Published by Associazione Teriologica Italiana
Hystrix, the Italian Journal of Mammalogy
Volume 24 (1): 75–83, 2013
Available online at:
http://www.italian-journal-of-mammalogy.it/article/view/6298/pdf
doi:10.4404/hystrix-24.1-6298



Research Article

Phenotypic trajectory analysis: comparison of shape change patterns in evolution and ecology

Michael L. Collyer[a,*], Dean C. Adams[b]

[a] Department of Biology, Western Kentucky University, Bowling Green, KY 42101, USA
[b] Department of Ecology, Evolution, and Organismal Biology and Department of Statistics, Iowa State University, Ames, IA 50011, USA

Keywords:
phenotypic trajectory analysis
multi-state phenotypic change
Procrustes

Article history:
Received: 1 June 2012
Accepted: 18 January 2013

Acknowledgements
We thank A. Cardini and F.J. Rohlf for constructive comments on this manuscript. This work was supported in part by a Western Kentucky University Research and Creative Activities Program grant 12-8032 (to MLC) and an NSF grant DEB-1118884 (to DCA).

Abstract

Research using shape data from geometric morphometric (GM) methods in ecology and evolutionary biology is typically comparative, analyzing shapes and shape change over different points along ecological or evolutionary gradients. Whereas standard multivariate statistics procedures are fine for "static" variation – testing for location differences of groups in multivariate data spaces – they are limited for "dynamic" variation – testing specific differences in the ways groups change locations associated with changes in state along ecological, developmental or evolutionary gradients. In this paper, we show that continuous phenotypic change can be described by trajectories in multivariate data spaces. We describe the geometric attributes of phenotypic change trajectories (size, direction, and shape), specifically for GM data. We illustrate, with examples, how differences in such attributes can function as test statistics for comparative analyses in order to understand the mechanisms that produce dynamic differences in shape change. We demonstrate that analysis of such attributes – called phenotypic trajectory analysis (PTA) – is a general analysis that can be applied to various types of research questions concerned with measuring dynamic variation. Finally, we posit some challenges for the future for this novel analytical method.

Introduction

Most users of geometric morphometric (GM) methods are aware of the famous quote by D'Arcy Wentworth Thompson (1917), "In a very large part of morphology, our essential task lies in the comparison of related forms rather than in the precise definition of each". This preamble to Thompson's "Theory of Transformations" defines the essential task in comparative studies to be the description of morphological change between two forms rather than the mere descriptions of two separate forms. The fields of ecology and evolutionary biology often share the same objective; that is, description of a current ecological or evolutionary state is not as fascinating or important as knowing the mode and tempo (Simpson, 1944) of the change of the state. As reviewed by Adams et al. (2004, this issue), many advances in the field of geometric morphometrics in the last decade use the data from GM to identify patterns of phenotypic change in ecology and evolutionary biology. One recent advance – the result of an interesting convergence of shared principles in the fields of ecology, evolutionary biology, developmental biology, and geometric morphometrics – is phenotypic trajectory analysis, which is the description and comparison of geometric attributes of phenotypic change (Adams and Collyer, 2009). In this paper, we discuss a specific phenotype that is frequently of interest in ecology and evolutionary biology research – organismal shape – and describe how the principles on which GM is based can also be applied to the comparison of patterns of shape change in shape spaces.

On one hand, the analysis of shape change is not new. Multivariate tests for comparison of means, such as Hotelling's 1931 T^2 or multivariate analysis of variance (MANOVA) can be used to estimate the probability of shape change (difference) between two means, or the joint-change between multiple means, respectively, under a null hypothesis of no difference in means (Marcus, 1993). Multivariate tests of linear association (e.g., multivariate regression) estimate the probability associated with the amount of shape change per unit of change in a continuous independent variable, under a null hypothesis of no association (Rencher, 2002). Inherently, any statistical test for comparisons of shapes among groups or the linear associations of shape and other continuous variables is concerned with evaluating shape change relative to per unit changes in one or more independent variables. Standard multivariate test statistics (e.g., Wilks' Λ, Pillai's trace) evaluate the probability of effects – due to a single coefficient or a combination of coefficients – under a null hypothesis of no effect. For some analyses, this level of hypothesis testing is sufficient.

On the other hand, standard multivariate analyses alone are not sufficient for understanding how shape changes, or how these changes may be similar or different in distinct groups. That is, standard multivariate tests are not sufficient for understanding more precise reasons for rejecting a null hypothesis of consistency in shape change (Collyer and Adams, 2007). For example, one might analyse shape variation for four groups (e.g., species) of organisms exposed to four different treatments (e.g., different temperatures) in an experimental setting using a factorial MANOVA, including one factor for species, another factor for treatment-levels (temperature-exposure), plus an interaction of the two factors. If a significant species-treatment interaction is observed, what does this mean? For univariate response data, one might be able to make sense of the significant interaction by estimating and evaluating a response surface (Box and Draper, 1987). For multivariate response data, a significant interaction means that the coefficients for the interaction are obviously important to the factorial model, but a response surface is not possible nor will analysis of each coefficient, one by one, easily reveal how the four species differ in terms of their patterns of shape change across a temperature gradient. Nonetheless, the coefficients allow one to estimate the expected shapes that each group has at each temperature.

Another way to think about the problem is to visualize the shape patterns as trajectories in shape space (Fig. 1). Using the example described above, the four species are each represented by four points in shape space (and the corresponding tangent space; see Adams et al. this issue). The differences in positions of these four points de-

*Corresponding author
Email address: michael.collyer@wku.edu (Michael L. Collyer)

© 2013 Associazione Teriologica Italiana

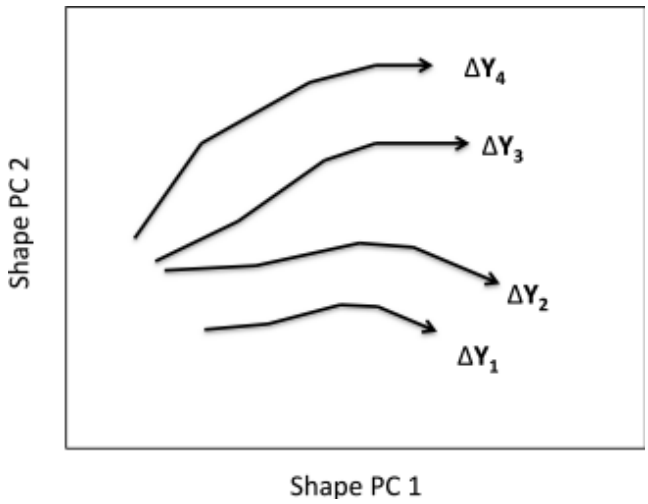


Figure 1 – Example of shape trajectories, projected onto two principal components of shape. These trajectories indicate the shape responses of four species to four different (experimental) temperature exposures. Each trajectory is denoted as $\mathbf{\Delta Y}$, comprised each of four $\mathbf{\Delta y}$ vectors (change in shape from one temperature to the next) from start point to end point (indicated by arrow direction). These trajectories illustrate differences in four geometric attributes. First, each trajectory is in a slightly different location (related to species differences in shape). Second, species 1 and 2 have trajectories of similar shape and direction, but differ in size. Third, species 2 and 3 have trajectories of similar size and shape, but differ in direction. Fourth, species 3 and 4 have trajectories of similar size and orientation, but differ in shape.

scribe the shape change of a group across a temperature gradient (four sequential, distinct temperature exposures) by the differences in their locations (in tangent space) along a trajectory. From this illustration it is easy to see why typical analyses do not describe all aspects of biological shape change. With standard multivariate analyses, one can test whether the species differ in shape (i.e., differences in their locations in tangent space), test whether shapes differ across temperature treatments, or whether there are differences in species-specific responses to particular temperatures. Yet none of these statistical components directly captures the phenotypic trajectories embodied by the shape data and shown in Fig. 1. Clearly, additional methods are required for this.

Viewing shape patterns as trajectories allows one to consider the phenotypic attributes that describe these trajectories. Specifically, trajectories in multivariate data spaces, much like landmark configurations in 2- or 3-dimensional coordinate spaces, can differ in four attributes: location, size, orientation, and shape. Differences in location between trajectories are analogous to general shape differences between groups. For the example above, a single-factor (one-way) MANOVA, followed by pairwise comparisons of species means, can be used to test for differences among species locations in shape spaces (Marcus, 1993). The other three attributes, however, describe species-specific shape change over discernable observational levels that are not encapsulated by standard multivariate analyses (Adams and Collyer, 2009). A factorial MANOVA might indicate significant variation among species-temperature means but does not implicitly test if species differ in the size, orientation, or shape of their shape trajectories. For the critical research mind, comparison of these trajectory attributes is an important and essential step in understanding why a factorial interaction is significant, and what it implies biologically.

In the following sections, we describe how such attributes of shape change can function much like test statistics to evaluate shape variation in comparative studies. We start with the simple comparison of shape change associated with an independent variable (e.g., size allometry). We then expand on this concept to introduce the comparison of shape change vectors, and finally the comparison of shape trajectories. In each section, we provide an illustrative example from empirical research. Our examples are necessarily brief and meant to highlight the conceptual advantage of examining differences in phenotypic (shape) trajectory attributes to test hypotheses. Further biological implications and specific analytical details can be found in the original sources that provide the data. Also, general analytical and statistical details can be found in Adams and Collyer (2007); Collyer and Adams (2007), and Adams and Collyer (2009).

Linear shape change associated with a continuous variable

It is common practice to compare the covariation of shape and other continuous variables collected on the same subjects (Rohlf and Corti, 2000), especially between two or more groups (e.g., Klingenberg 1996, 1998; Drake and Klingenberg 2008; Adams and Nistri 2010; Piras et al. 2010; Viscosi et al. 2010). For example, shape allometry comparisons between two or more groups – rather than comparisons of average shapes – seek to determine if groups differ in the way shape changes per unit change in size (e.g. the log of centroid size) (Klingenberg, 1996, 1998). The typical approach is to perform a "homogeneity of slopes" test. This is accomplished by comparing the log likelihoods of two different models: one model containing a group factor and a common (global) slope; another model containing a group factor and coefficients for independent slopes (i.e., including a group × size interaction; Rencher and Schaalje 2008). If there is a large difference in

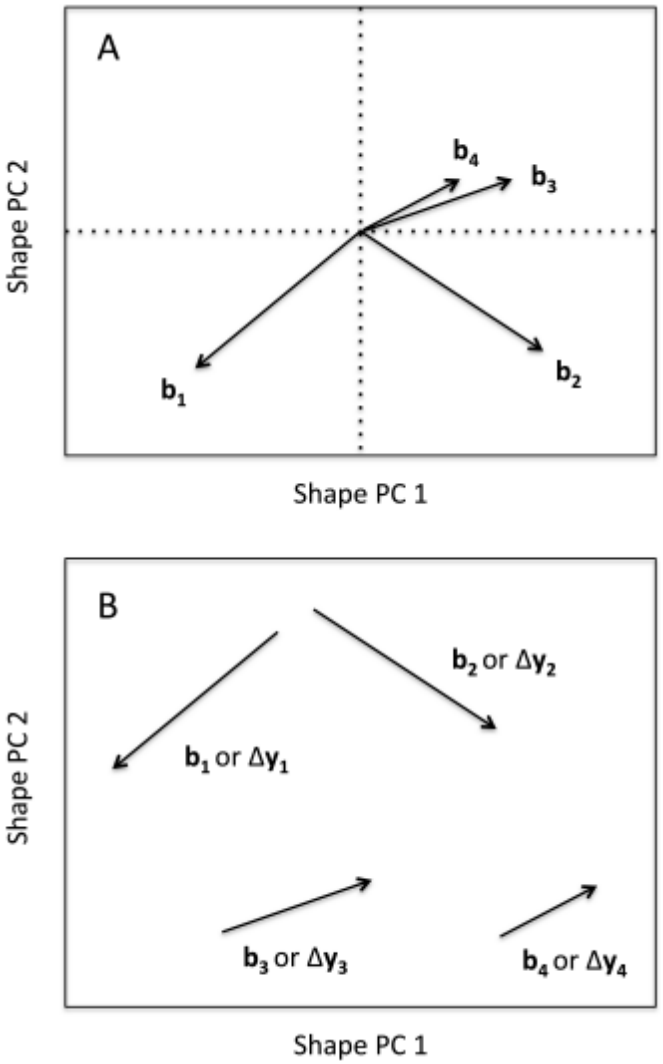


Figure 2 – Two ways to visualize shape allometry vectors. A) Allometry vectors ($\mathbf{b_i}$) as vectors of shape change from a predicted value (origin) per unit change in size. The length of the vector indicates the amount of shape change; the direction indicates the covariation of shape variables (projected onto principal components). In this case, groups 1 and 2 exhibit similar amount of shape change but the direction of shape change differs. Groups 3 and 4 change shape in similar ways with growth, but group 3 exhibits greater shape change. B) Allometry vectors ($\mathbf{b_i}$) as vectors of shape change between "small" and "large" sized estimates. Vectors are similarly scaled and orientated as vectors in A, but are located in different parts of the tangent space (based on group distinctions in shape). Such vectors could also describe shape change between discrete ecological or evolutionary states, rather than two points on a continuum ($\mathbf{\Delta y}$).

log-likelihoods between the two models (i.e., error is substantially reduced by having independent slopes), a null hypothesis of equal slopes among groups is rejected, meaning that shape allometries differ in some way. Determining significant heterogeneity in slopes is only part of the battle, as one should be compelled to understand how shape allometries differ.

For any shape defined by p shape variables, a $1 \times p$ vector of coefficients, $\mathbf{b}$, defines the linear change of each shape variable per unit change of an independent variable, X. As a vector, it has two geometric attributes: a length and a direction. Vector length describes how much shape change occurs per unit change of size; vector direction describes the relative covariations of shape variables per unit change of size (Fig. 2A). To compare these attributes between two or more groups, one can calculate the absolute difference in vector lengths (distances), $\Delta d = |d_1 - d_2|$, and the angle between vectors, θ (analytical details provided in Adams and Collyer 2007; Collyer and Adams 2007). These values can function as test statistics with expected values of Δd = 0 and θ = 0 radians or degrees, under a null hypothesis of equal allometries. The percentile of observed values of these geometric attributes from a distribution of the same values computed from a resampling experiment (i.e., generated from a null model), can be used as p-values for evaluation of the null hypothesis (see Krabbenhoft et al. 2009; Piras et al. 2010). An example based on the data from Krabbenhoft et al. (2009) is presented below. We feel it is important to point out that these attributes are calculated using all dimensions of the data space (such as the tangent space to shape space; see Adams et al. this issue) and that projection of vectors – as in Fig. 1 and 2 – onto principal components is for visual interpretation, only. Projection can distort angles and vector lengths by reducing either or both (see Collyer and Adams 2007). This phenomenon can be appreciated by envisioning an x, y, z Cartesian space containing two intersecting vectors of equal length that only differ in direction in the z dimension. An x-y projection would reveal two parallel, overlapping vectors of different length. Therefore, calculation of vector lengths and angles after projection into a space of fewer dimensions could produce erroneous results in terms of estimation, statistical evaluation, and interpretation (see also Mitteroecker et al. 2005).

It should be noted that an alternative method for describing attribute differences, which only rescales Δd and has no effect on θ, is to estimate shape at the same "small" and "large" sizes for each group (Fig. 2B). If $\hat{\mathbf{y}}$ is an $1 \times p$ vector of shape values (e.g., principal component scores) estimated from a linear regression model that describes the linear association between shape and size, then $\mathbf{b} = \hat{\mathbf{y}}_{\mathbf{large}} - \hat{\mathbf{y}}_{\mathbf{small}}$ is a vector that describes shape change between arbitrary large and small measures of size (which are consistent among groups). Differences in vector length and angles between vectors are calculated the same way. One advantage to using this approach is that changes in shape associated with growth can be shown at the locations of such points in principal component plots (Fig. 2B).

As an example, we use data originally reported in Krabbenhoft et al. (2009). These data contained landmark configurations for 868 fish, representing sister-species pairs for three genera. The three genera, *Etheostoma*, *Menidia*, and *Fundulus*, are broadly distributed in North America, but each genus contains one species endemic to the Pleistocene-originated Lake Waccamaw (North Carolina, USA) (Hubbs and Raney, 1946). These endemic species have substantially more slender body shapes in Lake Waccamaw compared to stream species occurring near Lake Waccamaw. The putative explanation for body-slendering is that higher predation in the shallow, clear waters of Lake Waccamaw is selection for streamlining, which is more energy efficient for swimming fishes (Hubbs and Raney, 1946). In order to estimate differences in linear allometric patterns, shape variables were estimated using GPA (Rohlf and Slice, 1990) performed on configurations of 12 landmarks per fish (Fig. 3A), followed by orthogonal projection and a principal component analysis. This procedure produced 20 shape variables (see Krabbenhoft et al. 2009 for further details).

One question addressed in this study was whether slender body shapes resulted from different shape allometries in Lake Waccamaw compared to streams. A homogeneity of slopes test indicated that shape allometries differed in some way (results not shown). Allometry vectors were calculated for all six species and Δd and θ were calculated between vectors for each sister-species (intra-genus) pair. Attribute differences were evaluated from sampling distributions generated from 10000 random permutations (based on a null model that lacked coefficients for independent allometries; see Krabbenhoft et al. 2009), and the null hypothesis Δd = 0 or θ = 0 was rejected if the p-value of the observed attribute difference was less than an acceptable type I error rate of α = 0.05.

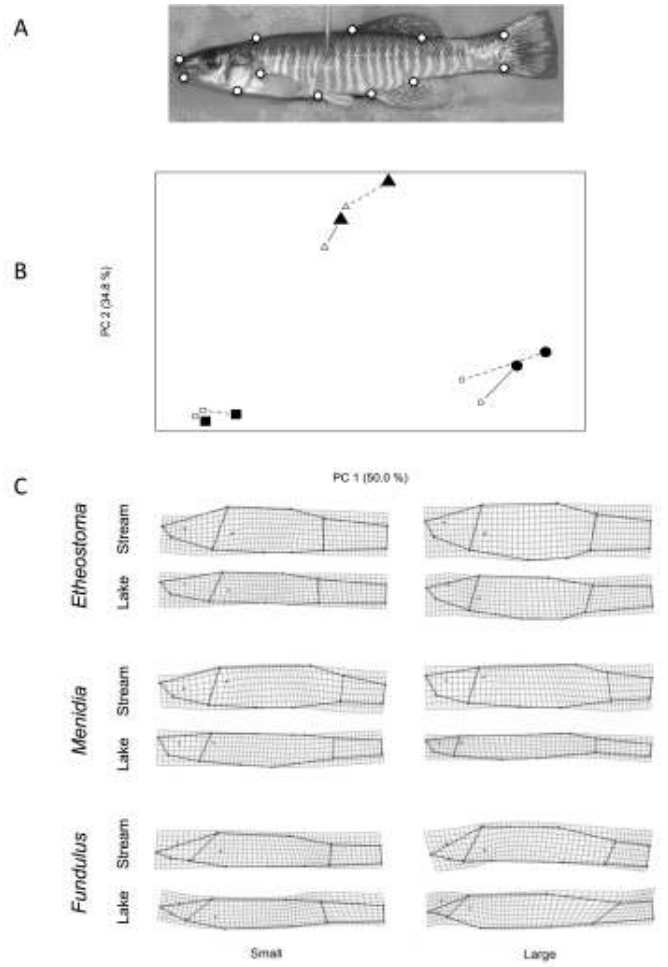


Figure 3 – Visualization of shape allometry for three species-pairs occurring in streams and Lake Waccamaw, North Carolina, USA (from Krabbenhoft et al. 2009). A) Twelve anatomical landmarks used to estimate shape (shown on a *Fundulus* specimen). B) A principal component plot (based on the covariance matrix of group means). Axes indicate the amount of inter-group variation explained by the principal component. Lines connect shape estimates at small (small, open symbols) and large (large, filled symbols) centroid sizes. Dashed lines are for Lake Waccamaw species; solid lines are for stream species. *Etheostoma* is shown as circles; *Fundulus* is shown as squares; and *Menidia* is shown as triangles. C) Transformation grids for corresponding points in the PC plot in B. Individual shapes not shown for ease of interpretation (but see Krabbenhoft et al. 2009).

It was found that genus-specific allometric differences were not consistent across genera (Fig. 3B). *Etheostoma* had a significantly higher rate of allometric shape change in the lake environment (Δd = 0.026, p = 0.0033) but allometry vectors did not significantly differ from parallel (θ = 43.65°, p = 0.5278). *Menidia* species differed significantly in vector direction (θ = 75.08°, p = 0.0198), but not length (Δd = 0.003, p = 0.9128). *Fundulus* species differed neither in length (Δd = 0.004, p = 0.3336) nor direction (θ = 30.96°, p = 0.0884). Although intergenus allometry comparisons were not statistically evaluated (as they occupied clearly different locations in tangent space), a principal component plot (Fig. 3B) illustrated that the three genera had different allometric patterns. Comparatively, *Fundulus* exhibited smaller rates of shape allometry, and *Etheostoma* and *Menidia* differed in general allometry directions (as defined by the first two PCs). Transformation grids confirmed that *Fundulus* exhibited comparatively little shape al-

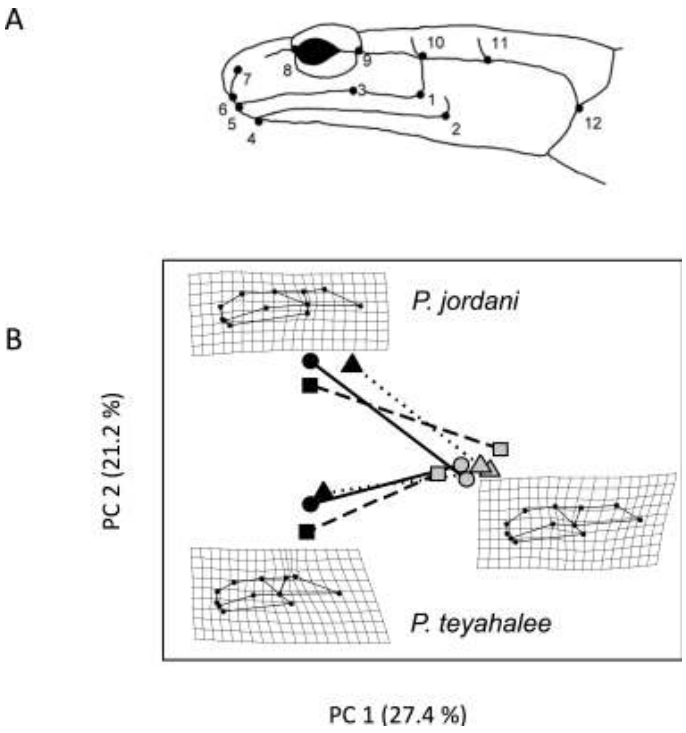


Figure 4 – Visualization of shape change vectors for three transects of repeated interspecific competition (from Adams 2010). A) The twelve anatomical landmarks used to estimate shape shown for a specimen of *Plethodon.* B) A principal component plot of shape variation (based on the covariance matrix of all individual shapes) for the two competing species of plethodontid salamanders: *Plethodon jordani* and *P. teyahalee* (percent variation explained by principal components is noted). The shape change vectors are lines in the plot; line and symbol differences correspond to different transects. Sympatric localities are represented as black symbols, allopatric localities as grey symbols. Transformation grids are shown to help facilitate an understanding of shape differences (discussed in further detail in Adams 2010). Individual shapes not shown for ease of interpretation (but see Adams 2010).

lometry, *Menidia* body shapes elongated, and *Etheostoma* body shapes deepened with growth. Within genera, body shapes were slenderer in the lake environment (Fig. 3C).

Shape change across two levels of an ordered ecological or evolutionary variable (gradient)

As indicated above, allometric vectors can either describe shape change per unit of size change, or they can describe the difference in shape between "small" and "large" sizes. The latter is simply a rescaling of the former (and a description of shape change at the location of average group shapes in a principal component plot). However, in some cases, two states are not simply points on a continuum, but represent rather a distinct change in category (i.e., a different categorical state of a qualitative independent variable). This is quite common in ecological or evolutionary biology studies. Examples include, but are not limited to, sex (e.g., Collyer and Adams 2007), experimental treatment (e.g., Hollander et al. 2006), environment type (e.g., predator/non-predator, as in Langerhans et al. 2004) and community type (e.g., allopatry/sympatry, as in Adams 2004).

For the purpose of clarity, we define a shape change vector as a vector of difference in shape between two states, a and b, $\mathbf{\Delta y} = \hat{\mathbf{y}}_\mathbf{b} - \hat{\mathbf{y}}_\mathbf{a}$, where $\hat{\mathbf{y}}$ is an estimated $1 \times p$ vector of shape, based on a categorical independent variable that describes an ecological or evolutionary gradient. (The shape change vector is the same as the phenotypic change vector, PCV, described by Adams and Collyer 2009) Conceptually, there is no difference between this vector and $\mathbf{b}$, an allometry vector, when $\mathbf{b}$ is defined for two distinct states of size (Fig. 2B). However, as a convention, we use different nomenclature to indicate that $\mathbf{\Delta y}$ represents shape change between two categorical states and $\mathbf{b}$ represents shape change associated with a continuous variable, even if defined for two fixed points on a continuum. Similar also to allometric vectors, heterogeneity in $\mathbf{\Delta y}$ is signified by a significant factor interaction, which can be determined by a likelihood ratio test that compares models that include and lack a factor interaction (Collyer and Adams, 2007). As a vector in tangent space $\mathbf{\Delta y}$, has both a length and a direction, corresponding to the amount of shape change andthe covariation of shape variables associated with a change in state for the independent variable, respectively (Adams and Collyer, 2007; Collyer and Adams, 2007). For comparison of $\mathbf{\Delta y}$ between two or more groups, Δd and θ are calculated as before, for all pairwise comparisons, and used as test statistics, which are evaluated with sampling distributions generated from null models.

As an example, we use data originally reported in Adams (2010). This study examined whether repeated interspecific competition generated parallel evolutionary divergence in phenotypes across different geographic populations of salamanders in the genus, *Plethodon*, in the Great Smoky Mountain National Park, USA. Two species, *P. jordani* and *P. teyahalee* occur in sympatry at mid-elevations in this region, where *P. jordani* is found in allopatry at higher elevations and *P. teyahalee* in allopatry at lower elevations. Various studies have documented character displacement (specifically in head shape) in this genus in sympatric populations (Adams, 2004, 2010; Adams and Rohlf, 2000; Adams and Collyer, 2007), and there is a strong genetic component to head shape (Adams, 2011), indicating that selection is capable of generating heritable, microevolutionary changes in this complex multi-dimensional trait. In the Adams (2010) study, an examination of the patterns of shape change associated with a change from allopatry to sympatry was conducted for the natural experiment of three replicated occurrences of allopatry-sympatry-allopatry gradients for these two species on different mountains.

A total of 336 salamanders comprising two species in three transects (i.e., six groups) were used in this study. Each of the six species-transect groups was found in one of two localities (sympatry or allopatry), and locality was the independent variable that described shape change within groups. Twelve anatomical landmarks were used on salamander heads and jaws (Fig. 4A) to quantify shape. The separate subset method (Adams, 1999) was used to separately perform GPA on head and jaw configurations. Shape variables from both subsets were found using the thin-plate spline and standard uniform components (Bookstein, 1991), which were combined to form a total set of 18 shape variables (see Adams 2010 for more details). (Principal components of these shape variables – referred to as relative warps [see Adams et al. 2004] – are the same as principal components of Procrustes residuals orthogonally projected into tangent space [see Adams et al. this issue]). Shape change vectors were estimated for all six sympatry-allopatry cases, and compared within species to test for differences in vector length and direction. Attribute differences were evaluated from sampling distributions generated from 10000 random permutations (based on a null model that lacked coefficients for a species-transect interaction), and the null hypothesis $\Delta d = 0$ or $\theta = 0$ was rejected if the p-value of the observed attribute difference was less than an acceptable type I error rate of $\alpha = 0.05$.

The attribute differences, Δd, and θ, for all pertinent pairwise comparisons (among transects, within species), were not significantly different from one another; though vectors between species were significantly different in direction. Thus, one can conclude that all allometry-sympatry vectors were similar in length and direction, within species (Fig. 4B; see Adams 2010 for details). Furthermore, based on locations of species-transect means in the tangent space, it was clear that allopatric localities were similar in head shape, regardless of species, but species diverged in head shape in sympatry (Fig. 4B). Together, these results demonstrate that not only did character displacement occur between these two species, but also the evolutionary mode of displacement within species was consistent among different transects (i.e., parallel evolution).

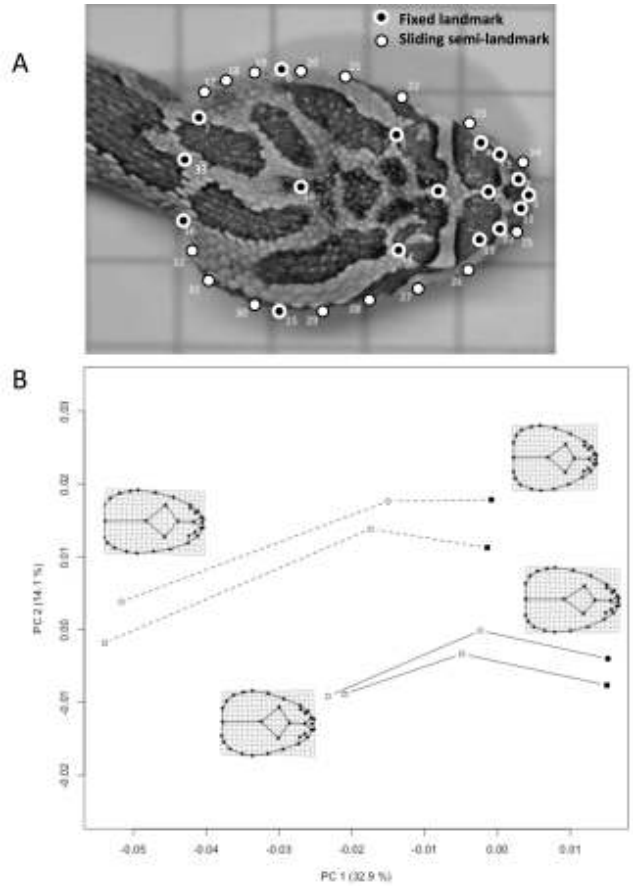


Figure 5 – Visualization of four shape trajectories for males and females from two different species of rattlesnakes (data from Davis 2012). A) The 33 landmarks used to estimate shape (fixed and sliding semi-landmarks are noted). B) Shape trajectories projected onto the first two principal components of between-group shape variation (i.e., based on covariance matrix of group means). Trajectories are shown as lines; *Crotalus viridis* trajectories are shown as dashed lines; *C. oreganus* trajectories are shown as solid lines. Filled symbols represent neonate stages (juvenile and adult can be inferred). Circles are female values; squares are male values. Transformation grids are added to facilitate an understanding of shape differences (corresponding to females end points on trajectories).

Shape change across multiple levels of an ordered ecological or evolutionary variable (gradient)

For studies that compare two or more groups that experience three or more levels of shape change associated with an ecological or evolutionary gradient, the set of geometric attributes of phenotypic trajectories increases, as well as the possible methods for describing their geometric attributes. However, the method of comparison among patterns of shape change is essentially the same (Adams and Collyer, 2009). For two ecological or evolutionary levels, shape change is described by a vector. When three or more levels are of interest, shape change forms a trajectory. These "phenotypic trajectories", in addition to location, size, and direction attributes, also have a shape (Adams and Collyer 2009; see Fig.1). (We use the term "phenotypic" rather than "shape" as a disambiguation between shape as phenotypic attribute of an organism and shape as a geometric attribute of trajectories. This disambiguation should become clear with the description of trajectory attributes below). The interpretations of differences in trajectory size and directional differences are also consistent with interpretations using shape change vectors. Trajectory size expresses the amount of shape change exhibited by a group associated with a change in ecological or evolutionary states; trajectory direction expresses the general covariation of shape variables associated with a change in ecological or evolutionary states (Adams and Collyer, 2009). The absolute difference in trajectory size and the angle between vectors of trajectory direction are values that can be evaluated from sampling distributions of random statistics (generated from null models).

Various measures of trajectory size could be used (see also the Discussion) and the choice of measure should consider the intent of the analysis. Adams and Collyer (2009) described trajectory path length – the summed lengths of vectors between sequential points in the trajectory – as a measure of trajectory size. Previous studies have also used path distance between sequential points (e.g., Dennis et al. 2011; Monnet et al. 2011; Turner et al. 2010; Frédérich et al. 2013). However, if one were interested in measuring the coverage of shape tangent space spanned by a trajectory, centroid size (Gower, 1971) or convex hull volume (Cornwell et al., 2006) might present logical alternatives for comparing the amount of shape space covered between multiple groups. (To our knowledge, no such method has yet been considered in studies that compare shape trajectories).

Adams and Collyer (2009) used the principal axis of variation among trajectory points (first principal component) as a description of trajectory direction. Large angles between principal components of compared groups indicate directional differences, which imply (under conditions that trajectories are located in similar regions of the tangent space) either ecological or evolutionary convergence or divergence (Stayton 2006; but see also Revell et al. 2007; Stayton 2006; Losos 2011; Dennis et al. 2011; Frédérich et al. 2013; Piras et al. 2012). Therefore, analyses of differences between size and direction attributes for phenotypic trajectories that describe shape change allow researchers to ascertain whether groups differ in e.g., the amount of morphological evolution, the direction of morphological evolution, or both.

The third geometric attribute of phenotypic trajectories (excluding location, as it does not describe shape change) is trajectory shape. Understanding the shape of phenotypic trajectories, and what differences in trajectory shapes describe is an area requiring additional research. Differences in trajectory shape are found as the Procrustes distances (Bookstein, 1991) between pairs of phenotypic trajectories (for details see Adams and Collyer 2009). Procrustes distances can function like test statistics to test for shape differences in trajectories among different groups. (We use D_p to denote Procrustes distance, to differentiate it from d, the distance [length] of either a shape vector or phenotypic trajectory). Unlike differences in trajectory size and orientation, differences in trajectory shapes are more challenging to interpret biologically. Differences in trajectory shape imply that, across ecological or evolutionary levels, changes in shape are accelerated or decelerated in one group relative to another, or are orientated in different directions, or both, in one portion or multiple portions of the trajectories (Fig. 1). Differences in trajectory shapes imply that there is a signal that some unique stage (or time) specific shape change is occurring. A potential useful exercise after concluding that trajectories differ in their shape is to perform qualitative pairwise comparisons of shape change vectors (or even a statistical test analogous to testing the attributes of shape change vectors), for the $k - 1$ vectors between the k points that sequentially comprise the trajectories.

As an example of phenotypic trajectory comparison, we use data originally reported by Davis (2012). These data comprise 3107 specimens of rattlesnakes in the genus, *Crotalus*. A total of 33 landmarks were digitized on the dorsal side of *Crotalus* heads from museum collections (Fig. 5A; see Davis 2012 for museum information). Although nine subspecies were analyzed in the original work, we only consider differences in ontogenetic trajectories (described for neonate-juvenile-adult sequences) between two species – prairie rattlesnakes (*C. viridis*) and western rattlesnakes (*C. oreganus*) – that have large, overlapping distributions in North America (i.e., all subspecies were pooled within species). We also consider whether there was sexual dimorphism in ontogenetic shape change in either or both species. Thus, there were four species-sex groups, each described by shape trajectories comprised of three ontogenetic stages. The three attributes of shape trajectories compared included path length, direction, and shape. The attribute differences, Δd, θ, and D_p, for all pertinent pairwise comparisons (between sex within species; between species within sex), were considered significant if the p-values from 10000 random permutations were less than an acceptable type I error rate of $\alpha = 0.05$.

Table 1 – Attribute differences, standardized scores, and *p*-values from the rattlesnake example. Pearson product-moment correlations between *p*-values and either attribute differences or standardized scores are also shown.

	Size			Direction			Shape		
Comparison	**Δd**	*Z*	*p*	**Δd**	*Z*	*p*	**Δd**	*Z*	*p*
FO-FV	0.0119	4.5368	0.0068	16.4264	5.4303	0.1342	0.2100	3.8535	0.0385
FO-MO	0.0060	3.0309	0.0662	14.4641	5.7164	0.0906	0.1568	3.5187	0.0680
FO-MV	0.0124	4.0199	0.0164	13.7274	3.9769	0.5334	0.2195	3.6074	0.0557
FV-MO	0.0178	6.6777	0.0001	23.9324	7.7021	0.0035	0.2949	5.2896	0.0033
FV-MV	0.0005	0.1660	0.9191	9.7181	2.5198	0.9890	0.0150	0.2351	0.9868
MO-MV	0.0184	5.9015	0.0007	22.7301	6.3988	0.0301	0.2971	4.8996	0.0059
correlation with *p*-value	-0.7926	-0.8542		-0.8165	-0.9388		-0.8898	-0.9386	

Both species were sexually dimorphic in terms of the amount of shape change associated with ontogeny (*C. oreganus*: Δd = 0.0060, p = 0.0001; *C. viridis*: Δd = 0.0005, p = 0.0005) (Fig. 5B). Males exhibited greater shape change in *C. viridis*, although the difference was small. Females exhibited a slightly larger amount of shape change in *C. oreganus*. Despite these results, males and females were more similar in shape at any stage of development compared to inter-stage variation in head shape, within sex (Fig. 5B). The only pertinent significant difference in directional shape change was between species, within males (θ = 22.73°, p = 0.0288), but this result was only slightly significant and the angle between principal directions was small. Therefore, the direction of shape changed was largely consistent both between males and females, and between species. *C. viridis* exhibited significantly more shape change during ontogeny (Females: Δd = 0.0119, p = 0.0069; Males: Δd = 0.0184, p = 0.0005), especially between juvenile and adult stages, confirming that differences among trajectory shapes were because of different ontogenetic patterns between species, within sex (Females: D_p = 0.21; p = 0.0405; Males: D_p = 0.21, p = 0.0048). Trajectories did not differ in shape between males and females, within species (*C. oreganus*: D_p = 0.16; p = 0.0643; *C. viridis*: D_p = 0.15; p = 0.9862).

In summary, sexual dimorphisms were small and only pertained to minor amounts of shape change, but species differed substantially in the amount and shape of shape change. These attribute differences correspond to accelerated shape change for *C. viridis* between juvenile and adult stages, compared to *C. oreganus*.

It is worth commenting that the three examples presented here were increasingly more complex in terms of the shape change gradient considered, but the analyses performed were all exactly the same. This analysis, called phenotypic trajectory analysis (PTA) is the pairwise comparison of geometric attributes – size, direction, shape – of phenotypic trajectories, and it is performed the same, irrespective of the number of phenotypic states in the trajectories. As discussed by Adams and Collyer (2009), two-state shape change is a simple case of multi-state shape change. The path length of a single vector is the vector length; the first principal component of a within-group covariance matrix is the vector that describes the difference between two states; and a vector has no shape (thus there are no trajectory shape differences among groups). Therefore, performing PTA to compare taxa that have two or more estimable shapes, corresponding to important ecological or evolutionary states, works the same for any number of sequential points in shape trajectories. Furthermore, as stated above, allometry vectors can be described as two-state shape change vectors. Although we did not discuss polynomial models (Rencher and Schaalje, 2008) as descriptions of shape allometry, one could imagine generating also multi-state trajectories to describe (potentially) non-linear shape allometries. Thus, analysis of the geometric attributes of shape trajectories is a generalized method for the comparison of any shape change associated with either qualitative or quantitative independent variables that describe important ecological or evolutionary gradients.

Qualitative comparison of geometric attributes: standardized scores of attribute differences

We have presented differences between geometric attributes of shape change as measures that function as test statistics, since the sampling distributions of the geometric attributes are created by a resampling procedure, which allows *p*-values to be estimated by the percentiles of observed attribute differences in the distributions. However, geometric attribute differences are not test statistics in the sense that they do not convey any information about the magnitude of the measure in relation to the variability of the measure. Much like many descriptive statistics can be expressed as standardized scores, geometric attribute differences can be converted to standardized scores using the standard deviations from their sampling distributions. For example, standardized scores of angles can be calculated as

$$Z_\theta = \frac{\theta - E[\theta]}{\hat{\sigma}_\theta} \quad (1)$$

where $\hat{\sigma}_\theta$ is standard deviation of angles between vectors, as estimated from the empirical distribution generated from a resampling procedure. Because the expected angle $E[\theta] = 0$ under the null hypothesis of parallel vectors, the standardized score simplifies to $Z_\theta = \theta/\hat{\sigma}_\theta$. Likewise, $Z_{\Delta d} = \Delta d/\hat{\sigma}_{\Delta d}$ and $Z_{D_p} = D_p/\hat{\sigma}_{D_p}$, for differences in trajectory size and shape, respectively. It should be noted that we use the variable, Z, as a convention because these statistics are similar to standard deviates, but we do not wish to imply that the standard normal distribution is used to find the probability of observed attribute differences. One might also think of Z as a measure analogous to Cohen's (1988) description of the standardized effect size (see also Sokal and Rohlf 2012), using the standard deviation of a sampling distribution of a statistic – i.e., the standard error – rather than a population standard deviation of a variable. Using the standard deviation of a sampling distribution (from the resampling experiment that generates random attribute differences) is necessary with the geometric attribute differences described here, because all attribute differences are scalars although the attributes, themselves, are either scalars (trajectory size), vectors (trajectory direction), or matrices (trajectory shape). Thus, although not directly analogous to Cohen's (1988) standardized effect size, the standardized scores are measures that are positively correlated to effect size. Finally, it should also be noted that we describe standardized scores based on the population of random permutations used to generate sampling distributions. One could argue that the number of resampling events is a sample of (the population of) all possible random outcomes, and standardized scores are, therefore, more analogous to t-statistics than z-statistics. However, this implication would only be a concern for unreasonably small sets of resampling iterations and would not change the important point that standardized scores are positively correlated with effect size.

Standardized scores allow one to qualitatively compare attribute differences within and between studies. This is especially useful when the value of a geometric attribute difference is unintuitive in terms of the outcome of a null hypothesis test. For example, large angles (like in the example above of comparison of allometry vectors for sister fish taxa occurring in streams and Lake Waccamaw) might not be significantly different from 0°, but small angles (like in the example above of comparison of shape trajectories between males of different species of *Crotalus*) can be significantly different from 0°, if the standard deviation of random angles is small. It is also useful if one wishes to qualitatively compare effect sizes among different geometric attributes of shape change, within the same study.

As an example, we calculated standardized scores for each of the six possible pairwise comparisons of the four groups in the *Crotalus*

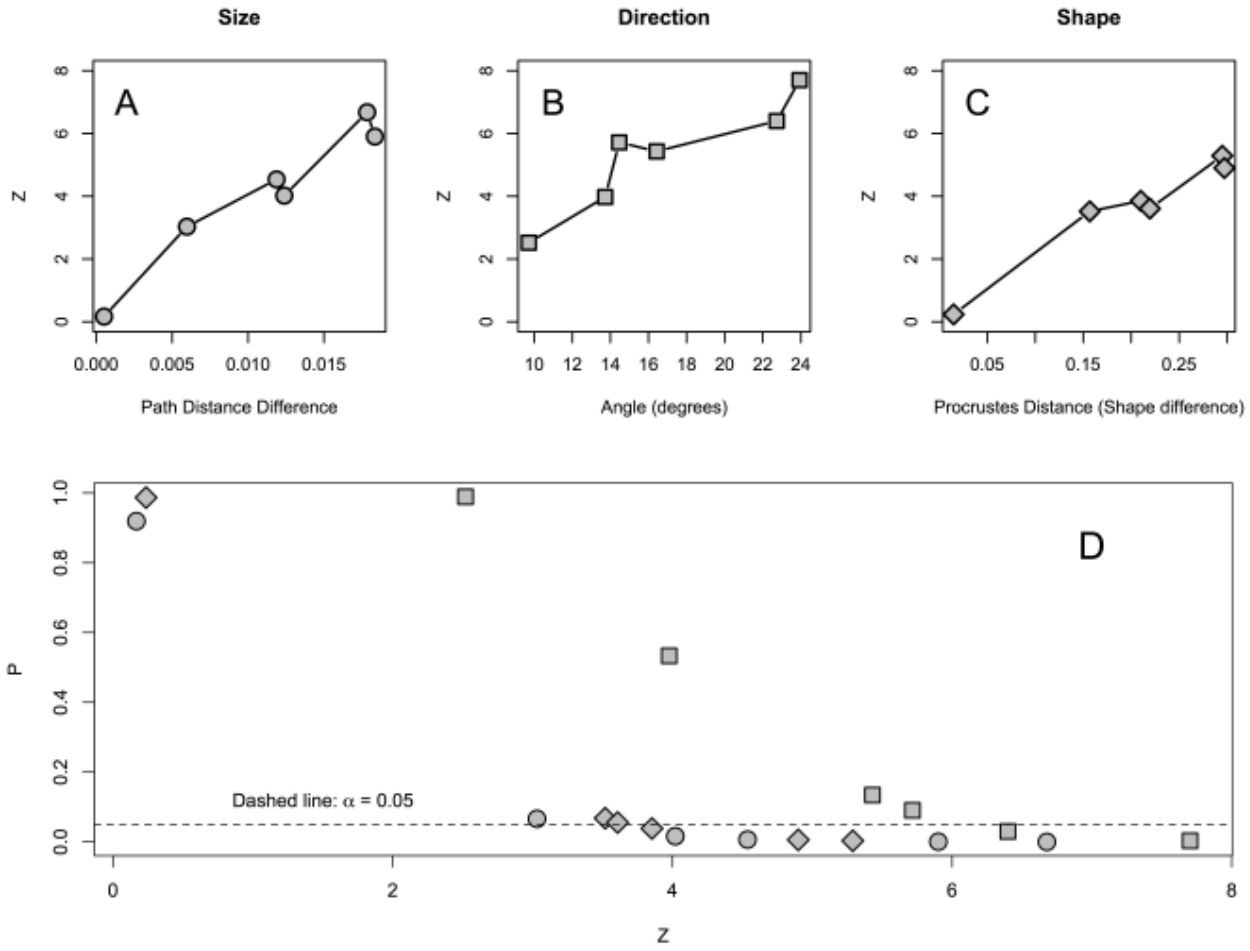


Figure 6 – Relationship between trajectory attribute differences and their standardized scores, Z, for the rattlesnake example (values given in Tab. 1). Standardized scores are plotted versus attribute differences in panels A-C. Lines connect the rank order of values for attribute differences. (A negative sloping line indicates a decrease in standardized score in spite of an increase in attribute difference). The same standardized scores are shown on the abscissa in the plot in panel D, with p-values from phenotypic trajectory analysis shown as the ordinate. Symbols in D correspond to the same symbols in A-C.

example of the previous section, for each of the geometric attributes considered. The standardized scores are shown alongside the original values and p-values in Tab. 1, and plotted in Fig. 6 to illustrate the association between standardized scores and original values, as well as their p-values. For this example, two things are clear. First, although larger attribute differences result in generally larger standardized scores, the rank orders of attribute differences and standardized scores were not exactly the same, meaning that in comparison, a larger standardized score can be found from a smaller attribute difference, or vice versa. The standardized scores presented here scale geometric attribute differences by the inverses of their standard deviations, as found from their sampling distributions. Thus, they function more as test-statistics that can be qualitatively compared within and between studies. In this example, standardized scores were more strongly correlated with p-values than attribute differences (Tab. 1), which suggests that there is some benefit to reporting them along with observed attribute differences. Second, although standardized scores might be more strongly correlated with p-values, one must use caution when comparing them between different attributes. In our example, standardized scores for directional diffrences were larger compared to size or shape differences, in general, but the null hypothesis was rejected fewer times. We suspect that this phenomenon results from using only the first principal components of randomly-generated within-trajectory covariance matrices, which might negatively bias variances of angular differences, thus inflating standardized scores. The relationship between standardized scores of geometric attribute differences and shape space dimensionality requires further research.

Discussion

As Thompson (1917) stated as an important prelude to his "Theory of transformations", in comparative studies it's the description of morphological change between two states that is intriguing rather than the precise definition of morphology in either state. Users of GM methods can appreciate that measuring shape change between two landmark configurations in terms of their Procrustes distance and the total shape change implied by it (as visualized by the thin-plate spline, Bookstein 1989), is intuitively more appealing to interpret than the meaning of the position of landmarks in one landmark configuration. Likewise, coefficients from a linear model, which describes shape variation as a function of one or more independent variables, measure the amount of expected shape change per unit change of an independent variable. As we have shown, that shape change forms a path in shape space, and has two important attributes: a size and a direction. Size is the amount of shape change associated with a per unit change in an independent variable; direction is the covariation of shape variables associated with change in an independent variable. When linear models also contain factors to describe ecological or evolutionary gradients, shape change is the cumulative sequence of vector changes in shape associated with the gradient, forming a trajectory. Therefore, a third attribute, trajectory shape, describes shape change for shape that is measured or estimated for more than two states. (Our examples all included shape data, but equally viable examples could have used other phenotypic data).

Most standard multivariate statistical analyses used in ecological and evolutionary research use linear models to estimate parameters of phenotypic change. However, standard hypothesis tests do not specifically evaluate attributes of phenotypic change, which are viable test statist-

ics themselves. Using the four species-four experimental temperatures example presented in the introduction, a factorial MANOVA evaluates the species-temperature interaction by estimating the probability of a test statistic for the comparison of two log-likelihoods for two different linear models. One model contains coefficients for just the species and temperature factors; the other contains coefficients for these factors, plus coefficients for their interaction. A large difference in log likelihoods indicates that model error is significantly reduced by including the extra coefficients (Manly and Rayner, 1987; Rencher and Schaalje, 2008). This implies that coefficients for the interaction are important, and a null hypothesis of no interaction effect is rejected (i.e., the interaction is a significant source of variation). For analyses that seek to test differences in location (e.g., species differences in shape), there is a direct connection between these test statistics and null hypotheses. We refer to such tests as tests of "static" variation – tests that evaluate variation among group locations in multivariate data spaces. By contrast, geometric attributes of phenotypic trajectories are test statistics for tests of "dynamic" variation – tests that evaluate variation among group changes in location in the multivariate data space. Standard multivariate tests might imply that dynamic variation is meaningful (e.g., a factor interaction is significant); but tests using the geometric attributes, themselves – PTA – are more direct and explain why dynamic variation is significant.

It is a safe assertion that many hypotheses in ecology and evolutionary biology describe patterns of phenotypic change, yet the standard approach to assessing these patterns is incomplete. For example, in the *Plethodon* salamander example above, a factorial MANOVA examining patterns of head shape variation revealed that a species × transect × locality interaction was significant (see Adams 2010). But how does this result confer any knowledge about the repeatability of evolutionary divergence? Alone, it does not. In fact, identifying a significant interaction term in a MANOVA is insufficient to discern among the many alternative explanations that may have generated this observed pattern. For instance, a significant interaction could be observed if there was heterogeneity among transects within species. However, by linking phenotypes across the ecological gradient, one forms trajectories of phenotypic change (in this case from allopatry to sympatry). Then, an explicit comparison of trajectory attributes using PTA allowed for direct tests of shape change and therefore of its underlying biological processes; in this case, identifying evolutionary parallelism of phenotypic change (Adams 2010, for examples of convergence see: Adams and Nistri 2010; Piras et al. 2010).

Another advantage of PTA is that it offers great flexibility and can be used for any sequence of shapes in tangent space that form a trajectory. Although trajectories tend to have logical sequences associated with ecological or evolutionary gradients, PTA works as efficiently for "configurations" of shapes in tangent space that are less logical as a sequence, but still correspond across groups. For example, one might study the shape responses of groups of organisms in an experiment or observational study due to different predator types (no predator, predator species A, predator species B, etc.). Here imposing an order or sequence to the levels is arbitrary, yet PTA can be used on the original unordered data to determine whether phenotypic responses to predators differ among groups (e.g., Hollander et al. 2006). Additionally, in many circumstances, PTA provides a complementary approach for describing patterns of change where alternative measures are more commonly used. For example, in community ecology, various dispersion metrics have been proposed to describe patterns in stable isotope data, as well as for food webs (Layman et al., 2007). Yet in these cases, trajectory analysis has proven to be a valuable complementary tool for identifying spatial and temporal patterns in isotope data spaces unexamined by standard approaches (e.g., Turner et al. 2010).

Despite the clear advantages of phenotypic trajectory analysis, PTA is still a recently developed tool. Thus, a number of issues remain to be addressed in terms of how it is implemented under particular circumstances. We feel that the following three issues will present the greatest challenges and the most interesting discoveries in the coming years. First, for certain types of data, different research designs, or different ecological or evolutionary gradients, are better alternative measures for trajectory attributes available? The initial conception of PTA was motivated by using sequential points along ecological and evolutionary gradients; thus the most appropriate size measure for such designs was path distance (Adams and Collyer, 2009). The benefit of this measure is that it produces large values from oscillatory shape changes, something centroid size or convex hull volume might fail to identify. Alternatively, if one were more interested in the amount of data space covered by a species over an ecological gradient, centroid size (Gower, 1971) or convex hull volume (Cornwell et al., 2006) might be better alternatives. With respect to directional differences, an alternative could be to analyse the angle between vectors that describe the difference between starting and end points, within groups, rather than comparing the first principal components of each. We anticipate that assessments of alternative trajectory attribute measures will be valuable in the future, especially as additional research questions prompt the need for alternative measures.

Second, for shape change across continuous variables, can PTA be used with polynomial regression or non-linear regression analyses? If so, how would such an analysis be optimized? In the discussion of allometry, we indicated that allometry vectors could also be described as two-state change vectors (between "small" and "large" size). Clearly, this formulation is a linear transformation (rescaling) of the original vector, so sampling distributions of attribute test statistics remain unchanged when the underlying trajectory is linear. However, ontogenetic trajectories can be decidedly non-linear (Mitteroecker et al., 2004), and in these circumstances, it is less obvious what quantitative representation should be utilized. For instance, one could use estimates of shape from polynomial models to form a trajectory for nonlinear allometry considerations, or describe trajectories in size-phenotype spaces (e.g., Mitteroecker et al. 2004), but how many points along the polynomial regression should the trajectory contain? Further, it remains unknown how changing the number of trajectory points would alter the sampling distributions of test statistics, and thereby affects biological interpretation. Finally, we could envision that PTA could be combined in some way with other methods for quantifying ontogenetic trajectories (e.g., the common allometric component and its residual shape variation: *sensu* Mitteroecker et al. 2004), where PTA would provide statistical tests for if, and in what manner, ontogenetic trajectories vary. Additional work is needed in this area.

Finally, what do trajectory shapes, or more specifically, differences in trajectory shapes, mean biologically? From our experience, trajectory shape differences are interpretable if trajectories have few points and are located in similar locations. Also, differences in trajectory shapes can be interpretable in the case of motion analysis (Adams and Cerney, 2007), as they describe how a motion is performed (e.g., straight-arm movements versus back-and-forth movements will generate trajectories of different shapes in shape space). However, interpreting differences in trajectory shapes when trajectories have many points or when trajectories are located in largely different regions of data spaces makes inferences more challenging (Collyer and Adams, unpublished data). Because trajectories are sequences of vectors, one has to question whether a significant trajectory shape difference implies that the entire trajectory differs in shape, or whether such patterns are isolated to particular portions of the trajectory. For instance, across trajectories representing multiple ecological or evolutionary levels, differences in trajectory shape may represent accelerated or decelerated shape change in one group relative to another, or shape changes orientated in different directions, and these may be found in one portion or multiple portions of the trajectories. Thus, for many studies, identifying significant differences in trajectory shape may represent the first step in a more in-depth and pairwise assessment of where trajectories differ, and how. As such, the shape of a trajectory might not be so much a global trajectory attribute, as it is a summary of size and directional attributes of the vectors that comprise it. More work is needed in this area.

In looking toward the future, regarding development of methods for analysis of patterns of shape change, it is perhaps instructive to look

back over the past few decades of development of GM methods. The field has certainly changed. In the past twenty years, the use of GM methods has dramatically increased, as the "Procrustes paradigm" has developed from one alternative approach of shape quantification into a rigorous discipline (Adams et al., this issue). New users of GM methods now take for granted, for example, that relative warps are typically principal components of unweighted partial warp scores (Zelditch et al. 2004, p. 423) or simple principal components of Procrustes residuals orthogonally projected into tangent space. They might not realize or appreciate the amount of attention once paid to weighting principal warps prior to generating partial warps and relative warps to develop alternative "biomathematical" strategies (e.g., Bookstein 1996). Likewise, we suspect that if another synthetic volume of progress in GM development is created ten years from now, some or all of the challenges listed in the previous paragraph will have since been rectified. Rather, the body of work on analyses of patterns of shape change – including PTA and alternative methods – will be much more comprehensive, and the analytical idiosyncrasies yet to be resolved will be fleshed out by the research questions that implore further discovery.

References

Adams D.C., 1999. Methods for shape analysis of landmark data from articulated structures. Evol. Ecol. Res 1: 959–970.

Adams D.C., 2004. Character displacement via aggressive interference in Appalachian salamanders. Ecology. 85: 2664–2670.

Adams D.C., 2010. Parallel evolution of character displacement driven by competitive selection in terrestrial salamanders. BMC Evol. Biol. 10: 1–10.

Adams D.C., 2011. Quantitative genetics and evolution of head shape in *Plethodon* salamanders. Evol. Biol. 38: 278–286.

Adams D.C., Cerney M.M., 2007. Quantifying biomechanical motion using Procrustes motion analysis. J. Biomech. 40: 437–444.

Adams D.C., Collyer M.L., 2007. The analysis of character divergence along environmental gradients and other covariates. Evolution 61: 510–515.

Adams D.C., Collyer M.L., 2009. A general framework for the analysis of phenotypic trajectories in evolutionary studies. Evolution 63: 1143–1154.

Adams D.C., Nistri A., 2010. Ontogenetic convergence and evolution of foot morphology in European cave salamanders (Family: Plethodontidae). BMC Evol. Biol. 10: 1–10.

Adams D.C., Rohlf F.J., 2000. Ecological character displacement in *Plethodon*: biomechanical differences found from a geometric morphometric study. Proceedings of the National Academy of Sciences, USA 97: 4106–4111.

Adams D.C., Rohlf F.J., Slice D.E., 2004. Geometric morphometrics: ten years of progress following the "revolution". It. J. Zool. 71: 5–16.

Adams D.C., Rohlf F.J., Slice D.E., 2013. A field comes of age: geometric morphometrics in the 21st Century. Hystrix 24(1) (Online First) doi:l0.4404/hystrix-24.l-6283

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

Bookstein F.L., 1996. Biometrics, biomathematics and the morphometric synthesis. Bulletin of Mathematical Biology 58: 313–365.

Bookstein F.L., 1991. Morphometric tools for landmark data: Geometry and Biology. Cambridge Univ. Press, New York.

Box G.E.P., Draper N.R., 1987. Least squares for response surface work. In: Box G.E.P., Draper N.R., 1987. Empirical model building and response surfaces. John Wiley & Sons, New York. 34–103

Cohen J., 1988. Statistical power analysis for the behavioural sciences, 2nd edition. Erlbaum, Hillsdale, NJ.

Collyer, M. L., and D. C. Adams. 2007. Analysis of two-state multivariate phenotypic change in ecological studies. Ecology 88:683-692.

Cornwell W.K.,Schwilk D.W., Ackerly D.A., 2006. A trait-based test for habitat filtering: convex hull volume. Ecology 87: 1465–1471.

Davis M.A., 2012. Morphometrics, molecular ecology, and multivariate environmental niche define the evolutionary history of the western rattlesnake (*Crotalus viridis*) complex. Ph. D. Dissertation, University of Illinois, Champaign, IL.

Dennis S.R., Carter M.J., Hently W.T., Beckerman A.P., 2011. Phenotypic convergence along a gradient of predation risk. Proc. Roy. Soc. B-Biol. Sci. 278: 1687–1696.

Drake A.G., Klingenberg C.P., 2008. The pace of morphological change: Historical transformation of skull shape in St Bernard dogs. P. Roy. Soc. B-Biol. Sc. 275:71–76.

Frédérich B., Sorenson L., Santini F., Slater G.J., Alfaro M.E., 2013. Iterative ecological radiation and convergence during the evolutionary history of damselfishes (Pomacentridae). Am. Nat, 181: 94–113.

Gower J.C., 1971. Statistical methods of comparing different multivariate analyses of the same data. In: Hodson F.R., Kendall D.G., Tautu P. (Eds.) Mathematics in the archaeological and historical sciences. Edinburgh Univ. Press, Edinburgh. 138–149.

Hollander J., Collyer M.L., Adams D.C., Johannesson K., 2006. Phenotypic plasticity in two marine snails: constraints superseding life-history. J. Evol. Biol. 19: 1861–1872.

Hotelling H., 1931. The generalization of the student's ratio. Ann. Math. Stat. 2: 360–378.

Hubbs C.L., Raney E.C., 1946. Endemic fish fauna of Lake Waccamaw, North Carolina. Misc. Pu. Univ. of Mich. Museum Zool. No. 65. Ann Arbor, MI.

Klingenberg K.P., 1996. Multivariate allometry. In: Marcus L.F., Corti M., Loy A., Naylor G.J.P., Slice D.E. (Eds.) Advances in morphometrics. Plenum Press, New York. 23–49.

Klingenberg K.P., 1998. Heterochrony and allometry: the analysis of evolutionary change in ontogeny. Biol. Rev. 73: 79–123.

Krabbenhoft T.J., Collyer M.L., Quattro J.M., 2009. Differing evolutionary patterns underlie convergence on elongate morphology in endemic fishes of Lake Waccamaw, North Carolina. Biol. J. Linn. Soc. 3: 636–645.

Langerhans R.B., Layman C.A., Shokrollahi A.M., DeWitt T.J., 2004. Predator-driven phenotypic diversification in *Gambusia affinis*. Evolution 58: 2305–2318.

Layman C.A., Arrington D.A., Montana C.G., Post D.M., 2007. Can stable isotope ratios provide for community wide measures of trophic structure? Ecology 88: 42–48.

Losos J.B., 2011. Convergence, adaptation, and constraint. Evolution 65: 1827–1840.

Manly B.F.T., Rayner J.C.W., 1987. The comparison of sample covariance matrices using likelihood ratio tests. Biometrika 74: 841–847.

Marcus L.F., 1993. Some aspects of multivariate statistics for morphometrics. In: Marcus L.F., Bello E., Garcia-Valdecasas A. (Eds.) Contributions to morphometrics. Monografías del Museo Nacional de Ciencias Naturales 8, Madrid. 95–130.

Mitteroecker P., Gunz P., Bernhard M., Schaefer K., Bookstein F.L., 2004. Comparison of cranial ontogenetic trajectories among great apes and humans. J. Hum. Evol. 46: 679–698.

Mitteroecker P., Gunz P., Bookstein F.L., 2005. Heterochrony and geometric morphometrics: a comparison of cranial growth in *Pan paniscus* versus *Pan troglodytes*. Evol. Develop. 7: 244–258.

Monnet C., De Baets K., Klug C., 2011. Parallel evolution controlled by adaptation and covariation in ammonoid cephalopods. BMC Evol. Biol. 11: 115.

Piras P., Colangelo P., Adams D.C., Buscalioni A., Cubo J., Kotsakis T., Meloro C., Raia P., 2010. The Gavialis-Tomistoma debate: the contribution of skull ontogenetic allometry and growth trajectories to the study of crocodylian relationships. Evol. Develop. 12: 568–579.

Piras P., Sansalone G., Teresi L., Kotsakis T., Colangelo P., Loy A., 2012. Testing convergent and parallel adaptations in talpids humeral mechanical performance by means of geometric morphometrics and finite element analysis. J. Morphol. 273: 696–711.

Rencher A.C., 2002. Methods of multivariate analysis, 2nd edition. Wiley Series in Probability and Statistics, John Wiley and Sons, Hoboken, NJ.

Rencher A.C., Schaalje C.B., 2008. Linear models in statistics, 2nd edition. John Wiley & Sons, Hoboken, NJ.

Revell L.J., Johnson M.A., Schulte J.A., Kolbe J.J., Losos J.B., 2007. A phylogenetic test for adaptive convergence in rock-dwelling lizards. Evolution 61: 2898–2912.

Rohlf F.J., Corti M., 2000. The use of partial least-squares to study covariation in shape. Syst. Biol. 49: 740–753.

Rohlf F.J., Slice D.E., 1990. Extensions of the Procrustes method for the optimal superimposition of landmarks. Syst. Zool. 39: 40–59.

Simpson G.E., 1944. Tempo and mode in evolution. Columbia Univ. Press, New York.

Sokal R.R., Rohlf F.J., 2012. Biometry, 4th edition. W. H. Freeman, New York.

Stayton C.T., 2006. Testing hypotheses of convergence with multivariate data: morphological and functional convergence among herbivorous lizards. Evolution 60: 824–841.

Thompson D.W., 1917. On Growth and Form. Cambridge, London.

Turner T.F., Collyer M.L., Krabbenhoft T.J., 2010. A general hypothesis-testing framework for stable isotope ratios in ecological studies. Ecology 88: 2227–2233.

Viscosi V., Loy A., Fortini P., 2010. Geometric morphometric analysis as a tool to explore covariation between shape and other quantitative leaf traits in European white oaks. In: Nimis P.L., Vignes L.R. (Eds.) Tools for identifying biodiversity: progress and problems. EUT Edizioni Università di Trieste. 257–261.

Zelditch M. L., Swiderski D.L., Sheets H.D., Fink W.L., 2004. Geometric morphometrics for biologists. Elselvier, San Diego, CA, USA.

Associate Editor: A. Loy