

## Polytomies in Comparative Analyses of Continuous Characters

ANDY PURVIS<sup>1</sup> AND THEODORE GARLAND, JR.<sup>2</sup>

<sup>1</sup>*Department of Zoology, University of Oxford, South Parks Road, Oxford OX1 3PS, England*

<sup>2</sup>*Department of Zoology, University of Wisconsin, Madison, Wisconsin 53706, USA*

Felsenstein's (1985) method of phylogenetically independent contrasts was the first statistically valid parametric method for correlation and regression analysis of continuous variables measured on multiple species. One requirement of the method is that the true phylogeny linking the species in the data set be known in full, including the lengths of the branches in units of expected variance of change for the character(s) being studied (see Harvey and Pagel, 1991; Martins and Garland, 1991; Garland et al., 1992). In practice, we cannot be certain of the true phylogeny for any set of species. Instead, we must use an estimate of it. This working phylogeny (*sensu* Grafen, 1989) might differ from the real phylogeny in three ways: (1) taxa depicted as sister groups in the working phylogeny might not be so in reality, causing the wrong taxa to be compared when computing contrasts; (2) the working phylogeny might not depict all of the real sister-group relationships, i.e., a series of bifurcations in the true phylogeny might be represented as a polytomy in the working phylogeny; and (3) the lengths of branches between nodes (including tips) of the working phylogeny might be incorrect.

Differences of the first type clearly must be minimized because they violate the assumptions of Felsenstein's method to an unknown and possibly very serious degree. Differences of the third type also weaken the method, but apparently not fatally (see Grafen, 1989; Martins and Garland, 1991). However, we still want the branch lengths on the working phylogeny to reflect the truth as accurately as possible

(Garland et al., 1992). Differences of the second type—incomplete resolution of the working phylogeny—raise some interesting conceptual and analytical issues. Unless we are prepared to assume that speciation is always a strictly dichotomous process, then we must recognize that some polytomies in working phylogenies will represent true cases of multiway speciation events; these are termed "hard" polytomies (Maddison, 1989). In the original description of the comparative method of phylogenetically independent contrasts, Felsenstein (1985:10) alluded to the correct way to analyze hard polytomies. However, other polytomies in working phylogenies simply reflect our ignorance of the true dichotomous branching pattern ("soft" polytomies, Maddison, 1989). Felsenstein's suggestion to set some internode branch lengths to zero when computing standardized independent contrasts leads to unbiased estimates of correlations and slopes for both hard and soft polytomies but does not indicate how many degrees of freedom are available for hypothesis testing with the latter. Grafen (1989), Harvey and Pagel (1991), and Pagel (1992) have offered other more complicated ways of dealing with soft polytomies that allow only one degree of freedom per node. Table 1 presents the differences between the two positions. We want to recognize any hard polytomies as such to claim the maximum degrees of freedom associated with them.

### HARD AND SOFT POLYTOMIES

Independent contrasts approaches use phylogenetic information to transform

TABLE 1. How the assumption of "hard" versus "soft" polytomies affects independent contrasts methods. When polytomous nodes are present, the number of contrasts computed and used to estimate the relationship between two traits should always be  $n - 1$ , as in a fully resolved  $n$ -taxon tree, whereas the associated degrees of freedom are reduced when soft polytomies are present. This table presumes bivariate regression (or correlation) through the origin; in multiple regression through the origin, one additional degree of freedom would be lost for each additional independent variable.

	All polytomies hard	All polytomies soft
Contrasts at an $n$ -taxon polytomy	$n - 1$	$n - 1$
Contrasts in a fully resolved $n$ -taxon tree	$n - 1$	$n - 1$
Contrasts in an $n$ -taxon working phylogeny that contains some unresolved nodes	$n - 1$	$n - 1$
Degrees of freedom for an $n$ -taxon polytomy	$n - 1$	1 <sup>a</sup>
Degrees of freedom for an $n$ -taxon working phylogeny that contains $p$ nodes (including the root)	$n - 1$	$p - 1$ <sup>a,b</sup>
Test for significance	regression through origin	regression through origin
Degrees of freedom in test	$n - 1 - 1$	maximum = $n - 1 - 1$ , minimum = $p - 1$

<sup>a</sup> Corresponds with the maximally conservative assumption of procedures proposed by Grafen (1989), Pagel and Harvey (1989, 1992), Harvey and Pagel (1991), and Pagel (1992).

<sup>b</sup> Except when the tree is completely unresolved (i.e., a star), in which case the minimum is 1, not 0, df. See also Figure 3.

values for terminal taxa to eliminate statistical nonindependence. The working phylogeny specifies what pattern of nonindependence we expect or, more specifically, the expected variances and covariances of the phenotypes of all terminal taxa. Hard and soft polytomies represent different specifications of phylogenetic pattern and so must be treated differently for hypothesis testing.

#### *Instantaneous Radiations: Hard Polytomies*

Felsenstein's (1985) original paper mentioned how hard polytomies should be an-

alyzed: they "can always be represented as a series of bifurcations having some branch lengths zero" (Felsenstein, 1985:10). At an  $n$ -taxon node, then, Felsenstein advocated that the polytomy be resolved arbitrarily into a series of bifurcations separated by branches of zero length. For each bifurcation, a contrast is computed and standardized to unit variance by dividing it by the square root of the sum of its branch lengths. On first consideration, it might seem that the result of the analysis will depend on how the polytomy is resolved; a node with  $n$  taxa can be resolved in  $(2n - 3)!/2^{n-2}(n - 2)!$  different ways (Felsenstein, 1978) (e.g., for eight taxa, 135,135 different resolutions are possible). However, all the different possible sets of contrasts give the same result when analyzed by regression through the origin; the slope and correlation coefficient, and hence the significance of the relationship between two variables as judged with  $n - 2$  degrees of freedom, are the same in every case. Figure 1 shows the three possible resolutions of a three-way polytomy, the plot of the standardized independent contrasts from each of them, and the nonphylogenetic regression (in which species values are taken to be independent and degrees of freedom are  $n - 2$  because an intercept is estimated). This congruence between phylogenetic and nonphylogenetic methods should not surprise us because the evolutionary models underlying the two approaches are identical: the subtaxa are all assumed to have radiated simultaneously from their common ancestor and hence to be statistically independent (Harvey and Pagel, 1991). The independent-contrasts method gives the correct estimates of correlation and  $P$  values even when applied to a star phylogeny; thus, it does not mislead, contrary to suggestions in Gittleman and Luh (1992:401).

#### *Lack of Information: Soft Polytomies*

In practice, working phylogenies often contain polytomies that definitely do not represent simultaneous radiations. For example, the bat genus *Rhinolophus* contains 64 extant species (Corbet and Hill, 1991), and little is known of its intrageneric phy-

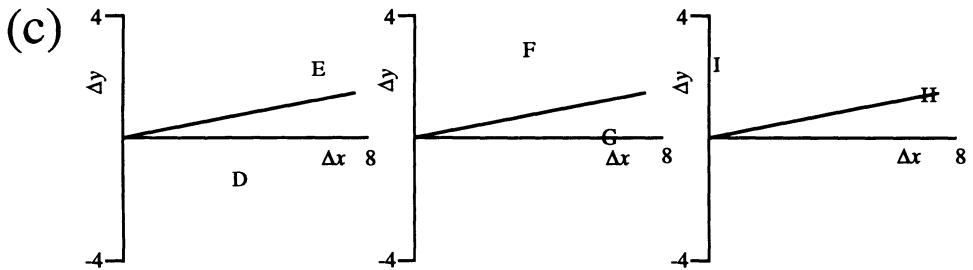
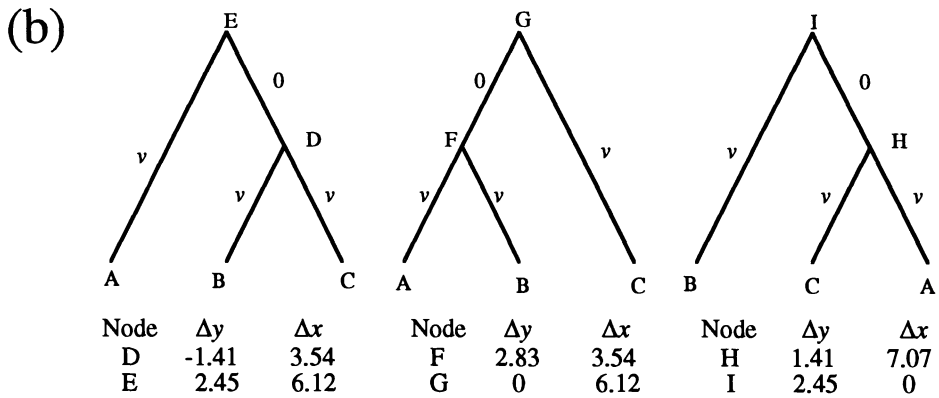
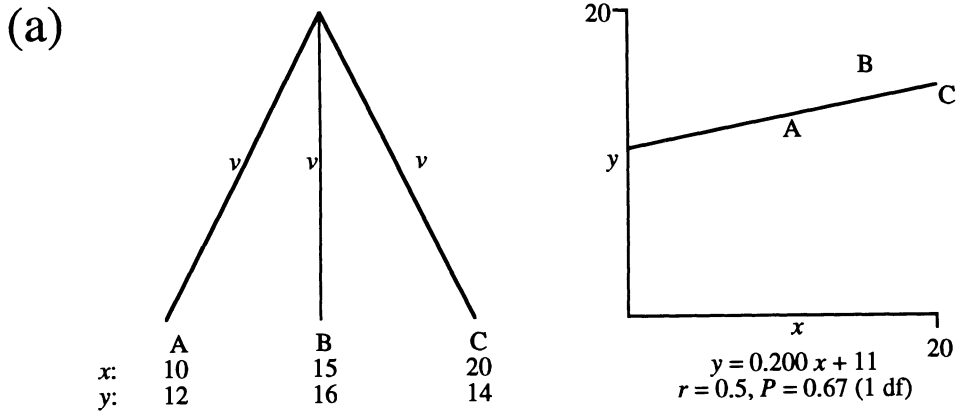


FIGURE 1. A single lineage split instantaneously, giving rise to three modern descendants, A, B, and C. (a) Present-day values of two characters,  $x$  and  $y$ , for each species and the regression of species values. (b) Three ways that the polytomy can be resolved into bifurcations separated by a branch of zero length (the other branches still have length  $v$ ) and the standardized independent contrasts ( $\Delta x$  and  $\Delta y$ ) calculated according to Felsenstein (1985). Note that for each contrast the direction of subtraction is arbitrary and thus so are the signs. (c) All three resolutions give exactly the same slope, correlation, and  $P$  value when the contrasts are analyzed by regression through the origin. Furthermore, these estimates are the same as those given by the regression of species values shown in (a).

logeny. Treating the phylogeny of this genus as a star leading to 64 tips most likely would not be a good representation of biological reality. Rather, some groups of species within the genus are closely related, but we do not know which ones. Such sister-group relationships that are not represented in the working phylogeny are known as "unrecognized phylogeny" (Grafen, 1989), which is essentially a synonym for soft polytomy. With respect to hypothesis testing, unrecognized phylogeny invalidates the procedure described above for independent contrasts approaches and for any other methods that assume all polytomies to be hard (e.g., Cheverud et al., 1985; Bell, 1989; Gittleman and Kot, 1990; Lynch, 1991) in exactly the same way that phylogeny in general invalidates non-phylogenetic regressions; some groups of taxa share character states by common descent rather than by convergent evolution, so their values are not statistically independent. Figure 2 highlights the difference between a hypothetical working phylogeny with soft polytomies and the associated "true" phylogeny. Garland et al. (1993) discussed how polytomies affect comparative methods that use computer-simulated null distributions for hypothesis testing (e.g., Martins and Garland, 1991).

Grafen (1989) suggested a way around the impasse of unrecognized phylogeny when using independent contrasts (Felsenstein, 1985) for multiple regression analyses. He assumed that hard polytomies are very much the exception in nature, so polytomies in the working phylogeny must be expressions of ignorance. Because of the statistical nonindependence among the  $n$  daughters of a node, it would be invalid to make  $n - 1$  comparisons among them. Instead, Grafen's method collapses the information at a node into a single "comparison" (technically a linear contrast: see Grafen [1989, 1992] for a definition and details). By taking only one piece of information from each node and hence only one associated degree of freedom for hypothesis testing, Grafen's (1989) approach avoids inflated type I error rates due to unrecognized phylogeny. This principle that each

node should contribute only one item of information to the statistical test was subsequently adopted in other linear contrasts approaches (i.e., modifications of Felsenstein's [1985] original presentation: Pagel and Harvey, 1989, 1992; Harvey and Pagel, 1991; Pagel, 1992). When applied to the working phylogeny shown in Figure 2a, these methods would compute only three contrasts: one within the taxon ABCDE, one within FGHI, and one between these two taxa.

The distinction between statistical estimation and hypothesis testing is important here. The various works by Grafen, Harvey, and Pagel have emphasized hypothesis testing. The primary aim was to avoid overestimating the degrees of freedom associated with an incompletely resolved working phylogeny. Table 1 lists the degrees of freedom associated with resolved versus unresolved nodes and with working phylogenies containing an arbitrary number of nodes ( $p$ ; including the root node), regardless of whether they are dichotomous or polytomous, in a way fully consistent with the procedures of Grafen, Harvey, and Pagel. When unresolved nodes are present, it is only possible to bound the degrees of freedom; they will be no more than  $n - 2$  (i.e., the number of contrasts minus one, when all polytomies are hard) and no fewer than the number of nodes minus 1 (or 1, as opposed to 0, in the special case of a completely unresolved star phylogeny). Significance tests will be liberal with  $n - 2$  degrees of freedom and conservative with  $p - 1$  degrees of freedom. Statistical power will be affected accordingly. Figure 3 illustrates an example for a phylogeny of eight species. When reporting such uncertainty about  $P$  values, it may help to indicate the fewest degrees of freedom that would be required for a given relationship to be considered statistically significant, e.g., at  $\alpha = 0.05$  (cf. Walton, 1993: table 3).

We differ from Grafen, Harvey, and Pagel as to how many contrasts should be computed for an unresolved node and hence how one should estimate an evolutionary relationship. Computing only a

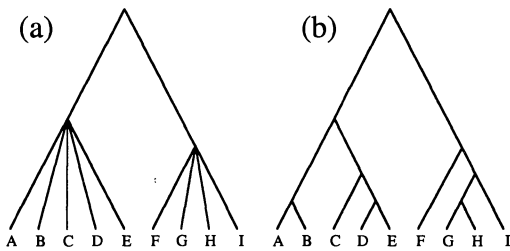


FIGURE 2. Two possible representations of the relationships among nine taxa. (a) Typical working phylogeny that might be taken from a taxonomic classification. Nodes in classifications commonly have many daughter taxa. (b) The true phylogeny includes many sister-taxon relationships not shown in the classification (unrecognized phylogeny), so it would be wrong to regard all daughters of a node in (a) as being statistically independent.

single contrast for each unresolved node will lead to cases in which only one (e.g., Fig. 3f) or a few (e.g., Figs. 2a, 3e) contrasts are computed for an entire tree, resulting in trimodal distributions of slopes or correlations (e.g., Gittleman and Luh, 1992) and poor estimates of evolutionary correlations (cf. Martins and Garland, 1991). We therefore recommend that the full  $n - 1$  contrasts should always be computed and used for estimation by correlation or regression through the origin. This procedure is simple and less arbitrary than previous proposals, and it may lead to better estimates (e.g., lower mean squared error), although this has not yet been studied.

#### Near-simultaneous Radiations: "Firm" Polytomies

In some cases, an unresolved node may be fully dichotomous in reality but may actually resemble a hard polytomy. For example, the topology illustrated in Figure 4 is the same as that in Figure 2b, but the tree looks much more like that in Figure 2a because the shared branches within ABCDE and within FGHI are very short. Because the shared branches (unrecognized phylogeny; they are not present in Fig. 2a, which is being used as the working phylogeny) are so very much shorter than the unshared ones (independent evolution), the working phylogeny shown in Figure 2a is actually a rather good descrip-

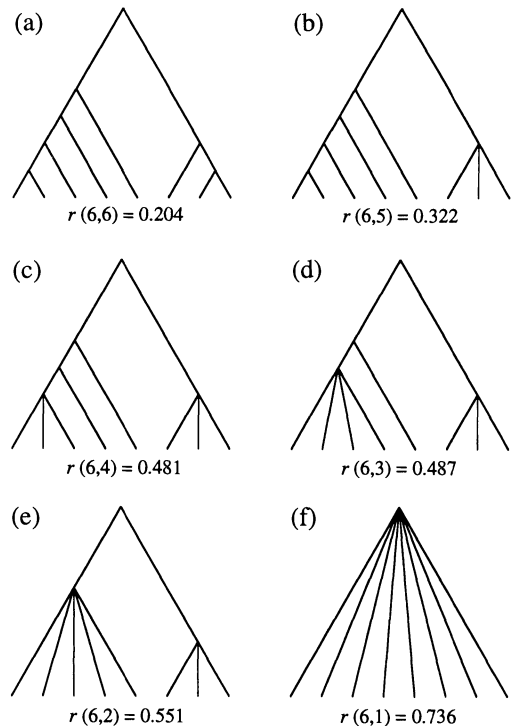


FIGURE 3. Illustration of how degrees of freedom vary in relation to the number of unresolved nodes in a working phylogeny. In this example, branch lengths have been set to arbitrary values suggested by Grafen (1989, 1992). The height of a node is one less than the number of tips descending from it. Note that the number of contrasts computed and used to estimate the evolutionary slope through the origin is always the number of terminal taxa minus one, i.e., seven. Degrees of freedom for hypothesis testing, however, vary between a maximum of the number of contrasts minus one and a minimum of the number of nodes minus one (except when the tree is completely unresolved, in which case the minimum is 1 df, not zero). We used the following arbitrary data for two characters for the eight tip species (from left to right): 1, 35; 2, 10; 3, 65; 4, 30; 5, 75; 6, 63; 7, 95; 8, 69. For each working phylogeny,  $r$  is the estimate of the correlation using standardized independent contrasts, and the maximum and minimum associated degrees of freedom for hypothesis testing are shown in parentheses. For (f), the independent contrasts estimate of the correlation is identical to the nonphylogenetic Pearson product-moment correlation coefficient (not computed through the origin). Only the estimate of 0.736 would be judged significant with 6 df (two-tailed critical value = 0.707), and none of the estimates would be significant with 5 df (critical value = 0.755) or fewer.

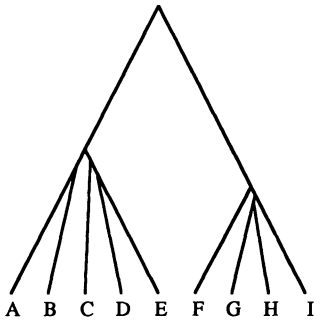


FIGURE 4. The same topology as Figure 2b but with much less unrecognized phylogeny because the shared branches within taxa ABCDE and FGHI are very short.

tion of the correct expectations for variances and covariances of the tip species' phenotypes. For instance, although G and H are truly more closely related to each other than either is to F or I, they are only very slightly so, and we would therefore not expect their phenotypes to be noticeably more similar. In this case, then, we might treat the polytomies ABCDE and FGHI as (almost) hard and claim (almost) the maximum of seven degrees of freedom for the entire tree rather than the minimum of two, without incurring much of an increase in type I error rate when hypothesis testing with independent contrasts. Consistent with this suggestion, Martins and Garland (1991: fig. 3a) performed simulations along a tree for which all species had radiated almost, but not quite, instantaneously from a common ancestor (i.e., almost a star phylogeny) and found that an ordinary nonphylogenetic Pearson correlation of tip values yielded acceptable type I error rates for estimating a correlation (Martins and Garland, 1991: table 6).

If we can identify some polytomies as being near-simultaneous radiations, as in Figure 4, then we can reasonably claim more than the minimum degrees of freedom (see Table 1 and Fig. 3) and hence gain extra statistical power without making our tests invalid (i.e., unduly inflating the type I error rate). It will, however, always be up to the practitioner to decide for each polytomy whether it should be treated as soft, firm, or hard. No strict rules

are possible, but three sorts of evidence suggest that a polytomy represents a nearly simultaneous radiation and can be treated as (almost) hard.

1. The polytomy is based on one or more cladograms with many characters per taxon but no synapomorphies resolving the polytomy. The absence of characters suggesting an internal branch implies that this branch persisted for only a short time at most.
2. The polytomy is based on molecular distance analyses using full data matrices (all species contrasted to all others). If so, then branches not shown in the working phylogeny cannot be longer than the confidence limits associated with the estimates of node ages.
3. Fossil evidence shows the radiation to have been rapid.

In contrast, the following features are danger signs and indicate that a polytomy must be treated as soft to avoid inflated type I error rates.

1. The polytomy comes from a standard taxonomy featuring taxonomic levels.
2. The polytomy is from a molecular distance analysis where the data matrix is incomplete. If only some species are used as tracers, polytomies will inevitably result unless the tracers happen always to be the most distal taxa (cf. Bledsoe, 1987; Lanyon, 1992).
3. The polytomy is taken from one or more cladograms where the ratio of informative characters to taxa is low: resolution in such cladograms is likely to be poor whatever the true phylogeny (Henderson et al., 1989; Penny et al., 1991).

When in doubt as to the "firmness" of a given polytomy, assume the worst—that it hides lots of unrecognized phylogeny—and test hypotheses with the fewer degrees of freedom, as indicated in Table 1.

#### CONCLUSION

Comparative tests must be based on an estimate of phylogeny, which is often incompletely resolved. Most polytomies are

soft, hiding an unknown bifurcating structure (unrecognized phylogeny). For the method of phylogenetically independent contrasts, we consider how to treat soft polytomies for both estimation and hypothesis testing. Studies of interspecific allometry, for example, often aim to estimate the form of an evolutionary relationship between characters. Some previous workers have suggested complicated ways to extract a single independent contrast from a given soft polytomy. Instead, we argue (consistent with Felsenstein's [1985] original suggestion) that soft polytomies should be resolved arbitrarily by setting branch lengths between internal nodes to zero. This procedure will yield a single slope or correlation coefficient, irrespective of how the polytomy is arbitrarily resolved, and unlike previous suggestions, it does not depend on the tip data being analyzed. When using independent contrasts to test hypotheses, however, unrecognized phylogeny must be taken into account or the test will invalidly claim too many degrees of freedom, thus yielding inflated type I errors. When soft polytomies are present, it is possible to bound the degrees of freedom available for hypothesis testing. Our pragmatic guidelines can be followed for assessing just how soft a polytomy might be so that the maximum reasonable degrees of freedom can be claimed.

#### ACKNOWLEDGMENTS

We thank Michael Bulmer, Alan Grafen, Paul Harvey, and Arne Mooers for helpful discussion and David Baum, Paul Harvey, and Arne Mooers for comments on this or earlier versions of this manuscript. We especially thank Joe Felsenstein for the motivation and for his insight. A.P. is funded by the U.K. Natural Environment Research Council (GR3/8151) and T.G. is funded by the U.S. National Science Foundation (BSR-9157268, DEB-9220872).

#### REFERENCES

- BELL, G. 1989. A comparative method. *Am. Nat.* 133:553-571.
- BLEDISOE, A. H. 1987. Estimation of phylogeny from molecular distance data—The issue of variable rates. *Auk* 104:563-565.
- 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.
- CORBET, G. B., AND J. E. HILL. 1991. A world list of mammalian species, 3rd edition. Natural History Museum Publications, London.
- FELSENSTEIN, J. 1978. The number of evolutionary trees. *Syst. Zool.* 27:27-33.
- FELSENSTEIN, J. 1985. Phylogenies and the comparative method. *Am. Nat.* 125:1-15.
- GARLAND, T., JR., A. W. DICKERMAN, C. M. JANIS, AND J. A. JONES. 1993. Phylogenetic analysis of covariance by computer simulation. *Syst. Biol.* 42:265-292.
- GARLAND, T., JR., P. H. HARVEY, AND A. R. IVES. 1992. Procedures for the analysis of comparative data using phylogenetically independent contrasts. *Syst. Biol.* 41:18-32.
- GITTLEMAN, J. L., AND M. KOT. 1990. Adaptation: Statistics and a null model for estimating phylogenetic effects. *Syst. Zool.* 39:227-241.
- GITTLEMAN, J. L., AND H.-K. LUH. 1992. On comparing comparative methods. *Annu. Rev. Ecol. Syst.* 23:383-404.
- GRAFEN, A. 1989. The phylogenetic regression. *Philos. Trans. R. Soc. Lond. B* 326:119-156.
- GRAFEN, A. 1992. The uniqueness of the phylogenetic regression. *J. Theor. Biol.* 156:405-423.
- HARVEY, P. H., AND M. D. PAGEL. 1991. The comparative method in evolutionary biology. Oxford Univ. Press, Oxford, England.
- HENDERSON, I. M., M. D. HENDY, AND D. PENNY. 1989. Influenza viruses, comets and the science of evolutionary trees. *J. Theor. Biol.* 140:289-303.
- LANYON, S. M. 1992. Review of *Phylogeny and Classification of Birds: A Study in Molecular Evolution*, by C. G. Sibley and J. E. Ahlquist. *Condor* 94:304-307.
- LYNCH, M. 1991. Methods for the analysis of comparative data in evolutionary ecology. *Evolution* 45:1065-1080.
- MADDISON, W. P. 1989. Reconstructing character evolution on polytomous cladograms. *Cladistics* 5:365-377.
- MARTINS, E. P., AND T. GARLAND, JR. 1991. Phylogenetic analyses of the correlated evolution of continuous characters: A simulation study. *Evolution* 45:534-557.
- PAGEL, M. D. 1992. A method for the analysis of comparative data. *J. Theor. Biol.* 156:431-442.
- PAGEL, M. D., AND P. H. HARVEY. 1989. Comparative methods for examining adaptation depend on evolutionary models. *Folia Primatol.* 53:203-220.
- PAGEL, M. D., AND P. H. HARVEY. 1992. On solving the correct problem: Wishing does not make it so. *J. Theor. Biol.* 156:425-430.
- PENNY, D., M. D. HENDY, AND M. A. STEEL. 1991. Testing the theory of descent. Pages 155-183 in *Phylogenetic analysis of DNA sequences* (M. M. Miyamoto and J. Cracraft, eds.). Oxford Univ. Press, New York.
- WALTON, M. 1993. Physiology and phylogeny: The evolution of locomotor energetics in hylid frogs. *Am. Nat.* 141:26-50.

Received 15 March 1993; accepted 21 June 1993