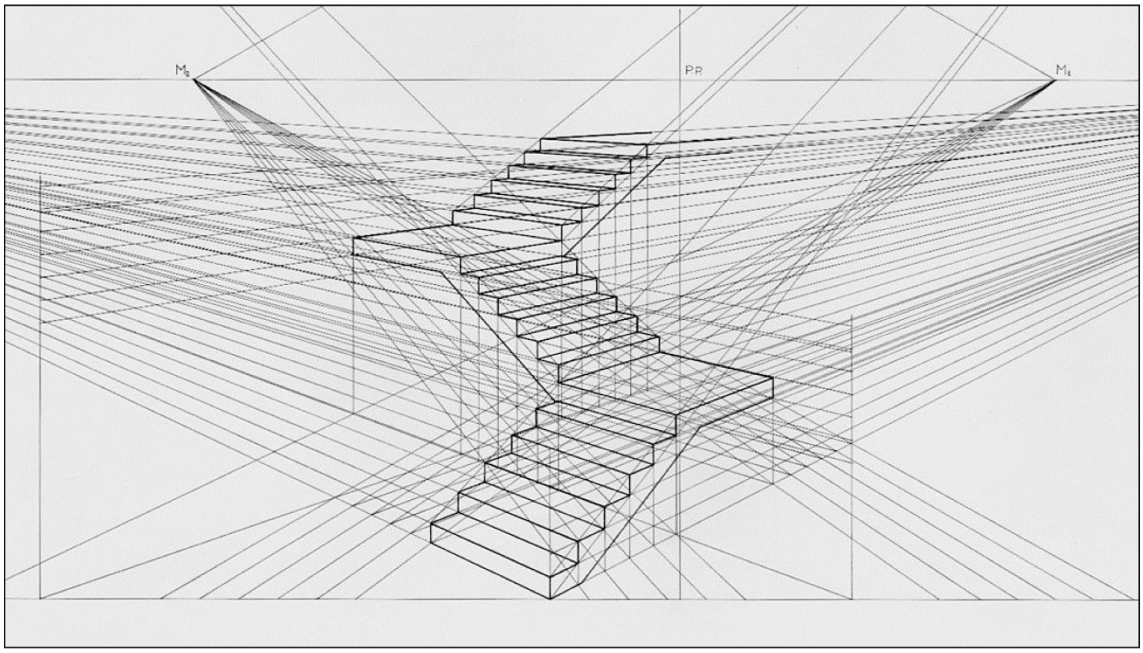

Figure 1. Drawing of a staircase using simple two-point perspective with the construction lines left in. Renaissance artists such as Lorenzetti, Brunelleschi and Alberti used graphical methods such as these to represent the forms of three-dimensional objects on two-dimensional surfaces — an early type of graphic morphometric analysis.

Proceeding from these artistic origins, the deformationist approach to morphometric analysis was used by the classical scholar and mathematician D'Arcy Wentworth Thompson (1917, 1945) to express biological shape variation as deformations of corresponding or topologically homologous points located across the bodies of closely related species in the manner of a mathematical grid (Fig. 3). Thompson's noteworthy discovery was that many seemingly complex morphological transformations could be produced through the application of rather simple geometric deformations to all regions of the body. Thompson used this observation to argue that evolutionary transformations between species might be generated by geometrically simple alterations in ontogenetic and phylogenetic developmental programmes and that the physical character

of the materials from which organismal bodies are constructed may place constraints on the modes of deformation or transformation that can be realized (Thompson, 1917, 1945). While Thompson's deformation-grid approach, like Brunelleschi's procedures for employing perspective and Düer's geometric systems of human facial-form variation, were strictly graphical constructs, by the latter half of the 20th century mathematical procedures were being employed to generalize and operationalize them into sophisticated computational tools for the analysis of morphological variation. These tools included trend-surface analysis (Sneath, 1967), biorthogonal grids (Bookstein, 1977,1978), and thin plate splines (Bookstein, 1989, 1991).

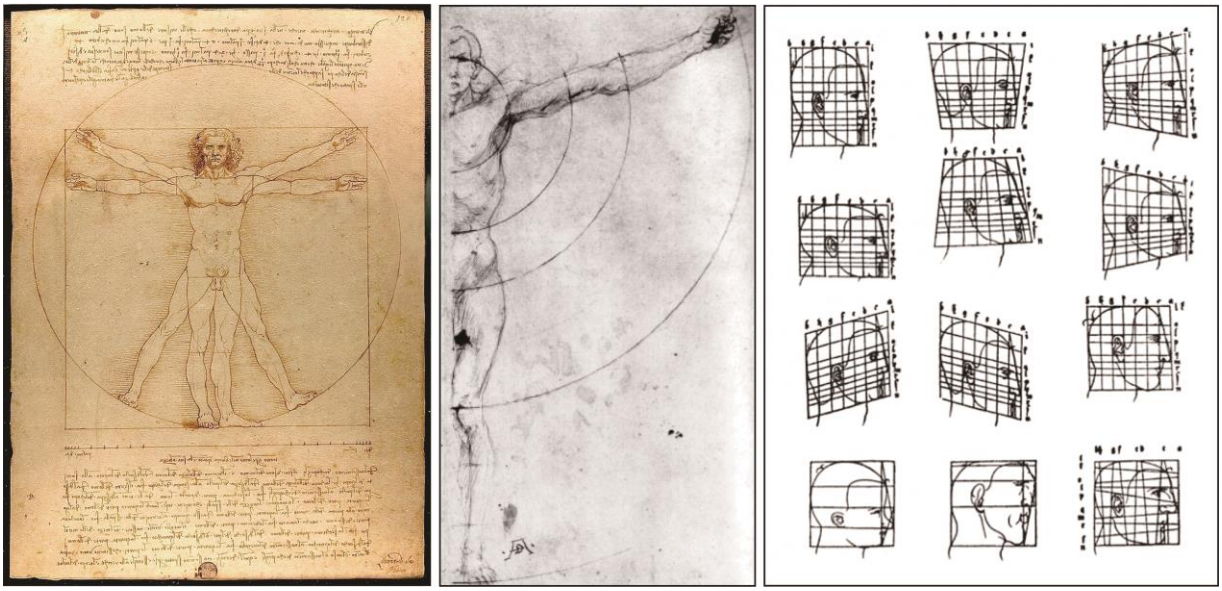

Figure 2. Examples of graphical morphometric analysis. Left. Leonardo daVinci's *Virtuvian Man* which is a graphical analysis of ideal or average human proportions (c. 1480). Center. Albrecht's Düer's sketch of ideal human arm proportions relative to other regional of the body (c. 1524). Right. Albrecht's Düer's sketches of human facial proportions as a system of isotropic and anisotropic deformations, from *Vire Bücher von Menchlicher Proportion* (1524).

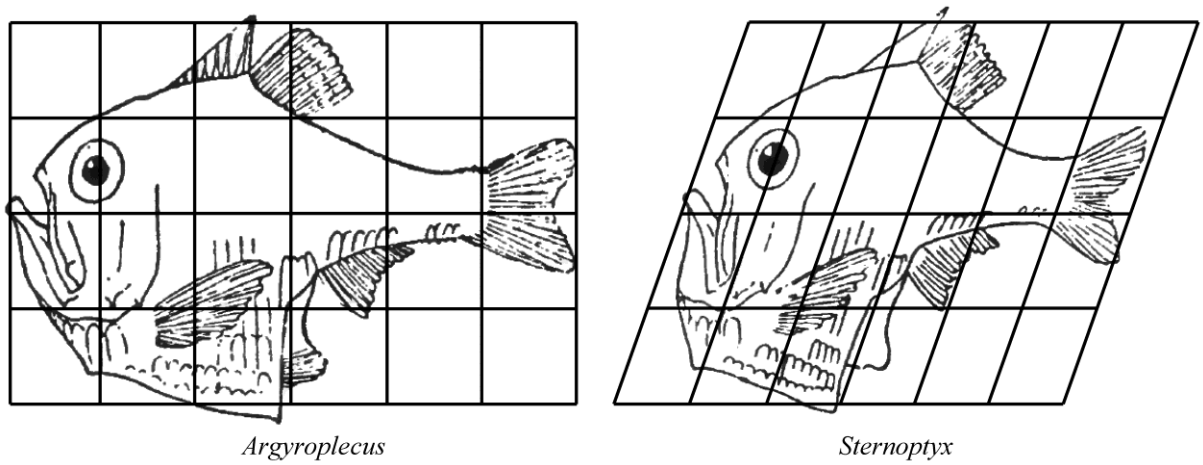

Figure 3. D'Arcy Thompson's (1917) deformation grid analysis of morphological variation for the transformation between *Argyroplecus* and *Sternoptyx*. Although these diagrams appeared intriguingly mathematical in design Thompson created them using qualitative techniques only. Redrawn from Thompson (1917).

The other traditional approach to morphometric analysis was statistical and derived from a biometric interest in social order and inheritance by way of human demographics (e.g., Graunt, 1662; Petty, 1690). The work of Adolphe Quetlet, particularly his 1835 book *Sur l'Homme et le Développement de ses Facultés, ou Essai de Physique Sociale*, is noteworthy with regard to its characterization of the "l'homme moyen" (average man) as the mean values of measured variables that follow a normal distribution. Within this school of morphometric analysis measurements taken on specimens were often

summarized graphically by plotting their values in a Cartesian space which was then used to guide and interpret various statistical summaries of form similarity and difference.

For example, the founder of modern biometry, Francis Galton, used the heights of 928 children as the basis for his seminal 1886 work on inheritance which introduced the concept of linear regression and laid the groundwork for Karl Person's (1901) development of principal components analysis (PCA). Robert A. Fisher (1936) used Edgar Anderson's (1935, 1936) sepal and petal length measurements taken on three species of *Iris* in his development of linear discriminant analysis. Julian Huxley (1924) in England and Georges Teissier (1928) in France developed the concept of relative growth, or allometry, using linear distance measurements collected from the crab *Unca pugnax* and the omatidia lengths of various insect species respectively (see Gayon, 2000). Sewell Wright (1921) used morphometric measurements collected from domestic chickens to illustrate his method of path analysis, variants of which are equivalent to multiple regression analysis, factor analysis, canonical correlation analysis, discriminant analysis, and a wide variety of variance and covariance analyses (e.g., MANOVA, ANOVA, ANCOVA). Finally, Calyampudi R. Rao (1948) used the generalized statistical distance (Mahalanobis, 1936) between sets of linear distance measurements to construct a topological representation of form similarities and differences between human crania.

Given the extended history of serious interest in the range of topics now subsumed under the heading "morphometrics", it is curious to note that this term was not coined until 1957 by Robert E. Blackith who used linear distance measurements to study polymorphism in locusts. Reyment and Blackith published the first book-length treatment of the topic in 1971 with a second edition (co-authored with N. A. Campbell) appearing in 1984. Since 1984 the growth of interest in, and publications on, applications of morphometric methods have been exponential (Adams *et al*., 2004, 2013).

2 Traditional approaches to morphometric data collection

While most of the technical morphometric literature is devoted either to describing new data-analysis techniques or the results of the application of morphometric data analysis to a sample of organisms, a decidedly neglected and confused topic fundamental to all morphometric analyses is the design of a set of measurements that can be taken on a sample of specimens (or parts thereof) in order to address some particular scientific problem of interest.

2.1 Linear distances

Following examples set by R. A. Fisher (1936) and C. R. Rao (1948), statistical morphometric measurement-design strategies have relied primarily on the collection of two-dimensional linear distances from extremal points located on a set of specimens and/or major subregions thereof. These measurements are typically oriented parallel and/or at right angles to the subject's axis of symmetry (Fig. 4). This sort of measurement strategy has also been used to calculate so-called "form factors" or "shape factors" in non-organic particle analysis and microscopy (e.g., Merkus, 2009; Olsen, 2011a, b; Rodriguez *et al*., 2013).

As noted in these references, an exceedingly large number of "form" and "shape" factors have been developed by engineers and sedimentologists in their efforts to characterize, classify, and/or grade rock particles for size, roundness, angularity and degrees of sorting. The problem with such indices when used in the context of biological morphometric analyses, however, is that linear distances (incl. perimeters, areas and volumes) are scalar magnitudes that can be used to represent size, but contain no shape information whatsoever. Even when such distances are constrained to adopt fixed orientations relative to one another (e.g., be perpendicular) and combined so that they specify classes of possible shapes rather than individual shapes themselves (see below), their utility for characterizing specific shapes in an unambiguous manner is limited. Moreover, if combined to form ratios (e.g., to standardize for "size") such indices exhibit the additional problem of each measurement being composed of an aspect of size variation *and* an aspect of shape variation. Since these aspects of form variation are compounded differentially for each specimen, no ratio is able to represent pure size or pure shape, and so be used compare the component of size (or shape) to any other variable in a simple and straightforward manner (Mosimann, 1970; Bookstein *et al*., 1985; Albrecht *et al*., 1993). This complication also applies to the representation of shape as a residual from the regression of any "size" variable on a set of linear distance variables (Bookstein *et al*., 1985; Albrecht *et al*., 1993). In addition, the representation of specimens via the recording of gross dimensions tends to result in overly redundant systems of measurements that do not achieve an even assessment of all aspects of the forms in question (see Fig. 4), and so veers toward the production of data matrices that exhibit high colinearities (but see Strauss and Bookstein, 1982 for an interesting linear measurement-design approach that minimizes this tendency).

Within Rao's multivariate morphometric approach to shape analysis, it is common to find research reports that employ

a PCA of a covariance matrix calculated from a set of linear distance variables as a means for separating size and shape variation under the assumption that the first principal component represents a multivariate size axis (e.g., Jolicouer & Mosimann, 1960). Despite its common assertion in peer-reviewed biological morphometric literature, this is not a correct interpretation of such results. As has been pointed out by Strauss (1985) and Somers (1986), for PC-1 to represent a true isometric size vector all eigenvector loading coefficients would need to be positive and equal (= pure or isometric size change). Burnaby (1966), Somers (1986), Bookstein *et al*. (1985) and Rohlf and Bookstein (1987) present modifications of the standard PCA procedure that will ensure PC-1 adopts this correct orientation thereby allowing the higher PCA eigenvectors to represent aspects of pure shape variation.

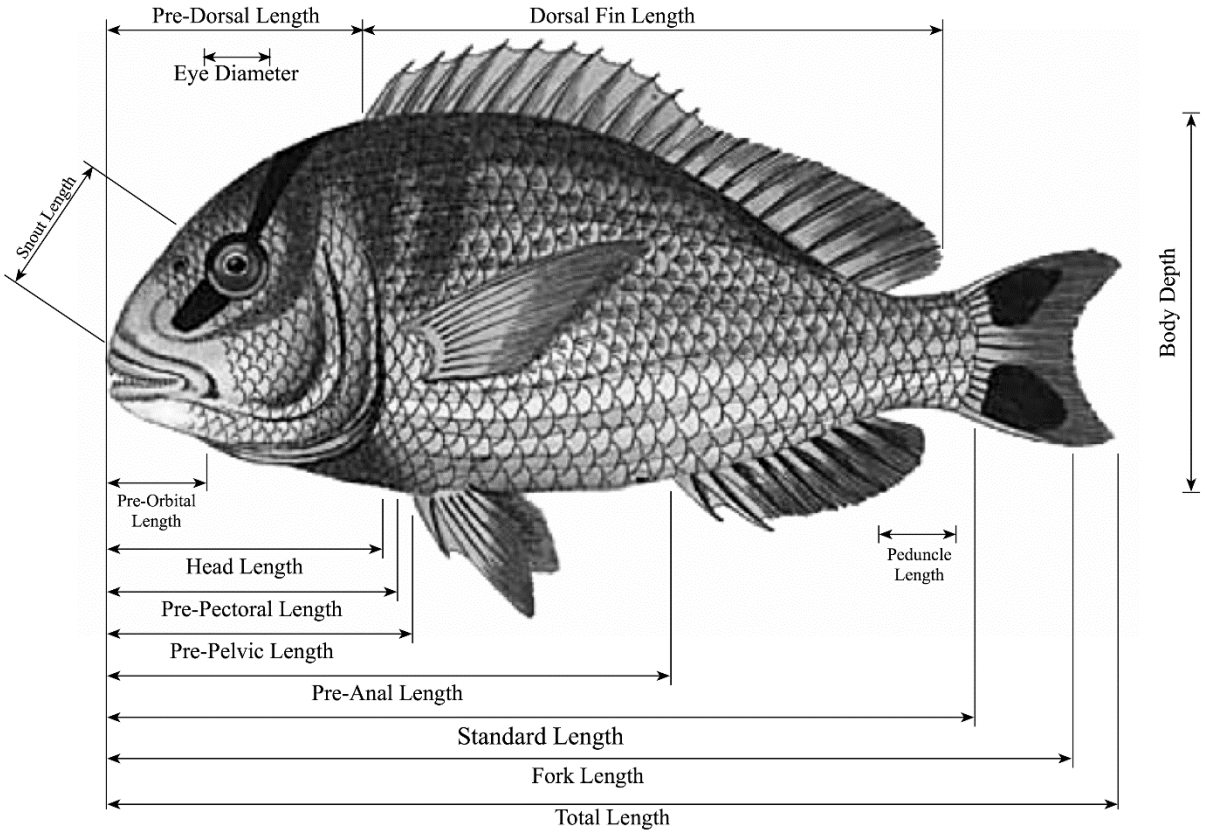

Figure 4. A selection of standard linear distance measurements used to quantify form variation in fish. Note the level of redundancy and uneven coverage of the form inherent in these measurements as well as the fact that the endpoints (= landmarks) that define these distances are not always located at topologically corresponding locations on biologically homologous structures. Fish image courtesy of The Natural History Museum (London, image no. BMNH 413511).

2.2 Boundary outlines

During the same historical interval in which these developments were occurring a similar sense of ferment was evident among those who focused on the analysis of object and organismal outlines. For reasons that will become clear later presently, outline-based morphometric studies have been largely ignored in many previous reviews of contemporary morphometrics (see Adams *et al*., 2004; Mitteröcker & Gunz, 2009; Adams *et al*., 2013). Nevertheless, the analysis of outlines is an important aspect of morphometrics, both in terms of its history and of its practice in contemporary research programmes.

The morphometric analysis of object outlines began as an attempt to use the well-established mathematical method of Fourier analysis to describe alpha-numeric characters for the purpose of automated character identification (Cosgriff, 1960; Fritzsche, 1961). In the field of natural history this approach was first applied to sedimentary grains (Ehrlich & Weinberg, 1970) and then to the shells of fossil ostracodes (Kaesler & Waters, 1972), pollen grains (Christopher & Waters, 1974), and bryozoan apertures (Cheetham & Lorenz, 1976). Each of these early implementations of Fourier analysis were based on sets of equiangular radius vectors that were used to sample the outline curve at regular intervals such that the resultant vector sequence defined a periodic function. Fourier analysis was then used to decompose this periodic function into sets of

form/shape descriptors: harmonic amplitudes and phase angles. This procedure imposes a set of limitations on the range of outline-based morphological problems that can be addressed in that it (1) can only be used to analyze closed, single-valued outline curves, (2) is dependent on accurate and stable location of the curve's centroid, and (3) assumes that all curve segments exist at more-or-less at the same distance from the curve's centroid lest gross differences in spatial resolution characterize different regions of the outlines under consideration.

In order to overcome these limitations Raudseps (1965), Zahn and Roskies (1972), and Bookstein (1978) experimented with various tangent-angle functions that could represent any boundary curve accurately, irrespective of its form and complexity, as a periodic mathematical function composed of a series of equal-length steps around the outline without requiring the location of a center. While Raudseps (1965) and Zahn and Roskies (1972) originally intended their functions to be used in the context of a Fourier decomposition, Bookstein (1978) recognized that this approach to outline-curve characterization was sufficient in and of itself to serve as the basis for subsequent analysis using a variety of different procedures; an approach that culminated in the development of Lohmann's eigenshape analysis (1983, see also Lohmann & Schweitzer, 1990). Around this same time Kuhl and Giardinia (1982) developed elliptical Fourier analysis in another attempt to extend the range of the Fourier approach as a tool for generalized form characterization and analysis.

Bookstein *et al*. (1982) criticized the use of radial Fourier analysis in biological morphometric contexts on the basis of the inability of this (and all) Fourier approaches to match landmark locations across forms using information derived from a consideration of biological homology. In addition, these authors noted that most practitioners of radial Fourier analysis only included part of the information provided by a Fourier decomposition (the harmonic amplitude coefficients) in subsequent data-analysis procedures (e.g., cluster analysis, PCA). A group of prominent practitioners of radial Fourier analysis in biological morphometrics responded to this challenge by noting that biological homology was not the only matching criterion that could be used to compare biological forms and/or test biological hypotheses (Ehrlich *et al*., 1983). In addition, they produced an example analysis in which the phase angle information did not appear to be necessary in order to achieve a reasonable representation of form similarities and differences within a sample of foraminifer test (= shell) outlines. Later, Kaesler (1997) accepted the principle that the morphometric use of radial Fourier analysis could be defended on purely utilitarian grounds, but rejected the idea that the Ehrlich *et al*. (1983) demonstration was sufficient to justify discarding phase-angle information in all cases.

Rohlf (1986) compared Fourier analysis (radial and elliptical), eigenshape analysis, and the direct analysis of outline coordinates, and concluded that, since all represent "orthogonal rotations of the same points" (p. 845), the results of such operations are mathematically equivalent provided all information produced by their intermediate steps are included in the computations (e.g., harmonic amplitudes and phase angles in the case of Fourier analysis, eigenvectors in the context of eigenshape and the direct analysis of outline coordinates). This conclusion is somewhat misleading insofar as Fourier analysis is commonly understood to involve only the redescription of a set of boundary outline points in terms of a set of harmonic equations involving amplitude and phase angle coefficients, whereas eigenshape analysis is commonly regarded as involving the redescription of a set of boundary outline as a tangent-angle function *and* the decomposition of sets of such functions via either PCA or singular value decomposition (SVD). Accordingly, Fourier analysis *sensu stricto* (radial and/or elliptical) is equivalent only to the curve-quantification step of an eigenshape analysis whereas eigenshape analysis is equivalent to a PCA or SVD decomposition of the complete set of Fourier harmonic amplitudes and phase angles (see MacLeod, 2011a, b, 2012a). Regardless, very few practitioners of Fourier analysis retain all harmonic amplitude and phase angle coefficients in any subsequent analyses of their data (Bookstein *et al*., 1982) and it is commonplace for those using PCA or SVD to summarize form similarities and differences only to employ the first few eigenvectors — usually just the first two or three — as the basis for comparison. In a practical sense then, the ordinations produced by the PCA or SVD analysis of Fourier-based redescriptions of object outlines are often not equivalent to those produced by eigenshape analysis because the complete set of data resulting from intermediate curve-characterization steps is rarely retained. In addition, eigenshape analysis often has to deal with data matrices that contain a large number of variables; often much larger than the number of specimens under analysis. This complication can usually be dealt with through the use of SVD (rather than PCA) as the basis for eigenvector decomposition and has benefitted practically from the increase in computation speeds realized by improvements in computer design.

2.3 Landmarks

An alternative to morphometric form characterization via linear distances or boundary outlines was proposed by Gower (1975), Kendall (1977, 1981, 1984) and Siegel and Benson (1982) in the form of combining the Cartesian coordinate format of boundary outline studies with the flexibility of distance measurement by recording the Cartesian coordinates of the end-points of traditional linear dimensions. These sets of coordinates point locations could be standardized for (uninteresting)

differences in their position, separation and orientation using the mathematical tool of Procrustes analysis to produce comparable constellations or configurations of points in either two dimensional (2D) the three-dimensional (3D) space. Once transformed into this position/size/orientation normalized format, the coordinate values that result express differences in pure shape that could be summarized using various multivariate procedures (e.g., cluster analysis, PCA, see Benson, 1982; Benson *et al*., 1982; Chapman, 1990).

Bookstein (1986, 1991) showed how it is possible to separate size and shape variation in such "landmark" data and proposed a size index — centroid size — that remains the most commonly used metric of size variation for all landmark-based morphometric analyses. Bookstein (1986) also proposed an alternative procedure for standardizing configurations of landmark points for differences in size, position and orientation relative to a landmark-defined baseline, the free coordinates of which he termed "shape coordinates", as well as proposing statistical tests (in addition to those that already had been proposed by Kendall 1977, 1981, 1984) that were designed specifically to test for allometric patterns of shape variation. While Bookstein's landmark-alignment procedure has largely been superceded by the Procrustes shape coordinate alignment method favoured by Kendall (see also Bookstein, 1991), Bookstein shape coordinates are still used for various special-purpose comparisons (see Webster *et al*., 2001; Kim *et al*., 2002).

Bookstein (1991), as well as others (e.g., Zelditch *et al.*, 2004), have argued that this landmark-based approach to the quantification of form is superior to any other because it allows analysts to incorporate the principle of biological homology into their landmark data. In some instances, this assertion is undoubtedly true (e.g., triple junctions of the sutures between bone of the vertebrate cranium which specify a point location, intersections of veins in insect wings and of wing veins with the wing periphery). But generally speaking, the majority of landmarks used in the majority of landmark-based morphometric investigations are topological rather than biological homologues (see Rohlf & Bookstein, 1990; Marcus *et al*., 1993, 1996 and chapters therein) and so no different in principle from a sparse set of boundary outline coordinates in many instances. The variety of criteria — both biological and topological — that can be used to locate landmarks was reviewed, by Bookstein (1991) in his classification of landmark types. Landmarks are useful for location the positions of structures relative to other structures across a form, but less useful for quantifying the shapes of these structures (see Fig. 5) with attempts to do so often grading into the boundary outline approach to the quantification of morphology. These caveats aside, since coming to the attention of the community of morphologists between 1986 and 1991, most morphometric analyses have been conducted using landmarks as the sole or primary means of quantifying morphology and morphological variation (see Adams *et al*. 2004, 2013). At present the dominance of landmarks across the field of geometric morphometrics is so complete that many now regard their use (inappropriately in this author's view) as being synonymous with "morphometrics" in general and "geometric morphometrics" in particular.

2.4 Digital images

Aside from morphometricians' classical focus on the representation of specimens as configurations of linear distances, semilandmark-delimited boundary outlines, or landmarks, a new trend in morphometric data collection is now emerging that focuses on the direct analysis of digital images. Sirovich and Kirby (1987) were the first to employ PCA in the analysis of digital images, in their case for the investigation of form variation in the human face. In this, and subsequent applications (e.g. Turk & Pentland, 1991a, b; Belhumer *et al*., 1997), collections of digital images were standardized to occupy a consistent pixel frame and then (usually) down-sampled to reduce the redundancy of the pixels included in the dataset. Next, these image matrices were reformatted into column vectors, assembled into a data matrix, and analyzed using either PCA, principal coordinate analysis (PCOORD, see MacLeod, 2006) or linear discriminant methods. The original Sirovich and Kirby (1987) study employed PCA operating directly on a m x m covariance matrix of pixel brightness values (where m is equal the total number of pixels comprizing the image frame) to produce a series of "eigenpictures" that expressed the major orthogonal patterns of image-set variation in a strictly hierarchical manner. Turk and Pentland (1991a, b) showed how a method, mathematically equivalent to singular value decomposition (SVD), could be used to reduce the dimensionality of image-set characterization problem and, in so doing, generalize Sirovich and Kirby's approach. These authors distinguished their image-set decompositions from those of Sirovich and Kirby (1987) by referring to them as "eigenfaces".

Whereas both Sirovich and Kirby (1987) and Turk and Pentland (1991a, b) were interested primarily in using the eigenvector-based ordination spaces they created to serve as a basis for face recognition (e.g., either new images of the same individuals whose images were used in the original sample or to distinguish other types of sampled image groupings such as distinguishing males from females or individuals wearing or not wearing glasses), neither PCA nor SVD incorporates any information about subgroups that might exist within the sample into their calculations. As a result, the set of eigenvectors resulting from these operations will not necessarily be optimal for achieving group-level identifications unless between-group morphological distinctions are aligned with the major axes of morphological variation within a sample. In order to

address this issue Belhumer *et al*. (1997) employed R. A. Fisher's (1936) linear discriminant analysis to create an ordination space that optimized differences between groups of face images. In order to distinguish these group-optimized image-set decompositions from those of Sirovich and Kirby (1987) and Turk and Pentland (1991a, b), Belhumer *et al*. (1997) referred to them as "fisherfaces".

Tests of the eigenface and fisherface technique on a sets of images in which lighting and facial expression were varied in a systematic manner showed that latter approach was markedly less susceptible to errors caused by both sources of variation (Belhumer *et al*., 1997). MacLeod (2015a) has carried the approaches pioneered by these authors forward by combining the Turk and Pentland (1991a, b) approach as a technique to reduce the dimensionality of the image analysis/recognition problem, and augmenting those results with the multivariate extension of LDA (= canonical variates analysis, CVA) when the characterization of subgroups or automated identification of unknown images is necessary; a procedure he has referred to by the generic term: "eigenimage" analysis in order to emphasize its similarity to morphometric eigenshape analysis.

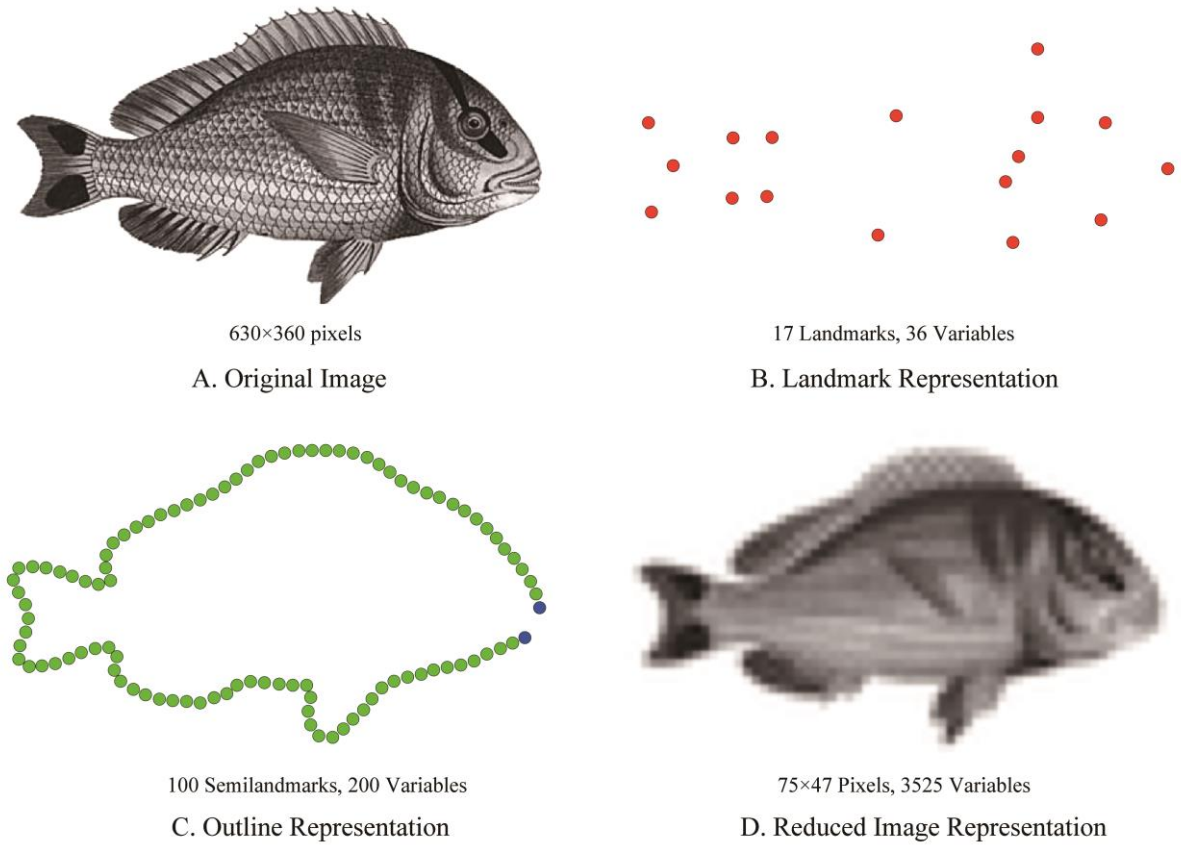

Figure 5. Alternative datasets that can represent the geometric aspects of an object or specimen's form. Landmarks shown in red. Boundary-outline curve termination landmarks shown in blue. Boundary outline semilandmarks shown in green. Note differences in the completeness, degree of abstraction, focus of representation, and sizes of these alternatives. Original image is a greyscale rendering of the porkfish (*Anisotremus virginicus)* published originally as a color figure in Shaw and Nodder (1799).

Eigenimage analysis (also eigenpictures, eigenfaces, fisherfaces) inverts the more familiar Cartesian approaches to morphometric characterization by regarding each pixel that comprizes an image frame as a topologically homologous landmark point (Fig. 5). These points may be assigned a variety of monochrome or color brightness values depending on the type of images selected for analysis. Whereas the configurations of landmark and boundary outline datasets are constructed via specification of the 2D or 3D coordinate values of the point locations of interest, the geometric configuration of an eigenimage dataset is simply the (reformatted) set of pixel values comprizing the digital image itself. While this configuration represents a very high-dimensional feature space, the use of SVD to decompose this feature space ensures the efficiency of the analysis which, in most instances, will have the same dimensionality as the number of specimens included in the sample. Perhaps more importantly, since each image pixel is assigned a brightness or color value irrespective of each

specimen's preservational state, incomplete specimens, or specimens that have exhibited minor to moderate amounts of damage, can be utilized in an eigenimage analysis. In addition, use of the entire image precludes the need to make a priori assumptions or decisions regarding which aspects of the specimens under consideration might be important for characterizing a sample or discriminating between groups residing therein. As a consequence, eigenimage analyses can often be used more efficiently and more effectively than landmark or boundary outline analyses in exploratory, as well as confirmatory, investigations; especially those whose primary purpose is to achieve group-level discrimination.

Concerns do arise with regard to ensuring that the images in eigenimage analyses are rendered geometrically comparable in terms of the sizes and orientations within the image frame and in terms of the elimination/standardization of sources of variation extraneous to the purpose(s) of the investigation. However, there a wide variety of image-standardization and image-processing routines that can be employed to ensure this comparability, just as there are for landmark and boundary outline data. In these senses then there is little or no practical distinction between the boundary outline, landmark and image pixel representations of the geometric configurations of specimens of interest. All three types of data can be used to express, summarize, analyze and represent geometric patterns of variation among and between specimens comprizing a sample using the same set of geometric morphometric data-analysis techniques.

3 The geometric morphometric synthesis

Geometric morphometrics is a generic term usually applied to the set of procedures that grew out of various researchers' attempts to synthesize the diverse styles of morphometric analysis into a single, unified, and conceptually coherent approach to the analysis of form. This synthesis is ongoing as new procedures are identified and aligned with geometric morphometric formalisms (see below). But the major work that forged the basic synthesis was published between 1981 and 1991 with reviews appearing at regular intervals thereafter.

Different morphometricians have described the conceptual breakthrough that created geometric morphometrics in different ways (see Bookstein, 1993; Reyment, 2010). Most geometric morphometricians would agree that the fundamental insight that made this synthesis possible was contributed by David Kendall (1981, 1984, see also Kendall *et al*., 1999) who published a the first comprehensive and fully generalized description of the geometric shape space. The revolutionary character of Kendall's contribution can be illustrated by way of an example.

Figure 6 shows three groups of triangles with identical height and width dimensions. All three groups plot at identical locations within the geometric spaced defined by the height and width variables and would plot at identical locations in any orthogonal linear space formed out of these variables (e.g., a PCA ordination space). Nevertheless, all three groups exhibit distinctly different forms. Thus, when sets of objects are characterized using simple scalar distances between corresponding points the ambiguity inherent in the system of measurements being used to describe their form renders quantitative comparisons between forms ambiguous. In this case, the height and width variables define the dimensions of the rectilinear box within which each triangle would fit. But the same box can, in principle, contain triangles of many different shapes. At best, such data define families of possible forms. As this example makes plain, such sets of variables a are unable to express true form differences completely, even in simple cases.

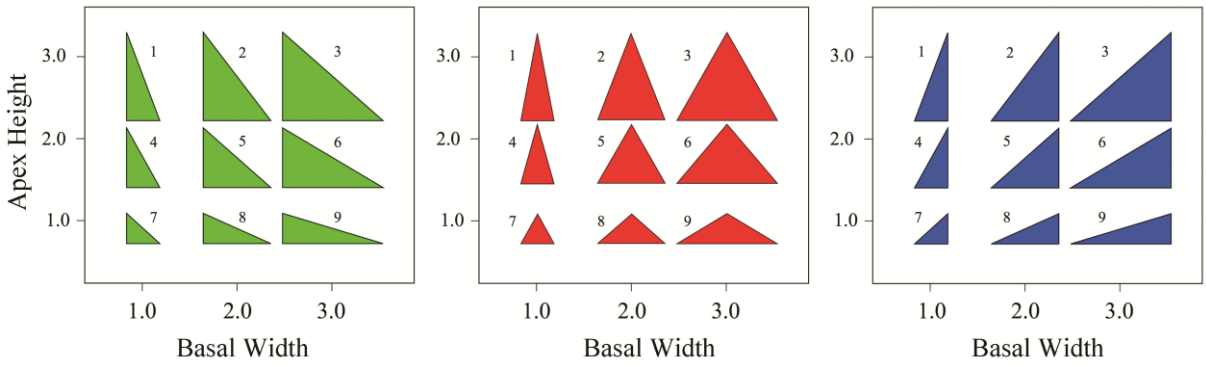

Figure 6. Three sets of triangles with identical apex height and basal width dimensions.

Figure 7 shows the result of a Procrustes alignment of vertex coordinates for the groups of triangles shown in Figure 6. This alignment expresses the least squares deviation of each triangle from the mean triangle which, for this dataset, happens

to be an equilateral triangle. Unlike the vertex height and basal width variables that were used to place the triangles in the width/height space (Fig. 6), each triangle now appears to occupy a unique set of coordinate positions or configurations in the space of these aligned, Procrustes (shape) coordinate locations.

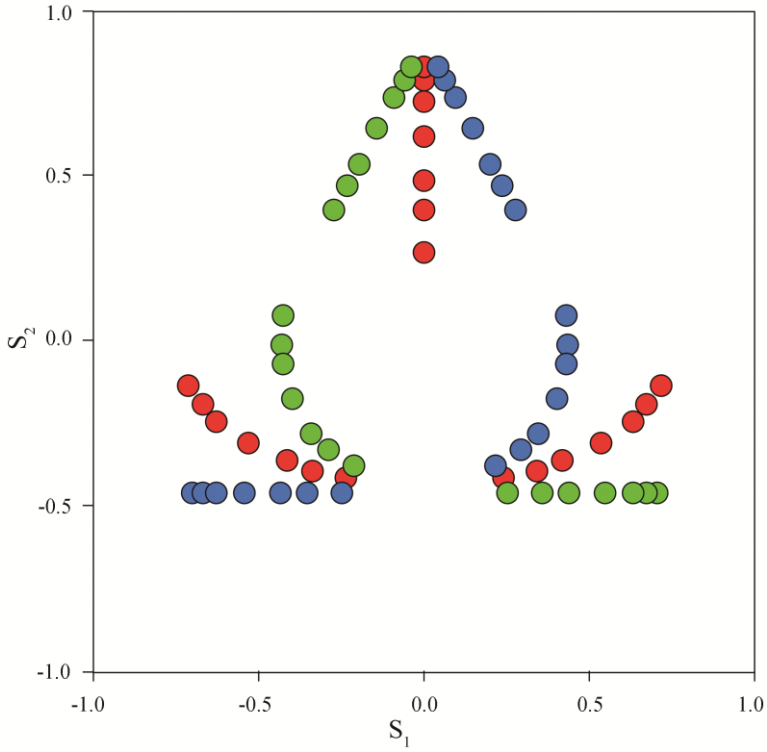

Figure 7. Procrustes superposition of the 27 sets of vertices for the triangles shown in Figure 6.

Figure 8 shows results of a covariance-based PCA of these Procrustes aligned shape-coordinate data both in terms of the first two (left) and a complete set of three (right) principal components.[1] Each point in this space represents a unique configuration of the six coordinate values that, together, comprise any triangle. But unlike the width/height space (or a PCA-determined representation thereof, see Fig. 6), all but three triangles in each group exhibit unique coordinate locations in the

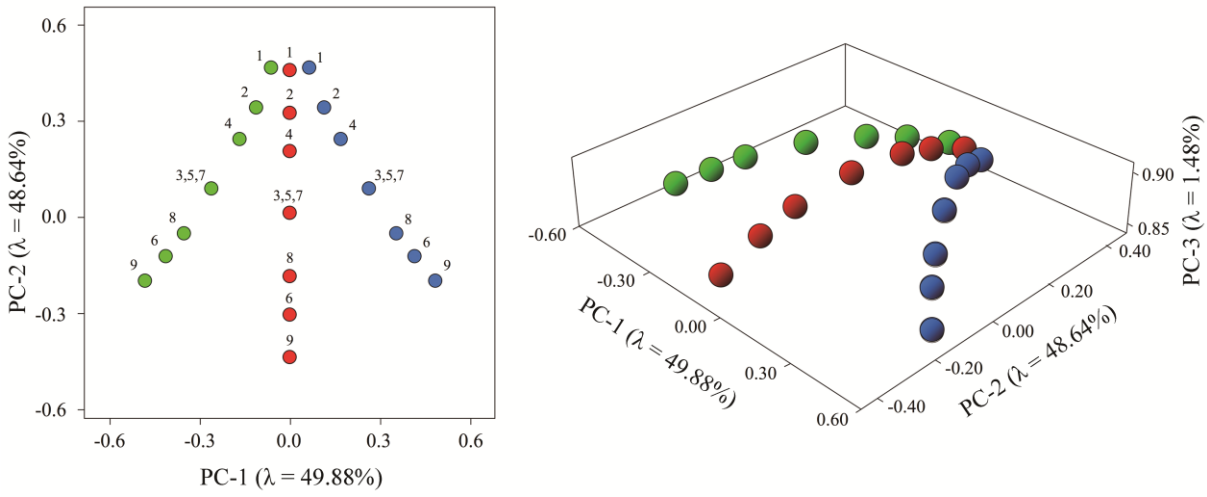

Figure 8. Principal components ordinations of the Procrustes-aligned triangle configurations shown in Figure 6.

[1] Only three principal components have positive eigenvalues owing to the reduction in variation caused by the Procrustes alignment of size, position and 2D rotation.

PCA-decomposed Procrustes shape-coordinate space. These unique locations reflect and summarize their unique shapes.

Note that triangles 3, 5 and 7 for each group shown in Figure 6 lie on the 45° diagonal in the original height/width space. As the scales for all axes in these plots are identical, the rate of change in the height and width variables is identical along this diagonal trend. Accordingly, all triangles located along these diagonals exhibit variation in size, but no variation in shape (= isometric growth). The positions of these three sets of isometric triangles in the PCA space shown in Figure 7 reflects this shape similarity exactly since differences in size among these triangles have been discarded during the Procrustes alignment procedure. Note also that the equilateral triangles in this series (triangles 3, 5 and 7 in the red group of Fig. 5) plot at the origin of the PC-1 vs. PC-2 coordinate system (Fig. 8, left) and at the unit coordinate along PC-3. This is the pole position of a unit hemisphere, the surface of which contains all the other triangles — and all conceivable triangles — arrayed in a systematic manner that reflects their intrinsic and geometrically organized differences accurately. All triangular configurations of landmark points are located on the surface of this hemisphere because differences in their sizes have been discarded during the Procrustes alignment procedure.

This simple mathematical experiment illustrates Kendall's (1981, 1984) conclusion that all polygonal shapes exist on the surfaces of hyperdimensional mathematical manifolds of dimension 2*K*-3 (where *K* is the number of polygon vertices) and in which the radius of the manifold reflects the size of the vertex-defined polygon. In the case of triangles (which are the simplest geometric objects that have a shape) this manifold is a sphere. As can be seen in Figure 9 (left), all conceivable triangles (incl. co-linear triangles) whose baselines have been aligned and whose sizes have been adjusted to a unit value can be located on the surface of the three-dimensional triangle shape sphere of constant radius. The two hemispheres of this sphere separated by the "equator" of co-linear triangles contain sets of triangles that are mirror images of each other with the apex vertex reflected across the basal chord. During Procrustes alignment all three corresponding vertices are rotated to corresponding positions, which reduces the full triangle shape sphere to a Procrustes aligned shape hemisphere. Confirmation of this theoretical proposition can be obtained by aligning *n* random triangles using the Procrustes procedure and plotting these in the space of the first three covariance-based principal components (Fig. 9, right).

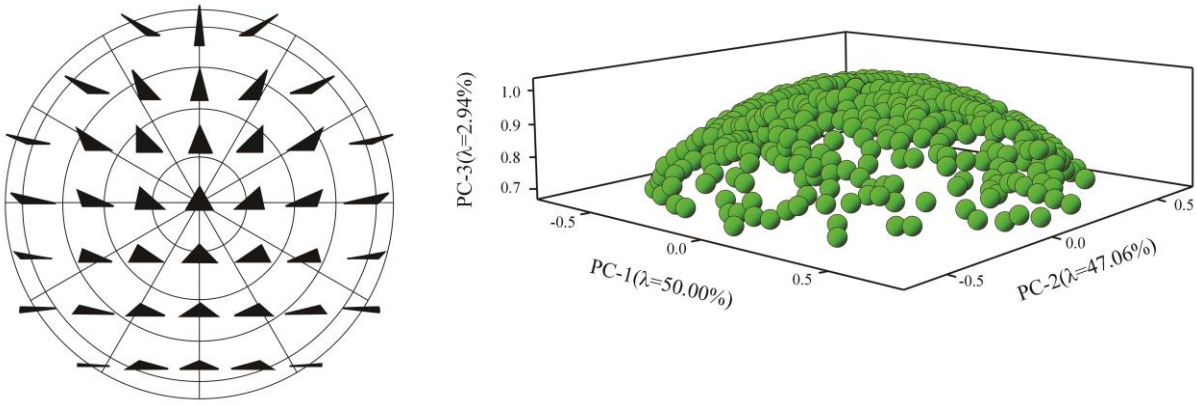

Figure 9. Left. Representation of the surface of the Kendall shape space for triangles with the pole position occupied by an equilateral triangle. Each point on the surface of the sphere represents a unique configuration of 3 2D coordinate points as illustrated by the triangle diagrams. Note isosceles triangles occupy a meridian, right triangles that verge left or right occupy two meridians, and co-linear triangles occupy the equator. Triangles in the lower hemisphere of this shape space are reflections across the baseline of triangles in the upper hemisphere. These two hemispheres collapse under Procrustes alignment as the lower hemisphere triangles are rotated to match the shapes of the upper hemisphere triangles. Right. Principal components ordination of 500 Procrustes-aligned random triangle configurations.

Kendall's (1981, 1984) discovery of shape manifolds opened the door to an appreciation of the true geometric character of shape and its relation to size. In addition, it clarified the way forward for devising statistical summaries of shape variation, characterizations of shape change, and tests of shape similarity. Since linear distance data and boundary outline data can be expressed in terms of configurations of Cartesian point coordinates, Kendall's insight effectively unified the three traditional approaches to morphometric analysis by demonstrating this unity.

Statistical tests for differences between shapes consistent with Kendall's concept of the shape manifold were developed by Goodall (1983, 1991), Bookstein (1984, 1986, 1991), Mardia and Dryden (1989, 1998), Goodall and Mardia (1991), Dryden and Mardia (1998) and others. In addition, Bookstein (1989) introduced a graphical interpolation procedure — the thin plate spline — that realized the long sought-after goal of providing a mathematical language within which Albrecht

Düer's and D'Arcy Thompson's shape deformation grids could be transformed from illustrative devices to a geometric data-analysis tool.

The physical metaphor of the thin plate spline (TPS) is that of a uniform sheet of metal in which the amount of resistance to bending is inversely proportional to the spatial separation of points over which the bending force is applied. In other words, the hypothetical thin (metal) plate that represents the pattern of shape deformation requires little "force" to be deformed into a broad, generalized bend at points that lie at great distances from one another, but is far more resistant to more localized deformations among closely spaced points in a manner that scales uniformly across the entire plate.

This procedure begins with the specification of a composite matrix (L) which quantifies the landmark coordinates of a reference shape (usually the mean shape of a sample) and a complete set of inter-landmark distances for these data. Taking the inverse of the L matrix effectively establishes a series of inter-landmark weight coefficients that quantify this spline's metaphorical pattern of resistance to bending across the reference landmark configuration. Post-multiplication of this inverse matrix by a target configuration (e.g., the landmark configuration of another specimen) can be used to interpolate the geometric pattern of displacements required to deform the reference shape into the target along with the total amount of "force" (= bending energy) required to do so. Figure 10 presents a graphic illustration of the outcome of this procedure for s set of landmarks collected from two fish illustrations that appeared originally in d'Arcy Thompson's 1917 treatise *On Growth and Form* and have been a staple of biological and morphometric textbooks ever since (see Bookstein, 1991; Zelditch *et al*., 2004 and MacLeod, 2010a for detailed presentations of the mathematics involved in calculating a thin plate spline).

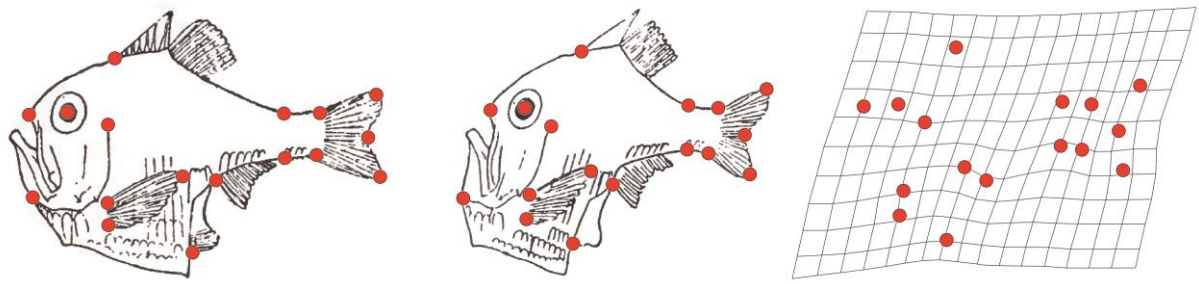

Figure 10. Landmarks and the thin plate spline resulting from a comparison of D'Arcy Thompson's (1917) drawings of *Argyroplecus* (left) and *Sternoptyx* (right) after GLS Procrustes alignment of their landmark configurations. In this calculation *Argyroplecus* was designated as the reference configuration and *Sternoptyx* the target configuration. It is instructive to compare the calculated thin plate spline representation of this deformation with Thompson's conceptual transformation grid for these same images (see Fig. 4). The total bending energy (= anisotropic deformation) is calculated as the sum of bending energies specified by the reference-to-target deformation at each landmark location and represents a useful means of quantifying shape variation within the sample. For this comparison the total bending energy is 1.25.

Note that, unlike D'Arcy Thompson's (1917, 1945) own graphical analysis of the morphological transition between *Argyroplecus* and *Sternoptyx* (see Fig. 3), a thin plate spline calculated on the basis of 17 corresponding landmark locations and aligned using the generalized least squares (GLS) Procrustes procedure incorporates both Thompson's isotropic shear deformation as well as a substantial component of anisotropic deformation. The latter effectively expresses deviations between three more-or-less well demarcated clouds of landmark points: anterior-dorsal (maxilla, eye, dorsal fin based and dorsal operculum terminus), posterior (peduncle and tail) and anterior-ventral (anal, pelvic and pectoral fin bases, operculum ventral terminus, pectoral fin tip).

The anisotropic (also non-uniform and/or non-affine) component of shape variation is calculated by multiplying the matrix of Procrustes-aligned deformations from reference to the target shapes by the inverse of the L matrix (L^{-1}, also termed the bending energy matrix). This operation weights each spatial deformation over the entire set of contrasts between landmarks by the inverse of the spatial scale over which the deformation occurs. Various approaches to calculation of the uniform (= isotropic) component of shape variation have been offered (Bookstein, 1991, 1996a; Rohlf, 1993; Rohlf & Bookstein, 2003). Figure 11 illustrates thin plate spline interpolations for the separate isotropic and anisotropic components of shape deformation for the *Argyroplecus* → *Sternoptyx* comparison.

While the thin plate spline method of shape deformation has many intriguing properties, it should be used with caution in developing interpretations of shape variation across a sample. This method of expressing shape variation derives from the spatial analysis technique called kriging which was developed to estimate the likely values of variables (e.g., ore concentrations) located at a distance from known samples (Krige, 1951; Cressie, 1988, 1990; Davis, 2002). Like many

multivariate procedures, kriging is based on linear regression analysis. But unlike standard linear regression analysis, kriging does not assume the dependent variable is either completely random or distributed deterministically with respect to independent spatial variables. Rather, it assumes the dependent variable is regionalized (Matheron, 1971).

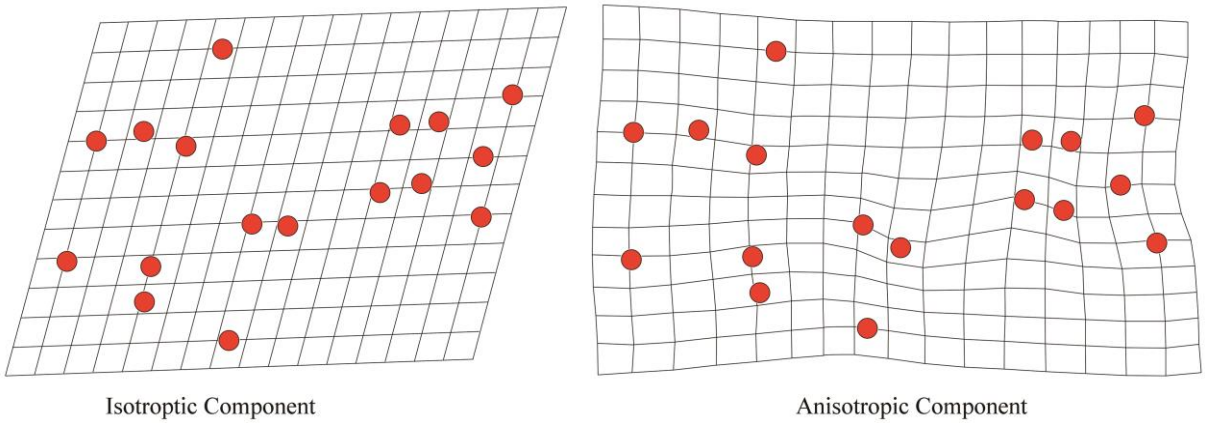

Figure 11. Thin plate spline representations of the isotropic (= uniform) and anisotropic (= non-uniform) components of shape variation implicit in the comparison between D'Arcy Thompson's (1917) drawings of *Argyroplecus* and *Sternoptyx* after GLS Procrustes alignment the landmark configurations. For this comparison the non-uniform component was calculated using the complement procedure described by Rohlf and Bookstein (2003). By convention there is no bending energy associated with the uniform component of shape deformation. D'Arcy Thompson identified the dominant uniform dorsal-ventral shear component of this shape transition correctly, but did not identify (or discuss) the non-uniform component, which he may have regarded as unstructured or random variation.

The idea of regionalized variables is a fundamental concept in spatial statistics of which morphometrics may, in certain senses, be regarded as part (Davis, 2002). Regionalized variables are variables that exhibit properties intermediate between those of a random variable and a deterministic variable. Such variables are continuous from point-to-point throughout the space over which they are defined and can exhibit high correlations (= structure) over short distances. However, the apparent consistency in the structure of their variation is inversely related to the distance between sampled locations. As a result, it is usually not possible to determine the rules by which variation is governed across the entire space.

In geological or geographical mapping situations regionalized variables are estimated (ideally) on the basis of a reasonable sample of patterns of variation in independent variables as assessed by observations made at specific locations. Regression analysis-like strategies are then used to estimate localized substructures in the dependent variable, and then to join these substructures into a single, continuous, global model of dependent variable variation (Stein, 1999). The typical morphometric situation, however, often falls far short of this ideal, especially insofar as the number and position of unambiguously corresponding and relocatable landmarks is usually not under the researcher's control.

In the case of the *Argyroplecus* → *Sternoptyx* shape, comparison the 17 landmarks used to quantify shape variation are located, for the most part, along the specimen's outline. Accordingly, thin plate spline-interpolated deformations specified for the internal head and trunk areas cannot be assumed to be accurate despite the fact that some deformation maxima may be inferred to occur in these areas. Moreover, the "bending energy" conceptually associated with these thin, uniform, metal plates represents a mathematically elegant, but biologically questionable proxy for the heterogeneous composition, structure, and complex systemic organization of organic bodies. In considering this issue from the vantage point of contemporary evolutionary biology it is ironic to note that, in his enthusiasm for exploring and explaining his approach to the analysis form variation in species as a geometric deformation, D'Arcy Thompson — as well as many of his contemporary disciples — may have overlooked the implications of Thompson's other great insight into the nature of form variation; that the physical properties of the materials out of which organismal bodies are constructed impose their own constraints on the structure and patterns of variation that can be realized by those bodies (Thompson, 1917, 1945). Since organismal bodies are not homogeneous in their composition it is doubtful that homogenous, system-wide deformations of the sort specified originally by D'Arcy Thompson and carried forward by the thin plate spline metaphor/technique could be anything other than statistical interpolations of highly localized patterns of landmark displacement. If the reference shape upon which the magnitude of these localized landmark deformations are calculated is changed, or if a different set of landmarks is selected for analysis, the thin plate spline representation of shape deformation can (and usually does) change radically.

Nevertheless, the theoretically satisfying separation of isotropic from anisotropic patterns of deformational variation

does not exhaust the range of operational possibilities inherent in the thin plate spline approach to describing the character of organismal variation in geometric morphometrics. Bookstein (1986, see also 1991 and Rohlf 1993) recognized originally that, owing to the fact that the bending energy matrix is square, it was possible to decompose it into orthogonal modes of variation using eigenanalysis. In this case the first few bending energy eigenvectors will be aligned with maximum bending energy deformations which (counterintuitively for those used to interpreting eigenanalysis in the context of PCA) will usually correspond to small displacements between landmarks located close to one another. Similarly, the last few bending energy eigenvectors will be aligned with minim bending energy which will, in most cases, will correspond to relatively large displacements between widely separated landmarks. In other words, these mathematically calculated aspects of shape variation — which Bookstein (1989) termed 'principal warps' — can be used to describe the anisotropic aspect of shape deformation recorded by the contrast of any target configuration of landmarks relative to a reference configuration in terms of a hierarchical series progressively more generalized orthogonal components which take the form of bends or warps when expressed in the coordinate system of the thin plate spline (Fig. 12).

For the 17 landmarks shown in Fig. 10, and taking *Argyroplecus* as the reference shape, the first principal warp identifies variation in a single peduncle landmark as the most localized (= highest bending energy) mode of shape variation. This can be contrasted with one of the most general principal warps (PW-13) which specifies a general, anterio-posterior bend in the thin plate spline organized around a lateral compression of the posterior landmarks accompanied by a localised dorso-ventral shear displacement in this region of the form. An intermediate principal warp, such as PW-8, specifies an intermediate mode of deformation, in this case one focusing on the expansion and asymmetrical deformation of the head and anterior trunk landmarks that appears to verge in a strongly dorso-posterior direction in the region marked by the operculum's dorsal terminus.

Bookstein (1991) hoped his principal warps, along with the projections of the data for an entire sample onto these mathematical axes (including the *z* axis for 3D data), which he termed "partial warps", would find use as signposts for the identification of morphometrically defined biological characters, the qualitative specification of which has long been used by taxonomists, ecologists, functional morphologists and phylogeneticists for placing organisms into similarity groups. Several researchers went beyond Bookstein's original, indicative interpretation of partial warp results and advocated their use as actual taxonomic, ecological, functional and phylogenetic characters (e.g., Zelditch *et al*., 1995; Burke *et al*., 1996; Lynch *et al*., 1996). However, as Rohlf (1998) noted in his critical review of this interpretation of partial warps results, partial warps are poor choices for taxonomic characters owing to their inherent dependence on the choice of the reference landmark configuration as well as the absence of any information about patterns of shape variation characteristic of samples (not to mention populations, see also Adams & Rosenberg, 1998). This concern echoed Bookstein's (1996b) advice that " … Any finding that requires the use of partial warps is erroneous," (p. 146) as well as violating the principal that geometric shape spaces are of sufficient complexity that there are, essentially, an infinite series of possible axis rotations and specimen ordinations (Bookstein, 1994).

Resolution of this controversy was found in the realization that principal/partial warps, like Fourier amplitude and phase angle coefficients, are simply redescriptions of the original landmark data that have been optimised in terms of their ordering to express certain features of those data, but that, when considered in toto, include all the information that was present in the original set of shape coordinates. These redescriptions have limited utility in and of themselves. But when taken as a whole, and subjected to additional analyses designed to extract sample-based signals, patterns and trends, are unquestionably useful in taxonomic, ecological, functional, phylogenetic and other contexts.[2]

The PCA ordinations of traditional morphometrics can be related geometrically to Kendall's concept of the Procrustes shape space in a straightforward manner, as projections of the locations of individual shape configurations existing on the complex, curved surfaces of Procrustes shape manifolds into a linear spaces of lower dimensionality; rather like the projections of globe locations onto a flat map (Fig. 13). To achieve this projection, the Procrustes shape coordinates, expressed as a set of deviations from a specified reference shape, may be submitted to a covariance-based PCA or SVD directly or in the form of the complete set of their partial warp scores on the complete set of principal warps. After rearrangement of these data such that each configuration is represented as a row vector of concatenated *xy*, or *xyz* values, a linear decomposition of that data matrix's covariance matrix will result in the specification of *m* positive eigenvalues where $m = 2K\text{-}4$ for two-dimensional data and $m = 3K\text{-}7$ for three-dimensional data.[3] In the simplest case — two-dimensional data for the Procrustes-aligned configurations of three-landmarks — the manifold surface has a dimensionality of 2 even though

[2] See MacLeod & Forey, 2002, Catalano *et al*., 2010, Klinkenberg & Gidaszewski, 2010, Goloboff & Catalano, 2011 for alternative explorations of the use of morphometrics in taxonomic and phylogenetic contexts.

[3] As above, this reduction in dimensionality occurs because of the standardizations for position, size and rotation that have been applied to the raw coordinate data to transform them into the Procrustes shape space (see Bookstein, 1991; Zelditch *et al*., 2004).

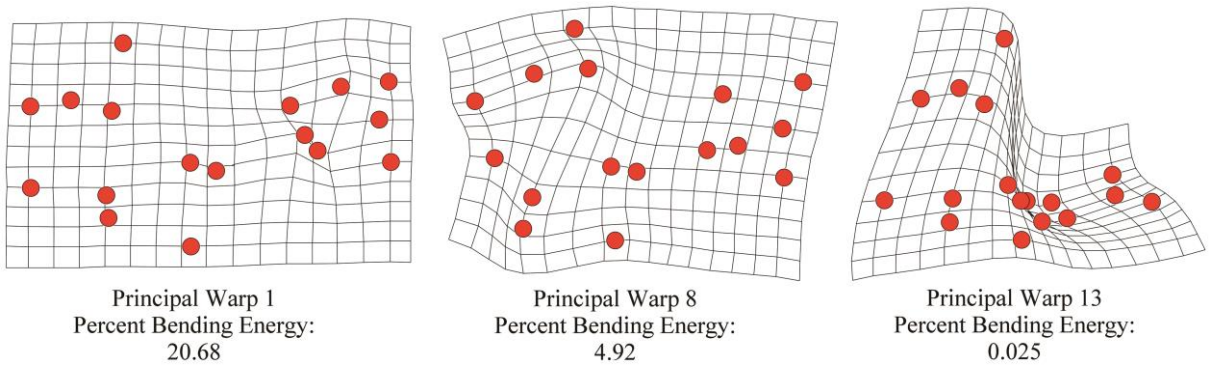

Figure 12. Thin plate spline interpolations of selected principle warps for *Argyroplecus* illustrating the increasing spatial generality, and decreasing bending energy, associated with these orthogonal components of non-uniform deformation. See text for discussion.

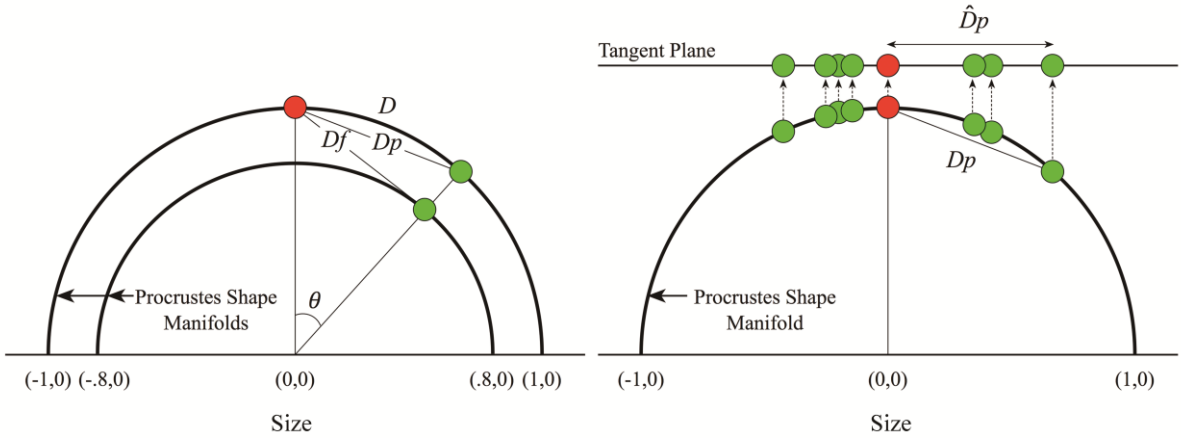

Figure 13. Cross sections through the Procrustes shape space illustrating various spatial relations between the reference (red) and target (green) configurations. All Procrustes-aligned landmark configurations exists as point locations on the surfaces of hyperdimensional manifolds the radius of which is determined by the configurations' size. Left. Between any two configurations the Procrustes Distance (*D*) is the length of the shortest arc between the configurations points on the shape manifold's surface. The Partial Procrustes Distance (*Dp*) is the linear distance between these points. For point comparisons for which the arc angle (θ) is small *Dp* will provide a good estimate of *D*. Interestingly, the closer match between the target and reference configurations points can be found by relaxing the constraint of size similarity. The Full Procrustes Distance (*Df*) is the minimal distance between the reference and the target configurations when the size of the target is allowed to vary, but not its shape. Right. The distribution of a shape or series of shapes with respect to a reference shape may be portrayed in the linear spaces that have become traditional in pre-geometric morphometrics by orthogonally projecting their positions on the Procrustes shape manifold to a linear plane. This operation can be accomplished via covariance-based PCA of the complete set of partial warp scores (= relative warps analysis) or direct covariance-based PCA of the Procrustes shape coordinates expressed as deviations from the reference shape. Under this convention the projection plane will be tangent to the shape manifold at the point of the reference shape and exhibit a reduced dimensionality from that of the shape manifold's surface. If the total range of shape variation present within the sample is small the apparent Procrustes distances ($\hat{D}p$) between point locations in this tangent space will be good estimates of both the Partial Procrustes Distance (*Dp*) and (by extension) the Procrustes Distance (*D*). Redrawn from an original diagram by F. J. Rohlf.

the curved surface of this manifold exists in a three-dimensional volume. Provided the range of shapes included in a sample of triangles is not too great, their pattern of geometric similarity on the shape manifold's surface can be represented accurately on a two-dimensional plane. Bookstein (1991) called this procedure relative warps analysis, though it has also been called Procrustes PCA (MacLeod, 2010b).

Orientation of the projection plane with respect to the shape space manifold is established by the reference shape configuration insofar as all shape coordinates are expressed as deviations from this reference shape. Since, in most cases, morphometricians are interested in analyzing collections or samples of forms, the obvious choice for this reference is the mean landmark configuration for the sample under consideration. Selection of this mean as a reference minimizes the Procrustes distance of each configuration from the reference, and so minimizes the distortion inherent in projection of configurations from the surface of the shape manifold to the linear tangent plane (Fig. 13). Some have advocated selection

of particular specimens as the reference for special purposes (e.g., one end of a developmental sequence of shapes, see Fink & Zelditch, 1995; Marcus *et al*., 2000; Zelditch *et al*., 2004). However, selection of any reference other than the sample mean will inevitably lead to an increase in the distortion of the projected ordinations and should be treated with caution (Rohlf, 1998).

4 Example analyses

As an example of the range of results that can be obtained from the application of geometric morphometric data-analysis procedures to different types of geometric morphometric observations a small set of images of common marine fish species was assembled and used for the extraction of landmark, outline and image data (Plate 1). As shown in Figure 5, a total of 17 landmark points were used to quantify the positions of the distal tip of the maxilla (1), anterior intersection between the dorsal fin spines and the trunk (2), dorsal and ventral extent of the peduncle (3, 4, 8, 9), dorsal and ventral tips of the caudal fin (5, 7), the caudal fin notch/medial point (8), anterior intersections of the anal and pelvic fins with the trunk (10, 11), ventral and dorsal termini of the operculum (12, 14), position of the eye (13), dorsal and ventral intersections of the pectoral fin and the trunk (15, 16) and pectoral fin tip (17, Fig. 5B). These are all standard landmark positions that have been used to quantify fish body form in previous geometric morphometric analyses (e.g., Zelditch *et al*., 2004). The outline dataset consisted of 100 semilandmark points spaced equally around the silhouette of the body (including the dorsal, caudal, anal and pelvic fins) starting at the anterior tip of the maxilla and terminating at the anterior tip of the mandible (Fig. 5C). This sequence of 100 points is sufficient to capture the length of each silhouette's perimeter to an accuracy of 95 percent or greater thus ensuring the accurate and complete representation of each fish's body-form outline. Finally, each digital image included in Plate 1 was processed to standardize the framesize, image size and image orientation in the frame, converted to a greyscale color palette and down-sampled to fit a 75 by 47 pixel matrix (Fig. 5D). The greyscale brightness values for each pixel in each image frame were then rearranged into a row vector and assembled into a standard data matrix for analysis.

The landmark and outline datasets were both subjected to identical Procrustes landmark procedures to correct for variation due to position, size and rotation (see Rohlf & Slice, 1990). The image-processing transformations specified for the image pixel dataset are computationally identical to a GLS Procrustes superposition and so may also be regarded as also occupying a Procrustes alignment space. All three datasets were subjected to identical SVD-based eigenanalysis in order to extract a set of orthogonal shape-configuration vectors that summarize the structure of patterns of shape similarity and differences as recorded by those data. These vectors were then used to construct a three-dimensional ordination space within which to represent the major aspects of this structure for the purposes of comparison and contrast.

The major dimensions of variation for the example fish dataset as assessed by the landmark data are illustrated in Figure 14. Together these axes account for 64.72 percent of the shape variation recorded by this dataset. The dominant shape-variation trend in these data involves the transition from narrow-bodied forms exhibiting rounded or truncate caudal fins (e.g., *J. carutta*, *E. merra*) to deep-bodied forms exhibiting deeply forked caudal fins (e.g., *T. autralis*, *G. erythrourus*). These modifications primarily involve the posterior, outboard migration of the caudal fin-tip landmarks (5, 7) anterior migration of the central caudal fin landmark (6), and dorso-ventral migration of the dorsal fin landmark (2) with increasing PC-1 score. Subsidiary displacements occur in the proximate peduncle fin (3, 9) and pectoral fin tip landmarks (17).

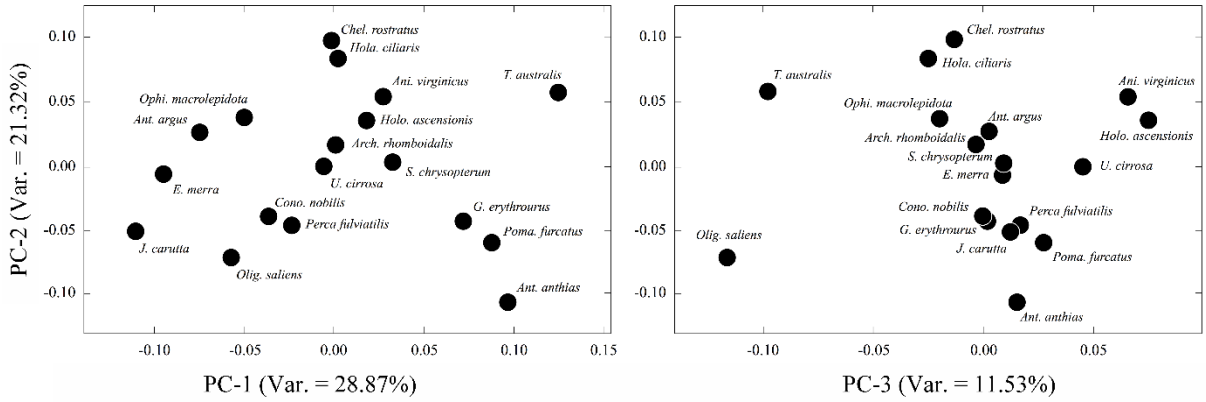

Figure 14. Major structure of shape variation for the fish sample as assessed the landmark data (Fig. 5B). See text for discussion.

The shape contrast captured by PC-2 involves the converse transition between narrow-bodied forms exhibiting deeply forked caudal fins (e.g., *Ant. anthias*, *Oligo. saliens*) and deep-bodied forms exhibiting rounded or truncate caudal fins (e.g., *Chel. rostratus*, *Ophi. macrolepida*). These modifications primarily involve displacements of the same landmarks as PC-1, but with an opposite polarity of movement in the caudal fin landmarks (5, 6, 7) with increasing PC-2 score. In addition, subsidiary (ventral) displacements occur in the anal and pelvic fin landmarks (10, 11).

The pre-penultimate orthogonal mode of shape variation (PC-3) records a twisting of the anterior region in which forms that project to low positions along this axis exhibit anterior eye placements with pectoral fins and dorsal fins that are set well back along the body (e.g., *Oligo. saliens*, *T. australis*) in contrast to forms whose eyes are placed in a more posterior positions along the head, as well as exhibiting longer pectoral fins, and a more anterior dorsal fin placement (e.g., *Hola. ascensionis*, *Ani. virginicus*). These transitions are focused on the anterior migration of the maxilla tip and dorsal fin landmarks (1, 2) with the eye landmark (2) itself remaining more-or-less stationary, accompanied by the posterior migration of the anal fin and pectoral fin-tip landmarks (10, 17) with increasing PC-3 score. Subsidiary variation also occurs with the posterior, inboard migration of the anal and pelvic fin landmarks (10, 11).

Examining the ordination spaces formed by these three axes a variety of shape groups are evident. The short-faced, deep-bodied, forked-caudal fin form of *T. australis*, and the strongly fusiform, narrow peduncle and forked tailed *Oligo. saliens*, are clearly identified as shape outliers for these landmark data. A subsidiary cluster of shapes formed by *G. erythororus*, *Poma. furcatus* and *Ant. anthias* all exhibit a characteristic combination of fusiform bodies and strongly forked caudal fins whose separation from the main group of forms is most evident in the PC-1 vs PC-2 subspace. By the same token, the deep-bodied and truncated caudal finned *Chel. rostratus* and *Holo. cilaris* also constitute a subgroup that occupies a position high on the PC-2 axis. With regard to the species that comprise the central cluster of shapes, two clear discontinuities exist which involve contrasts evident in the PC-1 vs PC-2 and PC-3 vs PC-2 subspaces. Based on these interpretations it would appear that, for this sample of forms as assessed by these landmark data, the distribution of shapes in both complex and discontinuous. But this is not the only — or perhaps even the most appropriate — summary of shape variation possible for this sample.

Figure 15 illustrates the major directions of shape variation for these same forms as assessed by the outline semilandmark dataset. Here the first three eigenvectors summarize 77.71 percent of the observed outline shape data, a substantial increase over the landmark dataset's result. For these data the major axis of shape variation is dominated by the distinctiveness of *Hola. ciliaris* — a form characterized by very long, trailing dorsal and anal fins — with respect to all other body morphologies included in the sample. Indeed, the outline shape of *Hola. ciliaris* is so distinctive with respect to the remainder of the sample that little more can be said with confidence about the major morphological trends in this sample from this plot. In order to obtain a clearer picture of shape-variation trends in the remainder of the sample the *Hola. ciliaris* data would need to be removed from consideration.

Along PC-2 this first subordinate orthogonal morphological trend can be seen to contrast broad, subquadrate body shapes with truncated caudal fins (e.g., *Chel. rosatrus*, *Holo. ascensionis*) with narrow, fusiform body shapes exhibiting deeply forked caudal fins (e.g., *Poma. furcatus*, *Anth. anthias*) as the projected score on this eigenvector axis increases. Subsidiary decreases in the relief of the dorsal, anal and pelvic fins with increasing PC-2 score also have roles in this shape-transformation mode. The PC-3 shape transformation axis contrasts elliptical body shapes with rounded caudal fins and comparatively low dorsal, anal and pelvic fins (e.g., *E. merra*, *Anth. argus*) with fusiform body shapes exhibiting deeply

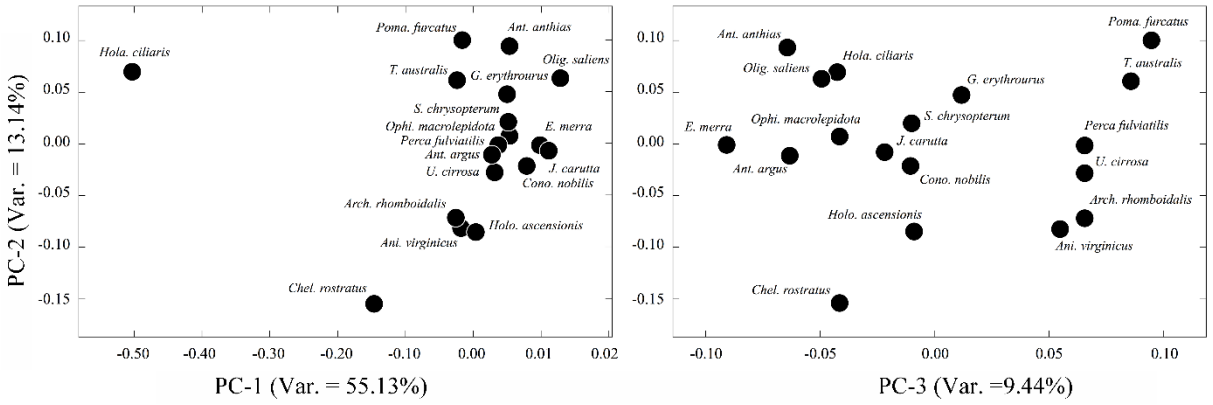

Figure 15. Major structure of shape variation for the fish sample as assessed the boundary outline semilandmark data (Fig. 5C). See text for discussion.

forked caudal fins and prominent dorsal, anal and pelvic fins (e.g., *T. australis, Poma. furcatus*).

Looking at the distribution of shapes in the PC-1/PC-2/PC-3 subspace of these outline data, the grouping structure appears much more clear-cut and distinctive than it was for the landmark data results. Both of the deep-bodied, subquadrate species (*Hola. ciliaris* and *Chel. rosatrus*) are identified as outline-shape outliers with the remaining taxa being placed into two broad outline shape groups based on the prominence of their dorsal, anal and pelvic fin shapes.

Obviously these outline data suggest a very different structure of shape similarities and differences in this fish dataset from those implied by the landmark data. This results primarily from the fact that, in addition to recording the gross aspects of the body, and fin location across these forms (all of which were quantified to a greater or lesser extent by the landmark data), these outline data also record information pertaining to aspects of the shapes of the dorsal, caudal, anal, and pelvic fins themselves. Under the standard ichthyological morphometric practice of placing landmarks only on the anterior intersections between the dorsal, anal, and pelvic fins with the body trunk, the shapes of these fins were not assessed by the landmark sampling scheme (see Fig. 5). Thus, the landmark dataset, if anything, records an underestimate of the extent of shape similarity and distinction present in this sample.

Of course it would be a simple matter to have included one or two landmarks on the dorsal, anal, and pelvic fins to enable the inclusion of aspects of their forms within the landmark dataset. But owing to the diversity of fin shapes exhibited by these species it would require a large number of landmarks to represent this diversity adequately, in which case the landmark dataset would be, effectively, transformed into an outline dataset. Even more importantly, the gross geometry of these fins is sufficiently diverse to raise questions over what constitutes a comparable topological location of individual anatomical points on different forms.

Take the dorsal fin, for example (Fig. 16). The dorsal fin of *Hola. ciliaris* is characterized by a reduction in the number of dorsal fin spines accompanied by development of the soft dorsal fin into an eccentric triangular form whose free vertex trails beyond the posterior extent of the caudal fin. *Holacanthus ciliaris* is the only species included in this dataset with such an extreme contrast between the spiny dorsal fin and the soft dorsal fin. In comparing this morphology with the dorsal fin of (say) *Holo. ascensionis*, it can be seen that the latter is characterized by a broadly prismatic form subequally divided into spiny and soft parts with the anterior dorsal fin spines being developed to greater lengths than the posterior spines or the soft fin rays. If these structures were to be represented by only a few landmarks substantial uncertainties would arise regarding the location of either biologically or topologically corresponding points across all forms included, with different decisions implying very different patterns of shape-similarity.

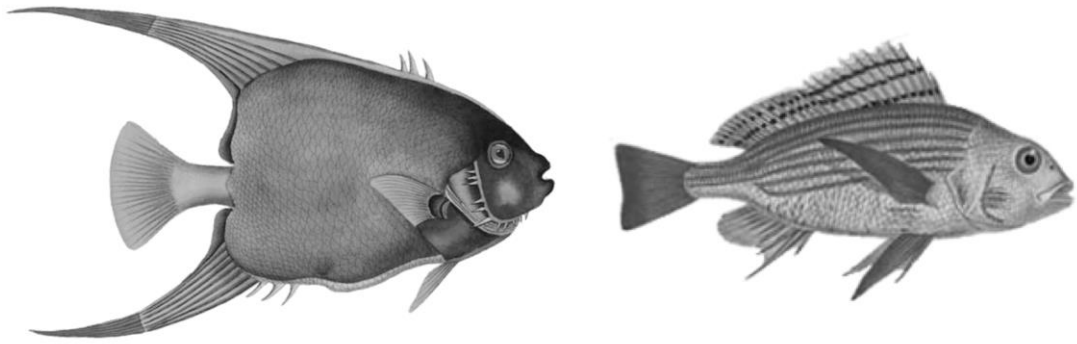

Figure 16. Gross dorsal fin shape differences between two species included in the example fish dataset. See text for discussion.

Bookstein (1991) regarded the uncertainty over these point-level correspondences as sufficient to exclude boundary outline semilandmark data from being considered in a geometric morphometric context. However, these dorsal fins, as well as other analogous structures on a wide range of different organisms, are present and are important to consider in the testing of a wide range of morphological, taxonomic and phylogenetic hypotheses. In all cases, however, these species have dorsal fins that are biologically homologous at the level of the structure and can be represented by comparable sequences of landmark-constrained semilandmark points (see below). But even these landmark and outline data do not exhaust the potential of geometric morphometrics to be of use in understanding patterns of morphological variation in nature.

The major dimensions of variation in the example fish dataset, as assessed by an unbiased analysis of a complete set of image pixels, are illustrated in Figure 17. Together, the first three eigenvectors summarize 48.91 percent of the observed shape data for this dataset, a substantially lower overall percentage than either the landmark or outline datasets. Nonetheless,

whereas the previous two datasets assumed that the features of interest among these images could be represented — and summarized — via the specification of configurations of mathematical points within a Cartesian space, this pixel-based dataset includes all the information contained in a 3,535-pixel rendering of the original images; information that not only captures aspects of the relative locations of major structural components of these fish and of their outlines, but also of the forms of individual structural components (e.g., eye size, eye shape, fin shapes) and of their body markings (e.g., shading variations, stripes, blotches). These data have been rendered comparable because (1) image sizes, positions and orientations variations have been standardized across the sample and (2) standardization of the image frame ensures that each pixel is topologically homologous in terms of its representation of morphological variation for each image as a whole.

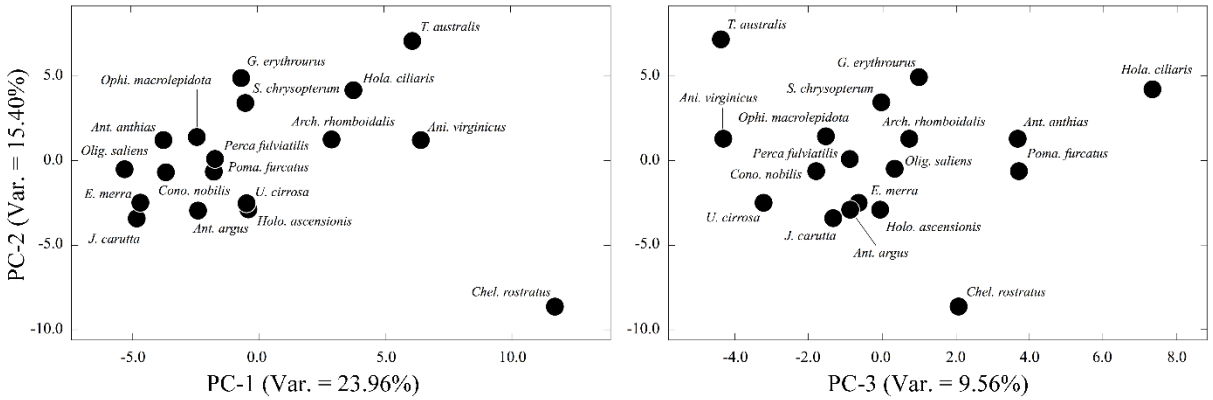

Figure 17. Major structure of shape variation for the fish sample as assessed the image pixel data (Fig. 5C). See text for discussion.

The major axis of shape configuration variation for these pixel data contrasts narrow elliptical or fusiform body forms with low-relief dorsal, anal and pelvic fins (e.g., *Olig. saliens*, *J. carutta*, *E. merra*) with broad fusiform and subquadrate body forms with high relief dorsal, anal and pelvic fins (*Arch. rhomboidalis*, *Ani. virginicus*, *Chel. rosatus*) as PC-1 score increases. In addition to these characteristics, it is interesting to note that fish images which project to high positions on the PC-1 axis all contain prominent dark regions on the body (e.g., the eye of *T. australis*, the eyespot of *Chel. rostatus*, the artefactual dark spot just above the pectoral fin of *Arch. rhomboidalis*) despite the fact that these features are not located at the same relative positions on the fish bodies. In contrast, the second PC axis contrasts narrow, elliptical-bodied forms with rounded caudal fins (e.g., *J. carutta*, *E. merra*, *Ant. argus*) with broad-bodied, fusiform species exhibiting deeply forked caudal fins (e.g., *T. australis*, *G. erythrourus*, *S. chrysopterum*) as projected PC-2 score increases. Despite its truncated tail, the broad bodied *Hola. ciliaris* also plots in this region. However, its gross similarity with these forms is understandable owing to the extreme morphologies of its dorsal and anal fins (which are used as secondary propulsion aids). Finally, the third PC axis separates forms exhibiting relatively large, well-developed, spiny dorsal fins and weak to moderately developed soft dorsal fins (e.g., *Ani. virginicus*, *Perca fluviatilis*) from those exhibiting weak to moderately developed spiny dorsal fins and large, well-developed soft dorsal fins (e.g., *Ant anthias*, *Poma. fucatus*, *Hola. ciliaris*).

In terms of the overall distribution of shapes in this three-dimensional image pixel subspace, two major groupings, in addition to several outliers, are present. The largest of these groupings occupies the central region of the subspace and consists of a miscellany of low elliptical to fusiform-bodied species with a variety of non-extreme fin shapes. Within the PC-3 vs PC-2 space there is a hint of a separation between equant elliptical-to fusiform, forked caudal fin species (e.g., *U. cirrosa*, *Cono. nobilis*, *Perca fluviatilis*) and low elliptical, round caudal-finned species (e.g., *J. cornuta*, *Ani. argus*, *E. merra*). This distinction is subtle and may be swamped by normal degrees of morphological variation in a larger sample. What is not subtle is the broad distinction between this large group of fish body forms and the smaller group of more extreme morphologies characterized by much more broadly elliptical and subquadrate body forms with relatively large, well-developed dorsal and anal fins (e.g., *Chel. rostratus*, *Hola. cillaris*, *Poma. furcatus*). It is interesting to note here that the somewhat narrow, fusiform-bodied *Ant. anthias* plots within this group at a position close to *Poma. furcatus* despite its much narrower body form. In all likelihood the aspect of shape similarity uniting this group also involves a strongly forked morphology in the posterior region which, in *Chel. rostratus* and *Hola. cillaris* is developed primarily through elaboration of the dorsal and anal fins and in *Poma. furcatus* and *Ant. anthias* is developed primarily through elaboration of the caudal fin. Surrounding these main groups is a small set of even more extreme, outlier forms including *Chel. rostratus* (extremely broad, elliptical body form, well-developed dorsal and anal fins, unusually elongate mandible/maxilla), *Holo. ciliaris* (unusually broad, trapezoidal body form, extremely elongate dorsal and anal fins, unusually short mandible/maxilla), *T.*

australis (broadly elliptical body form, unusually short and dorsally oriented maxilla, unusually large eyes) and *Ani. virginicus* (unusually broad, humped, fusiform body shape, unusually broad peduncle and an atypically long pectoral fin).

The results of this comparison among datasets collected from the same set of individuals — as well as others that have been published in the biological literature (see MacLeod, 2015b) — support the view that the conclusions obtained from any morphometric and/or geometric morphometric analysis are highly dependent on the manner in which the morphology is sampled. To put this another way, no set of morphometric results can be correctly regarded as being *the* summary of patterns of form or shape variation across a sample. All must be regarded as *a* summary of patterns of form or shape variation across a sample. Among the three sampling strategies examined above each has its own strengths and weaknesses.

Landmark data are most well-suited to the representation of form contrasts that are referenced to topologically corresponding locations on biologically homologous structures. But this focus usually comes at the expense of severely abstracting the shapes of the forms in question many aspects of whose morphology cannot be represented adequately or accurately by landmarks. Accordingly, landmarks should be used when the biological questions being assessed address themselves primarily to the relative placement of aspects of the morphology that can be represented adequately by landmarks. This constraint will inevitably result (and has inevitably resulted) in landmark sampling strategies being primarily used to summarize patterns of morphological variation within species and between species that are closely related to one another phylogenetically. It is within these cases that the largest number of comparable point locations that can be landmarked reliably will be found.

Boundary outline semilandmark data are often better able to capture morphological distinctions between a wider range of morphologies than is possible with landmark data. If used to assess patterns of variation within species and between species that are closely related to one another phylogenetically outline semilandmark data can take advantage of the inherent comparability of these forms to render small-scale differences location of topologically homologous points in the sequence of little consequence to the results of the analysis. But even in cases where there are substantial differences between the mathematical correspondence of points in a boundary outline sequence and the topological correspondences of points relative to biologically homologous structures, useful information and valid tests of shape similarity, difference, and sample-based characterizations of morphological structure can be made via reference to these data. Correspondences between outline sequences and biological features are often closer that has been suggested by a number of advocates of landmark analyses (MacLeod, 1999, 2012b) and can be improved substantially through the combination of landmark and outline approaches (see below). Boundary outline data also have the advantage of producing form/shape spaces and graphical summaries of form/shape results that are easier to interpret and communicate to non-quantitative audiences that landmark-based results.

Image pixel data have not be used extensively in geometric morphometric contexts to date though the use of images drawn from a wide variety of both biological and non-biological contexts is now common and growing as a result recent developments in machine learning and artificial intelligence (see Gaston & O'Neill, 2004; MacLeod, 2007, 2015a, b; Tarca *et al*., 2007; La Salle *et al*., 2009; Ranaweera *et al*., 2009; MacLeod *et al*., 2010; Selvam, 2010; Boyer *et al*., 2011; Kadir, 2015; MacLeod & Steart, 2015; Wilf *et al*., 2016; Li *et al*., 2016; MacLeod, in press a). Whereas the complete set of pixels comprising an image frame can be considered topologically homologous point locations at the level of a set of frame-standardized images, these point locations, of course, exhibit many fewer biological correspondences than either landmark or outline datasets. This having been said, if the subjects of analysis are individuals drawn from the same species or sets of species that are closely related to one another phylogenetically, degrees of biological correspondence between the image and the pixel sequence can be substantial. But even in the analysis of more heterogeneous sets of specimens/images, direct use of pixel data can be informative and sufficient to test a wide range of morphological hypotheses. In particular, if the purpose of the analysis is group identification the modelling of group variation and and/or the placement unknown specimens into group categories, image pixel data have a clear advantage over both landmark and boundary outline data insofar as wider range of information can be captured for use in the training of discrimination systems and because such data offer a means by which incomplete, poorly preserved and/or damaged specimens can be incorporated into an analysis.

5 Post-synthesis developments

There have been a number of developments in the core practices of geometric morphometrics since the mid-1990s when the core of the morphometric synthesis was essentially complete. If geometric morphometrics is considered to be a "toolkit" for the analysis of form — a popular metaphor among several prominent morphometricians — these developments have resulted in the addition of new tools to the kit so that a greater range of geometrically interesting and informative data can be treated in a manner consistent with the concepts and practices of the geometric approach.

Perhaps most prominent among these extensions has been the inclusion of data and techniques for the analysis of boundary outline curves as denoted by sets of semilandmarks as falling within the canon of geometric morphometric procedures considered appropriate and useful by most practitioners. Bookstein (1991) was highly critical of the idea of treating boundary outline data in a morphometric context owing to the impossibility specifying the precise position of corresponding point locations along two curves whose form differs. Since both size and shape differences in geometric morphometrics are strictly dependent on the point-by-point comparison of landmark locations on two or more forms, any ambiguity in the geometries of point matchings will lead to differences in the results obtained from the analysis of the same landmark configurations with no objective way of selecting the "correct" result from this (essentially infinite) set of alternatives. However, as can be seen in Figure 18, it is often the case that, even when the boundary outline semilandmark sequence is begun at a single landmark, the matching between boundary outline coordinates and aspects of topologically homologous locations on biologically homologous structures can be remarkably good and certainly of sufficient quality to enable a wide variety of morphological hypotheses to be tested.

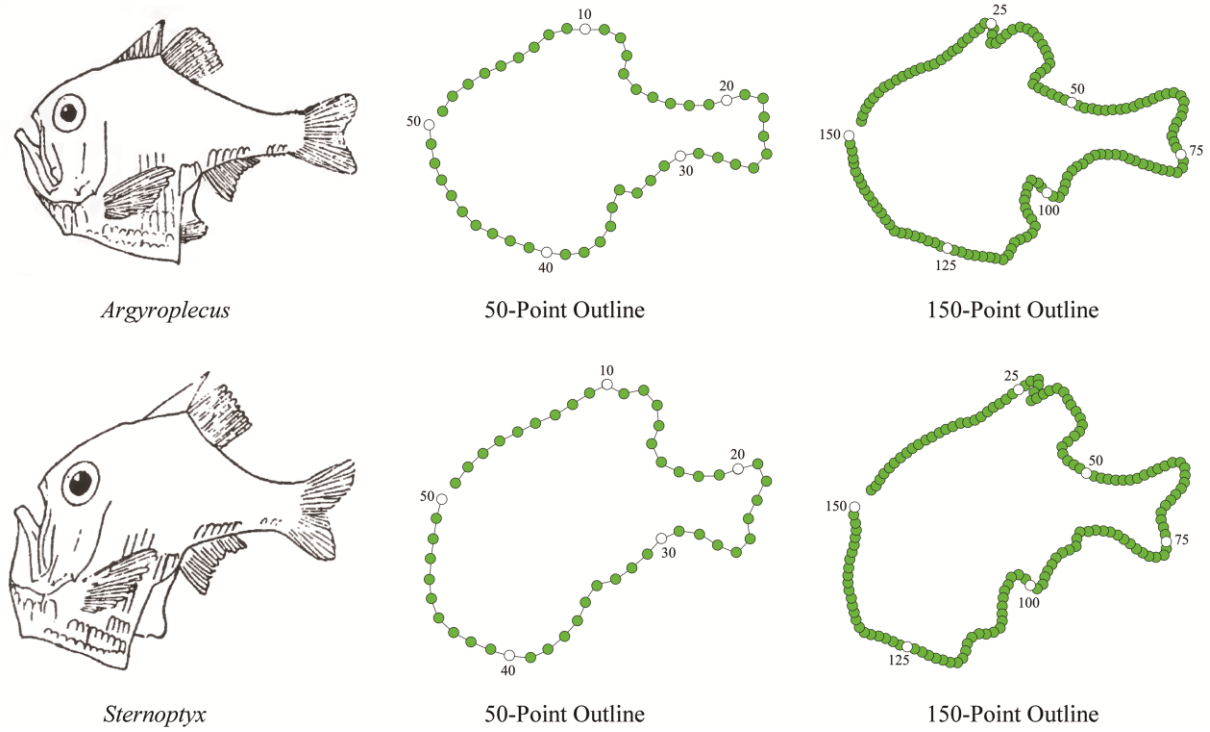

Figure 18. Equally-spaced boundary outline semilandmark coordinate representations of Thompson's *Argyoplecus* and *Sternoptyx* images using 50-point and 150-point sampling schemes. White icons represent every 10th semilandmark in the 50-point outline sequence and every 25th semilandmark in the 150-point outline sequence. Note how the detail of the outlines present in the original drawings has been smoothed in both schemes. Nevertheless, the representational quality is much greater in both schemes than for the landmark sample scheme shown in Fig. 5B. Note also that the level of mismatch between comparable points in both semilandmark sequences, with respect to biologically homologous structures, is remarkably good, though there is an obvious mismatch of points between these two outlines in the dorsal fin of the 150-point sampling scheme.

Traditionally this level of interpolation detail was determined by simply choosing a number of interpolated semilandmark coordinates arbitrarily (e.g., 100). Following this sampling step closed-boundary outline curves may be subjected to redescription using the elliptic Fourier or eigenshape procedures, or the set of boundary outline semilandmark coordinates may be aligned using the GLS Procrustes procedure and analysed via direct Procrustes PCA (= relative warps analysis) or a variety of other procedures (e.g., multivariate regression size for allometric analyses, CVA for group-based analyses). If the boundary of interest constitutes a closed curved with relatively sharp bends or an open curve some form of eigenshape analysis is recommended as elliptical Fourier analysis has difficulty characterizing the former and is not defined to enable analysis of the latter. Both boundary curves lying within the specimen outline and individual landmark points may also be added to boundary outline datasets and incorporated into a Procrustes PCA analysis (e.g., Figueirido *et al*., 2011)

If, for some reason, a small number of semilandmarks have been used to quantify the boundary of interest the issue of where to located these coordinate points along the curve of interest becomes important. Bookstein and Green (1993), Green

(1996), Bookstein (1997) and Sheets *et al*. (2004), and Perez *et al*. (2006) describe various procedures for migrating sparse semilandmarks to positions that optimize the characterization of boundary outline sequences in order either to minimize shape differences (= bending energy) between forms or result in the production of a series of curves that are maximally-smooth deformations of the average curve. This deformation is usually achieved by sliding the semilandmarks recursively along tangents to the curve at the position of the original semilandmark coordinate until the optimality criterion is met. Unfortunately, employment of tangents to facilitate the semilandmark sliding procedure results in production of a curve that was not an aspect of the original specimen and so represents a systematic distortion of the specimen's morphology in order to fulfil the assumptions that neither the number of semilandmarks, their spacing, their placement, or (apparently) the geometry of the curve they represent is of biological importance in the analysis geometric variation (Webster *et al*., 2001). Once such a deformed semilandmark configuration has been determined the locations of the slid semilandmark points can be projected back to points closer to the original outline, but the algorithms that enable this projection are arbitrary, and, as the sparse semilandmark points, in effect, define the outline, its "true" location following the sliding procedure is inherently uncertain. Regardless, when the number of semilandmark points sufficient to represent the geometry of a boundary or curve to a high degree of accuracy is employed application of the sliding semilandmark procedure rarely results in any substantial adjustment to the curve of interest because any displacement of landmarks located close to one another usually results in substantial increases in the bending energy. Indeed, in some cases even minor adjustments made by sliding of semilandmarks along tangents results in a less efficient characterization of shape-variation trends that would have been produced if the original high-quality semilandmark configuration had been retained (see MacLeod, 2015c for an example).

MacLeod (1999, 2012b) addressed the original concerns of all boundary-outline methods (incl. elliptical Fourier

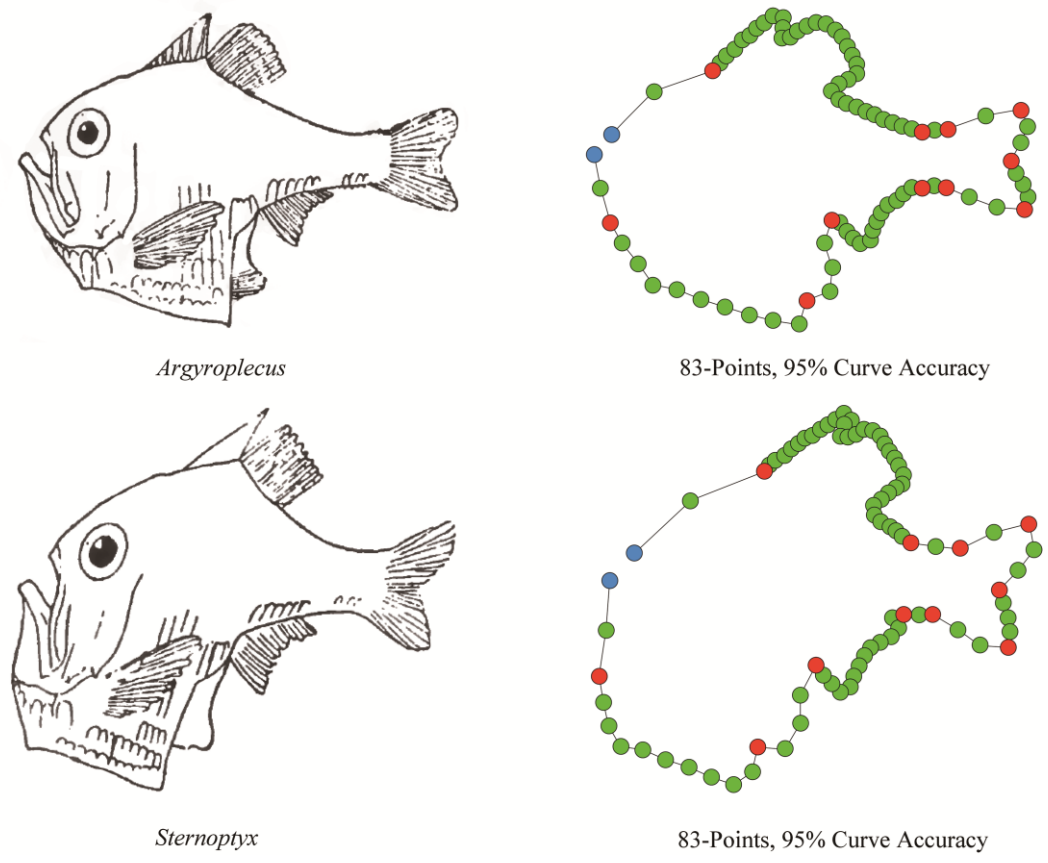

Figure 19. Results of an adaptive semilandmark interpolation analysis in which the *Argyroplecus* and *Sternoptyx* outlines, originally sampled with 250 semilandmark points, were interpolated using 13 landmark points located on the outline (blue = boundary outline termination landmarks, red = landmarks). Use of these landmark points breaks the outline into 12 segments which can then be interpolated separately using *n* equally spaced semilandmark points (green) where *n* equals the number of semilandmark points required to ensure a geometric fidelity fit of 95% of the length of the original (digitized) curve. This procedure ensures all aspects of the boundary outline are matched exactly at, and in the vicinity of the landmark points and ensures that all segments of the outline are represented to a consistent geometric representational standard. In order to meet these criteria over the entire outline for both forms 83 points were necessary. The eye, gill opercula dorsal terminus and the pectoral fin were excluded from this analysis, but could easily be added using identical criteria to constrain their representation, as could additional landmark points.

analysis) raised by Bookstein *et al*. (1982, see also Bookstein *et al.*, 1985) and Bookstein (1991) in terms of the ability of outline data to incorporate the concept of biological homology into their procedures by developing a variant of eigenshape analysis in which landmarks are used to subdivide the boundary outline into topologically homologous segments (Fig. 19, see MacLeod, 1999 for a detailed explanation of this procedure). The dimensionality issue discussed above can also be addressed under this "extended" formulation the eigenshape technique by adjusting the number of inter-landmark semilandmarks use to represent these landmark-delimited boundary outline curve segments based on the amount of shape variation displayed by each segment across a sample. In situations where such subdivision is possible employment of this strategy can improve the results of a geometric morphometric analysis of semilandmark-dominated datasets dramatically, even when a comparatively small number of landmarks are used to constrain the semilandmark interpolation. While some degree of ambiguity in the placement of semilandmarks located in the middle of these segments remains, semilandmarks located proximal to the true landmarks should be located accurately. This strategy also has the effect of allowing semilandmarks residing on different outline segments to be interpolated independently of one another, resulting in the analysis being able to achieve a more natural balance between outline segments exhibiting high and low degrees of shape variation.

Along more purely technological lines, the recent proliferation of relatively inexpensive three-dimensional (3D) digitizers has allowed morphometricians increasingly to undertake the morphometric analysis of 3D data. To some extent this development required little more than minor rewriting of various computer algorithms to handle the extra Cartesian variables as virtually all core geometric morphometric procedures had long been formulated to cover the 3D case (e.g., Bookstein shape coordinates, Procrustes shape coordinates). As a result, landmark and outline-based geometric morphometric analyses were little affected by the extension to 3D. But the arrival of 3D data has forced morphometricians to come to grips with the problem of characterizing featureless 3D surfaces within a geometric morphometric context.

Gunz *et al*. (2005, see also Gunz & Mitteröcker, 2013) and MacLeod (2008, see also Polly, 2008; Polly & MacLeod, 2008; Sievwright & MacLeod, 2012) have both published conceptually similar procedures for implementing such analyses. The Gunz *et al*. (2005) procedure takes an *a priori* set of 3D semilandmarks digitized a positions that are assumed to be comparable across a set of human crania and focuses on the problem of 3D semilandmark sliding along tangent planes to adjust individual semilandmark coordinates to positions that conform to external optimality criteria. In contrast, the MacLeod-Polly approach focuses on the problem of developing an automated algorithm that could objectively and accurately find the positions of a network or grid of semilandmarks that cover the surface in a complete, even, and comparable manner. This algorithm can be applied to the analysis of a wide range of morphologies, characters, taxonomic groups, etc. Under both approaches, once the semilandmark network has been finalized the coordinate data are processed using (3D) Procrustes superposition and the standard toolkit of secondary analytic procedures (e.g., PCA, CVA, multiple regression).

6 Future directions

To date morphometrics has proven to be an extraordinarily popular set of tools for a wide range of investigations within organismal biology and palaeontology. This popularity stems from the perennial scientific and practical interest in morphology as well as the intrinsic power of the geometric approach to form analysis. However, in this writer's opinion, there is no reason to suppose either that the geometric morphometric toolkit is complete, or that the use of its tools need stop at its application to organismal morphology. Situations in which the rapid, accurate, and geometrically comprehensive analysis of morphological variation occur frequently in many scientific, engineering and applied mathematical fields, and even in historical, cultural and artistic contexts.

Morphometrics also made strong inroads into the biological sciences initially because of the conceptual links between many of the data-analysis procedures employed by morphometricians with the numerical taxonomy movement of the 1950s, 1960s and 1970s (Sokal & Sneath, 1963; Sneath & Sokal, 1973). As I have noted above, a variety of criteria may be used to define the placement of landmarks and semilandmark sequences on images and 3D scans of specimens. Some of these reference the existence of corresponding locations on homologous biological structures and some do not. But increasingly, geometric morphometric concepts and methods are being applied to objects in which the concept of biological homology does not apply, but the concepts of form, size and shape variation do (e.g., Coles, 1992; Buchanan *et al*., 2007; Buchanan & Collard, 2010; Cardillo, 2010; Thulman, 2012; MacLeod *et al*., 2013; Buchanan *et al*., 2014; Seetah, 2014; Smith *et al*., 2014; Fox, 2015; MacLeod, in press b). Morphometricians should be unafraid to use the tools and skills that have been developed largely to support biological investigations in these (and other) novel contexts.

By the same token, morphometrics has, for the most part, relied on rather old-fashioned — though no less useful — methods of linear algebra and multivariate analysis to create summaries of patterns of form, size and shape variation. While these are perfectly serviceable in a wide range of applications and contexts, newer, more computation-intensive procedures are now available in the areas of non-linear multivariate analysis, computer vision, machine learning and artificial intelligence. Such procedures have yet to be integrated fully into the morphometric synthesis and/or developed for routine use within the geometric morphometric toolkit. Indeed, many morphometricians have come to define their field not by the set the scientific problems they find interesting, but by a set of (now) standard procedures that can be applied to standard types of data to yield largely predictable — but no less informative — results (e.g., Webster *et al*., 2001; Adams *et al*., 2013). Yet, the utility of these new procedures for addressing a wide range of problems that have heretofore been considered morphometric in character is obvious (see MacLeod, 2007 and chapters therein, MacLeod *et al*., 2010; Wilf *et al*., 2016). If morphometricians do not, or cannot, continue to develop the range, scope, flexibility and accuracy of their tools and approaches — and apply these to real-world problems that deliver interesting, practical, and useful results — the intellectual ferment and dynamism that has served the field so well over the past 30 years could be lost to other areas of applied mathematics and the practitioners thereof. Consolidation and strategic development of the field of morphometrics along these lines is just as important to its future as the equally necessary developments that are focused on the refinement, extension, and integration of morphometrics within its traditional mathematical and biological contexts (e.g., Mitteröcker & Bookstein, 2007; Klingenberg & Gidaszewski, 2010; Klingenberg, 2013; Adams & Collyer, 2015, 2016; Denton & Adams, 2015; Bookstein, 2016; Mitteröcker *et al*., 2016; Parr *et al*., 2016; Cardini, 2016a, b).

Acknowledgements I'd like to take this opportunity to thank Richard Reyment, Fred Bookstein and Jim Rohlf for all their many contributions to the field of morphometrics and for being my teachers and colleagues over the (now many) years I've been active in this field. As will be obvious from the text of this article, we don't always see eye-to-eye about every issue in morphometrics, but even in those areas where we disagree most of what I know about morphometrics I've learned by reading, listening, and thinking about the things they've written and said. An anonymous reviewer also read the manuscript in draft form and made several useful suggestions for modifications of the text. Finally, I'd like to thank The Natural History Museum for supporting my research in this area and providing a (mostly) congenial and interesting intellectual home over the greater part of my research career to date.

References

Adams, D.C., Collyer, M.L. 2015. Permutation tests for phylogenetic comparative analyses of high-dimensional shape data: what you shuffle matters. *Evolution*, 69(3): 823–829.

Adams, D.C., Collyer, M.L. 2016. On the comparison of the strength of morphological integration across morphometric datasets: *Evolution*. doi: 10.1111/evo.13045.

Adams, D.C., Rohlf, F.J., Slice, D.E. 2004. Geometric morphometrics: ten years of progress following the 'revolution'. *Italian Journal of Zoology*, 71(1): 5–16.

Adams, D.C., Rohlf, F.J., Slice, D.E. 2013. A field comes of age: geometric morphometrics in the 21st century. *Hystrix*, 24(1): 7–14.

Adams, D.C., Rosenberg, M.S. 1998. Partial warps, phylogeny, and ontogeny: a comment on Zelditch and Fink (1995). *Systematic Biology*, 47: 168–173.

Albrecht, G.H., Gelvin, B.R., Hartman, S.E. 1993. Ratios as a size adjustment in morphometrics. *American Journal of Physical Anthropology*, 91: 441–468.

Andersen, K. 2007. *The geometry of an art : the history of the mathematical theory of perspective from Alberti to Monge*. Springer, New York. 856 pp.

Anderson, E. 1935. The irises of the Grasp é Peninsula. *Bulletin of the American Iris Society*, 59: 2–5.

Anderson, E. 1936. The species problem in *Iris*. *Annals of the Missouri Botanical Garden*, 23(3): 457–509.

Belhumeur, P.N., Hespanha, J.P., Kriegman, D.J. 1997. Eigenfaces vs. fisherfaces: recognition using class specific linear projection. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 19(7): 711–720.

Benson, R.H. 1982. Deformation, Da Vinci's concept of form, and the analysis of events in evolutionary history. *In:* Gallitelli, E.M. (ed.). *Palaeontology, Essential of Historical Geology*. Mucchi, Modena, Italy. pp.241–277.

Benson, R.H., Chapman, R.E., Siegel, A.F. 1982. On the measurement of morphology and its change. *Paleobiology*, 8: 328–339.

Blackith, R.E. 1957. Polymorphism in some Australian locusts and grasshoppers. *Biometrics*, 13: 183–196.

Bookstein, F.L. 1977. The study of shape transformation after D'Arcy Thompson. *Mathematical Biosciences*, 34: 177–219.

Bookstein, F.L. 1978. *The Measurement of Biological Shape and Shape Change*. Springer, Berlin. 191 pp.

Bookstein, F.L. 1984. A statistical method for biological shape comparison. *Journal of Theoretical Morphology*, 107: 475–520.

Bookstein, F.L. 1986. Size and shape spaces for landmark data in two dimensions. *Statistical Science*, 1(2): 181–242.

Bookstein, F.L. 1989. Principal warps: thin-plate splines and the decomposition of deformations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11: 567–585.

Bookstein, F.L. 1991. *Morphometric Tools for Landmark Data: Geometry and Biology*. Cambridge University Press, Cambridge. 425pp.

Bookstein, F.L. 1993. A brief history of the morphometric synthesis.. *In:* Marcus, L.F., Bello, E., Garc á-Valdecasas, A. (eds). *Contributions to Morphometrics*. Museo Nacional de Ciendcias Naturales, Madrid. pp. 18–40.

Bookstein, F.L. 1994. Can biometrical shape be a homologous character? *In:* Hall, B.K. (ed.). *Homology: The Hierarchical Basis of Comparative Biology*. Academic Press, San Diego, California. pp. 197–227.

Bookstein, F.L. 1996a. Standard formula for the uniform shape component in landmark data. *In:* Marcus, L.F. Corti, M., Loy, A., Naylor, G.J.P., Slice, D.E. (eds). *Advances in Morphometrics. Proceedings of the 1993 NATO Advanced Studies Institute on Morphometrics in Il Ciocco, Italy*. Plenun Public Corp., New York. pp. 153–168.

Bookstein, F.L. 1996b. Combining the tools of geometric morphometrics. *In:* Marcus, L.F. Corti, M., Loy, A., Naylor, G.J.P., Slice, D.E. (eds). *Advances in Morphometrics. Proceedings of the 1993 NATO Advanced Studies Institute on Morphometrics in Il Ciocco, Italy*. Plenun Public Corp., New York. pp. 131–151.

Bookstein, F.L. 1997. Landmark methods for forms without landmarks: localizing group differences in outline shape. *Medical Image Analysis*, 1: 225–243.

Bookstein, F.L. 2016. The inappropriate symmetries of multivariate statistical analysis in geometric morphometrics. *Evolutionary Biology*, 43: 277–313.

Bookstein, F.L., Chernoff, B., Elder, R., Humphries, J., Smith, G., Strauss, R. 1985. *Morphometrics in Evolutionary Biology*. The Academy of Natural Sciences of Philadelphia, Philadelphia. 277 pp.

Bookstein, F.L., Green, W.D.K. 1993. A feature space for edgels in images with landmarks. *Journal of Mathematical Imaging and Vision*, 3: 213–261.

Bookstein, F.L., Strauss, R.E., Humphries, J.M., Chernoff, B., Elder, R.L., Smith, G.R. 1982. A comment on the uses of Fourier methods in systematics. *Systematic Zoology*, 31: 85–92.

Bowditch, S. 1828–38. *The Fresh-water Fishes of Great Britain.* Ackermann & Co., London.

Boyer, D.M., Lipman, Y., St Clair, E., Puente, J., Patel, B.A., Funkhouser, T., Jernvall, J., Daubechies, I. 2011. Algorithms to automatically quantify the geometric similarity of anatomical surfaces. *Proceedings of the National Academy of Science*, 108(45): 18221–18226.

Buchanan, B., Collard, M. 2010. A geometric morphometrics-based assessment of blade shape differences among Paleoindian projectile point types from western North America. *Journal of Archaeological Science*, 37(2010): 350–359.

Buchanan, B., Johnson, E., Strauss, R.E., Lewis, P.J. 2007. A morphometric approach to assessing late paleoindian projectile point variability on the southern high plains. *Plains Anthropologist*, 52(203): 279–299.

Buchanan, B., O'Brien, M.J., Collard, M. 2014. Continent-wide or region-specific? A geometric morphometrics-based assessment of variation in Clovis point shape. *Archaeological and Anthropological Sciences*, 6: 145–162.

Burke, R.L., Leuteritz, T.E., Wolf, J. 1996. Phylogenetic relationships of emydine turtles. *Herpetologia*, 52: 572–584.

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

Cardillo, M. 2010. Some applications of geometric morphometrics to archaeology. *In:* Elewa, A.M.T. (ed.). *Morphometrics for Nonmorphometricians*. Springer-Verlag. pp. 325–341.

Cardini, A. 2016a. Lost in the other half: improving accuracy in geometric morphometric analyses of one side of bilaterally symmetric structures. *Systematic Biology*, 65(6): 1096–1106. doi: 10.1093/sysbio/syw043.

Cardini, A. 2016b. Left, right or both? Estimating and improving accuracy of one-side-only geometric morphometric analyses of cranial variation. *Journal of Zoological Systematics and Evolutionary Research*. doi: 10.1111/jzs.12144.

Catalano, S.A., Goloboff, P.A., Giannini, N.P. 2010. Phylogenetic morphometrics (I): the use of landmark data in a phylogenetic framework. *Cladistics*, 26: 539–549.

Chapman, R.E. 1990. Conventional Procrustes approaches. *In:* Bookstein, F.L. and Rohlf, F.J. (eds). *Proceedings of the Michigan Morphometrics Workshop, Special Publication No. 2*. The University of Michigan, Ann Arbor, Michigan. 251–268 pp.

Christopher, R.A., Waters, J.A. 1974. Fourier analysis as a quantitative descriptor of miosphere shape. *Journal of Paleontology*, 48: 697–709.

Cheetham, A.H., Lorenz, D.M. 1976. A vector approach to size and shape comparisons among zooids in cheilostome bryozoans. *Smithsonian Contributions to Paleobiology*, 29: 1–55.

Cloquet, H. 1830. *Dictionnaire des Sciences Naturelles: Poissons et reptiles*. F. G. Levrault, Strasbourg. 200 pp.

Coles, P. 1992. Patterns in galaxy clustering. *In:* Feigelson, E.D., Babu, G.J. (eds). *Statistical Challenges in Modern Astronomy*. Springer-Verlag, New York. pp. 57–72.

Cosgriff, R.L. 1960. Identification of shape. *Ohio State University Research Foundation*, 820(11): 254–792.

Cressie, N.A.C. 1988. Spatial prediction and ordinary kriging. *Mathematical Geology*, 20(4): 405–421.

Cressie, N.A.C. 1990. The origins of kriging. *Mathematical Geology*, 22(3): 239–252.

Davis, J.C. 2002. *Statistics and Data Analysis in Geology (Third Edition)*. John Wiley and Sons, New York. 638 pp.

Denton, J.S., Adams, D.C. 2015. A new phylogenetic test for comparing multiple high-dimensional evolutionary rates suggests interplay of evolutionary rates and modularity in lanternfishes (Myctophiformes; Myctophidae). *Evolution*, 69(9): 2425–2440.

Dryden, I.L., Mardia, K.V. 1998. *Statistical Shape Analysis*. J. W. Wiley, New York. 376 pp.
Ehrlich, R., Pharr, R.B., Healy-Williams, N. 1983. Comments on the validity of Fourier descriptors in systematics: a reply to Bookstein *et al*. *Systematic Zoology*, 32(2): 202–206.
Ehrlich, R., Weinberg, B. 1970. An exact method for characterization of grain shape. *Journal of Sedimentary Petrology*, 40(1): 205–212.
Figueirido, B., MacLeod, N., Krieger, J., De Renzi, M., Pérez-Claros, J.A., Palmqvist, P. 2011. Constraint and adaptation in the evolution of carnivoran skull shape. *Paleobiology*, 37(3): 490–518.
Fink, W.L., Zelditch, M.L. 1995. Phylogenetic analysis of ontogenetic shape transformations: a reassessment of the piranha genus *Pygocentrus* (Teleostei). *Systematic Biology*, 44: 343–360.
Fisher, R.A. 1936. The utilization of multiple measurements in taxonomic problems. *Annals of Eugenics*, 7: 179–188.
Fox, A.N., 2015. A study of Late Woodland projectile point typology in New York using elliptical Fourier outline analysis. *Journal of Archaeological Science: Reports*, 4(2015): 501–509.
Fritzsche, D.L. 1961. A systematic method for character recognition. *Ohio State University Research Foundation*, 1222(4): 268–360.
Galton, F. 1886. Regression toward mediocrity in hereditary stature. *Journal of the Anthropological Institute*, 15: 246–263.
Gaston, K.J., O'Neill, M.A. 2004. Automated species identification—why not? *Philosophical Transactions of the Royal Society of London, Series B*, 359: 655–667.
Gayon, J. 2000. History of the concept of allometry. *American Zoologist*, 40(5): 748–758.
Goloboff, P.A., Catalano, S.A. 2011. Phylogenetic morphometrics (II): algorithms for landmark optimization. *Cladistics*, 27(1): 42–51.
Gombrich, E.H. 2007. *The Story of Art*. Phaidon Press, London. 688 pp.
Goodall, C.R. 1983. *The statistical analysis of growth in two dimensions*. Harvard University, Cambridge, Massachusetts.
Goodall, C.R. 1991. Procrustes methods in the statistical analysis of shape. *Journal of the Royal Statistical Society, Series B*, 53: 285–339.
Goodall, C.R., Mardia, K.V. 1991. A geometrical derivation of the shape density. *Applied Probability*, 23(3): 496–514.
Gould, S.J. 1971. D'Arcy Thompson and the science of form. *New Literary History*, 2(2): 229–258.
Gould, S.J. 2002. *The Structure of Evolutionary Theory*. Belknap/Harvard University Press, Cambridge, Massachusetts. 1433 pp.
Green, W.D.K. 1996. The thin-plate spline and images with curving features. *In:* Mardia, K.V., Gill, C.A., Dryden, I.L. (eds). *Proceedings in Image Fusion and Shape Variability Techniques*. Leeds University Press, Leeds. pp. 79–87.
Gower, J.C. 1975. Generalized Procrustes analysis. *Psyrohchometrica*, 40: 33–51.
Graunt, J. 1662. Natural and Political Observations mentioned in a following index, and made upon the Bills of Mortality With reference to the Government, Religion, Trade, Growth, Ayre, diseases, and the several Changes of the said City. Privately published, London.
Gunz, P., Mitteröcker, P., Bookstein, F.L. 2005. Semilandmarks in three dimensions. *In:* Slice, D.E. (ed.). *Modern Morphometrics in Physical Anthropology*. Kluwer Academic/Plenum Publishers, New York. pp. 73–98.
Gunz, P., Mitteröcker, P. 2013. Semilandmarks: a method for quantifying curves and surfaces. *Hystrix*, 24(1): 109–115.
Huxley, J.S. 1924. The variation in the width of the abdomen of immature fiddler crabs considered in relation to its relative growth-rate. *American Naturalist*, 58: 468–475.
Jolicouer, P., Mosimann, J.E. 1960. Size and shape variation in the painted turtle, a principle component analysis. *Growth*, 24: 339–354.
Kadir, A. 2015. Leaf identification using polar Fourier transform and linear Bayes normal classifier. *In:* Lestrel, P.E. (ed.). *Proceedings of the Third International Symposium on Biological Shape Analysis*. World Scientific, Singapore. pp. 40–49.
Kaesler, R.L. 1997. Phase angles, harmonic distance, and the analysis of form. *In:* Lestrel, P.E. (ed.). *Fourier Descriptors and Their Applications in Biology*. Cambridge University Press, Cambridge. pp. 106–125.
Kaesler, R.H., Waters, J.A. 1972. Fourier analysis of the ostracode margin. *Bulletin of the Geological Society of America*, 83: 1169–1178.
Kendall, D.G. 1977. The diffusion of shape. *Advances in Applied Probability*, 9: 428–430.
Kendall, D.G. 1981. The statistics of shape. *In:* Barnett, V. (ed.). *Interpreting Multivariate Data*. Wiley, Chichester, West Sussex. pp. 75–80.
Kendall, D.G. 1984. Shape manifolds, procrustean metrics and complex projective spaces. *Bulletin of the London Mathematical Society*, 16: 81–121.
Kendall, D. G., Barden, D., Carne, T.K., Le, H. 1999. *Shape and Shape Theory*. Wiley, New York. 328 pp.
Kim, K., Sheets, H.D., Haney, R.A., Mitchell, J.T. 2002. Morphometric analysis of ontogeny and allometry of the Middle Ordovician trilobite *Triarthrus becki*. *Paleobiology*, 28: 364–377.
Klingenberg, C.P. 2013. Cranial integration and modularity: insights into evolution and development from morphometric data. *Hystrix*, 24(1): 49–64.
Klingenberg, C.P., Gidaszewski, N.A. 2010, Testing and quantifying phylogenetic signals and homoplasy in morphometric data. *Systematic Biology*, 59(3): 245–261.
Krige, D.G. 1951. A statistical approach to some basic mine valuation problems on the Witwatersrand. *Journal of the Chemical, Metallurgical and Mining Society of South Africa*, 52: 119–139.
Kuhl, F.P., Giardina, C.R. 1982. Elliptic Fourier features of a closed contour. *Computer Graphics and Image Processing*, 18: 236–258.
La Salle, J., Wheeler, Q., Jackway, P., Winterton, S., Hobern, D., Lovell, D. 2009. Accelerating taxonomic discovery through automated character extraction. *Zootaxa*, 2217: 43–55.
Li, S.F., Jacques, F.M.B., Spicer, R.A., Su, T., Spicer, T.E.V., Yang, J., Zhou, Z.K. 2016. Artificial neural networks reveal a high-resolution climatic signal in leaf physiognomy. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 442: 1–11.

Lohmann, G.P. 1983. Eigenshape analysis of microfossils: A general morphometric method for describing changes in shape. *Mathematical Geology*, 15: 659–672.

Lohmann, G.P., Schweitzer, P.N. 1990. On eigenshape analysis. *In:* Rohlf, F.J., Bookstein, F.L. (eds). *Proceedings of the Michigan Morphometrics Workshop*. The University of Michigan Museum of Zoology, Special Publication No. 2, Ann Arbor. pp. 145–166.

Lynch, J.M., Wood, C.G., Luboga, S.A. 1996. Geometric morphometrics in primatology: craniofacial variation in *Homo sapiens* and *Pan troglodytes*. *Folia Primatologia*, 67: 15–39.

MacLeod, N. 1999. Generalizing and extending the eigenshape method of shape visualization and analysis. *Paleobiology*, 25(1): 107–138.

MacLeod, N. 2006. Minding your Rs and Qs. *Palaeontological Association Newsletter*, 61: 42–60.

MacLeod, N. 2007. *Automated Taxon Identification in Systematics: Theory, Approaches and Applications*. CRC Press, Taylor & Francis Group, London. 339 pp.

MacLeod, N. 2008. Understanding morphology in systematic contexts: 3D specimen ordination and 3D specimen recognition. *In:* Wheeler, Q. (ed.). *The New Taxonomy*. CRC Press, Taylor & Francis Group, London. pp. 143–210.

MacLeod, N. 2010a. Shape models II: the thin plate spline. *Palaeontological Association Newsletter*, 73: 24–39.

MacLeod, N. 2010b. Principal warps, relative warps and *Procrustes* PCA. *Palaeontological Association Newsletter*, 75: 22–33.

MacLeod, N. 2011a. Semilandmarks and radial Fourier analysis. *Palaeontological Association Newsletter*, 76: 25–42.

MacLeod, N. 2011b. The center cannot hold I: Z-R Fourier analysis. *Palaeontological Association Newsletter*, 78: 35–45.

MacLeod, N. 2012a. The center cannot hold II: elliptic Fourier analysis. *Palaeontological Association Newsletter*, 79: 29–42.

MacLeod, N. 2012b. Going round the bend II: extended eigenshape analysis. *Palaeontological Association Newsletter*, 81: 23–39.

MacLeod, N. 2015a. The direct analysis of digital images (eigenimage) with a comment on the use of discriminant analysis in morphometrics. *In* Lestrel, P.E. (ed.). *Proceedings of the Third International Symposium on Biological Shape Analysis*. World Scientific, Singapore. pp. 156–182.

MacLeod, N. 2015b. A comparison of alternative form-characterization approaches to the automated identification of biological species. *In:* Hamilton, A. (ed.). *Evolution of Phylogenetic Systematics*. University of California Press, Berkeley, California. pp. 214–245.

MacLeod, N. 2015c. Use of landmark and outline morphometrics to investigate thecal form variation in crushed gogiid echinoderms. *Palaeoworld*, 24(4): 408–429.

MacLeod, N. in press a. On the use of machine learning methods in morphometric analysis. *In*: Lestrel, P.E. (ed.), *Proceedings of the Fourth International Symposium on Biological Shape Analysis*. World Scientific, Singapore.

MacLeod, N. in press b. The quantitative assessment of archaeological object groups: beyond geometric morphometrics. *Journal of Archaeological Science*.

MacLeod, N., Benfield, M. and Culverhouse, P.F. 2010. Time to automate identification. *Nature*, 467(9): 154–155.

MacLeod, N., Forey, P. (eds). 2002. *Morphology, Shape and Phylogenetics*. Taylor & Francis, London. 308 pp.

MacLeod, N., Krieger, J., Jones, K.E. 2013. Geometric morphometric approaches to acoustic signal analysis in mammalian biology. *Hystrix*, 24(1): 116–125.

MacLeod, N., Steart, D. 2015. Automated leaf physiognomic character identification from digital images. *Paleobiology*, 41(4): 528–553.

Mahalanobis, P.C. 1936. On the generalized distance in statistics. *Proceedings of the National Academy of Science, India*, 12: 49–55.

Marcus, L.F., Bello, E., Garc á-Valdecasas, A. 1993. *Contributions to Morphometrics*. Museo Nacional de Ciencias Naturales 8, Madrid. 264 pp.

Marcus, L.F., Corti, M., Loy, A., Naylor, G.J.P., Slice, D.E. 1996. *Advances in Morphometrics*. Plenum Press, New York. 587 pp.

Marcus, L.F., Hingst-Zaher, E., Zaher, H. 2000. Applications of landmark morphometrics to skulls representing the orders of living mammals. *Hystrix*, 11: 24–48.

Mardia, K.V., Dryden, I.L. 1989. Shape distributions for landmark data. *Advances in Applied Probability*, 21: 742–755.

Mardia, K.V., Dryden, I.L. 1998. The statistical analysis of shape data. *Biometrika*, 76: 271–282.

Matheron, G. 1971. *The Theory of Regionalized Variables and Its Applications*. École National Sup érieure des Mines, Paris.

Merkus, H.G. 2009. *Particle Size Measurements: Fundamentals, Practice, Quality*. Springer, New York. 548 pp.

Mitter öcker, P., Bookstein, F.L. 2007. The conceptual and statistical relationship between modularity and morphological integration. *Systematic Biology*, 56: 818–836.

Mitter öcker, P., Cheverud, J.M., Pavlicev, M. 2016. Multivariate analysis of genotype-phenotype association. *Genetics*. doi: 10.1534/genetics.115.181339.

Mitter öcker, P., Gunz, P. 2009. Advances in geometric morphometrics. *Evolutionary Biology*, 36: 235–247.

Mosimann, J.E. 1970. Size allometry: size and shape variables with characterizations of the lognormal and generalized gamma distributions. *Journal of the American Statistical Association*, 65(330): 930–945.

Needham, J. 1959. *Science and Civilization in China, Volume 3: Mathematics and the Sciences of the Heavens and Earth*. Cambridge University Press, Cambridge. 877 pp.

Olsen, E. 2011a. Particle shape factors and their use in image analysis part I: theory. *Journal of GXP Compliance*, 15(3): 85–96.

Olsen, E. 2011b. Particle shape factors and their use in image analysis part II: practical applications. *Journal of GXP Compliance*, 15(4): 77–89.

Parr, W.C.H., Wilson, L.A.B., Wroe, S., Colman, N.J., Crowther, M.S., Letnic, M. 2016. Cranial shape and the modularity of hybridization in dingoes and dogs; hybridization does not spell the end for native morphology. *Evolutionary Biology*, 43(2): 1–17.

Perez, S.I., Bernal, V., Gonzalez, P.N. 2006. Differences between sliding semi-landmark methods in geometric morphometrics, with an application to human craniofacial and dental variation. *Journal of Anatomy*, 208: 769–784.

Pearson, K. 1901. On lines and planes of closest fit to a system of points in space. *Philosophical Magazine*, 2: 557–572.

Petty, W. 1690. *Political arithmetick*. Robert Clavel, London.

Polly, P.D. 2008. Adaptive zones and the pinniped ankle: a 3D quantitative analysis of carnivoran tarsal evolution. *In:* Sargis, E.J., Dagosto, M. (eds). *Mammalian Evolutionary Morphology: A Tribute to Frederick S. Szalay*. Springer, Dordrecht, The Netherlands. pp. 165–194.

Polly, P.D., MacLeod, N. 2008. Locomotion in fossil Carnivora: an application of the eigensurface method for morphometric analysis of 3D surfaces. *Palaeontologia Electronica*, 11(2): 13p.

Quetlet, A. 1835. *Sur l'homme et le développement de ses facultés, ou Essai de physique sociale*. Bachelier, Paris.

Ranaweera, K., Harrison, A.P., Bains, S., Joseph, D. 2009. Feasibility of computer-aided identification of foraminiferal tests. *Marine Micropaleontology*, 72(1–2): 66–75.

Rao, C.R. 1948. The utilization of multiple measurement in problems of biological classification. *Journal of he Royal Statistical Society, Series B*, 10(2): 159–203.

Raudseps, J.G. 1965. Some aspects of the tangent-angle arc length representation of contours. *Ohio State University Research Foundation*, 1801(6): 462–877.

Reyment, R.A. 2005. Aspects of applied morphometrics. *Zeitschrift der Deutschen Geologischen Gesellschaft*, 155(2–4): 263–274.

Reyment, R.A. 2010. Morphometrics: an historical essay. *In:* Elewa, A.M.T. (ed.). *Morphometrics for Nonmorphometricians*. Springer-Verlag, Berlin. 9–24 pp.

Reyment, R.A., Blackith, R.E., Campbell, N.A. 1984. *Multivariate Morphometrics (Second Edition)*. Academic Press, London. 231 pp.

Rodriguez, J. M., Edeskär, T., Knutsson, S. 2013. Particle shape quantities and measurement techniques–a review. *The Electronic Journal of Geotechnical Engineering*, 18(A): 169–198.

Rohlf, F.J. 1986. Relationships among eigenshape analysis, Fourier analysis, and analysis of coordinates. *Mathematical Geology*, 18(8): 845–854.

Rohlf, F.J. 1993. Relative warp analysis and an example of its application to mosquito wings. *In:* Marcus, L.F., Bello, E., García-Valdecasas, A., (eds). *Contributions to Morphometrics*. Museo Nacional de Ciencias Naturales 8, Madrid. pp. 131–160.

Rohlf, F.J. 1998. On applications of geometric morphometrics to studies of ontogeny and phylogeny. *Systematic Biology*, 47: 147–158.

Rohlf, F.J., Bookstein, F.L. 1987. A comment on shearing as a method for "size correction". *Systematic Zoology*, 36(4): 356–367.

Rohlf, F.J., Bookstein, F.L. (eds.). 1990. *Proceedings of the Michigan morphometrics workshop*. The University of Michigan Museum of Zoology Special Publication 2, Ann Arbor, Michigan. 380 pp.

Rohlf, F.J., Bookstein, F.L. 2003. Computing the uniform component of shape variation. *Systematic Biology*, 53: 66–69.

Rohlf, F.J., Slice, D. 1990. Extensions of the Procrustes method for optimal superposition of landmarks. *Systematic Zoology*, 39: 40–59.

Seetah, T.K. 2014. Geometric morphometrics and environmental archaeology. *In:* Smith, C. (ed.). *Encyclopedia of Global Archaeology*. Springer, New York. pp. 3029–3036.

Selvam, A.B.D. 2010. An automated or semi-automated identification system using venation pattern to delimit Indian leaf drugs: A proposal. *Pharmacognosy Research*, 2: 385–387.

Shaw, G., Nodder, F. 1799. *The Naturalist's Miscellany*. privately published, London.

Sheets, H. D., Kim, K., Mitchell, C.E. 2004. A combined landmark and outline-based approach to ontogenetic shape change in the Ordovician trilobite *Triarthrus becki*. *In:* Elewa, A.M.T. (ed.). *Morphometrics: Applications in Biology and Paleontology*. Springer Verlag, Berlin. pp. 67–82.

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

Sievwright, H., MacLeod, N. 2012. Eigensurface analysis, ecology, and modelling of morphological adaptation in the falconiform humerus (Falconiformes: Aves). *Zoological Journal of the Linnean Society*, 165: 390–415.

Sirovich, L., Kirby, M. 1987. Low-dimensional procedure for the characterization of human faces. *Journal of the Optical Society of America*, A4(3): 519–524.

Sokal, R.R., Sneath, P.A. 1963. *Principles of Numerical Taxonomy*. W. H. Freeman, San Francisco.

Smith, H., Smallwood, A., DeWitt, T. 2014. A geometric morphometric exploration of Clovis fluted-point shape variability. *In:* Smallwood, A.M., Jennings, T.A. (eds), *Clovis: On the Edge of a New Understanding.* Texas A&M University Press, College Station, Texas, pp. 161–180.

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

Sneath, P.H.A., Sokal, R.R. 1973. *Numerical Taxonomy: the Principles and Practice of Numerical Classification*. W. H. Freeman, San Francisco.

Somers, K.M. 1986. Multivariate allometry and removal of size with principal components analysis. *Systematic Zoology*, 35(3): 359–368.

Stein, M.L. 1999. *Interpolation of Spatial Data: Some Theory for Kriging*. Springer-Verlag, New York. 276 pp.

Strauss, R.E. 1985. Evolutionary allometry and variation in body form in the South American catfish genus *Corydoras* (Callichthyidae). *Systematic Zoology*, 34(4): 381–396.

Strauss, R.E., Bookstein, F.L. 1982. The truss: body form reconstructions in morphometrics. *Systematic Zoology*, 31(2): 113–135.

Tarca, A.L., Carey,, V.J., Chen, X.W., Romero, R., Drăghici, S. 2007. Machine learning and Its applications to biology. *PLoS Computational Biology*, 3(6): 953–963.

Teisser, G. 1928. Sur quelques dysharmonies de croissance des crustacés brachyures. *Comptes Rendus des Seances de la Societe de Biologie et de ses Filiales*, 99: 1934–1935.

Thompson, D.W. 1917. *On Growth and Form*. Cambridge University Press, Cambridge. 793 pp.

Thompson, D.W. 1945. *On Growth and Form*. Cambridge University Press, Cambridge. 1116 pp.

Thulman, D.K., 2012, Discriminating paleoindian point types from Florida using landmark geometric morphometric. *Journal of Archaeological Science*, 39(2012): 1599–1607.

Turk, M., Pentland, A. 1991a. Face recognition using eigenfaces. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, June 1991: 586–591.

Turk, M., Pentland, A. 1991b. Eigenfaces for recognition. *Journal of Cognitive Neurosciences*, 3(1): 71–86.

Webster, M., Sheets, H.D., Hughes, N.C. 2001. Allometric patterning in trilobite ontogeny: testing for heterochrony in *Nephrolemellus*. *In:* Zelditch, M.L. (ed.). *Beyond Heterochrony*. John Wiley & Sons, New York. pp. 105–142.

Wilf, P., Zhangb, S., Chikkerurd, S., Little, S.A., Wing, S.L., Serreb, T. 2016. Computer vision cracks the leaf code. *Proceedings of the National Academy of Science*, 113(12): 3305–3310.

Wright, S. 1921. Systems of mating. I. The biometric relations between parent and offspring. *Genetics*, 6: 111–123.

Zahn, C.T., Roskies, R.Z. 1972. Fourier descriptors for plane closed curves. *IEEE Transactions, Computers*, C-21: 269–281.

Zelditch, M.L., Fink, W.L., Swiderski. D.L. 1995. Morphometrics, homology, and phylogenetics: quantified characters as synapomorphies. *Systematic Biology*, 44: 179–189.

Zelditch, M.L., Swiderski, D.L., Sheets, H.D., Fink, W.L. 2004. *Geometric Morphometrics for Biologists: A Primer*. Elsevier/Academic Press, Amsterdam. 443 pp.

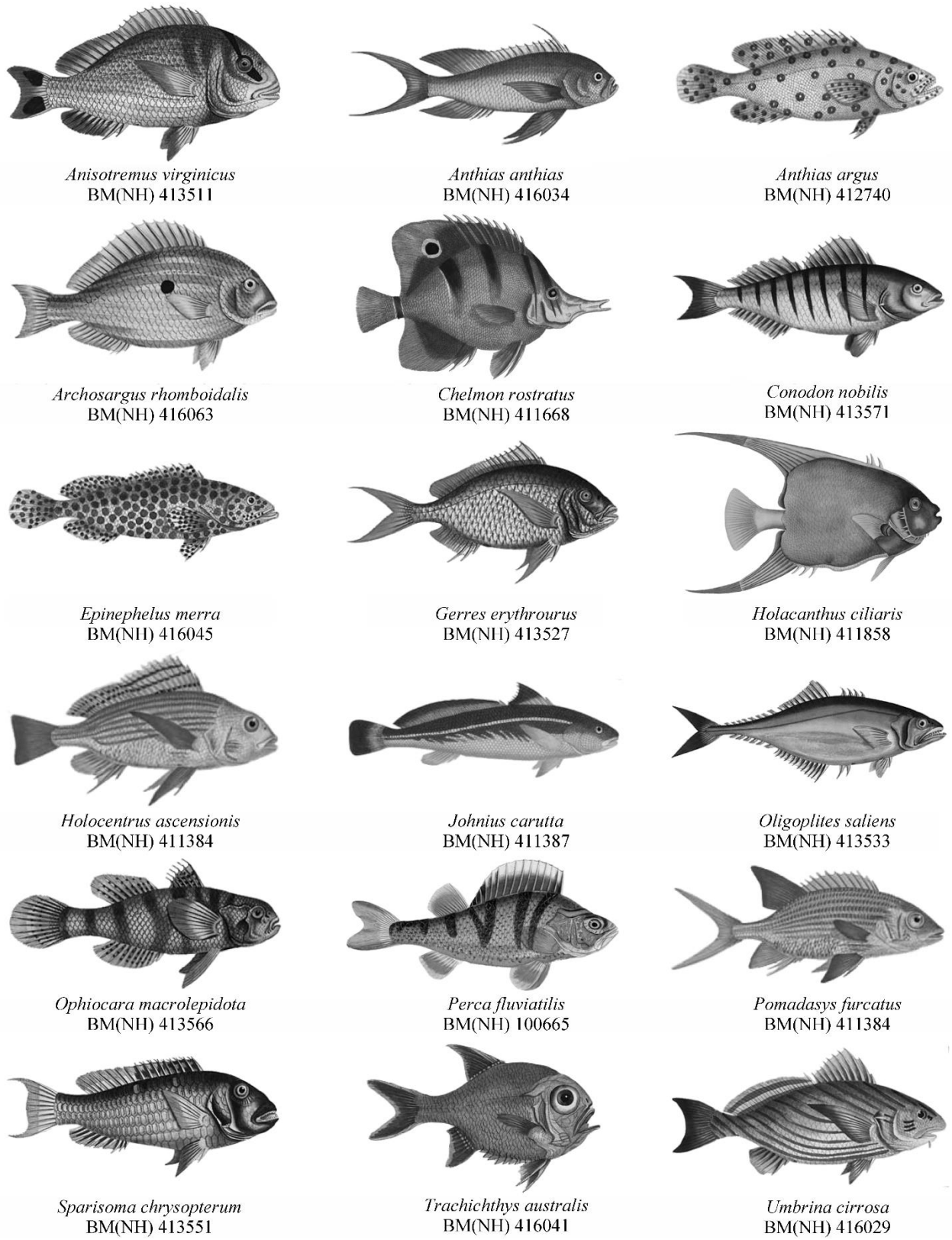

Plate 1. From Shaw and Nodder (1799): *Anisotremus virginicus, Anthias anthias*, *Anthias argus*, *Archosargus rhomboidalis*, *Chelmon rostratus*, *Conodon nobilis*, *Epinephelus merra*, *Gerres erythrourus*, *Holacanthus ciliaris*, *Oligoplites saliens*, *Ophiocara macrolepidota*, *Sparisoma chrysopterum*, *Trachichthys australis*, *Umbrina cirrosa.* From Cloquet (1830): *Holocentrus ascensionis*, *Johnius carutta*, *Pomadasys furcatus.* From Bowditch (1828–38): *Perca fluviatilis*. All images were published originally as color illustrations from live or recently caught specimens. The color has been removed from these reproductions in order to focus the image-based analyses on the patterns — as opposed to the colors — of scale patterns and body surface markings.