Syst. Biol. 51(6):873–880, 2002
DOI: 10.1080/10635150290155863

Phylogeny Shape and the Phylogenetic Comparative Method

EMÍLIA P. MARTINS[1] AND ELIZABETH A. HOUSWORTH[2]

[1]*Department of Biology, Indiana University, Bloomington, Indiana 47405, USA; E-mail: emartins@indiana.edu*
[2]*Departments of Mathematics and Biology, Indiana University, Bloomington, Indiana 47405, USA; E-mail: ehouswor@indiana.edu*

Abstract.—We explored the impact of phylogeny shape on the results of interspecific statistical analyses incorporating phylogenetic information. In most phylogenetic comparative methods (PCMs), the phylogeny can be represented as a relationship matrix, and the hierarchical nature of interspecific phylogenies translates into a distinctive blocklike matrix that can be described by its eigenvectors (topology) and eigenvalues (branch lengths). Thus, differences in the eigenvectors and eigenvalues of different relationship matrices can be used to gauge the impact of possible phylogeny errors by comparing the actual phylogeny used in a PCM analysis with a second phylogenetic hypothesis that may be more accurate. For example, we can use the sum of inverse eigenvalues as a rough index to compare the impact of phylogenies with different branch lengths. Topological differences are better described by the eigenvectors. In general, phylogeny errors that involve deep splits in the phylogeny (e.g., moving a taxon across the base of the phylogeny) are likely to have much greater impact than will those involving small perturbations in the fine structure near the tips. Small perturbations, however, may have more of an impact if the phylogeny structure is highly dependent (with many recent splits near the tips of the tree). Unfortunately, the impact of any phylogeny difference on the results of a PCM depends on the details of the data being considered. Recommendations regarding the choice, design, and statistical power of interspecific analyses are also made. [Comparative method; eigenvalues; eigenvectors; evolution; phylogeny; principal components; theory.]

In recent years, it has become generally accepted that statistical analyses of interspecific data should be conducted in a phylogenetic context using the phylogenetic comparative method (PCM; for description and reviews, see Martins and Hansen, 1996, 1997; Nunn and Barton, 2001). Whether incorporation of phylogenetic information will have a major impact on the results of a specific analysis, however, depends on the details of the phylogeny and of the data being analyzed. In some cases, this information will make a major difference; in others, the difference will be slight. Similarly, some researchers may be hesitant to apply a PCM because the phylogenetic information available for their measured taxa may be unreliable or not available. The actual effect of phylogenetic errors on the results of a PCM analysis depends in part on how those errors affect tree shape. Other researchers prefer not to make the restrictive assumptions necessary for application of some of the existing PCMs, especially given that phylogenetic information sometimes has only a minor impact on the results of an analysis (e.g., Price, 1997). But again, whether unreasonable assumptions will have an impact on the results of a PCM analysis depends on the details of the incorporated phylogeny. Here, we explore these issues mathematically and draw some general conclusions about the impacts of historical information on comparative analyses.

In general, we expect interspecific data measured from taxa that are more closely related on a phylogeny to be more similar to each other than they are to measures of more distant phylogenetic relatives simply because the closely related taxa evolved together for a longer period of time as a single common ancestor. However, most statistical procedures assume that the measured data are independent. Thus, as pointed out by Felsenstein (1985), if no extra phylogenetic information is included in a comparative analysis, the procedure assumes implicitly that the taxa are related by a "star" phylogeny, with all of the taxa emerging essentially instantaneously from a single common ancestor and evolving independently of each other after that point. Felsenstein's (1985) independent contrasts method and other PCMs fix this problem by allowing a researcher to specify alternatives to a star phylogeny. For example, Felsenstein's (1985) method combines phylogenetic information with a mathematical model of how phenotypes or characters are expected to evolve along that phylogeny (Brownian motion) and develops a prediction of how similar

the measured data should be given their shared evolutionary history. The method then uses this prediction to rescale the interspecific data, making them statistically independent of each other and thus appropriate for most standard statistical procedures (e.g., ancestral state estimation, regression, factor analysis).

From this perspective, whether incorporation of phylogenetic information will have an impact on the results of a comparative analysis depends in part on the difference between the way the true phylogeny is incorporated into the model of phenotypic evolution and the way the implicitly assumed star phylogeny is used in a nonphylogenetic approach. If none of the taxa are closer phylogenetic relatives than others, it may not be important to incorporate phylogenetic information. Similarly, for analyses that explicitly require phylogenetic information (e.g., ancestor reconstruction), whether particular phylogenetic errors will invalidate the results of an analysis will depend in part on any difference between the true phylogeny and the incorrect one used in the analysis and specifically how those errors interact with the data. Even the impact of PCM assumptions depends on the details of the analysis. For example, whether or not the Brownian motion assumption underlying Felsenstein's (1985) method is too unrealistic to be useful depends on the specific details of the evolution of the particular phenotypes of interest.

Phylogeny Shape and Relationship Matrices

All of the above issues require a comparison of correct and incorrect phylogenetic information within the context of a PCM. In most phylogenetic comparative analyses, the phylogeny is incorporated into the underlying evolutionary model as a relationship matrix describing the expected similarity (or difference) between the phenotypes of phylogenetically related taxa due to their previous shared evolution (e.g., Martins, 1995). For many PCMs, the relationship matrix is explicit. For example, the phylogenetic generalized least squares approach (Grafen, 1989, 1992; Martins and Hansen, 1997) and the phylogenetic mixed model (Lynch, 1991) require a matrix describing the expected amount of phenotypic change occurring between each pair of taxa on a phylogeny. Similarly, the spatial autoregressive model (Cheverud et al., 1985) uses a phylogenetic distance matrix to partition interspecific variation into phylogenetic and "specific" effects. Even those PCMs that do not explicitly apply a relationship matrix can usually be translated into matrix terms. For example, although Felsenstein's (1985) contrasts method was originally described as an algorithm, it can equivalently be described as a restricted maximum likelihood procedure in which we begin by finding an $n \times (n-1)$ matrix $\mathbf{K}$, in which $\mathbf{K^T G K} = \mathbf{I}$, where $\mathbf{I}$ is the $(n-1) \times (n-1)$ identity matrix and $\mathbf{G}$ is a relationship matrix describing similarities due to phylogenetic relatedness of the measured taxa (Rohlf, 2001). We then use this $\mathbf{K}$ matrix to transform the data into phylogenetically independent contrasts.

Thus, to determine the impact of phylogenetic information on a PCM, we begin by comparing relationship matrices, specifically the correct one and the one assumed by a particular analysis. In general, the rows and columns of the relationship matrix ($\mathbf{G}$) correspond to the measured taxa, whereas the elements of the $\mathbf{G}$ matrix describe the expected similarity (or distance) between the phenotypes of each pair of taxa due to the shared evolution of taxa along a common phylogeny. These elements are usually described as variances and covariances and can be transformed into correlations by factoring out a common variance term, thereby scaling the entries to a range between 0 and 1 (equivalent to scaling branch lengths on the phylogeny so that the maximum length equals 1). Here, we assume that $\mathbf{G}$ is a correlation matrix.

Under most microevolutionary models, the expected similarity between two taxa depends on the relative amount of time they evolved together versus apart (Hansen and Martins, 1996). For example, under a simple Brownian motion model of evolution, the element of $\mathbf{G}$ corresponding to the relationship between two taxa can be obtained by determining the length of the phylogeny from the root of the tree (when the taxa evolved together as a single common ancestor) to the most recent common ancestor of the two taxa (when they diverged and began to evolve independently of each other), using a phylogeny scaled so that the maximum

distance from root to tips equals 1. The diagonal elements of **G** are 1 because each taxon shares all of its evolutionary history with itself.

For a star phylogeny, the **G** relationship matrix is the identity matrix, **I**, with diagonals of 1 and off-diagonals equal to 0. Off-diagonals are 0 because taxa in the study have shared no phylogenetic history beyond the root. Thus, unless the **G** matrix differs substantively from the identity matrix, PCMs are likely to yield results very similar to those obtained using a nonphylogenetic analysis. Similarly, whether errors in the phylogeny or in the underlying model of phenotypic evolution are likely to affect the results of a PCM depends on the differences between the true **G** and the available, but erroneous, **G**. Phylogenetic errors that result in only small differences between the two matrices are less likely to have a major impact than are errors resulting in major differences.

PHYLOGENY RELATIONSHIP MATRICES

As pointed out by Piazza and Cavalli-Sforza (1974), all phylogeny relationship matrices have a distinctive block-type structure (Fig. 1) imposed by the bifurcating nature of the phylogeny. We can use this blocklike nature to compare phylogeny matrices on a common and evolutionarily relevant scale. Each bifurcation in the phylogeny, beginning at the root, splits the descendant taxa into two subclades. Because each row of the **G** matrix corresponds to the expected similarity between a subject taxon and every other taxon in the phylogeny, the bifurcating split at the root of the phylogeny splits the **G** matrix into four blocks. The upper left block of the **G** matrix corresponds to expected similarities between taxa within the first subclade (AB; Fig. 1), and the lower right block corresponds to expected similarities between taxa within subclade CD. The upper right and lower left correspond to expected similarities between taxa across the two subclades (e.g., between AB and CD) and are (for most PCMs) scaled to 0 for this initial bifurcation at the root of the tree. Each subclade is defined by a set of common ancestors within the subclade and by not sharing common ancestors with taxa outside the subclade subsequent to the bifurcation event. Thus, further subclade divisions also translate directly into subblocking of the **G** matrix.

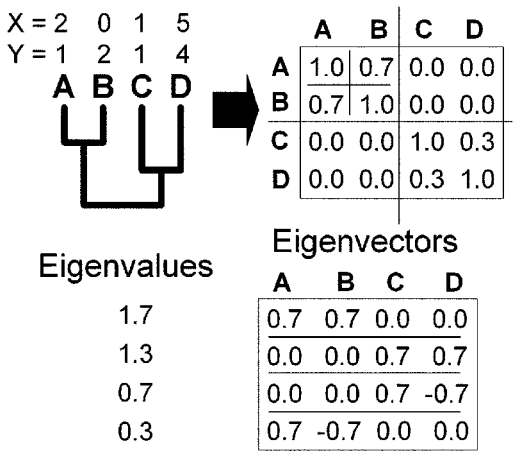

FIGURE 1. The blocklike nature of a phylogeny relationship matrix (**G**) and how that matrix can be described by its eigenvectors (**T**) and eigenvalues (λ, such that $\mathbf{T}'\lambda\,\mathbf{T} = \mathbf{G}$). **X** and **Y** are fictitious phenotypic measures of the four taxa used for an example in the text.

The phylogeny matrix and its hierarchical blocking structure can be described succinctly in terms of its eigenvalues and eigenvectors. Any $n \times n$ relationship (correlation) matrix, **G**, will have n eigenvalues (λ_i) and n eigenvectors ($\mathbf{T}_i$) satisfying $\mathbf{G}\mathbf{T}_i = \lambda_i \mathbf{T}_i$. Differences in the phylogenetic topology affect primarily the eigenvectors, whereas small differences in the branch lengths are reflected primarily in the distribution of the eigenvalues. Each eigenvector (row of the eigenvector matrix in Fig. 1) identifies a particular bifurcation event on a phylogeny, and the signs of the eigenvector entries describe the branching events that comprise an unordered phylogenetic topology (see Piazza and Cavalli-Sforza, 1974). Some changes in the phylogeny structure will have a greater impact on the eigenvectors of a phylogenetic relationship matrix than will others. For example, movement of a single taxon from one point to another within a subclade (Fig. 2) is likely to have less impact on the pattern of eigenvectors than movement of a taxon across the root of the phylogeny. Thus, errors in the fine branching structure of a phylogeny are likely to have less of an impact on the results of a PCM than would the erroneous placement of a single taxon on the wrong side of the root.

Each eigenvector has a corresponding eigenvalue that in a rough sense describes

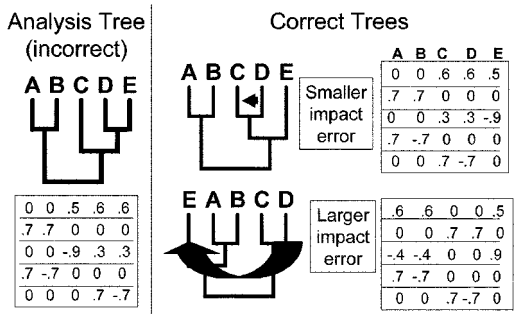

FIGURE 2. How a phylogeny error involving a single taxon can affect the eigenvectors of a phylogenetic relationship matrix. Movement of the taxon within a clade has less of an impact on the eigenvectors than does movement across the root.

the lengths on the phylogeny or the weights associated with that bifurcation event. For a starlike phylogeny (Fig. 3, top), the **G** matrix will be the identity matrix (**I**), with all eigenvalues of roughly equal size (1). For other phylogenies, eigenvalues corresponding to deeper splits in the trees (nearer the root) will be larger than those for splits nearer the tips of the tree (Piazza and Cavalli-Sforza, 1974). Thus, a phylogeny with one deep split followed by many more recent speciation events (e.g., Fig. 3, bottom) will result in a **G** matrix with two large eigenvalues (reflecting the two sides of the most ancient split) and several smaller ones (corresponding to the more recent splits). When there is a marked difference between large and small eigenvalues, most of the expected similarity between interspecific data is explained by the deepest splits on the tree.

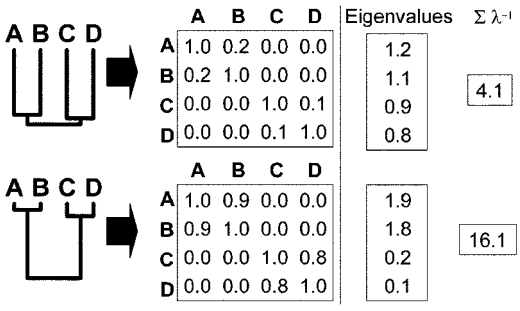

FIGURE 3. Distribution of eigenvalues obtained for two phylogenies with the same topologies but dramatically different branch lengths. The sum of the inverse eigenvalues can be used as a relative index of the expected independence of taxon phenotypes.

HOW SIMILAR IS SIMILAR?

Matrices are complex structures, and it is not obvious how to compare them. In general, the distribution of eigenvalues lies on a continuum, and there are several indices that could be used to describe the location of any specific **G** matrix on that continuum. For example, if the sum of the inverse eigenvalues of the **G** matrix is close to the total number of taxa in the study (n), the relationship matrix resembles that for a star phylogeny. In this case, the results of a PCM using this phylogeny are unlikely to differ much from those of a nonphylogenetic analysis. When the sum of the inverse eigenvalues is larger (e.g., n^2), measures of extant taxa are expected to be very similar to each other because of shared phylogenetic history, and incorporation of the phylogeny is likely to have a far greater impact on the results of PCM analysis.

This sort of index, however, will be most useful in comparing phylogenies with different branch lengths and does not capture possible variation in the phylogeny matrix eigenvectors (which summarize the topological structure). Sometimes, differences in the eigenvalues and eigenvectors will also interact. For example, smaller perturbations in the phylogenetic topology will have a greater impact on the eigenvectors and eigenvalues of the relationship matrix if the phylogeny has a highly dependent structure (i.e., with long branches near the root of the tree) than will similar perturbations if the phylogeny is more independent (i.e., with long branches leading to the extant taxa).

Although several randomization procedures for comparing matrices are available (e.g., Smouse and Long, 1992; Legendre et al., 1994; Phillips and Arnold, 1999), none of these seems fully adequate to the task of comparing phylogenetic relationship matrices for use with PCMs. Because of similarities among all relationship matrices, we expect any two phylogenies to share some aspects of their matrices simply because they are phylogenies. Thus, simply shuffling the elements along the rows or columns of a matrix may not be adequate. For example, Phillips and Arnold (1999) recently developed randomization tests to compare genetic relationship matrices using the Flury hierarchy. Their procedure involves sequentially determining whether two matrices have common principal components, are fully proportional,

or are actually equal. The principal components of a matrix are the eigenvectors multiplied by the square root of their associated eigenvalues. Because of the shared block-like structure of relationship matrices, we expect any two genealogies (or phylogenies) to share some aspects of their principal components. Thus, Phillips and Arnold (1999) developed a randomization test for Flury comparisons of genetic relationship matrices based on shuffling families within populations so that the two matrices are compared with a null background of matrices sharing a genealogical relationship structure. Similar randomization procedures involving models of speciation (e.g., Martins, 1996; Housworth and Martins, 2001) could also be developed for phylogeny matrix comparisons.

Unfortunately, direct comparison of two phylogeny matrices will not provide a final answer because the impact of a phylogeny on the results of a comparative analysis also depends on the details of the data and how they interact with the PCM of choice. For example, consider the phylogenetic generalized least squares approach (PGLS; Grafen, 1989, 1992; Martins and Hansen, 1997), which in its simplest form gives answers that are mathematically equivalent to those produced by Felsenstein's (1985) method. In this case, the phylogeny enters as a relationship matrix, **G**, describing expected similarities due to phylogenetic relatedness of the measured taxa. Imagine that we are using the independent contrasts method to estimate the relationship between two traits (X and Y). Using PGLS, we can estimate the regression slope between two traits while taking phylogenetic information into account using

$$b = [\mathbf{X}'\mathbf{G}^{-1}\mathbf{X}]^{-1}\mathbf{X}'\mathbf{G}^{-1}\mathbf{Y}$$
$$\mathrm{Var}(b) = \sigma^2[\mathbf{X}'\mathbf{G}^{-1}\mathbf{X}]^{-1},$$

where **X** and **Y** are the data, σ^2 is a scaling constant (variance in the residuals), and **G** is, again, a matrix describing the expected similarities between all pairs of taxa due to shared evolution along a phylogeny. If we do not incorporate phylogenetic information, **G** is the identity matrix, and these equations reduce to their usual least-squares versions.

Thus, the actual impact of **G** on estimates of the regression slope depends in a complex way on the data represented in **X** and **Y**. For example, errors in the small branches near the tips of the phylogenies are unlikely to have much of an impact on the results of the comparative analysis. Consider estimating the relationship between two traits (X and Y) as a regression slope, b, using the data in Figure 1 and a simple PGLS method (similar to Felsenstein's contrast method) to incorporate phylogenetic information. We estimate a slope of 18.5 (SE = 0.17), which (assuming normality) is significantly greater than 0. Imagine that there was an error in the original phylogeny such that taxa A and B actually share 80% rather than 70% of their phylogenetic history. Redoing the calculations with our new relationship matrix, we find the same conclusion of a strong, positive linear relationship: $b = 16.8$; SE = 0.17.

We could invent a data set, however, in which even minor changes in **G** will have a major impact on the resulting regression slope. Specifically, whenever there is a conflict between the covariance structure of the data and the similarities represented by the phylogeny, small changes in the **G** matrix could interact with the data to have large impacts on the results. For example, imagine that $X_D = 1.0$ (instead of 5, as in Fig. 1), such that despite the phylogenetic distance between taxa C and D, these two taxa share the same value for trait X ($X_C = X_D = 1$). By changing one data point, we have introduced a severe conflict between the phenotypic data and the phylogeny. When we estimate the regression slope (using the phylogeny in Fig. 1), we find that the traits continue to be positively related but to a lesser extent ($b = 2.3$, SE = 0.33). Imagine though that we again find the problem with the phylogeny, and note that taxa A and B are related by 80% rather than 70%. Suddenly, $b = -1.0$ (SE = 0.28), indicating a significant negative relationship between the traits. A minor change in the phylogeny has led to a major change in the conclusions and final interpretation.

We can understand this phenomenon by rephrasing the **G** matrix in terms of its eigenvectors ($\mathbf{T}_i$) and eigenvalues (λ_i) in the above variance equation. In doing so, we find that the variance is in the form of a harmonic mean:

$$\sigma^2/\mathrm{Var}\,(b) = (\mathbf{X}'\mathbf{T}_1)^2{\lambda_1}^{-1} + (\mathbf{X}'\mathbf{T}_2)^2{\lambda_2}^{-1} + (\mathbf{X}'\mathbf{T}_3)^2{\lambda_3}^{-1} + (\mathbf{X}'\mathbf{T}_4)^2{\lambda_4}^{-1} + \cdots .$$

Each term in this harmonic mean corresponds to a particular node on the phylogeny, beginning with the largest eigenvalue (λ_1) at the root and continuing on through the smaller eigenvalues near the tips (e.g., λ_4). This harmonic mean provides further justification for the use of the sum of inverse eigenvalues as an index of tree shape in considering relationships among phenotypic measures. When closely related taxa are phenotypically divergent (e.g., taxa C and D in Fig. 1 for trait **X**), the data vector is perpendicular to the **G** matrix eigenvector at the conflicted node (T_{Ct}), making the product, $\mathbf{X}'T_{Ct} = 0$ (in our example, the product $= [X_A * 0] + [X_B * 0] + [1 * -0.7] + [1 * 0.7] = 0$ for the conflicted node), thus eliminating the impact of this conflicted node on the variance estimator. Summing over all nodes leads to an increase in the relative importance of the other nodes, particularly those with the smallest eigenvalues (near the tips of the phylogeny). Thus, because of the peculiarities of this data set, minor changes in the nodes or branch lengths at the tips of the phylogeny can produce major changes in the results of a PCM.

Practical Recommendations

This mathematical consideration of phylogeny shape on PCMs leads us to several general conclusions and practical recommendations.

1. Small errors in the phylogeny involving, for example, shuffling of sister taxa within a clade may have little if any impact on the results of PCM analyses. Changes in the phylogeny that involve moving taxa across the root of the tree are more likely to have a major effect.
2. Smaller perturbations will have a greater impact if the phylogeny has a highly dependent structure (i.e., with long branches near the root of the tree) than if it is independent (with long branches leading to extant taxa).
3. The sum of the inverse eigenvalues of the phylogeny relationship matrix (ranging from 0 to the total number of taxa in the study) may provide a useful metric for comparing phylogenies or models of microevolutionary change in terms of their impact on PCM results.
4. Phylogenies are constrained structures that result in matrices with a particular blocklike structure. Randomization tests developed to compare matrices should involve procedures that generate comparable blocklike structures (e.g., using speciation models to generate random phylogenies). Simple shuffling of the rows or columns may not be sufficient.
5. The impact of phylogeny errors on the results of a PCM depends intrinsically on the data. For example, small phylogeny errors may have a large impact if there is considerable disagreement between phenotypic and phylogenetic similarities among taxa.

All of the above are rough generalizations, and the impact of changes in the phylogeny or underlying model of phenotypic evolution on each data set must be considered.

We can also use the eigenvectors and eigenvalues of a relationship matrix to help design a statistically more powerful or robust comparative analysis. For example, relationship matrices with a more even distribution of eigenvalues (sum of inverse eigenvalues closer to the number of taxa, n) indicate less dependence between taxon measures because of shared history. Thus, to maximize statistical power for estimating ancestral states or robustness against errors in the phylogeny when estimating correlations or regression coefficients, we might calculate the eigenvalues for different possible combinations of taxa and choose the one with the most even range (e.g., the smaller sum of inverse eigenvalues). For example, instead of conducting a full-fledged phylogenetic analysis, several authors have advocated using comparisons between several independent pairs of taxa (e.g., two bats, two whales; Felsenstein, 1985). Pairwise comparisons are sometimes viewed as providing a stronger evolutionary argument because they provide independent bits of evidence for the same phenomenon (e.g., Read and Nee, 1995). However, the pairwise comparison approach leads to a **G** matrix with several large eigenvalues corresponding to the deep splits between independent pairs and several small eigenvalues corresponding to the recent splits between taxa within each pair. The resulting distribution of eigenvalues is highly uneven (large sum of inverse eigenvalues), resulting in relatively low statistical power or robustness for the associated analyses. We can reduce the relative number

of small eigenvalues (and thereby increase statistical power) simply by considering only one taxon from each of several independent clades (e.g., one bat, one whale, one fox, and one mouse) but increasing the total number of clades.

6. Whenever possible, increase statistical power by including as many independent taxa as possible. Comparisons of pairs of taxa within each clade may provide better representation of those clades but may be ineffective in terms of increasing the effective sample size or robustness of the analysis.

The distribution of eigenvalues can also inform the choice of phylogenetic method or provide a better understanding of differences between methods. For example, Diniz-Filho et al. (1998) summarized a phylogeny in terms of the principal coordinates (eigenvalues) of a double-centered distance matrix and used only the most important of these to correct interspecific analyses. In theory, this method should perform well (robustly) with highly dependent phylogeny structures in which the two or three eigenvalues corresponding to deep splits in the phylogeny are much larger than the other eigenvalues of the relationship matrix. Moreover, the method should be largely insensitive to small errors near the tips of the phylogeny. As another example, Martins and Hansen (1997) proposed an extension of Felsenstein's (1985) method involving an exponential transformation of the **G** matrix to replace the Brownian motion assumption underlying the independent contrast method with a more general family of models that can be used to describe a wider family of microevolutionary scenarios (e.g., those summarized by Hansen and Martins, 1996). On statistical rather than evolutionary grounds, Diaz-Uriarte and Garland (1996) also proposed that we extend Felsenstein's (1985) contrasts method by using logarithmic or arcsine transformations of the data or branch lengths to improve the fit of the Brownian motion model to the data. Although all of these possible transformations will lead to increased evenness of the eigenvalues, they will do so at different rates and thereby translate into different evolutionary assumptions. By examining the rate of change in the sum of inverse eigenvalues for any particular phylogeny, we can compare the effects of these very different transformation procedures on PCM results.

7. Use caution when applying statistical transformations to comparative data or branch lengths to improve the fit of a PCM model. Transformations2 will affect the evolutionary assumptions underlying the method and are likely to increase the independence of phylogenetically related taxa at different rates, leading to major differences in the interpretation of any particular result.

We hope that future researchers will consider the use of eigenvectors and eigenvalues in both the design and analysis of comparative studies. Further work (both theory and simulations) is needed to assess the usefulness of the sum of the reciprocals of the eigenvalues in determining the robustness of phylogenetic comparative analyses to small uncertainties in the phylogeny. Gene trees and population-level networks may produce relationship matrices with very different eigenvector structures than are produced by bifurcating phylogenies, for example, because of horizontal gene transfer. More complex objects such as whole genomes and gene networks may also be better described by a relationship matrix than by a simply phylogeny. Understanding the eigenvalues and eigenvectors of both bifurcating trees and more complicated relationships will facilitate comparisons between the two. Eigenvectors and eigenvalues may also be useful in developing randomization procedures for generating phylogenies, perhaps, for example, leading to more efficient Markov chain Monte Carlo approaches (e.g., Kuhner et al., 1982; Yang and Rannala, 1997; Beerli and Felsenstein, 1999, 2001) for exploring the vastness of tree space.

REFERENCES

BEERLI, P., AND J. FELSENSTEIN. 1999. Maximum likelihood estimation of migration rates and effective population numbers in two populations using a coalescent approach. Genetics 152:763–773.

BEERLI, P., AND J. FELSENSTEIN. 2001. Maximum likelihood estimation of a migration matrix and effective population sizes in *n* subpopulations by using a coalescent approach. Proc. Natl. Acad. Sci. USA 98:4563–4568.

CHEVERUD, J. M., M. M. DOW, AND W. LEUTENEGGER. 1985. The quantitative assessment of phylogenetic constraints in comparative analyses: Sexual

dimorphism in body weight among primates. Evolution 39:1335–1351.

DIAZ-URIARTE, R., AND T. GARLAND, JR. 1996. Testing hypotheses of correlated evolution using phylogenetically independent contrasts: Sensitivity to deviations from Brownian motion. Syst. Biol. 45:27–47.

DINIZ-FILHO, J. A. F., C. E. RAMOS DE SANT'ANA, AND L. M. BINI. 1998. An eigenvector method for estimating phylogenetic inertia. Evolution 52:1247–1262.

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

GRAFEN, A. 1989. The phylogenetic regression. Philos. Trans. R. Soc. Lond. B 326:119–157.

GRAFEN, A. 1992. The uniqueness of the phylogenetic regression. J. Theor. Biol. 156:405–424.

HANSEN, T. F., AND E. P. MARTINS. 1996. Translating between microevolutionary process and macroevolutionary patterns: The correlation structure of interspecific data. Evolution 50:1404–1417.

HOUSWORTH, E. A., AND E. P. MARTINS. 2001. Conducting phylogenetic analyses when the phylogeny is partially known: Random sampling of constrained phylogenies. Syst. Biol. 50:628–639.

KUHNER, M. K., J. YAMATO, AND J. FELSENSTEIN. 1982. Estimating effective population size and mutation rate from sequence data using Metropolis–Hastings sampling. Genetics 140:1421–1430.

LEGENDRE, P., F. J. LAPOINTE, AND P. CASGRAIN. 1994. Modeling brain evolution from behavior—A permutational regression approach. Evolution 48:1487–1499.

LYNCH, M. 1991. Methods for the analysis of comparative data in evolutionary biology. Evolution 45:1065–1080.

MARTINS, E. P. 1995. Phylogenies and comparative data, a microevolutionary perspective. Philos. Trans. R. Soc. Lond. B 349:85–91.

MARTINS, E. P. 1996. Conducting phylogenetic comparative analyses when the phylogeny is not known. Evolution 50:12–22.

MARTINS, E. P., AND T. F. HANSEN. 1996. The statistical analysis of interspecific data: A review and evaluation of phylogenetic comparative methods. Pages 22–75 *in* Phylogenies and the comparative method in animal behavior (E. Martins, ed.). Oxford Univ. Press, Oxford, U.K.

MARTINS, E. P., AND T. F. HANSEN. 1997. Phylogenies and the comparative method: A general approach to incorporating phylogenetic information into the anlaysis of interspecific data. Am. Nat. 149:646-667. (Erratum, Am. Nat. 153:448.)

NUNN, C. L., AND R. A. BARTON. 2001. Comparative methods for studying primate adaptation and allometry. Evol. Anthropol. 10:81–98.

PHILLIPS, P. C., AND S. J. ARNOLD. 1999. Hierarchical comparison of genetic variance-covariance matrices. I. Using the Flury hierarchy. Evolution 53:1506–1515.

PIAZZA, A., AND L. L. CAVALLI-SFORZA. 1974. Spectral analysis of patterned covariance matrices and evolutionary relationships. Pages 76–105 *in* Proceedings of the Eighth International Conference on Numerical Taxonomy. W. H. Freeman, San Francisco.

PRICE, T. 1997. Correlated evolution and independent contrasts. Philos. Trans. R. Soc. Lond. B 352:519–529.

READ, A. F., AND S. NEE. 1995. Inference from binary comparative data. J. Theor. Biol. 173:99–108.

ROHLF, F. J. 2001. Comparative methods for the analysis of continuous variables: Geometric interpretations. Evolution 55:2143–2160.

SMOUSE, P. E., AND J. C. LONG. 1992. Matrix correlation analysis in anthropology and genetics. Yearb. Phys. Anthropol. 35:187–213.

YANG, Z., AND B. RANNALA. 1997. Bayesian phylogenetic inference using DNA sequences: A Markov chain Monte Carlo method. Mol. Biol. Evol. 14:717–724.

First submitted 5 January 2002; revised manuscript returned 2 May 2002; final acceptance 15 August 2002
Associate Editor: Arne Mooers