Published by Associazione Teriologica Italiana
Hystrix, the Italian Journal of Mammalogy
Volume 24 (1): 25–32, 2013
Available online at:
http://www.italian-journal-of-mammalogy.it/article/view/6282/pdf
doi:10.4404/hystrix-24.1-6282

Research Article

Morphometrics and the comparative method: studying the evolution of biological shape

Leandro Rabello Monteiro[a,*]

[a] *Laboratório de Ciências Ambientais, CBB, Universidade Estadual do Norte Fluminense, Av. Alberto Lamego 2000, Pq. Califórnia, Campos dos Goytacazes, RJ, cep 28013-620, Brasil*

Keywords:
Adaptation
morphometrics
comparative methods
phylogenetic analysis
multivariate analysis

Article history:
Received: 27 May 2012
Accepted: 9 July 2012

Acknowledgements
The author would like to thank M.R. Nogueira, A. Cardini, J. Claude, and A. Loy for important comments and improvement suggestions on previous versions of the manuscript. Work by LRM is funded by the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) (bolsa de produtividade em pesquisa, processo 304118/2010-9) and Fundação Estadual de Amparo à Pesquisa do Rio de Janeiro (FAPERJ) (Cientista do Nosso Estado, processo E-26/103.044/2011).

Abstract

Phylogenetic comparative methods are one of the most important parts of the morphometric toolkit for studies of morphological evolution. The assessment of repeated independent events of evolution of phenotypic and associated ecological-functional traits is still a starting point for the study of adaptation, but modern comparative approaches go beyond correlative methods, allowing for the modeling of evolutionary scenarios and analyses of trait evolution patterns. The evidence for adaptive change due to ecological diversification is stronger (even if still circumstantial) if models that predict increases in diversification rate fit the data well and the morphological changes are associated with ecological and functional changes. A large body of literature is dedicated to methodological and theoretical aspects of comparative methods, but in the context of univariate traits. On the other hand, biological shape is a complex trait, and morphometric data is essentially multivariate. Whereas most comparative methods allow for direct multivariate extensions, dimension reduction is an almost certain requirement due to the high dimensionality of morphometric data sets and the large number of evolutionary parameters that need to be estimated by comparative methods. Objective methods with considerable statistical support to determine data dimensionality exist, but the applied literature usually relies on subjective criteria to assess how many shape dimensions should be retained. The most appropriate calculation and interpretations of principal components, the most popular dimension reduction method, are also topics that should be considered more carefully in applications. The maturity of comparative methods and the development of model-based approaches linking macroevolutionary patterns and microevolutionary processes provide an exciting perspective for the study of morphological evolution.

Introduction

The study of interspecific comparative data sets has a long history in evolutionary biology. In this context, the recurrent association between phenotypic and ecological traits in different lineages is often used as evidence, even if circumstantial, of adaptation (Harvey and Pagel, 1991; Losos, 2011a), defined non-historically as the phenotypes that present higher fitness than other phenotypes in a given environment (Martins, 2000; Lewens, 2007). The strength of the evidence is related to the number of times that a given combination of phenotype and environment arises independently during the evolution of a lineage (Harmon et al., 2005). Felsenstein (1985) called attention to the fact that trait values for species at the tips of a phylogenetic tree are not necessarily independent pieces of evidence in statistical analysis, because part of the trait's evolutionary history is shared by common descent. Felsenstein pointed out issues with previous attempts to deal with the statistical problem of non-independence and proposed an elegant method to deal with the problem, known as phylogenetic independent contrasts.

A period of effervescence followed, when a number of methods were proposed for the statistical analysis of phylogenetic non-independent data (Cheverud et al., 1985; Grafen, 1989; Garland et al., 1993; Hansen, 1997; Martins and Hansen, 1997; Diniz-Filho et al., 1998). Despite a sharp increase in popularity, the interpretation of results from comparative analyses, the assumptions being made, and the appropriate combination of methods were not always straightforward. Many authors discussed aspects of application, interpretation or whether comparative methods were useful at all (Westoby et al., 1995; Björklund, 1997; Losos, 1999; Freckleton et al., 2002; Freckleton, 2009). According to the jargon commonly used in the early period, the comparative methods would sometimes transform the original variables "correcting" for phylogenetic non-independence (phylogenetic independent contrasts) or decompose variation into "phylogenetic" (autocorrelation) or "heritable" (mixed models) components versus specific or residual components. Another possibility was to incorporate phylogenetic non-independence in the residuals using a generalised linear model (phylogenetic generalised least squares – PGLS). The analogies between statistical results from the different models with evolutionary processes started to appear in the literature, such as the decomposition of variation into a phylogenetic (reflecting "historical constraints") and a specific component (reflecting "adaptation") via autocorrelation (Cheverud et al., 1985). In these early contributions, phylogenetic non-independence was interpreted as evidence of phylogenetic constraint (a concept similar to the modern definition of phylogenetic inertia – see below), and only correlations among specific values were accepted as evidence for the action of selection, both interpretations no longer considered valid (Hansen and Orzack, 2005). The multitude of methods, applications and interpretations generated some conflicts, particularly when different authors used the same concept names with different definitions (as in phylogenetic inertia – see below). A number of reviews related statistical properties, grouping methods and pointing out similarities, differences and limitations (Martins, 2000; Rohlf, 2001; Freckleton et al., 2002; Martins et al., 2002; Hansen and Orzack, 2005). A book summarised some of the early approaches (Harvey and Pagel, 1991), but the methodological and conceptual developments have been so massive during the last decade that new books on the subject (both entry level and advanced) are urgently needed again.

*Corresponding author
Email address: lrmont@uenf.br (Leandro Rabello Monteiro)

doi:10.4404/hystrix-24.1-6282

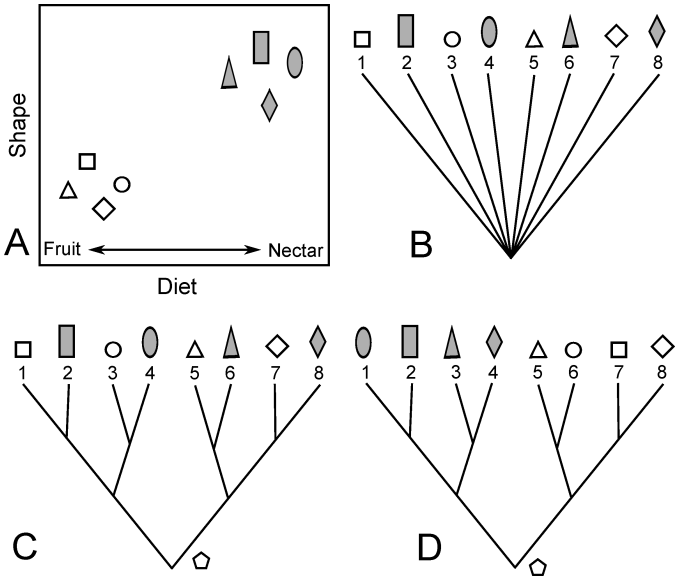

Figure 1 – Phylogenetic structure of a hypothetical correlational study, associating morphology and diet in a lineage of bats. A) Pattern of association between shape and diet, where one aspect of shape (elongation) is associated with dietary differences (relative contribution of fruit versus nectar, represented as shades filling the geometric figures). The axes relate to groups of multiple variables, but are represented in a bivariate scatterplot for simplicity. B) Star phylogeny pattern of covariance, where each OTU represents an independent piece of evidence. C) Phylogenetic structure where the evolutionary transition in diet and shape was observed four times independently. The assumed shape and diet of the ancestor (Frugivory, not elongated) is represented near the root of the tree. D) Phylogenetic structure where the evolutionary transition in diet and shape was observed only once.

Statistical approaches in comparative methods

The statistical problem tackled by comparative methods is represented in Fig. 1. When attempting to correlate two groups of variables (shape and diet) with values measured on OTUs (tips of the phylogeny – extant species), one might come to a result as in Fig. 1A, showing an association between shape (the elongation of the figures) and diet (relative importance of fruit and nectar). The correct interpretation of this pattern and its significance as evidence that natural selection might be responsible for the observed association depends on the phylogenetic covariance structure assumed for the residuals of this model. Common statistical methods (i.e. non-phylogenetic) assume that all observations are independent as in Fig. 1B. However, if the phylogenetic structure in Fig. 1C is assumed, the diet transition between frugivory and nectarivory was observed four times independently (the postulated ancestor was a frugivore), leading to a decrease in the number of degrees of freedom. The phylogenetic structure in Fig. 1D is what Felsenstein (1985) called a "worst case" phylogeny, where the diet transition was observed only once, and the four nectarivore species are not independent observations. The strength of circumstantial evidence for adaptation is greater if we assume the phylogeny in Fig. 1C than the one in Fig. 1D, but that is a statistical consequence of the reduced number of degrees of freedom. Even if we assume the phylogeny in Fig. 1D as true, it does not mean that the shape differences are "caused" or "constrained" by phylogeny.

Translating this problem into a working statistical model requires the definition of three components: hypotheses, sampling assumptions (random, independent?) and statistical assumptions (distribution of response variables) (McPherson, 1990). Of greater concern here are the sampling assumptions, as phylogenetic non-independence is the motivation behind the use of phylogenetic comparative methods. Considering the phylogeny in Fig. 1C as the actual sampling structure, the method of phylogenetic independent contrasts transforms the original values of shape and diet in a way that the resulting observations (contrasts) are standardised differences among tips or estimated nodes descending from the same immediate common ancestor (Felsenstein, 1985). Correlations of contrasts for different traits were expected to arise independently of similarities due to common ancestry. A statistically equivalent approach would be to use a generalised linear model that incorporates the structure of phylogenetic covariance among observations into the error term (Martins and Hansen, 1997; Rohlf, 2001), as can be done for non-independent spatial or time series data. In this case, the expected phylogenetic covariance for any pair of species would be proportional to the sum of branch lengths leading from the root to their last common ancestor. Using branch lengths to directly estimate phylogenetic covariances assumes that the response variables evolved according to a Brownian motion model (BM), where differences among species are proportional to the branch lengths leading to their most recent ancestor. This assumption, however, can be flexible, and other evolutionary models can be used by transforming branch lengths or species covariances to reflect alternative models, such as stabilising selection or allowing for varying rates of character evolution (Martins and Hansen, 1997; Freckleton et al., 2002; Blomberg at al., 2003).

Most applications of comparative methods assume that the species means used are representative and that within-species variation is negligible when compared to among-species variation (Garamszegi and Moller, 2010). Whereas this assumption might hold true for a number of studies, it has been shown that ignoring intraspecific variability can lead to increased type I error rates when intraspecific variances are large and sample sizes are small (Harmon and Losos, 2005). One concern for multivariate morphometric studies is the downward bias shown to influence estimates of correlation and regression coefficients when measurement error is large (Ives et al., 2007). Specific modifications of well known methods (phylogenetic independent contrasts, PGLS) were proposed to take measurement error (and within-species biological variability) into account when estimating parameters (Ives et al., 2007; Felsenstein, 2008), and the potential for these methods in morphometrics is great, both providing less biased estimates of covariances and correlations or in using maximum likelihood to compare observed and hypothetical covariance matrices or levels of variation (intraspecific, interspecific) (Felsenstein, 2008).

Early attempts to use quantitative genetics theory to discriminate between evolutionary processes came from the expectation of among-population covariances under genetic drift proposed by Lande (Lande, 1979, 1980) and the associated tests, comparing among and within population covariance matrices (Lofsvold, 1988; Ackermann and Cheverud, 2002), and more recently even incorporating phylogenetic non-independence (Revell, 2007). Modern comparative approaches use model selection to infer best fitting evolutionary models or scenarios from expected OTU means instead of among-group covariances (Butler and King, 2004; Hansen et al., 2008) and to estimate evolutionary parameters reflecting historical changes of evolutionary rates of trait diversification (O'Meara et al., 2006; Revell at al., 2012). Going back to the example discussed, one might then ask, given the data in Fig. 1A and the tree in Fig. 1D, whether the likelihood of the pattern of differences observed is higher under a Brownian motion model (which encompasses both neutral evolution or random fluctuations of adaptive optima over time) or under a stabilising selection model with two optima (for eating fruit or nectar), that function as attractors in the model and can be interpreted as fixed adaptive peaks. In this model-based approach, the focus moves from the number of independent branches showing the evolutionary transitions of interest, to the maintenance of adaptive optima under stabilising selection (Hansen, 1997). Because the evolutionary mechanisms are modeled directly, the methods also provide a way to estimate parameters relating postulated processes to observed interspecific patterns in the evolution of quantitative traits, such as the strength of selection towards an adaptive optimum for a character (Hansen and Orzack, 2005), or evolutionary divergence rates (a measure of phenotypic variation in the phylogeny) that might change along different branches (O'Meara et al., 2006; Revell at al., 2012).

Conceptual issues with comparative methods

The maturity of comparative methods have also provided clearer definitions for concepts that have been confused and misinterpreted for a long time, such as phylogenetic signal (or effect) and phylogenetic inertia, considered synonyms by some authors (Losos, 1999). Phylogen-

etic signal is a pattern of statistical non-independence, where phenotypic similarity is associated with phylogenetic relatedness (Revell et al., 2008; Klingenberg and Gidaszewski, 2010), whereas phylogenetic inertia is the tendency of a trait to resist a current adaptive force (Blomberg and Garland, 2002; Hansen and Orzack, 2005). Phylogenetic inertia can be caused by constraints in development or variation (canalisation, trade-offs due to correlations with lower fitness traits), and is itself recognised as one of the causes behind phylogenetic effects (Harvey and Pagel, 1991), along with adaptive explanations. The confusion between the concepts of phylogenetic inertia and signal led to a misconception of phylogenies as sources of variation in statistical models (Losos, 2011b), and the wrong interpretation of phylogenetic signal as evidence of "constraint". Phylogenetic inertia and adaptation are not mutually exclusive explanations for interspecific differences, nevertheless, they can only be properly assessed in comparisons where their effects on each other are jointly controlled for (Hansen and Orzack, 2005; Hansen et al., 2008). The phylogenetic generalised least squares (PGLS) models provide a useful framework for such analyses. Adaptation can be tested and estimated by the main effects in a PGLS model and inertia can be tested and estimated from the phylogenetic signal in the residual (or error) term of the same PGLS model (Hansen et al., 2008).

Several measures and tests of phylogenetic signal strength have been proposed (Blomberg and Garland, 2002; Freckleton et al., 2002; Blomberg at al., 2003; Klingenberg and Gidaszewski, 2010; Diniz-Filho et al., 2012). Interpretations relating specific values of these phylogenetic signal statistics to evolutionary processes have been proposed. For example, Blomberg's K statistic is expected to be one if trait evolution behaves as expected by a Brownian motion model, whereas it is less than one if species are less similar than expected by phylogenetic relatedness (too many convergences?) and greater than one if species are more similar than expected by neutral evolution (stabilising selection?). However, the link between evolutionary patterns and processes has been challenged by simulation studies (Revell et al., 2008), showing that different models can lead to similar patterns of species similarity within a phylogeny. The model-based approach discussed before provides a more robust basis for evolutionary process inference in the context of comparative data. Measures of phylogenetic signal have also been proposed as a means to determine whether it is necessary to use a comparative method or not, implying that traits with low or non-significant phylogenetic signals would not require the phylogenetic non-independence to be included in the model (Losos, 1999; Klingenberg and Gidaszewski, 2010; Losos, 2011b). This approach is no longer recommended, because the statistical models make assumptions about the distribution and independence of the error (differences between observed and predicted), not the raw data (Revell, 2010). Testing for phylogenetic signal on the traits directly to decide whether to use a phylogenetic comparative method is an error equivalent to testing for normality and homoscedasticity in raw data, before a linear model is fitted (Hansen and Orzack, 2005). The strength of phylogenetic signal can be jointly estimated with the statistical model to determine the most appropriate form of the covariance error matrix (assumptions regarding the evolutionary model), using a number of alternative measures of phylogenetic signal (Martins and Hansen, 1997; Blomberg at al., 2003). As a result, if the error structure of the response variable does not show strong phylogenetic signal, the phylogenetic covariance matrix will approach the identity matrix assumed when errors are expected to be independent. Because the comparative methods are flexible regarding the evolutionary model and the strength of phylogenetic signal, it is advisable to always include phylogenetic information in the statistical models.

An important issue that is frequently overlooked is the fact that studies showing correlations between ecological and morphological changes are by no means considered direct evidence of causation (Martins, 2000). A number of reasons (alternative to direct causation) can be invoked to explain correlations (Losos, 2011a), for example, selection on body size (associated with dietary differences) has probably led to correlated allometric shape differences in the skull of new world monkeys (Marroig and Cheverud, 2005). When many traits are correlated with the same fitness differences, it can be complicated to discern which ones are under selection just from correlation results. In spite of the problems, correlation studies are an important source of patterns that require further investigation by in depth studies measuring the strength of selection or the experimental link between performance, morphology, and ecology (Losos et al., 1997; Winter and von Helversen, 2003; Langerhans and DeWitt, 2004; Langerhans et al., 2004; Nogueira et al., 2009; Losos, 2011a). The problem of inferring evolutionary origin from correlation is by no means exclusive to phylogenetic comparative analyses, but a more general issue in evolutionary biology (Martins, 2000), and the subject of a deep philosophical debate (Sober, 1993; Lewens, 2007). The difficulty of inferring causation from correlation is also firmly rooted in Popper's principle of falsificationism, where it is much easier to falsify a hypothesis than to prove it (Paipneau, 2003). A single negative finding is sufficient to disprove a hypothesis, whereas no number of supporting findings will be considered a conclusive proof.

Comparative methods and morphometric data

Shape is defined as all the geometrical information that remains after location, scale and rotational effects are filtered out from an object (Dryden and Mardia, 1998), and the study of shape is essentially a multivariate undertaking. The simplest morphometric data structure is a triangle of landmarks (Bookstein, 1991), requiring two variables to describe its shape variation. As a result, morphometric studies of comparative data are presented with the additional difficulty of either extending the statistical models to multiple response variables when feasible, or using a method to reduce dimensionality to one or just a few variables. A review of published papers that presented results of morphometric analyses of comparative data provides a summary of the diversity of approaches as discussed below. In this review, I did not discriminate among geometric morphometric studies using landmark coordinates or "traditional" morphometric studies using distances measured among landmarks as data, because the relevant multivariate methods are the same for both kinds of data.

Most comparative methods are readily extensible to multivariate data, but comparative studies seldom use shape variables (as Procrustes aligned coordinates or partial warp plus uniform component scores) directly in the statistical models. They reduce the dimensionality of the data set first (using their principal components also called relative warps). It is also important to note that shape is not necessarily the response variable, and the position of shape variables in the models will depend on the hypotheses being tested. Although comparative studies mostly associate shape variation with ecological or functional variables, morphological diversification within lineages can be associated with amounts of speciation (Adams et al., 2009), or the evolution of specific morphologies can be associated with shifts in cladogenesis (Fitzjohn, 2010). One possibility for correlational studies is to use PGLS models to fit regressions between matrices of shape variables and functional and ecological variables (Rüber and Adams, 2001; Clabaut et al., 2007; Meloro et al., 2008; Raia et al., 2010). Phylogenetic independent contrasts can also be calculated for each shape variable before associating them with contrasts for ecological variables, either by multivariate regression (Figueirido et al., 2010) or partial least squares (PLS) (Klingenberg and Ekau, 1996). This approach should be statistically equivalent to PGLS, as long as the among-taxa covariances are exactly proportional to the branch lengths (assuming a Brownian motion evolutionary model) in the phylogenetic tree (Rohlf, 2001). Another possibility is to correlate morphometric distances with ecological and phylogenetic distances using a matrix correlation method (Harmon et al., 2005; Young et al., 2007; Astua, 2009; Monteiro and Nogueira, 2010). This distance matrix-based approach is less informative than other comparative methods (PGLS, model-based approaches), and is considered a less powerful option (Peres-Neto and Jackson, 2001). The multivariate models discussed above have been mostly used for significance testing, whereas the visualisation of shape variation patterns has been almost exclusively dependent on principal components analysis (as in the phylomorphospace discussed below).

The sets of shape variables can be considerably large, and the most common approach by far is the reduction of shape variables (where the observations are the species means) to a set of principal components (PCs) and fitting comparative statistical models using one or a few PCs. Dimension reduction by principal components analysis (PCA) is a technique commonly used in multivariate analysis to reduce multidimensional data sets to a small number of interpretable axes that retain a maximum amount of variation (Jolliffe, 2002). These shape PCs can be associated (using comparative methods) with ecological variables (or PCs of these) that might explain the patterns of variation among species, but they need to be carefully interpreted, because although they correspond to axes of major variation in shape space, they are not designed to maximise correlation or covariance with any set of ecological variables (as partial least squares would). This is important when interpreting non-significant results, because not finding a clear association between shape PCs and ecological variables does not mean a significant association does not exist for a different linear combination of shape variables and does not exclude multivariate significance (the ecological variables are just not associated with the main axes of interspecific variation).

Studies that reduced the original set of variables into principal components for comparative analysis have used them as univariate variables (one PC at a time) or smaller multivariate data sets to calculate regressions or correlations of phylogenetic independent contrasts (Bergmann and Irschick, 2011; Brusatte et al., 2012), fitting PGLS regression models (Nogueira et al., 2009), and evaluating alternative evolutionary scenarios with the model-based approach (Bergmann et al., 2009; van Buskirk, 2009; Collar et al., 2009). Meloro (2012) used a slightly different approach to evaluate the association between shape and functional variables in the mandible of carnivores, employing partial least squares (PLS) to extract pairs of linear combinations within both sets of variables that explain most cross-covariance between sets (Rohlf and Corti, 2000). The PLS axes were then "validated" by PGLS regression of the shape (as dependent) on diet and size. In this case, the PLS will not construct a linear combination of shape variables that maximises the variation among species, but that should be associated with the variables of interest. One possible problem with this approach, as pointed out by Revell (2009), is that the phylogenetic non-independence needs to be incorporated in the estimates of covariances themselves, not just in *a posteriori* statistical tests (further discussed below).

The principal components are more appropriately used as a multivariate set in comparative analyses (Losos, 1990; Monteiro and Nogueira, 2011). If any subset of PCs will be used as variables (as opposed to the set explaining 100% of total variation), a criterion is needed to determine how many PCs should be retained. A popular criterion in morphometrics is the broken-stick, where the percentage of variance explained by each principal component is compared with a null distribution of expected percentages when the total variation is randomly distributed among principal components (Harmon et al., 2005; Morgan, 2009; Brusatte et al., 2012). However, the vast majority of studies use subjective criteria, such as the number of PCs that sums up more than 70 or 80% of total variation (considered reasonable amounts), or retaining the PCs that account for more than 10 or 5% by themselves. On the other hand, well defined criteria to assess dimensionality of a data set do exist and extensive simulation studies have pointed to a number of well-performing stopping rules for PCs (Peres-Neto et al., 2005). One consensus method in the statistical literature for best performance in recovering subjacent dimensionality is Horn's parallel analysis (Franklin et al., 1995; Dinno, 2009), where a number of data sets with random uncorrelated variables are generated and the distribution of random eigenvalues compared with the observed values (see the R package `paran` – http://CRAN.R-project.org/package=paran – for an implementation of the method. A simplified R code for covariance matrices is also available from the author). Using a well defined criterion to determine the number of overdispersed PCs (the ones that account for more variance than expected in random uncorrelated samples) should prove more informative in determining shape space dimensionality than subjective criteria based on absolute amounts of variance explained. One caveat of parallel analysis is that it is unable to determine dimensionality of spaces larger than $\frac{p}{2.5}$, where p is the number of variables (A. Dinno, *pers. comm.*), which is not a relevant issue for most morphometric applications. A more relevant issue is that because parallel analysis attempts to estimate the "true" number of latent dimensions using statistical inference, a sample size/variables ratio larger than 3 is recommended (although the effect of high dimensional low sample size data is still to be determined with simulations).

The estimation of principal components from interspecific data is usually performed without taking the phylogenetic non-independence into account. Revell (2009) pointed out that when the observations are phylogenetically related, this procedure is not optimal, and proposed a method to estimate principal components that take phylogenetic non-independence into account. The procedure uses the among-taxa $(n \times n)$ phylogenetic covariance matrix $\mathbf{C}$ (the same used in PGLS, for example) to allow for the estimation of the eigenstructure. The elements of the $\mathbf{C}$ matrix can be flexible, depending on the evolutionary model assumed. If a Brownian motion model of evolution is assumed, the elements of $\mathbf{C}$ will be γt_{ij}, where t_{ij} is the distance along the phylogenetic tree between the root and the last common ancestor for species pair ij, and γ is a parameter related to the magnitude of trait variation (Martins and Hansen, 1997), also referred to in the literature as the "rate of evolution" along a branch (Eastman et al., 2011). A more flexible model that allows for stabilising selection constraints would have the elements of $\mathbf{C}$ estimated as $\gamma \exp[-\alpha t_{ij}]$, where α is interpreted as the strength of selection towards an adaptive optimum (Martins and Hansen, 1997). Alternatively, the off-diagonal elements of $\mathbf{C}$ can be multiplied by λ $(0 \leq \lambda \leq 1)$, estimated from the data, to take into account varying strengths of phylogenetic signal (Freckleton et al., 2002). Starting with a data matrix $\mathbf{X}$ (n taxa and p shape variables), one first estimates a vector $\mathbf{a}(p \times 1)$ of phylogenetic means, corresponding to the estimated ancestral shape (for morphometric data) at the root:

$$\mathbf{a} = [(\mathbf{1}^\top \mathbf{C}^{-1} \mathbf{1})^{-1} \mathbf{1}^\top \mathbf{C}^{-1} \mathbf{X}]^\top, \quad (1)$$

where $\mathbf{1}$ is a $n \times 1$ vector of ones. The evolutionary variance-covariance matrix among variables is calculated as:

$$\mathbf{R} = (n-1)^{-1} (\mathbf{X} - \mathbf{1}\mathbf{a}^\top)^\top \mathbf{C}^{-1} (\mathbf{X} - \mathbf{1}\mathbf{a}^\top). \quad (2)$$

The covariance matrix $\mathbf{R}$ is also called the evolutionary rate matrix, because it estimates a matrix of Brownian rate parameters (Revell and Harmon, 2008; Revell and Collar, 2009). The covariances in the off-diagonal indicate associated differences relative to an estimated ancestral shape inversely weighted by the phylogenetic covariances in $\mathbf{C}$ (the influence of neighbouring branches is removed by negative weights). As in PGLS models, multiplication by the inverse of $\mathbf{C}$ ensures that the correct error structure (phylogenetic non-independence) is used in estimating the covariances. The evolutionary covariance matrix above is equivalent to a covariance matrix calculated from phylogenetic independent contrasts, if $\mathbf{C}$ assumes a Brownian motion evolutionary model (Revell, 2009). The PCA of the evolutionary variance-covariance matrix can be performed with the spectral decomposition $\mathbf{R} = \mathbf{V}\mathbf{D}\mathbf{V}^{-1}$, and the PC scores are computed in the original space as

$$\mathbf{S} = (\mathbf{X} - \mathbf{1}\mathbf{a}^\top)\mathbf{V}. \quad (3)$$

These phylogenetic PCs will indicate directions of maximum evolutionary rates in shape space because the shape changes are standardised by branch lengths. Revell (2009) strongly emphasised that this procedure does not produce "phylogenetic corrected" PC scores. It just estimates the eigenstructure properly (as judged by simulations of multivariate data evolving under a Brownian motion model), taking the phylogenetic non-independence into account. If phylogenetic non-independence was present in the original data, the phylogenetic PC scores will be non-independent as well. The assumption of evolutionary model is included in the derivation of matrix $\mathbf{C}$ elements (phylogenetic covariances), which can be changed to reflect different underlying evolutionary models (Martins and Hansen, 1997; Freckleton et al., 2002; Blomberg at al., 2003). The R package `phytools` (Revell, 2012)

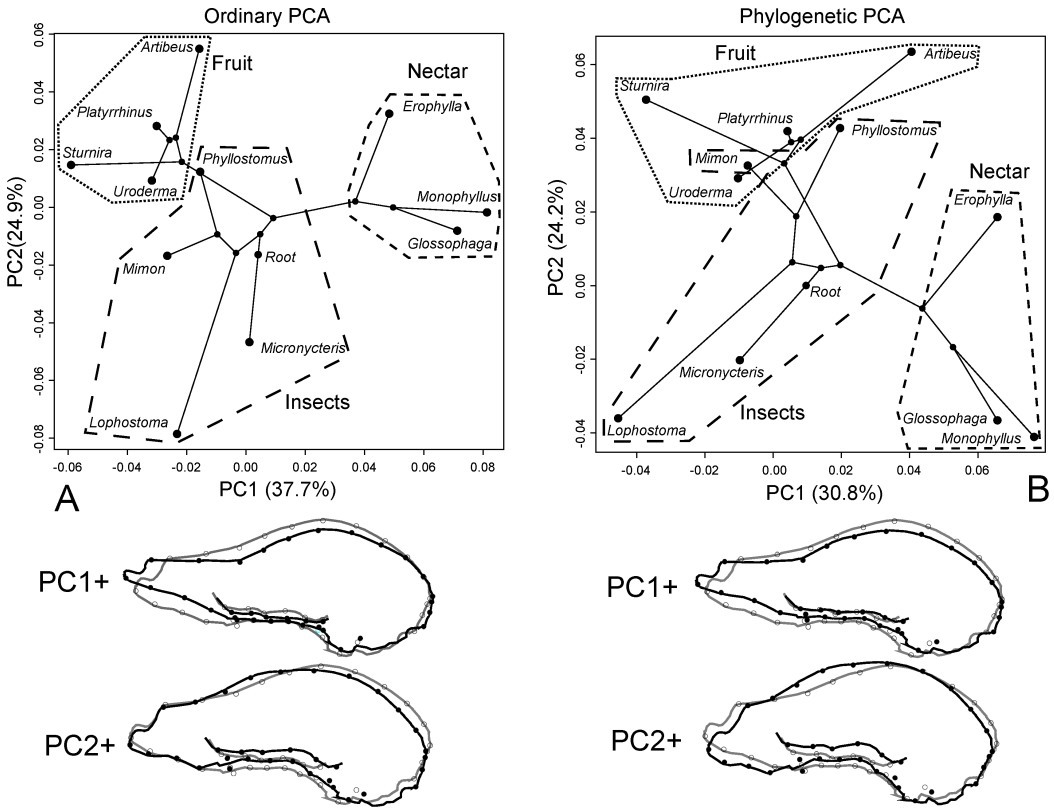

Figure 2 – Phylomorphospaces for principal components of shape, ordinating 11 species of phyllostomid bats based on lateral skull landmarks (Nogueira et al., 2009). The upper panel depicts a scatterplot for PC scores, whereas the lower panels depict the shape change along the positive direction for each axis as warped outlines (average shape as gray line, positive deviation as black line). Species with similar diets are enclosed by polygons with different line patterns. The main diet item is listed near the respective groups. The phylogenetic tree is mapped on the shape subspace. The large circles correspond to the tips and the smaller circles correspond to the nodes (estimated ancestors). A) Scatterplot for the first two PCs from ordinary PCA. B) Scatterplot for the first two phylogenetic PCs (based on the evolutionary rate matrix, see text for explanation).

provides a function to estimate phylogenetic principal components and allows for an estimate of $\mathbf{C}$ by a Brownian motion model or by the lambda transformation (which will be equivalent to Brownian motion if $\lambda = 1$). For the visualisation of shape changes associated with phylogenetic PCs, the eigenvectors can be plotted as shape changes from a mean starting shape (Claude, 2008).

This simple procedure can be easily extended to other types of multivariate analysis, such as Partial Least Squares (PLS) or Canonical Correlation Analysis (CCA). Revell and Harrison (2008) proposed the extension of the method to Canonical Correlation Analysis (CCA), where the data matrices for different sets were transformed before the analysis, using the phylogenetic covariance matrix $\mathbf{C}$ and the phylogenetic mean vector $\mathbf{a}$. The method is then carried out using regular algorithms. For PLS, the same transformed matrices for each block of variables can be used to calculate the cross-covariances in submatrix $\mathbf{R}_{12}$, and the singular value decomposition to estimate pairs of vectors explaining most covariance (Rohlf and Corti, 2000). Significance testing for singular values with permutations (for a null hypothesis of no association between blocks of variables) might still be carried out in the analyses. As in the phylogenetic PCA, the interpretation needs to take into account that the transformed data are not the shapes themselves, but shape differences from the estimated root standardised by branch lengths (similar to what one would get with phylogenetic independent contrasts). An alternative method based on maximum likelihood for comparing hypotheses concerning cross-covariances has also been proposed by Felsenstein (2008), incorporating the influence of the within-species phenotypic covariances on the among-species covariances.

A visualisation method for comparative morphometric data is mapping a phylogeny on the shape subspace composed by the major axes of variation among species (PCs). The position of tree nodes is estimated by ancestral character reconstruction and the branches are drawn as lines connecting the species and their immediate ancestors. This procedure was first proposed by Rohlf (2002), and became popular in applied studies (Nicola et al., 2003; Clabaut et al., 2007; Sidlauskas, 2008; Figueirido et al., 2010; Klingenberg and Gidaszewski, 2010; Monteiro and Nogueira, 2011), receiving the name phylomorphospace (Sidlauskas, 2008). Ancestral character reconstruction can be performed by maximum likelihood (under a Brownian motion evolutionary model) or squared-change parsimony, depending on software, but the approaches are mathematically equivalent (Sidlauskas, 2008). A least-squares approach as used in the calculation phylogenetic independent contrasts is also possible (Felsenstein, 1985). The estimation of ancestral values is controversial and probably inaccurate (Cunningham et al., 1998; Losos, 1999; Webster and Purvis, 2002), unless the real evolutionary process was similar to a Brownian motion (Martins, 1999), and should be used mostly for illustration, so the phylogenetic tree can be mapped on the ordination scatterplot. This visualisation is particularly interesting to detect possible convergences among species in different branches, to get evidence on accelerated morphological evolution, or to analyse morphospace occupation by different lineages (Sidlauskas, 2008), as long as the number of relevant PCs are determined by an objective method and the shape subspace is neither ignoring relevant PCs nor including irrelevant ones (i.e. PCs that do not explain more variation than expected at random).

An example of application of the phylomorphospace and the phylogenetic PCA is shown in Fig. 2. The data set correspond to landmarks placed on a lateral view of the skull of 11 species of phyllostomid bats, studied in Nogueira et al. (2009). The panels on the left (Fig. 2A) show an ordinary PCA with a phylogenetic tree mapped on the ordination of the first two PCs (phylomorphospace). The branch leading to the nectarivore species is aligned with the first PC, whereas the divergence between frugivores and insectivores is aligned with the second PC. The phylomorphospace plot shows that each branch diverged morphologically towards different regions of the PC shape subspace. Looking at the ordination in Fig. 2A, one might feel tempted to attribute the closeness of *Phyllostomus* with frugivore species to convergence. In fact, the species of *Phyllostomus* used (*P. hastatus*) has a mixed diet where

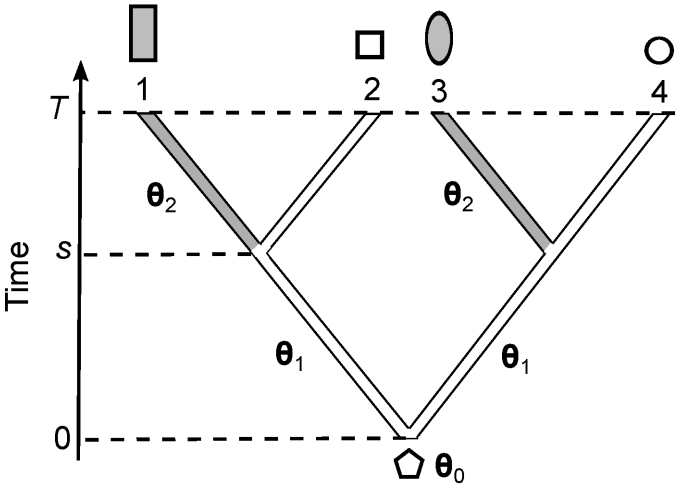

Figure 3 – Hypothetical phylogeny with four species showing a change in selective regime after a speciation event (marked by s). The ancestral phenotype is depicted as θ_0, whereas different selective optima are depicted as θ_1 and θ_2. The total sum of branch lengths leading from root to tip is given by T.

the contribution of fruit is almost as large as the one from insects. However, Horn's parallel analysis suggested that 3 PCs were over-dispersed (had larger observed eigenvalues than expected from simulations), and the third PC is mainly a contrast between the larger insectivores – *Mimon* and *Phyllostomus* and the remaining species (not shown in the example). Therefore, their branch might not be considered convergent if we consider all relevant shape dimensions. The phylogenetic PCA (Fig. 2B) is shown on the right panel. The ordination and the shape changes are similar to the ordinary PCA, in fact, the angles between the corresponding vectors are more similar than expected from a distribution of angles between uncorrelated uniform vectors (Claude, 2008). However, the qualitative interpretation of the ordination pattern could be different, and morphological divergence among closely related species is given greater weight in the phylogenetic PCs.

The phylogenetic PCs are a rigid rotation of the original shape space (and as a matter of fact of the ordinary PCs). Any statistical analysis that is based on the scores of the entire set of PCs will achieve the same quantitative results regardless of using phylogenetic of ordinary PCs (Polly et al. 2013, this volume). However, when a subset of the first PCs are used or displayed, one will find differences in the ordination and due to slight changes in rotation. For example, although the shape change associated with nectarivory is clearly associated with PC1 of the ordinary PCA, in the phylogenetic PCA the same shape change would be a combination of the positive changes along PC1 and negative changes along PC2. This finding also visually highlights one important recommendation: to consider the entire multivariate set of over-dispersed PCs when fitting a comparative statistical model is more informative and appropriate than analysing each PC separately as univariate variables. For the phylogenetic PCA, Horn's parallel analysis indicated 4 over-dispersed PCs, and the divergence of *Phyllostomus* from the frugivore branch is relegated to the 4th PC, but the general qualitative conclusions are similar between methods.

Model-based approaches and dimensionality

Moving beyond the phenotype-function-ecology correlations, comparative methods allow for the estimation of evolutionary parameters for trait diversification and measure the agreement between models of evolutionary processes and data (Butler and King, 2004; Freckleton and Harvey, 2006; Hansen et al., 2008). Model-based comparative analyses of multivariate data are still mostly confined to univariate analyses of PCs (Bergmann et al., 2009; van Buskirk, 2009; Collar et al., 2009; Harmon et al., 2010), possibly because both of software package limitations and restrictions of sample size imposed by the number of parameters needed in more complex models with multivariate data. However, due to the potential pitfalls with univariate analyses of PCs, a multivariate model-fitting procedure should be favoured with morphometric data, as currently available in the R package `OUCH` (King and Butler, 2009).

Monteiro and Nogueira (2011) fitted multivariate models based on Brownian motion and several postulated adaptive landscapes based on dietary differences among phyllostomid bat species, to a data set corresponding to five PCs obtained from mandible shape variables. Considering that p is the number of variables in a multivariate set, a Brownian motion model (the simplest evolutionary model possible) requires the estimation of an average vector with p elements and a square variance-covariance matrix (Sigma squared matrix) with $p \times (p+1)/2$ elements (the lower triangular matrix, including the diagonal) that measures the strength of genetic drift (Butler and King, 2004). Stochastic models for the evolution of a multivariate phenotype consistent with stabilising selection around adaptive optima, such as the Ornstein-Uhlenbeck (O-U) model (King and Butler, 2009), assume the form of the differential equation

$$d\mathbf{x}(t) = \mathbf{A}(\boldsymbol{\theta}(t) - \mathbf{x}(t))dt + \mathbf{S}dB(t), \quad (4)$$

where $\mathbf{A}$ and $\mathbf{S}$ are $p \times p$ square symmetric matrices of parameters measuring the strength of stabilising selection and random drift, respectively. The vector $\boldsymbol{\theta}(t)$ is the optimum phenotype corresponding to a particular selection regime and $B(t)$ is the standard Wiener (Brownian motion) process. Fig. 3 shows a hypothetical phylogeny with four species where branch colours indicate selection regime, based on example in Butler and King (2004). The evolution of the multivariate phenotypes can be seem as weighted sums of selective optima in each branch leading to each specific extant species. In this context, the expected phenotype for species 1 depends on the ancestral phenotype (θ_0) and the estimated optima (θ_1, θ_2) for each branch leading to its current position. According to the multivariate Ornstein-Uhlenbeck model, the expected mean trait vector for species 1 is

$$\begin{aligned} E[\mathbf{x}_1(T)] =& \theta_0 \mathbf{Q} e^{-DT} \mathbf{Q}^{-1} \\ &+ \theta_1 \mathbf{Q} e^{-D(T-s)} \mathbf{Q}^{-1} (\mathbf{I} - \mathbf{Q} e^{-D(s)} \mathbf{Q}^{-1}) \\ &+ \theta_2 (\mathbf{I} - \mathbf{Q} e^{-D(T-s)} \mathbf{Q}^{-1}), \end{aligned} \quad (5)$$

where the thetas are $1 \times p$ vectors of optima, and the multiplying matrices are equivalent to matrix exponentials as $e^{\mathbf{A}} = \mathbf{Q} e^{D} \mathbf{Q}^{-1}$, for $\mathbf{A} = \mathbf{Q}\mathbf{D}\mathbf{Q}^{-1}$, where $\mathbf{Q}$ and $\mathbf{D}$ are the eigenvectors and eigenvalues of $\mathbf{A}$, respectively, and e^{D} is a diagonal matrix with elements e^{λ}, exponential functions of the eigenvalues of $\mathbf{A}$. The random walk around the estimated optima is determined by the covariance matrix Sigma ($\mathbf{S}$), and along with the selection strength matrix $\mathbf{A}$, is used to estimate the covariance matrix for the expectations. The thetas and expected values for each species in the tree can be calculated via generalised least squares and the matrices $\mathbf{A}$ and $\mathbf{S}$ can be optimised during GLS iterations. The procedure described above is implemented in package `OUCH` for R (King and Butler, 2009). Each model requires the estimation of $2 \times p \times (p+1)/2 + p \times n_\theta$ parameters (the number of degrees of freedom – DOF for the model), where n_θ = the number of adaptive optima. The number of parameters that can be estimated is limited by the number of taxa available, and multivariate data sets will be particularly demanding on sample sizes. For example, an O-U model with 5 shape variables and three adaptive optima will require the estimation of 45 parameters. As a result, the confidence on estimation will be reduced and such models can be severely penalised by information criteria used in model comparisons (Butler and King, 2004), as the Akaike Information Criterion will usually be in the general form $AIC = -2 \times \text{logLikelihood} + 2 \times DOF$, and models with smaller AIC fit better the data.

The assumptions of these multivariate models can be better understood if we visualise them as a restricted random walk on a multidimensional fitness landscape. The variances and covariances determining the Normal distribution of the random walk will be the same for all selective regimes and branches (single $\mathbf{S}$ matrix), and the strength of the selective pull will also be the same for all optima (single $\mathbf{A}$ matrix). These assumptions might be unrealistic, and further advances have been proposed as expansions of the original O-U model with different

variances and selective strengths for each adaptive optimum (Beaulieu et al., 2012), demanding separate $\mathbf{A}$ and $\mathbf{S}$ matrices for each selective regime and greatly increasing the number of parameters estimated. The model-based comparative approaches will usually require the reduction of dimensionality via principal components, but the limitations and suggestions discussed above regarding interpretation and choice of number of components should be kept in mind when using this solution, and the dimensionality of the shape space carefully considered to avoid excluding relevant dimensions.

The model-based approaches can use the information from macroevolutionary patterns of morphological diversification among phylogeny tips (OTUs) to make inferences about the underlying evolutionary processes. Recent improvements in model-fitting, such as Bayesian estimation of model parameters seem to be more accurate in determining evolutionary mechanisms generating data (at least from simulations) (Eastman et al., 2011). Instead of looking at the correlation of phenotypic and ecological variables, these models detect changes in rates of morphological diversification to detect bursts or variation in tempo and mode of the evolution of continuous characters among lineages as evidence of past adaptive radiations, periods of continuous gradual change or periods of stasis (Harmon et al., 2003; O'Meara et al., 2006; Harmon et al., 2010; Venditti et al., 2011; Thomas and Freckleton, 2012). Because most of these models have large numbers of parameters to be estimated, the need for dimensionality reduction and the sample size requirements are again a concern for morphometric data sets. It is also important to realise that the estimation of evolutionary model parameters and inferences about the correlation of trait blocks are complementary approaches (Paradis, 2012), providing mutual support for the understanding of morphological evolution. A claim of evidence for an adaptive radiation due to ecological diversification is made stronger if models that predict increases in diversification rate fit the data well and the morphological changes are associated with ecological and functional changes.

Conclusions

The development of morphometric methods went through an exponential phase just before the end of the 20th century with the geometric morphometric revolution (Adams et al., 2004) and started stabilising during the last 10 years. As the methods matured, the most appropriate methodological combinations, theoretical implications and concepts have become more clear, for instance, the superiority of Procrustes superimposition and associated spaces in statistical shape analysis, as compared to many alternative methods considered before (Rohlf, 2000). It is now relatively easier for a beginning researcher to find relevant advice or guidance around the literature and jargon. The comparative methods are going through a period of exponential production, where a large number of methods and alternative models are proposed, and the detailed statistical properties and relevance of different methods are being clarified through simulations and theoretical contributions (Freckleton et al., 2002; Martins et al., 2002; Hansen and Orzack, 2005; Rohlf, 2006; Revell et al., 2008). There are good introductory texts focused on practical aspects (Butler et al., 2008; Paradis, 2012), but a larger theoretical book summarising recent developments is in demand by an ever increasing community of users. One difficulty of the early days was the availability of software packages to perform the analyses, the formatting differences of all types of data (phylogenies, traits), and software bugs. The widespread use of the R environment and the large number of packages for phylogenetic and comparative analyses (Paradis, 2012) available in this system (31 packages implementing comparative method functions, checked in 31 July 2012) has greatly improved access to almost all published methods. The most comprehensive package `APE` (Paradis et al., 2004) provides not only the most common methods (contrasts, PGLS), but also phylogenetic tree edition and visualisation tools also used by most other phylogenetic packages. See also the continuously updated package descriptions on the phylogenetics CRAN task view maintained by Brian O'Meara (http://cran.r-project.org/web/views/Phylogenetics.html). The different functions and packages show considerable compatibility among each other and the authors of most new methods being proposed provide R packages and functions. A good source for help and information is the email list for the Special Interest Group `R-sig-phylo` (https://stat.ethz.ch/mailman/listinfo/r-sig-phylo), where most authors and package developers make announcements, answer questions and provide assistance with the use of comparative methods. An alternative software that can perform comparative analyses in two-dimensional (previously aligned) morphometric data is Mesquite (Maddison and Maddison, 2011), but R provides a more flexible and complete statistical environment.

The complexity of morphometric data is increasing with the greater availability of 3D data collection devices, and the models in comparative analysis are also becoming more complex in number of parameters. As a consequence, the assessment of appropriate shape space dimensionality should be a matter of great concern. It is an exciting time to be a morphometrician working with comparative data.

References

Ackermann R.R., Cheverud J.M., 2002. Discerning evolutionary processes in patterns of tamarin (genus *Saguinus*) craniofacial variation. Am. J. Phys. Anthropol. 117: 260–271.

Adams D.C., Berns C.M., Kozak K.H., Wiens J.J., 2009. Are rates of species diversification correlated with rates of morphological evolution?. Proc. R. Soc. B. 276: 2729–2738.

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

Astua D., 2009. Evolution of scapula size and shape in Didelphid marsupials (Didelphimorphia: Didelphidae). Evolution 63: 2438–2456.

Beaulieu J.M., Jhwueng D., Boettiger C., O'Meara B.C., 2012. Modeling stabilizing selection: expanding the Ornstein-Uhlenbeck model of adaptive evolution. Evolution 66: 2369–2383.

Bergmann P.J., Irschick D.J., 2011. Vertebral evolution and the diversification of squamate reptiles. Evolution 66: 1044–1058.

Bergmann P.J., Meyers J.J., Irschick D.J., 2009. Directional evolution of stockiness coevolves with ecology and locomotion in lizards. Evolution 63: 215–227.

Björklund M., 1997. Are "comparative methods" always necessary? Oikos 80: 607–612.

Blomberg S.P., Garland T., 2002. Tempo and mode in evolution: phylogenetic inertia, adaptation and comparative methods. J. Evol. Biol. 15: 899–910.

Blomberg S.P., Garland T., Ives A.R., 2003. Testing for phylogenetic signal in comparative data: behavioral traits are more labile. Evolution 57: 717–745.

Bookstein F.L., 1991. Morphometric tools for landmark data: geometry and biology. Cambridge University Press, Cambridge, UK.

Brusatte S.L., Sakamoto M., Montanari S., Harcourt-Smith E.H., 2012. The evolution of cranial form and function in theropod dinosaurs: insights from geometric morphometrics. J. Evol. Biol. 25: 365–377.

van Buskirk J., 2009. Getting in shape: adaptation and phylogenetic inertia in morphology of Australian anuran larvae. J. Evol. Biol. 22: 1326–1337.

Butler M.A., King A., 2004. Phylogenetic comparative analysis: a modeling approach for adaptive evolution. Am. Nat. 164: 683–695.

Butler M.A., O'Meara B.C., Pienaar J., 2008. Comparative methods and data analysis in R. Available from http://www2.hawaii.edu/~mbutler/Rquickstart/Rcomparative.pdf.

Cheverud J.M., Dow M.M., Leutenegger W., 1985. The quantitative assessment of phylogenetic constraints in comparative analyses: sexual dimorphism in body weight in primates. Evolution 39: 1335–1351.

Clabaut C., Bunje P.M.E., Salzburger W., Meyer A., 2007. Geometric morphometric analyses provide evidence for the adaptive character of the Tanganyikan cichlid fish radiations. Evolution 61: 560–578.

Claude J., 2008. Morphometrics with R. Springer, New York.

Collar D.C., O'Meara B.C., Wainwright P.C., Near T.J., 2009. Piscivory limits diversification of feeding morphology in centrarchid fishes. Evolution 63: 1557–1573.

Cunningham C.W., Omland K.E., Oakley T.H., 1998. Reconstructing ancestral character states: a critical reappraisal. Trends Ecol. Evol. 13: 361–366.

Diniz-Filho J.A.F., Rangel T.F., Santos T., Bini L.M., 2012. Exploring patterns of interspecific variation in quantitative traits using sequential phylogenetic eigenvector regressions. Evolution 66: 1079–1090.

Diniz-Filho J.A.F., Sant'Ana C.E.R., Bini L., 1998. An eigenvector method for estimating phylogenetic inertia. Evolution 52: 1247–1262.

Dinno A., 2009. Exploring the sensitivity of Horn's parallel analysis to the distributional form of random data. Multivariate Behav. Res. 44: 362–388.

Dryden I.L., Mardia K.V., 1998. Statistical shape analysis. John Wiley & Sons, New York.

Eastman J.M., Alfaro M.E., Joyce P., Hipp A.L., Harmon L.J., 2011. A novel comparative method for identifying shifts in the rate of character evolution on trees. Evolution 65: 3578–3589.

Felsenstein J., 1985. Phylogenies and the comparative method. Am. Nat. 125: 1–15.

Felsenstein J., 2008. Comparative methods with sampling error and within-species variation: contrasts revisited and revised. Am. Nat. 171: 713–725.

Figueirido B., Serrano-Alarcon F.J., Slater G.J., Palmqvist P., 2010. Shape at the crossroads: homoplasy and history in the evolution of the carnivoran skull towards herbivory. J. Evol. Biol. 23: 2579–2594.

Fitzjohn R.G., 2010. Quantitative traits and diversification. Syst. Biol. 59: 619–633.

Franklin S.B., Gibson D.J., Robertson P.A., Pohlmann J.T., Fralish J.S., 1995. Parallel analysis: a method for determining significant principal components. J. Veg. Sci. 6: 99–106.

Freckleton R.P., 2009. The seven deadly sins of comparative analysis. J. Evol. Biol. 22: 1367–1375.

Freckleton R.P., Harvey P.H., 2006. Detecting non-Brownian trait evolution in adaptive radiations. PLOS Biology 4: 2104–2111.

Freckleton R.P., Harvey P.H., Pagel M., 2002. Phylogenetic analysis and comparative data: a test and review of evidence. Am. Nat. 160: 712–726.

Garamszegi L.Z., Moller A.P., 2010. Effects of sample size and intraspecific variation in phylogenetic comparative studies: a meta-analytic review. Biol. Rev. 85: 797–805.
Garland T., Dickerman A.W., Janis C.M., Jones J.A. 1993. Phylogenetic analysis of covariance by computer simulation. Syst. Biol. 42: 265–292.
Grafen A., 1989. The phylogenetic regression. Philos. Trans. R. Soc. Lond. B 326: 119–157.
Hansen T.F., 1997. Stabilizing selection and the comparative analysis of adaptation. Evolution 51: 1341–1351.
Hansen T.F., Orzack S.H., 2005. Assessing current adaptation and phylogenetic inertia as explanations of trait evolution: the need for controlled comparisons. Evolution 59: 2063–2072.
Hansen T.F., Pienaar J., Orzack S.H., 2008. A comparative method for studying adaptation to a randomly evolving environment. Evolution 62: 1965–1977.
Harmon L.J., Kolbe J.J., Cheverud J.M., Losos J.B., 2005. Convergence and the multidimensional niche. Evolution 59: 409–421.
Harmon L.J., Losos J.B., 2005. The effect of intraspecific sample size on type I and type II error rates in comparative studies. Evolution 59: 2705–2710.
Harmon L.J., Losos J.L., Davies T.J., Gillespie R.G., Gittleman J.L., Jennings W.B., Kozak K.H., McPeek M.A., Moreno-Roark F., Near T.J., Purvis A., Ricklefs R.E., Schluter D., Schulte J.A., Seehausen O., Sidlauskas B.L., Torres-Carvajal O., Weir J.T., Mooers A.O., 2010. Early bursts of body size and shape evolution are rare in comparative data. Evolution 64: 2385–2396.
Harmon L.J., Schulte J.A., Larson A., Losos J.B., 2003. Tempo and mode of evolutionary radiation in iguanian lizards. Science 301: 961–964.
Harvey P.H., Pagel M., 1991. The comparative method in evolutionary biology. Oxford Univ. Press, Oxford, UK.
Ives A.R., Midford P.E., Garland T., 2007. Within-species variation and measurement error in phylogenetic comparative methods. Syst. Biol. 56: 252–270.
Jolliffe I.T., 2002. Principal component analysis. Springer-Verlag, New York.
King A.A., Butler M.A., 2009. ouch: Ornstein-Uhlenbeck models for phylogenetic comparative hypotheses (R package). Available from: http://ouch.r-forge.r-project.org.
Klingenberg C.P., 2011. MorphoJ: an integrated software package for geometric morphometric morphometrics. Mol. Ecol. Resour. 11: 353–357.
Klingenberg C.P., Ekau W., 1996. A combined morphometric and phylogenetic analysis of an ecomorphological trend: pelagization in Antarctic fishes (Perciformes: Nototheniidae). Biol. J. Linn. Soc. 59: 143–177.
Klingenberg C.P., Gidaszewski N.A., 2010. Testing and quantifying phylogenetic signals and homoplasy in morphometric data. Syst. Biol. 59: 245–261.
Lande R., 1979. Quantitative genetic analysis of multivariate evolution, applied to brain:body size allometry. Evolution 33: 402–416.
Lande R., 1980. Genetic variation and phenotypic evolution during allopatric speciation. Am. Nat. 116: 463–479.
Langerhans R.B., DeWitt T.J. 2004. Shared and unique features of evolutionary diversification. Am. Nat. 164: 335–349.
Langerhans R.B., Layman C.A., Shokrollahi A.M., DeWitt T.J., 2004. Predator-driven phenotypic diversification in *Gambusia affinis*. Evolution 58: 2305–2318.
Lewens T., 2007. Adaptation. In: Hull D.L., Ruse M. (Eds.). The Cambridge companion to the philosophy of biology. Cambridge University Press, Cambridge, UK. 1–21.
Lofsvold D., 1988. Quantitative genetics of morphological differentiation in *Peromyscus*. II. Analysis of selection and drift. Evolution 42: 54–67.
Losos J.B., 1990. Ecomorphology, performance capability, and scaling of West Indian *Anolis* lizards: an evolutionary analysis. Ecol. Monogr. 60: 369–388.
Losos J.B., 1999. Uncertainty in the reconstruction of ancestral character states and limitations on the use of phylogenetic comparative methods. Anim. Behav. 58: 1319–1324.
Losos J.B., 2011a. Convergence, adaptation, and constraint. Evolution 65: 1827–1840.
Losos J.B., 2011b. Seeing the forest for the trees: the limitations of phylogenies in comparative biology. Am. Nat. 177: 709–727.
Losos J.B., Warheit K.I., Schoener T.W., 1997. Adaptive differentiation following experimental island colonization in *Anolis* lizards. Nature 387: 70–73.
Maddison W. P., Maddison D., 2011. Mesquite: a modular system for evolutionary analysis. Version 2.75. Available from: http://mesquiteproject.org.
Marroig G., Cheverud J.M., 2005. Size as a line of least evolutionary resistance: diet and adaptive morphological radiation in new world monkeys. Evolution 59: 1128–1142.
Martins E.P., 1999. Estimation of ancestral states of continuous characters: a computer simulation study. Syst. Biol. 48: 642–650.
Martins E.P., 2000. Adaptation and the comparative method. Trends Ecol. Evol. 15: 296–299.
Martins E.P., Diniz-Filho J.A.F., Housworth E.A., 2002. Adaptive constraints and the phylogenetic comparative method: a computer simulation test. Evolution 56: 1–13.
Martins E.P., Hansen T.F., 1997. Phylogenies and the comparative method: a general approach to incorporating phylogenetic information into the analysis of interspecific data. Am. Nat. 149: 646–667.
McPherson G., 1990. Statistics in scientific investigation: its basis, application, and interpretation. Springer-Verlag, New York.
Meloro C., 2012. Mandibular shape correlates of tooth fracture in extant Carnivora: implications to inferring feeding behaviour of Pleistocene predators. Biol. J. Linn. Soc. 106: 70–80.
Meloro C., Raia P., Piras P., Barbera C., O'Higgins P., 2008. The shape of the mandibular corpus in large fissiped carnivores: allometry, function and phylogeny. Zool. J. Linn. Soc. 154: 832–845.
Monteiro L.R., Nogueira M.R., 2010. Adaptive radiations, ecological specialization, and the evolutionary integration of complex morphological structures. Evolution 64: 724–744.
Monteiro L.R., Nogueira M.R., 2011. Evolutionary patterns and processes in the radiation of phyllostomid bats. BMC Evol. Biol. 11: 137.
Morgan C.C., 2009. Geometric morphometrics of the scapula of South American caviomorph rodents (Rodentia: Hystricognathi): form, function and phylogeny. Mamm. Biol. 74: 497–506.
Nicola P.A., Monteiro L.R., Pessoa L.M., von Zuben F.J., Rohlf F.J., Reis S.F., 2003. Congruence of hierarchical, localized variation in cranial shape and molecular phylogenetic structure in spiny rats, genus *Trinomys* (Rodentia: Echimyidae). Biol. J. Linn. Soc. 80: 385–396.
Nogueira M.R., Peracchi A.L., Monteiro L.R., 2009. Morphological correlates of bite force and diet in the skull and mandible of phyllostomid bats. Funct. Ecol. 23: 715–723.
O'Meara B.C., Ané C., Sanderson M., Wainwright P.C., 2006. Testing for different rates of continuous trait evolution using likelihood. Evolution 60: 922–933.
Papineau D., 2003. Philosophy of science. In: Bunnin N., Tsui-James E.P., (Eds.) The Blackwell Companion to Philosophy. Blackwell Publishing, Malden, MA. 286–318.
Paradis E., 2012. Analysis of phylogenetics and evolution with R. Springer-Verlag, New York.
Paradis E., Claude J., Strimmer K., 2004. APE: Analyses of phylogenetics and evolution in R language. Bioinformatics 20: 289–290.
Peres-Neto P.R., Jackson D.A., 2001. How well do multivariate data sets match? The advantages of a Procrustean superimposition approach over the Mantel test. Oecologia 129: 169–178.
Peres-Neto P.R., Jackson D.A., Somers K.M., 2005. How many principal components? Stopping rules for determining the number of non-trivial axes revisited. Comput. Stat. Data Anal. 49: 974–997.
Raia P., Carotenuto F., Meloro C., Piras P., Pushinka D., 2010. The shape of contention: adaptation, history, and contingency in ungulate mandibles. Evolution 64: 1489–1503.
Revell L.J., 2007. Testing the genetic constraint hypothesis in a phylogenetic context: a simulation study. Evolution 61: 2720–2727.
Revell L.J., 2009. Size-correction and principal components for interspecific comparative studies. Evolution 63: 3258–3268.
Revell L.J., 2010. Phylogenetic signal and linear regression on species data. Meth. Ecol. Evol. 1: 319–329.
Revell L.J., 2012. `phytools`: an R package for phylogenetic comparative biology (and other things). Meth. Ecol. Evol. 3: 217–223.
Revell L.J., Collar D.C., 2009. Phylogenetic analysis of the evolutionary correlation using likelihood. Evolution 63: 1090–1100.
Revell L.J., Harmon L.J., 2008. Testing quantitative genetic hypotheses about the evolutionary rate matrix for continuous characters. Evol. Ecol. Res. 10: 311–331.
Revell L.J., Harmon L.J., Collar D.C., 2008. Phylogenetic signal, evolutionary process, and rate. Syst. Biol. 57: 591–601.
Revell L.J., Harrison A.S., 2008. PCCA: a program for phylogenetic canonical correlation analysis. Bioinformatics 24: 1018–1020.
Revell L.J., Mahler D.L., Peres-Neto P.R., Redelings B.D., 2012. A new phylogenetic method for identifying exceptional phenotypic diversification. Evolution 66: 135–146.
Polly P.D., Lawing A.M., Fabre A.-C., Goswami A., 2013. Phylogenetic Principal Components Analysis and Geometric Morphometrics. Hystrix 24(1) (Online first) doi:l0.4404/hystrix-24.l-6383
Rohlf F.J., 2002. Geometric morphometrics in phylogeny. In: Forey P., Macleod N. (Eds.). Morphology, shape and phylogenetics. Taylor & Francis, London 175–193.
Rohlf F.J., 2000. On the use of shape spaces to compare morphometric methods. Hystrix 11: 9–25.
Rohlf F.J., 2001. Comparative methods for the analysis of continuous variables: geometric interpretations. Evolution 55: 2143–2160.
Rohlf F.J., 2006. A comment on phylogenetic correction. Evolution 60: 1509–1515.
Rohlf F.J., Corti M., 2000. Use of two-block partial least-squares to study covariation in shape. Syst. Biol. 49: 740–753.
Rüber L., Adams D.C., 2001. Evolutionary convergence of body shape and trophic morphology in cichlids from Lake Tanganyika. J. Evol. Biol. 14: 325–332.
Sidlauskas B.L., 2008. Continuous and arrested morphological diversification in sister clades of characiform fishes: a phylomorphospace approach. Evolution 62: 3135–3156.
Sober E., 1993. The nature of selection: evolutionary theory in philosophical focus. University of Chicago Press, Chicago, IL.
Thomas G.G., Freckleton R.P., 2012. MOTMOT: models of trait macroevolution on trees. Meth. Ecol. Evol. 3: 145–151.
Venditti C., Meade A., Pagel M., 2011. Multiple routes to mammalian diversity. Nature 479: 393–396.
Webster A.J., Purvis A., 2002. Testing the accuracy of methods for reconstructing ancestral states of continuous characters. Proc. R. Soc. Lond. B 269: 143–149.
Westoby M., Leishman M.R., Lord J.M., 1995. On misinterpreting the "phylogenetic correction". J. Ecol. 83: 531–534.
Winter Y., von Helversen O., 2003. Operational tongue length in phyllostomid nectar-feeding bats. J. Mammal. 84: 886–896.
Young R.L., Haselkorn T.S., Bayaev A.V., 2007. Functional equivalence of morphologies enables morphological and ecological diversity. Evolution 61: 2480–2492.

Associate Editor: A. Cardini