Mapping cladograms into morphospaces

J. R. Stone

Abstract

In cladistic analyses, taxa are grouped hierarchically into clades according to shared apomorphic character states to construct cladograms; cladograms are interpretable as phylogenetic hypotheses. In morphological space analyses, organism forms are represented as points in morphospaces; point proximities in morphospaces represent similarities that might be attributable to phenetic convergence and, consequently, may correspond inaccurately with hypothesized evolutionary relationships. A method for synthesizing phylogenetic results that are interpreted from cladistic analyses with phenetic results that are obtained from morphological space analyses is presented here; in particular, points that represent forms typifying taxa in morphospace are assigned as terminal nodes for appropriate cladograms that are mapped into morphospaces by positioning nonterminal nodes and orienting internodes according to a geometric algorithm. Nonterminal nodes may be interpreted as ancestors in phylogenetic hypotheses and occupy positions that represent particular organism forms in morphospaces. By mapping cladograms into morphospaces, therefore, evolutionary morphologists can reconstruct ancestral morphologies and test historical transformation hypotheses.

Dr J. R. Stone, Biology Department, Dalhousie University, Halifax, Nova Scotia, Canada B3H 4J1 and Biology Department, McMaster University, Hamilton, Ontario, Canada. E-mail: jrstone@is.dal.ca

Introduction
Since phylogenetic systematics revolutionized evolutionary biology (Hennig 1966; Hull 1988), cladistic analysis has become the preferred method for constructing taxonomic classifications. Information concerning characters is coded and stored in matrices from which taxa groupings, or ‘clades’, are deduced. Clade sets are represented by two-dimensional connected graphs (mathematical trees), or ‘cladograms’ (Hennig 1966; Wiley 1981; Fig. 1), which may be interpreted as phylogenetic hypotheses (i.e. connected graph branching patterns may be interpreted as phylogenetic event reconstructions; Brooks and McLennan 1991). Two dimensions are sufficient to summarize visually and efficiently the results that are obtained from cladistic analysis. However, there is no reason why cladograms cannot be extended into three (or more) dimensions (i.e. morphological evolution may be interpreted using parameters that complement relative historic divergence).

Since the concept ‘morphospace’ innovated organism morphology studies (Thompson 1917; Raup 1966), morphological space analysis has become the preferred method for analysing or modelling growth and form (Stone 1997a; McGhee 1999). Information concerning ontogeny, size, or shape is quantified and represented as points in mathematical spaces with axes that are delimited by parameters that are deduced from functional considerations, derived from statistical analyses, or defined according to mathematical models (Stone 1997a). Morphospaces may be represented using tables (e.g. Thomas and Reif 1993; Thomas et al. 2000) or contour diagrams (e.g. McGhee 1980), as two-dimensional co-ordinate systems (e.g. Gould 1984; Stone 1996), or as hyperspaces (or projections therefrom; e.g. Foote 1993, 1994). However, they often are represented as three-dimensional mathematical spaces (e.g. Raup 1966; Niklas 1994, 1997).

Formally synthesizing cladistic and morphological space analyses would enhance morphological evolutionary studies, by providing a means for reconstructing ancestral morphologies and testing historical transformation hypotheses. Bookstein et al. (1985) considered superimposing cladograms onto two-dimensional morphospaces but provided no theoretical basis for positioning nonterminal nodes. In the synthesis that is introduced here, methods for positioning nonterminal nodes and orienting internodes are described,
and a geometric algorithm for re-forming (and reforming) traditionally depicted two-dimensional connected graphs as three-dimensional branching patterns that are contained in three-dimensional morphospaces is presented.

Methods for positioning points that represent forms typifying taxa in morphospaces are well established (Stone 1997a), and cladogram terminal nodes may be considered as points in morphospaces. Therefore, provided that nonterminal nodes that are common to adjacent clades, or ‘sister groups’, can be positioned and internodes can be orientated, cladograms can be mapped into three-dimensional morphospaces. Nonterminal nodes should be positioned precisely because taxa that are represented as sister groups on cladograms (e.g. t2 and t1 in Fig. 1) may be interpreted in phylogenetic hypotheses as having shared a common ancestor that is represented by their common nonterminal node (e.g. n0 in Fig. 1).

Positioning internodes and orienting nonterminal nodes may be considered as the geometric analogue to polarizing character states in cladistic analysis. Polarization may be accomplished by implementing ‘functional outgroup analysis’ (Watrous and Wheeler 1981), wherein the relative nature with which character states are interpreted and deductive aspects that are inherent in cladistic analyses (i.e. information provided by each character state is tested against information provided by other character states) are utilized. For example, the binary character states that are associated with one character may be utilized to define a functional outgroup and polarize the multiple character states that are associated with another character (Fig. 1).

Materials and Methods

Positioning nonterminal nodes and orienting internodes

Positioning nonterminal nodes can be accomplished by implementing algebraic algorithms or optimization procedures. Parameter values are mapped onto terminal nodes and deduced at nonterminal nodes according to mathematical formulae or prescribed criteria (because these are established practices of cladistics, each is considered in only an abbreviated manner herein).

Algebraic algorithms. Algebraic algorithms (Cavalli-Sforza and Edwards 1967; Felsenstein 1985; Huey and Bennet 1987; Swofford and Maddison 1987, 1992; Coddington 1988; Donoghue 1989; Harvey and Pagel 1991; Garland et al. 1992; Frumhoff and Reeve 1994; Schultz et al. 1996; Schluter et al. 1997; reviewed in Cunningham et al. 1998) involve topology-dependent formulae for determining continuous variable values at nonterminal nodes. The appropriateness that is associated with the algebraic algorithm that is used in any particular character-state evolution reconstruction (e.g. the Brownian motion model of character-state transformation in Felsenstein 1985 or the independent contrasts method in Harvey and Pagel 1991) should be assessed critically.

Optimization procedures. Optimization procedures involve inferring character states at nonterminal nodes by considering character-state transformations according to prescribed criteria (e.g. parsimony or maximum likelihood). Camin–Sokal parsimony (Camin and Sokal 1965) involves assigning changes irreversibly to derived, or ‘apomorphic’, character states (i.e. after the origin of an apomorphic character state, transformation to relatively underived, or ‘plesiomorphic’, character states cannot occur in subsequent clades that originate further from the root); Wagner parsimony (Kluge and Farris 1969; Farris 1970) involves assigning changes
with respect to ordered character-state sequences (e.g. transformation from 0 to 2 contributes two character-state changes); Fitch parsimony (Fitch 1971) involves assigning all changes equally (e.g. transformation from 0 to 1 or 2 each contributes one character-state change); and Dollo parsimony (Farris 1977) involves assigning character-state changes uniquely (i.e. each transformation to a shared apomorphic, or ‘synapomorphic’, character state occurs once). Maximum likelihood techniques involve assigning character-state changes that yield the greatest likelihoods according to prescribed mathematical models (e.g. Markov chains).

**Geometric algorithms.** Alternatively, geometric algorithms may be used for positioning nonterminal nodes and, in addition, orienting internodes. For example, under an equivalence assumption, terminal nodes that represent sister groups in three-dimensional morphospaces (e.g. t2 and t1 in Fig. 2) can be joined to nonterminal nodes by the equivalent sides comprising right-angle isosceles triangles. Initially, the unique vertex in each right-angle isosceles triangle, the nonterminal node (n0 in Fig. 2), may be positioned anywhere about the unique circle that is situated between the terminal nodes and midway along and perpendicular to an imaginary straight line joining them (Fig. 2; each unique circle corresponds to the locus traversed by rotating the vertex in the right-angle isosceles triangle through $2\pi$ radians, whereas each imaginary straight line corresponds to the unique edge delineating the right-angle isosceles triangle; the imaginary line length determines the circle radius). Because three points may be used to define a plane, each right-angle isosceles triangle may be defined to lie in a plane that contains points that correspond to sister-group terminal nodes (e.g. t2 and t1 in Fig. 2) and the terminal node that represents their sister group (e.g. t3 in Fig. 2), which thereby serves as a geometric functional outgroup. This plane intersects the unique circle circumference at two points and therefore can be used to define the particular position on that circumference that is co-planar with the clade (e.g. t3, t2 and t1) and situated closest to the functional outgroup node (e.g. t3 in Fig. 2) to position the nonterminal node (e.g. n0 in Fig. 2).

To exemplify applying this geometric algorithm, a cladogram for strombid gastropod species (Fig. 3; Stone 1997b) was mapped into a three-dimensional morphospace (Fig. 4). Points that represent forms typifying taxa in morphospace were determined according to parameter values that are contained in a general mathematical shell model (Stone 1995) and describe shell forms comparatively (Stone 1997c): offset $O$ affects the coiling distance from the columella or umbilicus, the coiling ‘tightness;’ horizontal expansion $H$ affects the aperture lateral dimension, the aperture ‘width;’ vertical expansion $V$ affects the aperture longitudinal dimension, the aperture ‘height.’ Cladogram polychotomies were...
Results

Cladogram for strombid species in morphospace: reconstructing ancestral morphologies

In the three-dimensional OHV morphospace (Fig. 4), points representing *Lambis truncata* and nonterminal node n4 are connected to their common nonterminal node n5. Interpreted phylogenetically, n5 represents the ancestor that is common to all extant lambid gastropod species; therefore, its position in morphospace represents a ‘proto-lambid-pseudo-strombid’ ancestor form. Similarly, n4 represents the form for the ancestor that is common to the strombid group P2 comprising *L. millepeda*, *L. digitata*, *L. chiragra*, and another ancestor that is represented by nonterminal node n3; n3 represents the form for the ancestor that is common to *L. violacea* and another ancestor that is represented by nonterminal node n2; n2 represents the form of the ancestor that is common to *L. lambis* and another ancestor that is represented by nonterminal node n1; n1 represents the ancestor that is common to *L. scorpius* and the ancestor that is represented by terminal node P1, which represents polychotomy P1.

Discussion

Cladistic morphospaces

The geometric algorithm that is presented herein enables evolutionary morphologists to construct ‘cladistic morphospaces’: mathematical spaces containing points that represent taxa and comprise terminal nodes for cladograms that are contorted by a variety of techniques according to particular criteria. Implementing right-angles renders orthogonal terminal internode orientations; in phylogenetic hypothesis terms, this represents an assumption that speciation yields lineages that traverse perpendicular morphological evolutionary trajectories. Implementing straight lines guarantees that terminal internodes bridge minimal distances in morphospaces; in phylogenetic hypothesis terms, this is tantamount to a Pythagorean assumption that ‘species span shortest morphological transformations’. Implementing isosceles triangles renders isometric distances from nonterminal nodes to terminal nodes that represent sister groups; in phylogenetic hypothesis terms, this represents an ‘equivalent morphological modification between diverging lineages’ assumption.

Of course, any combination among these assumptions might be inappropriate for conducting any particular cladistic-morphospace analysis. To accommodate other situations (e.g. if ‘branch lengths’ were known), equilateral or scalene triangles, acute or obtuse vertices, or curved lines could be implemented with the geometric algorithm. For example, if two particular morphological transformations were perpendicular and followed shortest trajectories but were characterized by nonequivalent rates in the ratio 2 : 1, then the sides that oppose the $\pi/6$ and $\pi/3$ radian angles in a right-angle scalene triangle could be used to join sister group terminal nodes. Alternatively (e.g. if branch lengths were unknown), nonterminal nodes could be positioned so that internodes minimize entire cladogram lengths (Cavalli-Sforza and Edwards 1964, 1967).

Without information concerning relative orientations, trajectories, or morphological evolution rates, positioning nonterminal nodes using right-angle isosceles triangles is most convenient geometrically, conservative phylogenetically, and informative morphometrically. Geometrically, positioning nonterminal nodes using right-angles is accommodated very easily, because orthogonal line properties (e.g. cross-products = 0) may be utilized to define planes for positioning nonterminal nodes. Phylogenetically, taxa joined to nonterminal nodes by straight lines are hypothesized to have evolved morphologically minimally and unidirectionally with respect to ancestor morphologies. These morphological transformation hypotheses are very sensitive to falsification by including additional information (Popper 1987), because any datum that is consistent with a straight-line definition will be coincident with that line, whereas any datum that is discordant with a straight-line definition will lie on any among
the myriad other curves that could join terminal nodes to a non-terminal node. Morphometrically, terminal internode lengths provide metrics for morphological transformation: short lines imply little change, whereas long lines imply great change.

Cladistic morphospace for strombid species: testing historical transformation hypotheses

On the basis of terminal internode lengths, shells of *Lambis lambis*, *L. scorpites*, and members in P1 are similar and have changed insubstantially with respect to ancestor forms; whereas the shells of *L. truncata*, members in P2, and *L. violacea* are dissimilar and have changed substantially with respect to ancestor forms (Fig. 4; in particular, shells of *L. truncata* and *L. violacea* ‘unwound’ and accreted ‘taller’ apertures). Therefore, morphological evolution was most pronounced early in this strombid gastropod group’s history and diminished thereafter. This example demonstrates the utility that is conferred by cladistic morphospaces. The contorted cladogram provides a topology with which points that represent forms typifying taxa in morphospace may be connected and, thereby, integrates morphological variation measures with inferred evolutionary relationships in a comparative manner. This integration is requisite to comprehending morphological disparity (Gould 1991).

Acknowledgements

Selective, calculated, voluminous and stimulating comments were provided by J. Rising, J. Repka, M. Telford and R. D. K. Thomas. Financial support was provided by a Malacological Society of London Centenary Research Grant, a Conchologist of America Research Grant, Natural Science and Engineering Research Council of Canada (NSERC) Individual Grant 4696, an Ontario Graduate Scholarship, an NSERC Postdoctoral Fellowship, a Swedish Natural Sciences Research Council Postdoctoral Project Grant, and a Canadian Institutes of Health Postdoctoral Fellowship.

References


