Chapter 7

Phylogenetic signals in morphometric data

Norman MacLeod

Department of Palaeontology
The Natural History Museum
Cromwell Road, London SW7 5BD

ABSTRACT

Although many of the goals and concepts of qualitative morphological analysis and morphometrics are similar, systematists have largely rejected the use of morphometric methods in phylogenetic analysis on a variety of grounds. This review finds that (1) the concepts of a cladistic character and a morphometric variable are essentially identical, (2) morphometric methods can be instrumental in discovering and documenting new morphological character and character states, (3) prior objections to the use of morphometric variables because of their continuous nature confuse the issues of variable type with those surrounding the distributions of sets of observations, (4) morphometrics offers the best method of determining whether morphological observations are discontinuous (= can be coded as discrete character states) or continuous (= cannot be coded as discrete character states), (5) constellations of landmark-based morphometric variables represent adequate summaries of putative structure-level homologues for use in phylogenetic analyses, (6) partial warp analyses do not perform well in either simulated or actual phylogenetic systematic analyses because of their inherent instability and lack of adequate spatial localization, and (7) a new method of subdivided relative warp analysis (described herein) performs very well at recognizing simulated morphological character states and recovering a simulated morphological phylogenetic hierarchy. Based on these results it is concluded that the potential of morphometric data analysis methods (especially relative warp-based methods) to contribute to phylogenetic-systematic investigations should be explored further.

INTRODUCTION

The fundamental observation of biology is morphology. Morphological data form the basis of virtually all systematic descriptions. Morphological features define the basic units of biology: the species and other monophyletic taxa (Nelson 1989) and are used by all biologists—including geneticists and molecular systematists—to identify those groups in the overwhelming majority of cases. Indeed, even molecular data are morphological insofar as the chemical properties that enable particular molecules to function in biological processes—and so be maintained by natural selection—derive as much from the arrangement of atoms in each molecule’s structure (= its shape) as from its composition.

Morphological data are regarded as being of significance in systematics because morphological variation is believed to be characterized by gaps between taxa. The presence of these gaps makes each taxon uniquely diagnosable and their hierarchical structure reflects action of morphological change superimposed on the evolutionary process of ancestry and descent. These gaps may arise as a result of a number of evolutionary processes (see Otte and Endler 1989 for reviews), but their discovery, description, and interpretation represents the first and most basic task of all systematic research.

Morphometrics is the study of covariances between patterns of morphological variation and patterns of variation in other associated or causal variables (Bookstein 1991; MacLeod in press). As such, morphometrics and systematic biology share a common interest in the analysis of morphology, in assessing the nature of morphological variation, and in studying degrees of covariance with those patterns (e.g., taxonomic covariances, ecological covariances, functional covariances, phylogenetic covariances). Given that morphometrics also invariably incorporates strong elements of quantification and formal hypothesis testing, it would
seem natural for biological morphologists to regard morphometric tools as an integral part of their approach to systematics. This, however, has not been the case.

The reasons for the persistent lack of a strong connection between systematics and morphometrics are many. But, for contemporary systematists I believe they can be traced back to a sense of unease within the systematics community over historical connections between the systematic philosophy of phenetics and many morphometric procedures (e.g., Crowe 1994). Indeed, it often seems as though many systematists equate morphometrics with phenetics—even though this is demonstrably not the case (see Bookstein et al. 1985; Bookstein 1994)—and regard both as being beyond the bounds of accepted systematic practice. This, perhaps unrecognized, avoidance of morphometrics by systematists is mirrored within the morphometrics community which has, for the most part, avoided taking phylogenetic patterning into consideration in their interspecific data analyses despite many recent and clear demonstrations of the need to do so (Felsenstein 1985, 1988; Harvey and Pagel 1991; MacLeod in press).

The purpose of this chapter is to explore the past, present and future of relations between systematics, and morphometrics. In keeping with the theme of the volume, this exploration will be organized around the topic of morphological phylogenetic analysis, though the methods, discussion, and conclusions drawn should be applicable to other areas of systematics (e.g., biogeography, ecology) as well. In particular, it will consider the question of how and why morphometric data should be utilized in the context of descriptive and analytic phylogenetic systematics. These explorations will take the form of both a (re)consideration of systematic concepts and practices, as well as their demonstration via example analyses. By venturing into the ‘no-man’s land’ between systematics and morphometrics this study represents a gathering together of this topic’s disparate strands in an attempt to bridge the conceptual divides that presently separate large segments of the systematics community in an area that has traditionally stood at the heart of organism-centered biological investigations.

CHARACTERS AND VARIABLES

Any discussion of the relation between contemporary systematics and morphometrics should begin by examining issues surrounding the concepts of ‘systematic characters’ and ‘morphometric variables’. Farris et al. (1970, p. 172) defined a systematic character (= the transformation series of Hennig 1966), as “a collection of mutually exclusive states which (a) have a fixed state such that (b) each state is derived directly from just one other state and (c) there is a unique state from which every other state is derived.” Pimentel and Riggins (1987, p. 201) defined a ‘character’ as “a feature of organisms that can be evaluated as a variable with two or more ordered states.” Contrast these descriptions with the standard biometric concept of a variable (Zar 1974, p. 2) as “a characteristic that varies from one biological entity to another.” Similarly, Sokal and Rohlf (1981, p. 11) define a biometric variable as “a property with respect to which individuals in a sample differ in some ascertainable way.” If one strips away parts b and c from the Farris et al. (1970) definition—which are matters of theory and interpretation—and restricts the Zar (1974) and the Sokal and Rohlf (1981) definitions to morphological characteristics or properties, it can be appreciated that the operational concepts of ‘systematic character’ and ‘morphometric variable’ are essentially identical.

There are four basic types of variables: ratio-scale variables, interval-scale variables, ordinal-scale variables, and nominal variables. Ratio-scale variables represent continuous, infinitely divisible, numerical scales in which a unit difference represents the same quantity regardless of its location along the scale (e.g., measured heights, lengths, widths). These variables are typically represented by real numbers. Meristic variables are considered a special class of ratio-scale variables that can take only discrete, integer values (e.g., no. of eggs laid in a clutch, no. of eyes, no. of digits). Interval scale variables represent continuous numerical scales in which a unit difference represents different quantities depending upon its location along the measurement scale. The classic examples of this variable type are the Fahrenheit and Centigrade (but not the Kelvin) temperature scales (e.g., 100°C is not twice as hot as 50°C) since the zero point is set arbitrarily. Circular variable scales are also of this type. Ordinal-scale and nominal variables both represent discrete measurement scales, but along these scales there is no requirement that a unit difference represents the same quantity regardless of its location along the scale. These scales differ depending on whether the attribute can be represented as a ranked (ordinal) or non-ranked (nominal) sequence. Both ratio-scale and interval-scale variables can be transformed into ordinal-scale or nominal variables via application of appropriate rules.
The characters described and discussed in the theoretical systematic literature (e.g., Kitching et al. 1998) are almost always nominal-scale variables. In addition to the G,A,T,C nominal variables of molecular phylogenetics, systematic attributes typically treated as nominal variables include such standard examples as eye color (e.g., brown, hazel, blue), hair color (e.g., brown, black, blonde), egg type (e.g., aminote, non-aminote), and body covering (e.g., scales, feathers, hair). A standard symbology for a nominal systematic variable is A, A'. The convention of naming these variables with numerical symbols (0,1,2,...) does not make them interval-scale variables because the difference criterion essential to that variable type is violated. Nominal variables may be converted to ordinal-scale variables though application of a rule (e.g., outgroup comparison) that establishes the variable's order. Hennig's (1966) original concept of the transformation series represents an example of a systematic ordinal-scale variable. However, the common contemporary practice of submitting numerically coded state variables to phylogenetic analysis in the unordered mode means that most systematic character/character-state datasets are composed of nominal variables. This data type contrasts markedly with morphometric datasets, that are almost always composed of ratio-scale variables.

I suspect it is the fundamental differences between these two variable types, along with a desire to work as close to the abstract ideal of a systematic character as possible, that many systematists have in mind when they make categorical statements like the following.

“Continuously varying quantitative data are not suitable for cladistic analysis because there is no justifiable basis for recognizing discrete states among them.” (Pimentel and Riggins 1987, p. 201).

“...it would be inadmissible to use a [morphometric] length variable that had been arbitrarily divided into to states, one of lengths less than, the other of lengths greater than the median length, since it would be just as reasonable to choose any point along such a continuum at which to delimit states.” (Crisp and Weston 1987, p. 67).

“None of the authors on coding methods has yet faced the question of how we could test for the presence of underlying discrete states. Lacking such a test, there is no reason to discretize [sic] quantitative characters.” (Felsenstein 1988, p. 462).

Such statements would be uncontroversial if each organismal phenotype presented large sets of ratio-scale and nominal variables for analysis in that one could simply identify the variable type and exclude the ratio-scale subset. In practice, however, things are a bit more complicated. Many authors have pointed out that it is a common practice for phylogenetic systematic analyses to be performed on data matrices containing quantitative, ratio-scale variables that have been semantically ‘discretized’ into nominal variables (= characters, see Simon 1983; Almeida and Bisby 1984; Thorpe 1984; Archie 1985; Baum 1988; Goldman 1988; Chappill 1989; Stussey 1990; Stevens 1991; Theile 1993; Rae et al. 1996; Swiderski et al. 1998). Moreover, even many standard examples of ‘good’ cladistic (= nominal) variables fail to stand up to even casual scrutiny.

For example, Pimentel and Riggins (1987, p. 202) refer to flower-petal colour as a cladistic (nominal) variable in their discussion of character-state concepts. Colour, of course, is really a ratio-scale variable—hence the endless list of names for slightly different colors—based on the frequency spectrum of reflected light. The fact that systematists often find it convenient to describe colour as though it were a nominal variable changes neither the nature of the phenomenon that produces colour nor the arbitrariness of the various nominal scales used to describe the visible light spectrum. Based on the fact that ratio-scale variables are routinely used as characters in contemporary phylogenetic analysis, one cannot logically exclude morphometric variables from the list of useful systematic character types or covariates simply on the basis of their continuous measurement scales.

**USING MORPHOLOGY TO DISCOVER CHARACTER STATES**

Contemporary systematists are able to treat ratio-scale and interval-scale variables as ordinal-scale or, more typically, nominal variables because they regard the pattern(s) of variation they exhibit as being dis-
continuously distributed along the theoretically continuous measurement scale. In those instances where a discontinuous distribution of size or shape measurements can be documented for a sample there can be no objection to the morphometric definition of subdistribution limits and the relabeling of the subdistributions as nominal classes. This operation conforms exactly to the recognition of discrete states within a larger character (see Eldredge and Cracraft 1980; Nelson and Platnick 1981; Wiley 1981; Smith 1994; Kitching et al. 1998). Under this conceptualization the theoretically continuous variable axis represents the character and the discontinuously distributed clusters of observations arrayed along this axis represent fixed and mutually exclusive state classes. Indeed, explicit demonstration of the distributional discontinuities on which such state boundaries are based, their formal definition, and their illustration should be required of all peer-reviewed, morphological, systematic communications.

For instance, morphological size variables (e.g., length, area) are routinely used to subdivide the variation observed between presumed homologous structures into states. Typically these subdivisions are given qualitative names (e.g., feature size: small, large) and coded as nominal states (e.g., 0,1), often without any supporting data demonstrating discontinuity or defining state boundaries. Morphometric data analysis methods were originally formulated to provide such demonstrations and definitions. In this context it is difficult to understand how recourse to morphometric variables and analyses has been so consistently ignored by systematists. The fact that these simple principles of demonstration-documentation have not been followed in
the past has been responsible for much unnecessary confusion and over the nature of phylogenetic systematic analysis and the validity of particular analytic results.

The salient aspects of this process—and the advantages of the morphometric approach—can be illustrated with a simple example. In their phylogenetic analysis of Silurian encrinurine trilobites from the central Canadian Arctic, Adrain and Edgecombe (1997) employed 40 different morphological characters drawn from six different trilobite character complexes. Of these characters—all of which were nominally coded in the standard manner—25 represented continuous, ratio-scale morphometric variables (e.g., Height of Eye: not stalked [0], very tall, ‘stalked’ [1]; Size of Transverse Tubercle Row: subdued [0], prominently expressed [1]; Depth of Doublural Notch: shallow, broad [0], deep [1]). Figure 7.1 shows tracings of the ‘Depth of Doublural Notch’ character taken from the plates provided by Adrain and Edgecombe (1997) with the character-states assigned to these morphologies marked. Obviously, there is a wide range of morphological variation inherent in this feature. Just as obviously, there are notch outlines that could be described accurately as ‘shallow, broad’ (e.g., *Avalanchurus simoni*) and others that could be described as ‘deep’ (e.g., *Struzia epsteini*). Nevertheless, a wide range of intermediate morphologies also exist that do not seem to fit clearly into either nominal class (e.g., *M. deedei, S. dimitrovi, S. onoae*).

Is the variation exhibited by these morphologies distributed continuously or discontinuously? What are the class boundaries of Adrain and Edgecombe’s (1997) character-state classes? Are the character-state assignments correct? Are other characters lurking in this feature and could be used in phylogenetic systematic contexts? Systematists must have answers to these questions—along with analogous questions for all of the other morphometric characters used in this study—if they are to evaluate the validity of the conclusions reached or to attempt to reproduce the analysis on their own, or (perhaps most importantly) attempt to extend this analysis to other groups of encrinurine trilobites in particular, or to trilobites in general. The answers to these questions are very difficult to obtain, explain, or justify through simple, qualitative inspection of these figures. [Note: Adrain and Edgecombe (1997) did not provide a table of illustrations such as Figure 7.1 for this—or any other—character, but illustrated their characters via traditional plates of photographs within which many more sources of variation were present.] Obtaining answers to these questions is perhaps the most obvious area in which morphometrics can contribute to systematics.
Figure 7.3. Results of a shape coordinates analysis of the trilobite doublural notch landmarks shown in Figure 7.2. This plot represents the distribution of \(x\) and \(y\) values of the central landmark for each representative individual after the baseline (see Figure 7.2 caption) had been rescaled to unit length and standardized orientation. Note the lack of a pronounced discontinuity among these shapes along the \(y\)-shape coordinate. Such a discontinuity would be expected given Adrain and Edgecombe’s (1997) subdivision of the doublural notch character into two discrete character states. \(S. = Struszia.\ M. = Mackenzieurus.\ A. = Avalanchurus.\)

A variety of morphometric approaches to the analysis of this dataset could be selected. Perhaps the simplest geometric analysis method would be to characterize each notch by a triangle of landmarks, with two landmarks representing the end-points of the structure and the third representing the notch’s nadir. Once coordinate values for these landmarks had been obtained the curves could be compared to one another using the Bookstein Shape Coordinate (BSCoord) method (Bookstein 1986). Bookstein Shape Coordinates are well-suited to this analytic situation in that the Adrain and Edgecombe (1997) ‘Depth of Doublural Notch’ character seems concerned primarily with the depth of the notch nadir relative to the baseline formed by the notch endpoints.

Figure 7.2 shows the landmarks points selected to represent the doublural notch character and Figure 7.3 shows the results of BSCoord analysis. The distribution of the free coordinate (= notch nadir) suggests that, for this small dataset, shape variation is continuous from the shallowest to the deepest notch profiles. Given these results there seems no obvious morphological discontinuity at which to place a line of definition between Adrain and Edgecombe’s (1997) depth of doublural notch state classes.

In addition to raising questions about the definition of Adrain and Edgecombe’s character states, these results cast doubt on their character-state assignments. Although the four species they assigned to the ‘shallow, broad’ character-state class do represent the shallowest curves in this dataset, the variation exhibited by these curves along the \(y\)-shape coordinate axis (= curve depth sensu stricto) is greater or equal to the variation between the lower limit of this provisional subgroup and the next grouping of shape coordinates. Based on these results it is difficult to understand why “Struszia mccartneyi”’s notch is described as ‘shallow, broad’ while Struszia martini’s and Mackenzieurus deedei’s notches are described as ‘deep’.

Note that this BSCoord-based analysis reduced these somewhat complex curves to triangles of landmark points (Figure 7.2). In some cases the underlying morphology is indeed triangular (e.g., A. simoni, S. martini) and so would not suffer by representation as a geometric triangle. For the majority of the curves in this small dataset, though, representation as a triangle severely distorts the true nature of the morphology that is being used in the qualitative assessment of their variation.
Figure 7.4. Eigenshape results for the doubletural notch curves shown in Figure 7.1. A. Ordination of notch curves within the plane formed by eigenshapes 1 and 2. B. Ordination of notch curves within the plane formed by eigenshapes 2 and 3. By taking the geometry of the entire notch into consideration geometric discontinuities between the central cluster of morphologies and *Bilevittia admira*, *Mackenzieurus ceejayi*, and *Avalanchurus simoni* are evident along Eigenshape 1. A similar shape discontinuity exists between the two *Mackenzieurus deedeei* specimens and the remaining shapes along Eigenshape 2. Qualitative assessments of discontinuities such as these form the basis for traditional systematic character-state assignments. Morphometrics methods, such as eigenshape analysis, are available to help systematists evaluate observed shape distributions. Once such results have helped ‘sharpen the eyes’ of systematists they would be free to re-evaluate shape distributions in a traditional qualitative manner, or to explicitly use results like these to make character-state assignments. See text for additional discussion.

In addition, to these considerations the reliability of the biological correspondence with which the notch nadir landmark has been placed on the curve is open to question. While various geometric criteria can be advanced to help guide the placement of this landmark, none of these geometric criteria have anything necessarily to do with the underlying biological processes responsible for the notch structure. In the absence of much more biological information about the notch, the physio-chemical processes responsible for its formation, its comparative ontogeny, etc., it is questionable whether this third coordinate represents the same type of observation represented by the baseline landmark coordinates. Given this inherent biological uncertainty, coupled with the distortions imposed on the morphological system as a result of abstraction to just three landmarks, one might suspect that the negative results obtained by the BSCoord analysis may have as much to do with how the morphological variation was measured as with the nature of geometric variation among the original notch curves. Consequently, it seems reasonable to employ an alternative method of morphometric analysis in order to (1) test the BSCoord results for robustness to examination by different geometric methods and (2) determine whether the addition of geometric data to the measurement system supports refinements to the Adrain and Edgecombe (1997) system of morphological descriptors.

Once again, there are a variety of methods that could be employed to analyze these curves and draw conclusions about the manner in which their shapes are distributed. One available candidate is open-curve eigenshape analysis (MacLeod 1999). This method is similar to relative warp analysis (Bookstein 1991), but draws a distinction between landmarks and semi-landmarks (Bookstein 1997). In essence, open-curve eigenshape analysis uses formal landmarks to define the endpoints of the curve and (if necessary) subdivide the curve into segments. These segments are then represented by semi-landmarks that are accorded a sequence-level correspondence to one another between shapes within the sample.

It is important to note that there is no necessary implication of biologically homologous correspondence between semi-landmarks. The nature of their correspondence resides at the level of geometry and sequence order only. This type of correspondence is justified when it is the only level of correspondence assessment available on which to base morphological comparisons. Sequence-level correspondence is no different in principle from the qualitative assessments of between-curve-segment correspondence in the absence of additional biological information that are made routinely by systematists and that have been accepted as a basis for the description-comparison of morphological attributes for centuries. If one can base comparisons on point-to-point correspondences that have some larger biological-phylogenetic significance one should do so (but see discussion of morphological homology below). In the absence of such biological information,
Figure 7.5. Geometric models of the first three eigenshape axes shown in Figure 7.4. A. Model sequences computed at designated locations along the eigenshape 1, 2, and 3 axes. B. Along-axis or superposition-style representations of model sequences. Using these figures it is clear that the shape contrast represented along Eigenshape 1 corresponds to Adrain and Edgecombe’s (1997) qualitative states ‘shallow, broad’ and ‘deep’. Eigenshape 2 corresponds to a contrast between U-shapes and V-shaped notch geometries (a distinction missed by Adrain and Edgecombe, 1997). Eigenshape 3 corresponds to a contrast between curve symmetry or irregularity in the region of the nadir. Models such as these can aid in the interpretation of abstract morphometric ordinations—and in the recognition of new characters or character states—by portraying coordinate positions within the shape space as geometric figures that can be compared in a traditional, qualitative manner.

morphologies can still be quantitatively compared—and shape-distribution hypotheses tested—using the sequence-level correspondences inherent in the concept of semi-landmarks (Bookstein 1997).
In the present example the fifteen doublural notch outlines shown in Figure 7.1 were digitized and the two end-point coordinates designated as biological landmarks. These semi-landmark-defined outlines were then converted to Zahn and Roskies (1972) angular-deviation shape functions and submitted to a singular value decomposition in the manner that has become standard for eigenshape-based methods (see MacLeod 1999 and references therein). A plot of the notch-curve scores on (= covariances with) the first two eigenshape axes (Figure 7.4) suggests that discontinuities do exist in the distribution of shapes within this dataset. In particular, A. simoni, M. ceejayi, and Billeevittia adraini exhibit shape scores along the first eigenshape axis (ES-1) that are atypically low relative to the remainder of the dataset. Since Adrain and Edgecombe (1997) assigned these three species to their ‘broad, shallow’ character-state class, this grouping is not unexpected. However, “S”. mccartneyi occupies a position along the ES-1 axis on the other side of the morphological discontinuity delimited by A. simoni, M. ceejayi, and B. adraini. Since “S”. mccartneyi was also characterized as exhibiting a ‘shallow, broad’ doublural notch by Adrain and Edgecombe (1997), these eigenshape-based morphometric results are inconsistent with their qualitative assessment of morphological variation. Similarly, the distribution of shapes along the second eigenshape axis (ES-2) suggests that shape variation within the two M. deedeei specimens also exhibits a morphological discontinuity with respect to the remainder of the sample. This distinction in doublural notch morphology was not noticed by Adrain and Edgecombe (1997).

To obtain a better understanding of the types of morphological contrasts implicit in Figure 7.4, morphological models of ‘pure’ shape variation along these two eigenshape axes can be obtained through modeling (Figure 7.5, see MacLeod 1999 for an explanation of the modeling method). The series of shape models determined for the ES-1 axis confirm Adrain and Edgecombe’s (1997) original distinction between shallow and broad curves (= low scores) and narrow deep curves (= intermediate and high scores). But, contrary to the results of Adrain and Edgecombe’s (1997) qualitative analysis, this is not the only shape-variation mode that has potential for defining systematically interesting morphological discontinuities. The series of models determined for the (ES-2) suggest that the contrast between V-shaped notch outlines (= low and intermediate scores) and U-shapes notch outlines (= high scores) also has potential for characterizing M. deedeei in a manner consistent with traditional qualitative analyses. Inspection of Figure 7.1 after viewing these results shows that the two M. deedeei notch outlines do indeed exhibit characteristically more U-shaped notch profiles—much more steeply sided with broader, flatter bottoms—than the remainder of the dataset.

In this example, morphological character states were defined on the basis of pronounced gaps or discontinuities in the distribution of shapes within a sample of trilobite pygydia. It is important to note that even though the morphometric variables used to ordinate these shapes were continuous, ratio-scale variables, the eigenshape analysis revealed two different discontinuities within the distribution of observations along these continuous variables and that these quantitative gaps were later confirmed by a qualitative (re)inspection of the shapes. Consequently, the specification and analysis of continuous morphometric variables facilitated the recognition of discontinuous patterns of shape variation that could be retrospectively coded by the traditional method of qualitative inspection. This is consistent with the descriptions of standard contemporary systematic procedure (e.g., Farris et al. 1970; Pimentel and Riggins 1987; Felsenstein 1988).

Based on this example analysis it seems hard to argue that morphometric data analysis can, in principle, play no useful role in the systematic study of biotic character-state data. Indeed, these results suggest that morphometric approaches can represent nothing more than a direct—albeit more sophisticated—extension of traditional qualitative morphological analysis methods that simply draw the systematist’s eye to patterns that he or she might not have immediately recognized. Once the presence of such discontinuities within the sample has been made obvious by the morphometric analysis, they can be readily understood and utilized by more traditional qualitative inspection. Similar conclusions have been reached by a number of other systematists (Theile and Ladiges 1988; Theile 1993; Fink and Zelditch 1995; Zelditch et al. 1995; Rohlf 1998; Swiderski et al. 1998; Rae 1998), though none has provided a simple example of this process. [Note: see also Chappill 1989 and Theile 1993 for discussions of semantic confusion between the notions of continuous variables and discontinuous patterns of variation that often occur in the systematic literature.]

CONTINUITY: AXES VS. OBSERVATIONS

Obviously, not all patterns of variation between putative taxonomic groups will be characterized by pronounced discontinuities. For instance, morphological characters that are the result of polygenic suites will
Figure 7.6. Conceptual diagram illustrating the difference between continuous variables and continuous distributions of variable observations. A and B represent sets of frequency distributions along a continuous morphometric variable axis for discontinuously distributed (A) and continuously distributed (B) sets of observations. These observations have been gathered into two putative taxonomic groups, or phena. The gap in A represents a morphological discontinuity that serves to distinguish the phena from one another objectively and independently and that may be used to recognize one, or the other, or both as a monophyletic group. This situation poses no theoretical or practical problems in terms of character recognition of character-state coding, as evidenced by the widespread inclusion of such characters in systematic datasets (Scheme 1). The lack of an inter-phenal gap in B presents several systematic problems. Continuously distributed observations cannot be used to objectively or independently distinguish the phena from one another, because the limiting criterion is irredidually dependent on the recognition of one of the other phena, which are presumably diagnosed on the basis of other variables. Various schemes have been developed to arbitrarily subdivide such continua (e.g., gap coding, interval coding, see schemes 2 and 3), but all are dependent on a priori group recognition to work. Since character states defined in these ‘member-coding’ schemes cannot objectively and independently delineate phena they are of little use in phylogenetic systematics.

exhibit a progressive shift in the mean value for a population under directional selection; even if the selection pressure is intense (Falconer 1981; Felsenstein 1988). While the continuous patterns of variation characteristic of anagenetic evolution should not be confused with the morphological discontinuities implied by cladogenesis (Zelditch et al. 1995), it is the case that both modes of evolutionary change may be character-
ized by zone of between-group morphological overlap that are determined (at least in part) by group-

This observation has been used by a number of systematists to argue that continuously-distributed variables should be used, together with discontinuously-distributed variables, to reconstruct phylogenetic patterns. However, continuously-distributed ratio-scale or interval-scale variables cannot be transformed into the nominal or sometimes ordinal variables used by most phylogenetic inference algorithms except through the application of ad hoc rules. Gap coding (Mickevich and Johnson 1976; Archie 1985), segment coding (Colless 1980; Thorpe 1984; Chappill 1989) statistical ‘difference between means’ tests (e.g., homogeneous subset coding, Farris 1990; Theile 1993; Rae 1998); and statistical ‘overlap analysis’ (Almedia and Bisby 1984; Swiderski et al. 1998) all represent attempts to devise rule-based systems for subdividing continuous patterns of variation in morphological variables based on a consideration of frequency and/or a priori group membership.

The difficulty with these rule-based methods is that they—sometimes to a greater extent, sometimes to a lesser—alter the nature of the variable description from that of the individual to that of the group. Figure 7.6 illustrates this problem. Figure 7.6A shows the typical situation for a morphological character measured on a continuous, ratio scale. All values along the scale are possible. Above this scale the hypothetical frequency distribution of a set of morphometric measurements is illustrated. In this example the overall distribution is discontinuous with two well-separated subgroups evident along the continuous variable axis. Using morphometric methods it is a simple matter to determine the observed range of separation between the subdistributions (= discovery phase of analysis). Once this is accomplished, the morphological discontinuity discovered by the systematist can be used to define the range of variation for putative state classes of the morphological variable (Figure 7.6A, Scheme 1). Individuals can then be assigned to these state classes without needing to take putative group membership into consideration. As a result, the variable axis (= character) is objectively subdivided into segments (= state classes) by the observations with taxic groups emerging as a result, rather than as an assumption, of the analysis. Character states discovered and described in this manner will be genuinely independent in that they make no a priori reference to group membership.

Contrast this with the situation in Figure 7.6B which shows the same morphological scale above which a the frequency histogram for an alternative set of morphological observations is illustrated. In this case, the distribution is continuous and bimodal along the measurement scale. Since there is no morphological discontinuity that can be used to recognize and define phenotypic states unambiguously, partitioning must be made on the basis of group membership. Either the range of the overlap between putative phena is designated as a separate state from the non-overlapping ranges of the distribution (e.g., Archie 1985) or the means of the putative phena are tested for statistically significant separation and the entire pheno assigned to a state on the basis of the test results (e.g., Theile 1993; Rae 1998). This practice, which lies at the heart of all gap coding, gap weighting, segment coding, subset coding, overlap analysis, etc. procedures, might be described by the generic term ‘member coding’.

Member coding leads to the production of questionable character-state definitions. For example, since the definition of the group must be based on other criteria (presumably discontinuously distributed character states), the member-coded character state cannot be regarded as an independent descriptor of morphological variation. Logical consistency is also challenged by member coding. Member-coded character-state definitions cannot be recognized as ‘features of organisms’ because some proportion of the member specimens exhibit morphologies that are dependently distinguished from (Figure 7.6B, Scheme 2) or identical to (Figure 7.6B, Scheme 3) to those of other groups. Such procedures can even result in group members receiving a single character-state code as if all members of the group were morphologically distinct from all members of another member coded group regardless of whether this disjunction was actually observed. Member-coded character states codify assessments of group-based frequency trends or probability statements that are dependent on group diagnosis rather than objective and independent morphological descriptors. As such they represent a fundamentally different type of quasi-morphological variable that should not be mixed with real morphological variables (whether continuously distributed or discontinuously distributed) in either phylogenetic or morphometric analyses unless appropriate measures are implemented to take the introduction of such mixed-mode data into consideration.

Rather that spending time trying to rationalize the arbitrary subdivision of continuous morphological data so that these can be used in phylogenetic systematic analyses, systematists would be better advised to ac-
quaint themselves with tools of morphometrics since the use of these tools will greatly improve their chances of recognizing and correctly interpreting the morphological discontinuities that are present in their data. The doublural notch analysis discussed in the previous section contains a practical example of this situation. Despite the fact that Adrain and Edgecombe’s (1997) qualitative analysis failed to recognize the distinction between V-shaped and U-shaped notch morphologies, application of a more generalized morphometric procedure was able to reveal its presence in those data. The addition of a new character and set of discontinuous character-state distinctions—such as can be achieved by morphometric methods—is of far more use to systematics than the inconsistent and logically suspect results of member-coding procedures.

DISTANCES, LANDMARKS, AND HOMOLOGY

While morphometric variables cannot be rejected for systematic study because of their type—they are no different in this respect from traditional systematic characters, cladistic characters, statements or systematic keys—there may be other problems with this class of morphological descriptors that would limit their utility in some systematic contexts. Pimentel and Riggins (1987, p. 201) argued that in addition to being able to be ordered and independent, true systematic characters must be “homologous expressions of a feature found in the ingroup and outgroup”.

Morphometric variables have long been criticized as being intrinsically non-homologous representations of form on several different grounds. The most common of these is an objection to the reification of (usually multivariate) morphometric variables as organismal ‘attributes’ in any meaningful biological sense of that term (Pimentel and Riggins 1987). This criticism confuses the methodology of principal components (= eigenanalysis) with the data to which the method has been applied.

Until recently the most common type of morphometric data was linear distances between pairs of landmark points (see Blackith and Reyment 1971; Reyment et al. 1984; Bookstein et al. 1985; Reyment 1991). Eigenanalysis of covariance or correlation matrices derived from such data yield variables that are differentially weighted amalgamations of scalar magnitudes. Pimentel and Riggins (1987) point out correctly that such amalgamations do not correspond to the concept of biological homology, but this deficiency arises as a result of the lack of topological information in the scalar distance matrix rather than arising out of the eigenanalytic procedure itself. For example, if the necessary topological information is restored to the analytical system by keeping track of the relative orientations of the distance variables (e.g., the ‘truss analysis’ method of Strauss and Bookstein 1982; see also Bookstein et al. 1985) the results of such an analysis can be recombined into a model of morphological deformation that exhibits the properties of topological correspondence and spatial localization required of biological homologues (MacLeod in press). Such topologically-informed eigenanalysis-based procedures could be used, in principle, to identify and interpret discontinuities in shape distributions in a manner conceptually similar to qualitative morphological analysis procedures. Thus, the problem Pimentel and Riggins (1987) refer to in their criticism of morphometrically-defined characters is a problem of the manner in which morphological variation had been portrayed up to that point in time (as topology-free scalar magnitudes) rather than a problem that arises from the eigenanalytic methods used to summarize patterns of variation within those data.

At approximately the same time that Pimentel and Riggins (1987) made their criticism several morphometricians independently began to recognize and understand the origins of the same problem: that the absence of topological information from morphometric datasets severely constrained the interpretability of their analytic results. From 1986 through 1991 F. L. Bookstein, C. Goodall, D. G. Kendall, F. J. Rohlf, and others effectively synthesized and reformulated several disparate data and method-based schools or morphometric analysis into a single, unified ‘geometric morphometrics’ with topology at its center. This synthesis was achieved by refocusing attention on the coordinate positions of landmark points scattered over a structure and regarding deformations of those using those coordinate point constellations as morphometric variables. Operationally, this reformulation of morphometrics involves the submission of landmark constellations—variously adjusted to remove the effects of size and differential orientation—as input into (for the most part) eigenanalysis-based procedures. The results of such analyses produced mathematically elegant summaries of geometrical deformation patterns that could be expressed in either the abstract notation of mathematics (e.g., data matrices, deformation grids, scatterplots) or the geometric representation of morphological variation traditional in qualitative systematics (e.g., Bookstein 1991; MacLeod 1999). This was
possible because the magnitudes of Cartesian coordinates preserve information about the relative amount and directions—the topology—of landmark displacements between forms.

Since topological information has now been inextricably embedded into the corpus of geometric morphometrics, a reconsideration of Pimentel and Riggin’s (1987) claim that morphometric variables cannot express biological homology is necessary. In my view, this question has two aspects, (1) whether the constellations of landmark points defined on the basis of biological structures are homologous to other constellations of landmark points similarly defined (= structure-level homology), and (2) whether individual landmark points are homologous with other landmark points (= point-level homology).

Unfortunately, the morphometric literature makes little distinction between the concepts of geometric homology and biological homology. Landmarks are defined as relocatable coordinate positions on an object in a two-dimensional or three-dimensional Euclidean measurement space (Bookstein 1991, MacLeod in press). Since geometrical homology is defined on the basis of topological correspondence, corresponding landmarks are, by definition, geometrical homologues.

Biological homology begins with topological (= geometrical) homology, but extends its concept to embrace aspects of history and origin. Although the concept of the homologue was known to Aristotle (who used it to infer correctly that porpoises were more closely related to mammals than to sharks), Richard Owen (1843, p. 374) is responsible for the concept’s canonical formulation as “the same organ in different animals under every variety of form and function”. Darwin (1859) regarded his theory of common descent as providing a biological explanation for the difference between homologous and analogous structures, after which the former was re-defined as ‘similarity due to common ancestry’. Unfortunately this formulation led to the concept’s use in two different senses: (1) a transcendental or transformational sense as a sequence of idealistic modifications (e.g., fish jaw bones changing into mammalian ear ossicles) and (2) a taxonomic or taxic sense (e.g., tetrapods being recognized as a monophyletic lineage based the fact that all members exhibit four limbs, except for those in which the limb number has been reduced due to secondary loss). Patterson (1982) pointed out this duality, rejected the transformationalist conceptualization as being non-falsifiable, and equated taxic homology with the Hennigian concept of synapomorphy.

Thus, while the triangular dorsal fins of sharks, porpoises, and goldfish might be regarded as geometrically homologous, these structures are not biologically homologous because the implied taxic grouping is not supported by other morphological characteristics (e.g., developmental patterns, skeletal characteristics, soft anatomy). This stands in contrast to the popular description of these three organismal groups as belonging to the group ‘fish’ which is a phenetically-defined morphological concept based entirely on topological similarities in gross external morphology.

Since the concept of biological homology is logically tied to organic structures and incorporates the notion of topological similarity (Reippel 1980, 1994; Patterson 1982) morphometric variables that assess aspects of topological similarity between those structures can be used to delimit taxic groups in a manner consonant with the strictures of biological homology. This is nothing more, or less, than what is done in the qualitative assessment of morphological variation patterns within or between groups of organisms. Contra Zelditch et al. (1995), morphometric decompositions of landmark constellations are not biologically homologous by definition because (1) morphometrics measures topological similarity and topological similarity is only one aspect of biological homology, and (2) the taxic groups recognized by a morphometrical analysis of one structure may be falsified by other morphometric or qualitative results for other structures (this is also true of ‘standard’ systematic characters). Caution must also be exercised not to confuse the failure of a particular morphometric result to recognize a distinction between complex morphological structures with a failure of such a distinction to exist—morphometrics usually assesses patterns of topological similarity and difference between aspects of organisms shape, not the organisms themselves; see doublural notch example above). Nevertheless, the overall similarity between morphometric and qualitative procedures of analysis at the level of morphological structures (structure-level homology) seems clear and consonant.

The idea that individual landmark points represent biological homologues is logically separate from the issue of structure-level homology. Landmark points were originally described as ‘homologous’ in order to distinguish them from the geometrically-constructed boundary point locations used in most forms of morphometric outline analysis (Bookstein et al. 1986; Bookstein 1990, 1991). Bookstein (1991) identified three classes of biological landmarks: discrete juxtapositions of structures or tissues (Type 1), maxima of curva-
ture (Type 2), or extrema (Type 3). This classification focuses attention on the amount of information necessary to identify or relocate the landmark.

Type 1 landmarks may occur at any point on or within a form so long as that form is composed of different structures or tissue types. While these landmarks are constrained to exist on the boundaries (= outlines) of structural components or tissue-defined regions, their locations are not determined by any characteristics of the overall boundary or outline. Type 2 landmarks lie on the boundaries of single structures or regions and are defined by the nature of the curving surface of that boundary. Type 3 landmarks represent those coordinate locations on single structures (irrespective of whether the structure is composed of various substructures or regions) that represent the extremes of the structure’s boundaries. Like Type 2 landmarks these points are constrained to lie on the object’s outline.

No consideration has been traditionally given to the nature of any substructure or tissue when locating Type 3 landmarks. Their definition is dependent on the nature of the outline (= by the distribution of adjacent boundary coordinates), on the orientation of the object, and on the number of axes one wishes to locate extreme along. Because the nature of Type 3 landmarks is so variable and dependent on such a wide variety of conditions Bookstein (1997) revised his 1991 classification and termed this class of landmarks ‘semilandmarks.’ The category semi-landmarks includes the former Type 3 landmarks of Bookstein (1991) as well as the boundary coordinates used in outline morphometrics (e.g., Fourier analysis, eigenshape analysis, edgels).

Bookstein’s (1997) revised landmark classification describes the range of landmark-based morphometric observations more comprehensively and recognizes fundamental similarities between observational types more consistently. Since the newer landmark taxonomy abandons the older distinction between individual landmarks and boundary coordinates—which formed the rationale for labeling the former as ‘homologous’ in order to distinguish them from the latter—this appellation no longer serves any purpose. More importantly though, Bookstein’s (1997) revised landmark classification recognizes the fundamental unity of all landmark types as relocatable points that correspond across specimens in a geometrical sense. In other words, individual landmarks represent geometrical homologues. But, the ability to represent and summarize topological patterns among biologically homologous structures does not render corresponding landmarks themselves biologically homologous. To make such a conceptual leap is to confuse the idiosyncrasies of a representative with the characteristics of the group being represented.

In order to appreciate this distinction, in your mind’s eye visualize a familiar morphological feature. While it is acceptable to describe the alternative forms of a radius bone, a canine tooth, a genal spine, or a pectoral fin as ‘long’ or ‘short’, ‘elliptical’ or ‘subquadrate’, ‘pointed’ or ‘blunt’, etc.—because any reasonable set of morphometric measurements derived from sets of landmarks located on these objects exhibit non-overlapping distributions—it is quite a different matter to claim that the ‘geometrical midpoints’ or the ‘proximal and distal termini’ of differently shaped bones or teeth, or spines, etc. correspond with the ‘geometrical midpoints’ or the ‘proximal and distal termini’ of differently shaped bones or teeth, or spines, etc. of other specimens in any save a topological sense. Within reasonable limits it is irrelevant whether corresponding landmark points fall on precisely the same point of a feature because the level of topological similarity required by this decision is not part of the biological homology concept. Since, in the great majority of instances, a wide variety of alternative landmark pairs can pass the similarity, conjunction, and congruence tests of biological homologues (see below for an example), true homology—for there can only be one pair of landmark points that define a length on any biological structure that are biological homologues of another pair of landmark points on another homologous structure—cannot be separated from false homology at the level of the mathematical point by the tests available to systematists. Consequently, the entire question of ‘biologically homologous landmarks’ is moot. Wagner (1994) points out that cases may exist in which the concepts of biological and topological homology coincide (e.g., point-intersections between three skull bones representing a Type 1 landmark). However, such situations are conjectural at present and represent a distinct minority of landmarks currently used for morphometrical analyses. In the absence of highly-detailed developmental and phylogenetic evidence, the notion of mathematical ‘point homologies’ will likely remain either an assumption or an assertion for the foreseeable future.

Bookstein (1991) described another type of relation that bears on the issue of homology in morphometrics; deformational homology. Tracing the origins of this concept to Thompson (1917), deformational homology begins with a series of geometrically-homologous point-to-point mappings on two forms and postulates
sets of smooth deformations implied by a comparison of the forms. Often these deformations can be described by a single or a series of generalized deformational type(s) (e.g., pure inhomogeneous, quadratic, rigid motion involving several landmarks, spiral deformation; see Bookstein 1991 for examples). Unfortunately, Bookstein’s (1991) discussion is unclear as to whether he was referring to geometric or biological homology in advancing this concept of Thompsonian deformational homology. Thompson (1917) believed his deformational types (even though he never referred to them in those terms) resulted from the operation of basic physical laws. However, Thompson refrained from discussing his concept in evolutionary-phylogenetic terms because he rejected Darwin’s theory (see Mayr 1982).

Regardless of Thompson’s opinions on evolutionary theory, his concept of deformational homology underpins much of the ‘morphometric synthesis’ (Bookstein 1993) because the language of deformations is useful in summarizing and interpreting the results of geometrical morphometric data analyses. It is this concept of deformational homology, however, that Pimentel and Riggins (1987) implicitly refer to when they criticize morphometric variations as being capable of representing only transformational homology. However, as pointed out by Bookstein (1994), Rohlf (1998) and MacLeod (in press) this metaphorical linkage between the morphometric-geometric ‘language of deformations’ and transformational homology diverts uncautious readers from the main point of systematic morphometrics. The deformational graphic methods used to portray geometric morphometric results (e.g., thin plate splines) are only illustrative conventions; useful for visualizing geometric relationships. Just as written descriptions of transformational homologies can be recast rhetorically as statements of taxic homology (Zelditch et al. 1995), so too can ratio-scale morphometric variables—even complex, multivariate variables—be used to quantitatively define taxic groups on the basis of topological similarities or differences among a priori-defined putative homologous features. The relevant questions, then, are not whether individual landmarks can be declared homologous (they cannot), whether morphometric variables can be used to recognize groups of taxa on the basis of shared topological similarity between putative homologues (the evidence for this is quite overwhelming as demonstrated by over a century of morphometric analyses), nor whether morphometric variables represent some quality different from what is typically represented by a large number of qualitative morphological characters (they do not). Rather the relevant questions are (1) whether morphometrically-defined variables exhibit a hierarchical structure that can be logically represented on a cladogram and used to define congruently nested sets of taxa and (2) whether the taxic groups defined on the basis of morphometric analysis agree with groupings defined on the basis of more traditional morphological analyses.

PARTIAL WARP VARIABLES AS PHYLOGENETIC CHARACTERS: A TEST OF CONGRUENCE

Extending from their arguments regarding the homology of landmarks, Zelditch et al. (1995, see also Fink and Zelditch 1995) have recently advocated the use of partial warp-based morphological-deformation variables in phylogenetic analysis. Partial warps are calculated from principal warps, which are eigenvectors of the bending-energy matrix that express the ways a reference configuration of landmarks can be geometrically deformed (Bookstein 1991). The partial warp scores are computed by projecting the values of a Procrustes-aligned set of landmarks separately onto each of the principal warp vectors.

Zelditch et al. (1995) preferred the partial warps approach to all other multivariate methods of morphometric data analysis for systematic studies because they believe such summaries to represent non-arbitrary and spatially-localized features of a geometric-biological system that are unique to individual organisms (as opposed to being arbitrary summaries of populations or samples), and that variables so-defined can be used to recognized morphological characters and define character states in a manner that supports their analysis within a hierarchical data-analysis system (e.g., those used for phylogenetic inference). The primary problems these authors sought to solve by advocating the use of partial warp variables in systematic-phylogenetic contexts were (1) improved ways of discovering of new morphological characters to be used in phylogenetic analysis and (2) demonstrating that quantitative morphological characters were no different from qualitative morphological characters in systematic contexts.

The Zelditch et al. (1995) arguments have not been met with widespread agreement within the morphometrics or systematics communities. Bookstein (1994) questioned whether any morphometrically-defined shape variables could unambiguously order shape transformations. The Shape Nonmonotonicity Theorem shows that deformation-based shape variables can be created to support any ordering of end-member shapes.
Moreover, Bookstein (1994) argued that geometric deformations per se could not be used as systematic characters because they lack the property of commutativity. Fink and Zelditch (1995) and Zelditch et al. (1998), in very brief responses to the Bookstein (1994) article did not dispute any of his geometrical arguments, but attributed Bookstein’s disagreements with them over the use of morphometrical variables in systematics to disagreements over ‘semantics’.

Lynch et al. (1996) attempted to use the method of Zelditch et al. (1995) to create systematic characters, but were uncomfortable with the results. These authors recommended that simulation studies be used to validate the Zelditch et al. (1995) method. Naylor (1996) conducted such a study using a simulated fish phylogeny that was reported to contain no homoplasies and was based on only a single morphological character-state change per branch. His results showed that although parsimony-based analysis of the entire multistate-coded, partial-warp dataset did recover the correct tree topology, these morphometrically-defined characters exhibited an extraordinarily high degree of homoplasy (RI = 0.48). In addition, Naylor’s results indicated that none of the known character-state transformations—most of which could have been captured easily by traditional, qualitative analysis—were represented in the character-state matrix based on partial warp scores. Zelditch et al. (1998) dismissed Naylor’s (1996) results claiming that he used a different method form the one they proposed. [Note: since details of the methods advocated by Zelditch and Fink have differed in different studies (e.g., Fink and Zelditch 1995; Zelditch and Fink 1995; Zelditch et al. 1995) the differences Zelditch et al. (1998) are alluding to is unclear. I have been unable to recognize any substantive difference between the method described by Zelditch et al. (1995) and the one used by Naylor (1996).]

Rohlf (1998) also criticized several practical aspects of the Zelditch et al. (1995) method. These include (1) the arbitrariness in the Zelditch et al. (1995, see also Fink and Zelditch 1995) advocacy of using a single individual exhibiting an extreme landmark configuration (e.g., representative of an outgroup in the case of phylogenetic studies, and early developmental stage in the case of ontogenetic studies) rather than a Procrustes mean landmark configuration as the basis (= tangent point) for the principal warp decomposition; (2) the arbitrariness of using a method that makes no reference to patterns of shape variation present within a sample as a basis for summarizing shape patterns of shape variation within the sample; (3) the well-known sensitivity of the partial warps method to changes in the reference landmark configuration (Bookstein 1991); (4) the notion that partial warp decompositions are uniquely spatially localized (see also MacLeod in press); (5) the inevitability of even simple morphological changes being partialled out into a complex of distinct morphometric variables with consequent loss of interpretability; and (6) the unsuitability of partial warp-based results for subsequent statistical analysis due to lack of independence. Zelditch et al. (1998) did not dispute any of Rohlf’s (1998) geometric arguments, but appealed repeatedly to non-specific, putative failures on Rohlf’s (1998) part to understand their ‘biological logic’, ‘biological interpretation’, or ‘biological reasoning’. Variations on these criticisms have also been voiced by Adams and Rosenberg (1998) with a response by Zelditch and Fink (1998).

In order to continue the evaluation of the Zelditch et al. (1995) partial warp method, the comparative strategy was used to determine the level of congruence between a tree obtained from traditional, qualitatively-defined, morphological characters and one based on morphometrically-defined, partial warp characters with both datasets being obtained from a small sample of real organisms. Previous partial warp-based phylogenetic studies of real organisms have either mapped a few character-state transitions onto independently justified cladograms (e.g., Zelditch et al. 1995) or included them along with other qualitatively-assessed morphological data in a single analysis (e.g., Fink and Zelditch 1995). While former cannot distinguish between consistent and coincidental patterns of correspondence between coded morphological and morphometrical variables (Rohlf 1998), the latter cannot measure the unique contribution of the morphometric subset or support direct comparisons between alternative summaries of morphological variation. The more rigorous approach of Naylor (1996), which is similar to the approach used by Cranston and Humphries (1988) to evaluate the contribution of quantitative characters to tree resolution, accomplishes both of these tasks and has not, to my knowledge, been attempted previously with this type of morphometric data. This approach also avoids the simplicity and artificiality of simulations.

Patterns of hierarchical character-state variation were compared for two datasets collected from images 13 trilobite encrinurine species illustrated by Adrain and Edgecombe (1997, see Figure 7.7). These authors conducted a traditional, morphology-based parsimony analysis on a larger group of encrinurine trilobites, of which these 13 species form a subset. Between 33% and 71% of the 40 characters assigned to the character complexes used by Adrain and Edgecombe (1997) to infer phylogenetic relations among these trilobites
Figure 7.7. Encrinurine trilobite cranial morphology (upper figure, right), landmarks (upper figure, left) and morphological variation for 12 representative species.

were morphometric variables whose axes had been semantically subdivided and described-defined as nominal
Figure 7.8. A. Strict consensus tree for the two equally parsimonious cladograms that resulted from a branch-and-bound analysis (equal character weighting) of the 40 morphological characters coded for these 13 encrinurine trilobite species by Adrain and Edgecombe (1997). Tree statistics as follows: CI = 0.6593, RI = 0.6643, RC = 0.4380. Note general agreement between phylogenetic and taxonomic groupings, especially for the genus *Mackenzieurus*. B. Agreement subtree for the eight equally parsimonious cladograms that resulted from a branch-and-bound analysis (equal character weighting) of the 13 cranidial morphological characters coded for these same 13 trilobite species. Tree statistics as follows: CI = 0.6818, RI = 0.7021, RC = 0.4787. Note general agreement between the cranidial character agreement subtree and total character consensus tree topologies.

Character states. Figure 7.8 shows the position of these 12 species within a maximum parsimony cladogram calculated on the basis of the entire 40-character dataset (8A) and within an agreement subtree formed from the 9 equally-parsimonious cladograms calculated from the 12 Adrain and Edgecombe (1997) cranidial characters (8B). While there are differences between these two cladograms, the primary topology of *A. simoni* – *S. harrisoni* *Struzia* species + *F. bachae* *Mckenzienius* is stable, as is the unity of the *MacKenziurus* subclade. The topology of these trees, along with their ensemble consistency and retention indices (CI = 0.6593, RI = 0.6643, RC = 0.4380) indicate that a substantial degree of hierarchical structure in morphological characters exists within these two datasets.

In order to determine whether principal warp-based morphometric methods can recover phylogenetically informative characters these trees were compared with a tree derived from partial warp scores computed in the manner recommended by Zelditch *et al.* (1995). For this analysis a total of 10 landmarks representing consistency relocatable positions on the trilobite cranidium were collected from each specimen (Figure 7.7, upper drawing). Because trilobites are bilaterally symmetrical in dorsal view these landmarks were confined to the left side of the cranidium. Aspects of the glabella and fixed cheek—both of which supplied characters for the traditional analysis—were quantified by these landmarks.

Principal warps were calculated from the *B. adraini* landmark constellation (the outgroup used in the Adrain and Edgecombe 1997) and used to determine partial warp scores for the 12 ingroup taxa. Translation of these partial warp scores into a series of nominal character states was accomplished using scatterplots of the scores to look for gaps in the score distribution along the partial warp x and y axes (see Zelditch *et al.* 1995; Fink and Zelditch 1995). Figure 7.9 illustrates two examples of these, along with the corresponding thin-plate splines, for warps that represent different spatial scales of deformation. [Note: although this analysis differs from Zelditch *et al.* (1995) and Fink and Zelditch (1995) in that only a single representative of the species in question was used, this simplification does not change the principles involved; especially insofar as each of the specimens illustrated in Figure 7.7 exhibits the entire range of cranidial characters and character states used in the traditional, qualitative analysis.]

Interestingly, none of the partial warp plots could be regarded as resembling any of the nominal characters used in the traditional analysis. Whereas the latter are almost always confined to various subregions within the form (e.g., degree of glabellar elongation, angle formed by the posterior margin of the fixed cheek) that are independent structural units of the carapace, the former—by definition—represent patterns of variation over the entire landmark series irrespective of any structural subdivision. Of course, different landmarks within the series receive different weights within the partial-warp vectors (in the same way that different variables receive different weights in a PCA axis), but these spatially defined weight patterns do not respect
Figure 7.9. Representative partial warps and partial warp score plots for a principal warp analysis of the trilobite cranidial landmarks shown in Figure 7.1. A. Scatterplot of partial warp 2 scores (above) and corresponding principal warp (below). This primarily expresses shape variation in the region of the eye socket and fixed cheek. B. Scatterplot of partial warp 5 scores (above) and corresponding principal warp (below). This primarily expresses shape variation as a transverse antero-posterior compression of the cranidium. Zelditch et al. (1995) suggest that partial warp scatterplot axes such as these can be used as systematic characters. These authors recommend that gaps in the distribution of taxa along partial warp axes be used to define character states. For example, the pronounced gaps between *Struszia harrisoni* and the remaining trilobite species along both partial warps 2x and 2y (A, upper figure), and between *Mackenzieurus ceejayi* and the remaining trilobite species along Partial Warp 5y (B, upper figure), could be coded as separating these three putative characters into two states.

obvious structural boundaries. Moreover, the associated scores must be calculated from the entire set of landmark data. This means that the principal warps— and the partial warps of which they are a part—are not truly localized in the same sense that this term is applied in a traditional, qualitative, morphological analysis (see Rohlf 1998 and MacLeod in press for further discussions of this issue).

The agreement subtree derived from these partial warp-based morphometric variables is shown in Figure 7.10. Obviously, the partial warps analysis failed to recover a consistent hierarchical structure in these partial warp-defined variables. As a result, the cladogram based on a qualitative analysis of morphological characteristics (Figure 7.8) was not recovered. These results are consistent with the previous results of Naylor (1996) based on simulated patterns of morphological variation, and the warnings of Bookstein (1996) and Rohlf (1998) regarding the stability and consistency problems inherent in attempting to use partial warp-defined variables as taxonomic characters. The best that can be said of this result is that the partial warps method demonstrated very low discriminatory power for inferring adequately resolved hierarchical patterning from these landmarks. The poor performance of this character-definition method on these morphologies takes its place among the list of similar empirical failures to infer credible phylogenetic patterns from morphometric data (e.g., Cranston and Humphries 1988; Chappill 1989; Crowe 1994).
A RELATIVE WARP APPROACH TO QUANTITATIVE MORPHOLOGICAL ANALYSIS IN SYSTEMATICS: CONGRUENCE, INTERPRETABILITY, AND EXTENSION

Relative warps in systematics
As an alternative to the partial warp-based approach to morphometric character definition advocated by Zelditch et al. (1995), the method of relative warps was evaluated using the same comparative test. Relative warps differ from partial warps in being based on patterns of shape covariance between objects included within an empirical or reference sample. They are, in their simplest formulation, the results of an eigenanalysis of the reference sample covariance matrix where the objects from which the covariances are calculated represent a series of column vectors of landmark locations, in either 2-space (x,y) or 3-space (x,y,z).

While this relative warps approach has been advocated indirectly by Rohlf (1998, see also Bookstein 1996), it has, to my knowledge, never been tested empirically. Zelditch et al. (1995) criticized methods that employ an eigenanalysis of covariance-matrices, arguing that such methods (1) employ optimization criteria that do not correspond to putative principles of 'phylogenetic informativeness' (2) are inherently tied to particular samples (and so do not constitute independent descriptions of morphological state, see also Pimentel and Riggins 1987), and (3) do not produce variables that are spatially localized descriptions of morphological variation to a degree sufficient to conform to the biological concept of homology.

With respect to the first criticism, the purpose of a systematic analysis of organismal morphology is to search for discontinuities or gaps in the patterns of morphological variation. This is true whether the analysis is qualitative—as is the case in traditional systematic investigations—or quantitative. Small discontinuities or gaps will exist between each individual in a sample or population. Nevertheless, to be of systematic utility, the discontinuities or gaps between character states must be greater than the discontinuities or gaps that exist between individuals that exhibit the same character state. If morphological gaps exist between subsets of species in a sample their existence will contribute to the variance of any morphological descriptors (e.g., shape functions) that are sensitive to the presence of the gap. Since covariance-based eigenanalysis aligns multivariable vectors (= axes) with the directions of maximum variance within a dataset, it likely that such procedures will be of great use in locating any gaps that are of genuine systematic interest gaps within morphological datasets. In other words, we would expect that the discontinuities we seek as systematists would be reflected in specimen ordinations that are variance-optimized over the entire sample. This is precisely the sort of pattern we look for when we qualitatively assess morphology. It cannot be logical to accept such evaluation procedures in qualitative contexts, while, at the same time, denying their relevance in
quantitative contexts. In addition, exactly the same type of eigenanalysis-based optimization is employed by partial warp analysis (where a hypothetical bending-energy matrix is substituted for the covariance matrix) to define the principal-partial warps whose use Zelditch et al. (1995) advocate.

With respect to the second criticism, this too is inconsistent with accepted contemporary phylogenetic practice. While theoretical treatments of cladistic characters often make reference to the desirability of non-relative character definitions (e.g., Pimentel and Riggins 1987), in practice, the use of relative, sample-referenced, character definitions is commonplace (Chapill 1989; Stevens 1991; Thiele 1993). Moreover, Zelditch et al. (1992), Zelditch and Fink (1995), Fink and Zelditch (1995), Zelditch et al. (1995), and Zelditch et al. (1998) employ sample-referenced methods to obtain the ‘mean forms’ that they then use as the basis (= tangent point) for their partial warp analysis. Given the extreme instability of partial warps in the face of changes in the reference shape, their criticism of covariance-based eigenanalysis methods—which are, on the whole, much more robust to variation in appropriately constructed samples than partial warps—seems erratic.

This also holds for the Zelditch et al. (1995, as well as Pimentel and Riggins’ 1987) criticism that any alteration of a sample’s composition might significantly perturb the orientation of a variance-optimised morphometric variable axis such as those determined by PCA or relative warp analysis. Certainly this is true if a strongly atypical individual is included in the sample. But in appropriately selected samples the need to deal with substantially outlying individuals should be minimized. Inclusion of a ‘typical’ individual in a sample would not necessarily produce strongly divergent results (see below). Moreover, inclusion of a strongly contrasting individual in a parsimony-based morphological phylogenetic analysis or a likelihood-based molecular analysis also has the ability to perturb the results (e.g., Zelditch et al. 1998), yet this is obviously not regarded by most systematists as grounds on which to preclude the use of such methods in phylogenetic contexts. Indeed, in the same way that Zelditch et al. (1992, see other references above) employed a small, but representative, reference sample to establish the basis for their morphometric analysis of ontogeny, a small, but representative reference sample could be used to establish a basis for any covariance-eigenanalysis morphological analysis with additional individuals—that were not part of the sample used to estimate the population eigenvectors—being projected into the eigenvector-defined reference space (see MacLeod and Rose 1993; MacLeod in press for examples and discussion). While this issue deserves a much more in-depth treatment, once again, it should not be the case that particular morphometric methods are dismissed for being sensitive to certain analytic situations while, at the same time, the implications of such sensitivities for other methods are casually accepted elsewhere in systematic practice.

With respect to the third criticism, MacLeod (in press) has shown that the interpretation of partial warps is no more, and no less, spatially localized than the interpretation of relative warps (= coordinate point eigen-shapes) or PCA axes. As a result, these latter constructs provide as good a fit to the concept of biological homology as partial warp axes; which is to say, not a very good fit at all (see Distances, Landmarks, and Homology section above).

Interestingly, in the canonical examples of partial warp-based character-state descriptions the arrays of partial warp axes are calculated for a constellation of landmarks scattered over the entire body of the organism under investigation. This stands in striking contrast to the normal, qualitative systematic practice of subdividing a complex organic structure into a number of putatively homologous structures and then treating these structures as independent units of morphological variation. From their discussions of the partial warps method it is clear that Zelditch and Fink regard the eigenanalytic decomposition of the reference form’s bending-energy matrix as being the mathematical equivalent of the traditional systematist’s qualitative disassembly of an organism into quasi-independent character complexes. Nevertheless, the partial warps are never truly localized in that spatial information from all parts of the morphology (via the landmark positions) is used to compute all partial warp scores.

To achieve an analytic procedure that is closer to accepted contemporary practice in comparative morphology it would be necessary to first subdivide the organisms body into landmark-defined substructures, carry out separate morphometric analyses on these substructures, use those results to search for discontinuous patterns of variation, code those patterns in the standard manner, and then compute the results of a parsimony-based phylogenetic analysis. Of course, such a subdivided data analysis strategy could be undertaken for any morphometric procedure, including partial warps. However, the fact that no examples of this analytic variant currently exist, along with the emphasis on the putative homology of partial warps that
underpins these canonical examples, suggests that the original partial-warps-as-systematic-characters concept does not include such a subdivided data analytic strategy.
The Naylor simulated fish phylogeny revisited

To test the proposition that a subdivided relative warp morphometric data analytic approach can recover systematically useful character states better than a canonical partial warps approach the former was applied to the Naylor (1996) fish morphology simulations. Those results were then compared to the ‘phylogenetic’ derivation of the simulation models and the results of Naylor’s partial warp analysis of these same data. Naylor’s (1996) simulations are regarded as the most relevant set of test data for these comparisons because the ‘phylogeny’ is known and because the fish simulations represent a level of morphological complexity commensurate with the canonical examples of the partial warp approach.

Figure 7.11 shows the fish simulations used in the relative warps test. These images represent new drawings based on scans of the original Naylor (1996, figure 2) figures. Since these are not the original figures there may be some variance between the originals and these new drawings; especially insofar as the originals were reproduced at such a small size that the precise original landmark locations were very difficult to discern. Inferred landmark locations on the new drawings were quantified as pairs of Cartesian coordinates in the usual way.

In order to test the fidelity of the new drawings and the reproducibility of the Naylor (1996) results, a partial warps analysis was conducted on the new drawings and the resulting warp score distributions compared to Naylor’s figure 3 and table 1 [Note: for this analysis a single ‘eye’ landmark located in the middle of the eye ellipse was used (instead of the four eye landmarks shown in Figure 7.11) in order to render the landmark system comparable to that used by Naylor (1996).]. Although the magnitude of the scores differed (presumably as a result of modifications to the tpsSplin program used to compute the partial warp scores between 1995 and 2000), the relative patterning of scores was virtually identical (Figure 7.12). In addition to supporting the fidelity of the new drawings with respect to Naylor’s originals, this near identity makes a useful point about the nature of landmarks. Since it is very unlikely that exactly the same (= homologous) landmarks were chosen on the original and new drawings in all cases, the similarity of the old and new partial warp results demonstrates that there is no need for landmarks to be absolutely homologous in either biological or geometrical senses in order to represent the gross shapes of landmark-defined objects. The concept of a landmark as nothing more than a relocatable reference whose purpose is to locate the approximate relative positions of gross structural elements is sufficient to achieve remarkably consistent results provided all other aspects of the analysis remain constant. [Note: this result should not, however, be taken as supporting a claim that partial warps per se are robust to changes in the reference or basis shape, see Rohlf 1998.]

Prior to the relative warp analysis, these landmarks were combined into a series of 13 groups (Table 1) that effectively subdivided the overall morphology into a set of quasi-distinct, but biologically homologous,
Figure 7.13. Representative results of a relative warp-style analysis of morphological variation in four different character complexes based on the redrawn Naylor (1996) fish simulations. A. Pectoral fin region shape. B. Abdomen shape. C. Tail shape. D. Pelvic fin shape. E. Optic-branchial region shape. F. Pectoral fin shape. Inset landmark constellations illustrate representative character state morphologies for the morphometrically-defined taxic subgroups. These scatterplots represent patterns of shape variation among the outgroup form and the eight crown simulations of the simulated phylogeny within the plane of the two most important shape-discrepancy axes (relative warps 2 and 3). In each case a clear morphological discontinuity oriented along the single most important shape discrepancy axis (RW-2) was revealed by the analysis. It is suggested that these axes represent putative morphological characters, and the gaps represent the divisions between different states of these characters, that can be useful in systematic analysis. See Figure 7.11 for simulation morphologies.

Table 7.1. Characters and defining landmarks used in the relative warp analysis of Naylor's fish icons. See Figure 11 for landmark positions.

<table>
<thead>
<tr>
<th>n</th>
<th>Characters</th>
<th>Landmarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Mouth Region</td>
<td>1-2-29-28-27</td>
</tr>
<tr>
<td>2</td>
<td>Orbital-Branchial Region</td>
<td>3-2-27-26</td>
</tr>
<tr>
<td>3</td>
<td>Pectoral Region</td>
<td>4-3-26-25</td>
</tr>
<tr>
<td>4</td>
<td>Abdomen</td>
<td>11-10-7-6-22-21-18-17</td>
</tr>
<tr>
<td>5</td>
<td>Caudal Peduncle</td>
<td>12-11-17-16</td>
</tr>
<tr>
<td>6</td>
<td>Tail</td>
<td>16-15-14-13-12</td>
</tr>
<tr>
<td>7</td>
<td>Dorsal Fin</td>
<td>4-5-6</td>
</tr>
<tr>
<td>8</td>
<td>Adiopose Fin</td>
<td>7-8-9-10</td>
</tr>
<tr>
<td>9</td>
<td>Anal Fin</td>
<td>18-19-20-21</td>
</tr>
<tr>
<td>10</td>
<td>Pelvic Fin</td>
<td>22-23-24-25</td>
</tr>
<tr>
<td>11</td>
<td>Eye</td>
<td>32-30-31-33</td>
</tr>
<tr>
<td>12</td>
<td>Gill</td>
<td>34-35-36</td>
</tr>
<tr>
<td>13</td>
<td>Pectoral Fin</td>
<td>37-38-39-40</td>
</tr>
</tbody>
</table>

This step parallels the subdivision of complex morphologies that is universally applied in systematic practice. While this subdivision was arranged to reflect the characters used to construct the Naylor (1996) simulation, a complete measurement of the form required the inclusion of additional characters (e.g., dorsal fin shape, pelvic fin shape).

Once these subsidiary datasets had been assembled each set of landmark constellations was registered (= oriented and scaled) using the Generalized Least Squares (GLS) algorithm (Rohlf, 1990). These registered coor-
dinate data were then used as input for a series of thirteen separate relative warp analyses. Representative patterns of shape variation on the two most important relative warp shape difference axes for six of these characters is shown in Figure 7.13. In each instance the relative warp results separated the nine shapes into a series of mutually exclusive groups; usually into two groups, but in two cases (eye shape and pectoral fin shape) into three. Each of these taxic groupings save one was consistent with the pattern of morphological changes used by Naylor (1996) to construct the simulation.

This single exception was the Orbital-Brachial Region Shape (Figure 7.13E). Naylor’s (1996) change (3) – ‘lengthening of region containing gill and eye’ was purported to be shared by simulations III and VII in his simulation. The relative warp results for this character groups together simulations III, VII, and IX as exhibiting orbital-brachial region landmark constellations that are similar to one another and distinctly different from the remaining substructure (= character) constellations of the other simulations by a considerable degree along the most important shape difference axis (Relative Warp 2, RW-2). However, inspection of the fish shape simulations in Figure 7.11, and Naylor’s (1996) original figure 2, shows that Simulation IX does indeed exhibit an antero-posteriorly lengthened orbital-brachial region that is strongly reminiscent of the states for this character found in simulations III and VII. This interpretation is also supported by Character 12 (gill shape, Figure 13F), that suggests Simulation IX is apomorphic for this orbital-brachial region-related attribute.

Table 7.2. Relative warp-based character/character-state matrix. See Figure 11 for simulation morphologies. See Table 1 for character definitions and Figure 13 for character-state groupings.

<table>
<thead>
<tr>
<th>Simulations</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
</tr>
</thead>
<tbody>
<tr>
<td>III</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>IV</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>V</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>VI</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>VII</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>VIII</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IX</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>X</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Outgroup</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Although the relative warp analysis was quite successful in finding the same taxic groupings implied by Naylor’s simulated phylogeny, this does not necessarily mean that character-state assignments made of the basis of these morphometric data exhibit a consistent hierarchical structure that could be used to recover the simulated pattern of sister-group relationships. To conduct this test the distributions of taxa for the thirteen characters were divided into subgroups based on the presence of unambiguous discontinuities in the patterns of shape variation along either the RW-2 or RW-3 axes. These taxic subgroups were then assigned nominal character-state labels with state ‘0’ being assigned to the group containing the outgroup morphotype. The resultant morphometric character-state matrix is shown in Table 2.

Analysis of this character-state matrix eliminating simulations I and II (as ancestral internal nodes) to repeat the conditions of the Naylor partial warp analysis resulted in the location of one maximally parsimonious trees (Figure 7.14). Although the orientations of the relative warp axes between topology of this tree is identical to that of the original Naylor (1996) simulation. Of the 13 relative warp characters used in this analysis, only one (Orbital-Brachial Region Shape, see above) shows any homoplasy and this homoplasy represents a single character-state reversal. This reversal is regarded as a mistake in Naylor’s original coding of his morphological simulations. Taken as a whole, these relative warp-based characters represents a perfect match in terms of topology and tree descriptive indices to the Naylor (1996) simulated phylogeny. This relative warp-based tree also correctly predicts the character states of simulations I and II (the ancestral forms withheld from the phylogenetic analysis). Finally, contrary to the predictions of Zelditch et al. (1995), this
Figure 7.14. Maximum parsimony cladogram resulting from a branch-and-bound analysis (equal character weighting) of the 13 relative warp-based morphological characters coded for the nine Naylor (1996) simulated fish morphologies (see Figure 7.13 for examples). Tree statistics as follows: CI = 0.9286, RI = 0.9474, RC = 0.8797. Naylor’s (1996) previous partial warps analysis of these simulations succeeded in recovering this tree but yielded a much lower tree Retention Index (0.48). This suggests that the partial warps approach creates substantial amounts of homoplasy in morphometrically-based systematic datasets. See text for discussion.

relative warp-based character analyses is robust to the inclusion of additional specimens in sample (Figure 7.15).

While one successful analytic result—especially of a simulated dataset—does not definitively prove the case for the utility of morphometric approaches to character analysis, this relative warp-based result is by far the most successful example of morphometric variables being used to locate and define systematic-phylogenetic characters currently known. In the most successful partial warps analysis to date (Fink and Zelditch 1995) the tree based on only the morphometrically-defined characters differed from the tree produced by the combined morphometric + qualitative character dataset suggesting that these two datasets disagree. Tree statistics for the relative warp analysis of the Naylor (1996) simulations are overwhelmingly better than Naylor’s (1996) partial warps results and substantially better than for the (Fink and Zelditch 1995) partial warps characters.

SUMMARY

Goals of qualitative and quantitative morphological analysis in systematics are the same. Both are concerned primarily with representing patterns of morphological variation in organisms and relating these patterns to other patterns of variation in other variables. Systematists are particularly concerned with the documentation of morphological discontinuities that exist between groups of individual organisms because these discontinuities (may) reflect cladogenesis.

Many of the morphological patterns that interest systematists represent geometries that can, in principle, vary continuously. These are best expressed as ratio-scale variables; the same types of variables that are routinely used in morphometric analysis. Such variables are compatible with the standard systematic concept of the ‘character’ as defined by Farris et al. (1970) and by Pimentel and Riggins (1987). Systematists have often been confused by the fact that sets of observations or measurements along continuous, ratio-scale, morphometric variables can exhibit either continuous or discontinuous (= clustered) patterns of
Figure 7.15. Effect of including new morphologies on the Naylor (1996) fish simulation relative warp results. A. Relative warp ordinations within the plane of the two most important shape-discrepancy axes (relative warps 2 and 3) for simulations III through X and the ancestral form (A). B. Relative warp ordinations within the plane of the two most important shape-discrepancy axes (relative warps 2 and 3) for simulations I through X and the ancestral form (A). Inset landmark constellations illustrate representative character state morphologies for the morphometrically-defined taxic subgroups. Note that inclusion of the internal node simulation morphologies I and II did not appreciably change the separation between character-state groupings or the intra-state ordinations of morphotypes within this plane through the shape space. While the inclusion of additional (typical) individuals to a sample will engender some small alteration the geometry of the sample’s shape space, this alteration would not be expected to automatically obscure the presence and recognition of real morphological discontinuities. Moreover, it is possible to project individuals into the space defined by a statistically representative sample of morphotypes without altering the nature of the eigenvector-defined shape space at all (see MacLeod and Rose 1993 for an example). Accordingly, objections to the use of eigenanalysis-based methods in morphological systematics because of their so-called inherent instability in the face to additions to the reference sample (e.g., Zelditch et al. 1995) may, in many instances, be irrelevant to systematic practicalities. Geometric analysis should be able to duplicate any result obtained by traditional, qualitative methods. See Figure 7.11 for simulation morphologies.

variation. Sets of observations or measurement that exhibit the latter offer no particular difficulties in terms of devising objective rules for transforming morphometric ratio-scale variables (= characters) into the nominal variables (character states) required by contemporary phylogenetic analysis algorithms. Sets of observations or measurement that exhibit the former offer no hope of being able to delineate groups of taxa logically, consistently, or objectively, irrespective of various ‘member coding’ procedures that have been devised.

The fact that a large number of systematic studies employ characters that are ratio-scale morphometric variables that have been subdivided arbitrarily into discrete states emphasizes the practical and uncontroversial use of morphometric data in systematic contexts, as well as the routine systematic observation of discrete distribution of observations or measurements along such variable axes. Explicit morphometric analysis methods, however, offer systematists practical means of discovering, assessing, and describing the morphological gaps on which taxic diagnoses are based as well as more consistently coding observations or measurements along variable (= character) axes for systematic analysis.

Traditional objections to the use of morphometric data in systematic contexts because they do not conform to the concept of biological homology derive from a misunderstanding of morphometric variables (especially landmarks coordinates), an inconsistent approach to the specification of traditional, qualitative observations in systematics, and a lack of appreciation for the spatial limits implicit in the concept of biological homology. As has been pointed out by a pantheon of systematists stretching from Richard Owen to Colin Patterson, homologues are structures that exist at particular spatial scales and are most often recognized by either internal or external topological similarity with other such structures. Landmarks may represent homologous structures for the purposes of assessing topological similarity, but landmarks—as geometric points—are not intrinsically homologous (see MacLeod 1999 for examples of non-homologous systematic landmarks). Rather they represent abstractions of structures that may (or may not) represent true homologues.
Finally, the recent suggestion that morphometric partial warps of landmark configurations can be used as a new source of morphological characters and character-states was evaluated via comparison of trees resulting from a traditional qualitative, and partial warps-based quantitative analysis of trilobite cranial characters for a selection of 12 encrinurine species from the Canadian Arctic. Results showed that the partial warps based tree differed strongly from the tree derived from qualitative morphological analysis and that none of the characters used for the latter was ‘discovered’ by the former. These results are consistent with the theoretical and practical results obtained by others on both simulated and real taxa. In aggregate they raise serious concerns regarding the applicability of using partial warps in systematic contexts. Alternatively, a new procedure for using landmark-based morphometric analysis based on relative warps is described and applied to a simulated phylogeny. Results show that this new method (1) correctly ‘discovered’ the distinctions between simulated morphologies, (2) produced exactly the same tree as would have been produced by a qualitative analysis of the simulated morphologies, (4) did not introduce elevated levels of homoplasy to the phylogenetic analysis, and (3) resolved heretofore unsuspected ambiguities in the simulation.

Reliance on qualitative methods to recognize and document morphological discontinuities in systematic datasets has led, in many cases, to needless confusion and controversy over the validity of phylogenetic systematic results and appropriateness of interpretations based on those results. The descriptive and analytic rigor that would result from introducing morphometric methods into phylogenetic systematics would have substantial and positive implications for both fields. A period of active experimentation with these methods is now needed to further explore their compatibility.

ACKNOWLEDGEMENTS

I wish to thank the following for informative discussions regarding morphometric procedures and/or their use in systematic contexts: Fred Bookstein, Jim Rohlf, Chris Humphries, David Polly, and Donald Swiderski. This essay was reviewed by Peter Forey and Andrew Smith both of whom submitted comments and criticisms that greatly improved the arguments presented.

REFERENCES


