Syst Biol. 41(1):18-32, 1992

# PROCEDURES FOR THE ANALYSIS OF COMPARATIVE DATA USING PHYLOGENETICALLY INDEPENDENT CONTRASTS

THEODORE GARLAND, JR.,[1] PAUL H. HARVEY,[2] AND ANTHONY R. IVES[1]

[1]*Department of Zoology, University of Wisconsin, Madison, Wisconsin 53706, USA*
[2]*Department of Zoology, University of Oxford, South Parks Road, Oxford OX1 3PS, England*

*Abstract.*—We discuss and clarify several aspects of applying Felsenstein's (1985, Am. Nat. 125: 1-15) procedures to test for correlated evolution of continuous traits. This is one of several available comparative methods that maps data for phenotypic traits onto an existing phylogenetic tree (derived from independent information). Application of Felsenstein's method does not require an entirely dichotomous topology. It also does not require an assumption of gradual, clocklike character evolution, as might be modeled by Brownian motion. Almost any available information can be used to estimate branch lengths (e.g., genetic distances, divergence times estimated from the fossil record or from molecular clocks, numbers of character changes from a cladistic analysis). However, the adequacy for statistical purposes of any proposed branch lengths must be verified empirically for each phylogeny and for each character. We suggest a simple way of doing this, based on graphical analysis of plots of standardized independent contrasts versus their standard deviations (i.e., the square roots of the sums of their branch lengths). In some cases, the branch lengths and/or the values of traits being studied will require transformation. An example involving the scaling of mammalian home range area is presented. Once adequately standardized, sets of independent contrasts can be analyzed using either linear or nonlinear (multiple) regression. In all cases, however, regressions (or correlations) must be computed through the origin. We also discuss ways of correcting for body size effects and how this relates to making graphical representations of relationships of standardized independent contrasts. We close with a consideration of the types of traits that can be analyzed with independent contrasts procedures and conclude that any (continuous) trait that is inherited from ancestors is appropriate for analysis, regardless of the mechanism of inheritance (e.g., genetic or cultural). [Allometry; body size; branch lengths; comparative method; evolutionary rates; functional morphology; home range; statistics.]

One of the fundamental topics in evolutionary biology is the manner in which the evolution of different traits is correlated throughout a phylogenetic lineage (e.g., Ridley, 1983; Felsenstein, 1985; Donoghue, 1989; Maddison, 1990; Harvey and Pagel, 1991; Martins and Garland, 1991). Methods for incorporating phylogenetic information into statistical analyses of the correlated evolution of continuous traits are now receiving intensive study. Although new or improved methods continue to appear (e.g., Felsenstein, 1988; Bell, 1989; Grafen, 1989; Gittleman and Kot, 1990; Harvey and Pagel, 1991; Lynch, 1991; Martins and Garland, 1991; Pagel, 1992), Felsenstein's (1985) independent contrasts approach is now being widely employed, many of its statistical properties have been verified analytically and through computer simulation techniques (Grafen, 1989; Martins and Garland, 1991), and computer programs are available to do the necessary calculations (Grafen, 1989; Martins and Garland, 1991; J. Felsenstein's latest version of PHYLIP; A. J. Purvis, pers. comm., whose programs are available from the second author).

The independent contrasts method is straightforward in principle (Burt, 1989; Harvey and Purvis, 1991), but it does encompass several analytical complexities that may bewilder potential users. In particular, the possible sources and treatment of phylogenetic branch lengths have not been adequately discussed. We therefore examine several technical aspects of applying the independent contrasts method and illustrate these with an empirical example. Where complicated arguments or mathematical proofs are required, we have rel-

egated them to appendices. We assume the reader will have read Felsenstein's (1985) original description and possibly relevant portions of Grafen (1989), Harvey and Pagel (1991), and/or Martins and Garland (1991).

### What Does the Independent Contrasts Method Require?

Three types of information are required to use Felsenstein's (1985) method: (1) data for two or more phenotypic traits for a series of extant and/or extinct species, (2) the cladistic relationships of these species, and (3) phylogenetic branch lengths in units of expected variance of change. Although a fully dichotomous topology simplifies the analysis, it is not required. Grafen (1989, 1992), Harvey and Pagel (1991), Pagel (1992), and Pagel and Harvey (1992) explained how unresolved polytomies can be analyzed, although with some loss of information and of statistical power (see also Felsenstein, 1985:10). In essence, if our knowledge of topology is incomplete, then we are forced to lump some species together (or delete some from the analysis) when computing independent contrasts. A number of workers have used the independent contrasts method in the absence of complete topological information. Sometimes, they have used taxonomic information alone or in part to construct a topology, assuming that named taxa (e.g., each genus within a family) represent monophyletic groups (cf. Cheverud et al., 1985; see Grafen [1989], Harvey and Pagel [1991], and Pagel [1992] for further details and examples). Obviously, such a practice may be misleading if the taxonomy is not cladistic, but it does allow analysis of the data now, rather than waiting for actual phylogenetic information to become available. For the remainder of this paper, we assume that a fully dichotomous phylogeny is available and that its topology is known without error. In the process of designing a comparative study (i.e., deciding which species should be measured, given limited resources), it might be possible to avoid sets of species whose cladistic relationships were in doubt (cf. Felsenstein, 1985:13). Many comparative studies, however, involve data compiled from the literature, and most practitioners feel compelled to analyze all available data.

### Rationale and Brief Overview of the Independent Contrasts Method

Because species are descended in a hierarchical fashion from common ancestors, they generally cannot be considered as independent data points in statistical analyses (review in Harvey and Pagel, 1991:Ch. 2). This phylogenetic nonindependence reduces degrees of freedom available for hypothesis testing, lowers statistical power, and affects parameter estimation (Grafen, 1989; Harvey and Pagel, 1991; Martins and Garland, 1991). Felsenstein (1985) therefore proposed computing (weighted) differences ("contrasts") between the character values of pairs of sister species and/or nodes, as indicated by a phylogenetic topology, and working down the tree from its tips. This procedure results in $n - 1$ contrasts from $n$ original tip species. Insofar as the ancestral nodes are correctly determined, each of these contrasts is independent of the others in terms of the evolutionary changes that have occurred to produce differences between the two members of a single contrast (although phenomena such as character displacement might violate this assumption). Because the $n - 1$ contrasts are statistically independent, they generally can be employed in standard statistical analyses.

Contrasts involving longer periods of time are likely to be greater in absolute value and would, in effect, be given greater weight in statistical analyses such as correlation or regression. This increased weight would be wrong because it would negate the use of standard probability tables for hypothesis testing. (However, such a weighting may be desirable for estimation of certain types of evolutionary correlations [see Martins and Garland, 1991].)

The usual probability tables can be employed if independent contrasts are first standardized. Standardization customarily denotes subtraction of the mean and division by the standard deviation. The ex-

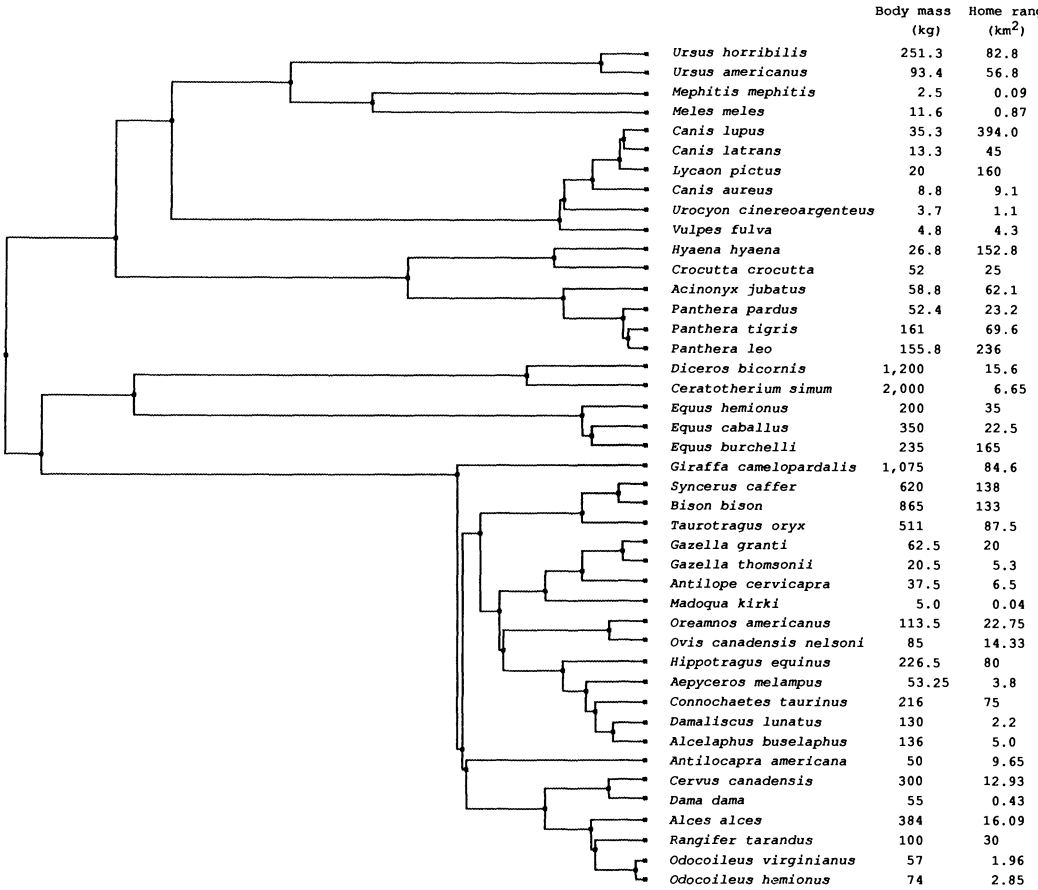


FIGURE 1. Phylogeny and data for body mass (kg) and home range area (km²) for 43 species of mammals (from Garland and Janis, 1992; Janis, in press). Branch lengths are in units of time, as estimated from first fossil appearances and from molecular clock information (sources in Garland and Janis, 1992). The shortest branch segment depicted is 0.5 million years; the basal split, between the Carnivora and the ungulates, is at 70 million years (complete specifications are available from the senior author). Although many other comparisons might be of interest, we note only the comparison of these two clades, which may generally be considered as predators and prey (Garland and Janis, 1992; Janis, in press).

pected mean of any set of contrasts is zero because the direction of subtraction is arbitrary (see below and Appendix 1), so only standard deviations are needed. If branch lengths in units of expected variance of change are available, then the standard deviation of a contrast is the square root of the sum of its branch lengths.

## Obtaining and Validating Branch Lengths: An Empirical Example

Branch lengths in units of expected variance of change are required for proper standardization of contrasts, but how branch lengths may be estimated is not obvious. Intuitively, estimates of divergence times would be most appropriate (e.g., Fig. 1). In fact, if evolution behaves like Brownian motion, with increases or decreases in a trait occurring randomly and independently of the value of the trait (Felsenstein, 1985, 1988), then time (or number of generations; cf. Sibley and Ahlquist, 1991) should be used. Whether character evolution ever behaves as pure Brownian motion is questionable.

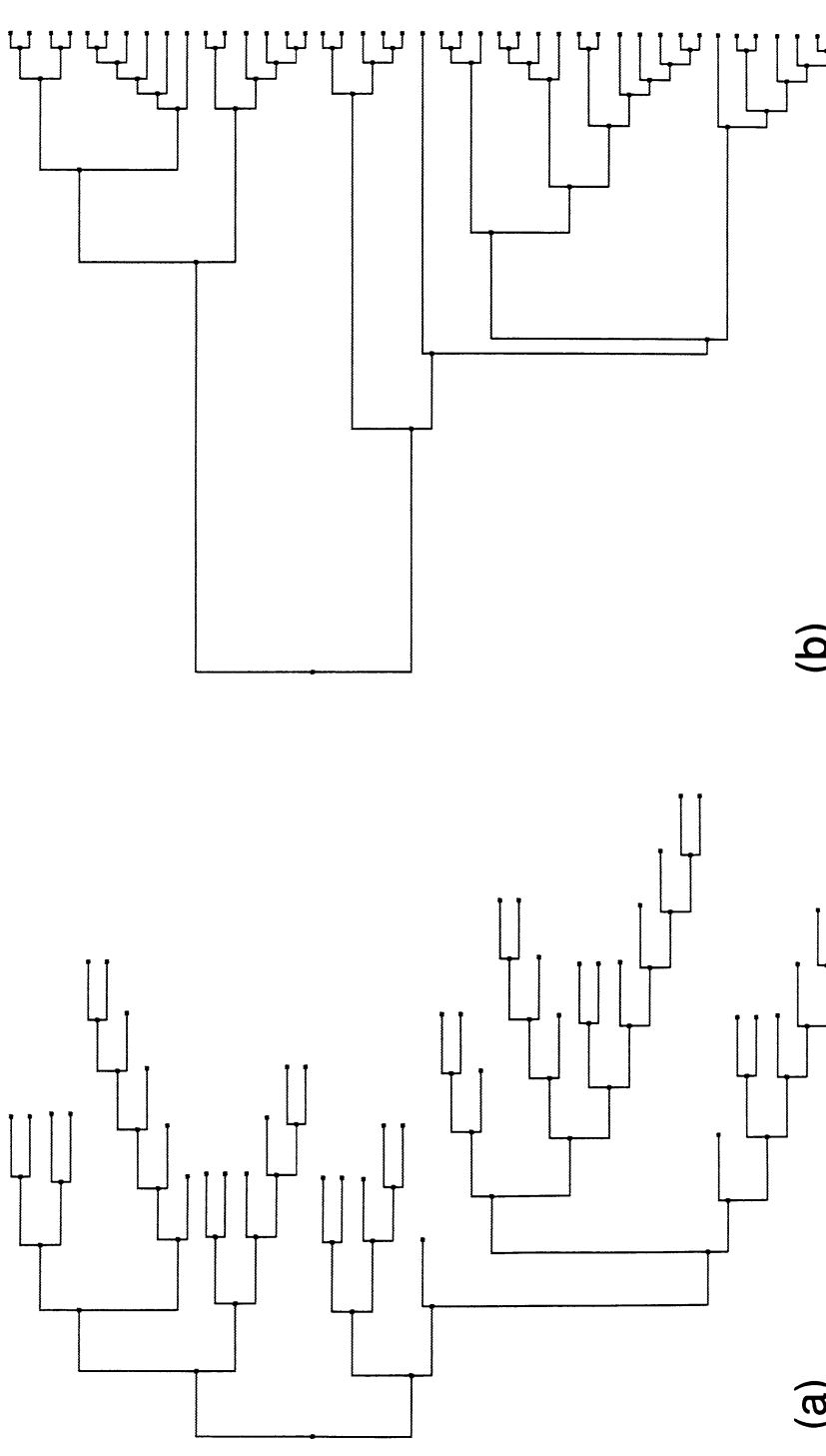


FIGURE 2. Topology as in Figure 1, but with branch lengths (a) $\log_{10}$ transformed and (b) set arbitrarily as described in Grafen's (1989) Figure 2. For (b), the depth of nodes is taken as one less than the number of species descended from them. Thus, the shortest branches at the right side of our tree are one unit long, and the basal node is at a depth of 42 units because it leads to 43 measured species.

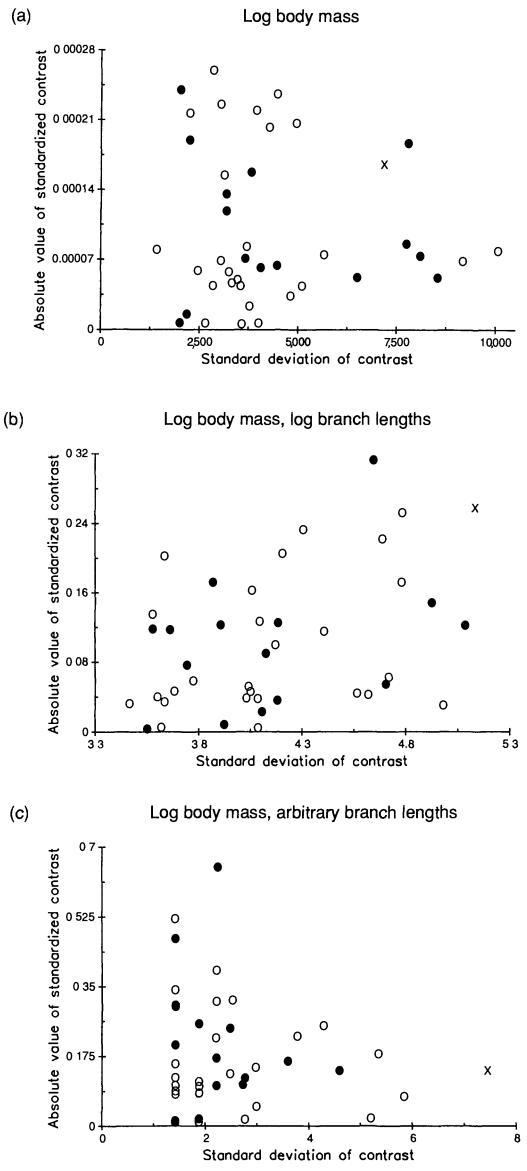


FIGURE 3. Absolute values of standardized independent contrasts in $\log_{10}$ body mass plotted versus their standard deviations (square roots of sums of branch lengths): (a) using the raw time branch lengths of Figure 1, (b) using the log-transformed branch lengths of Figure 2a, and (c) using the arbitrary branch lengths of Figure 2b. Open circles are contrasts within the ungulates, closed circles are contrasts within the Carnivora, and X represents the basal contrast. In (b), note that the basal contrast has the highest standard deviation but does not appear to involve the longest branch lengths depicted in Figure 2a. This apparent discrepancy occurs because all branch segments, except for those involving two tip species, are lengthened during the computation of standard deviations to reflect uncertainty in the estimation of nodes (see formulas in Felsenstein [1985:11]).

Aside from time, many other sources of possible branch lengths exist, such as genetic distances or perhaps the number of steps along a branch as indicated by a cladistic analysis of morphological characters (other than those being studied). Alternatively, given a rooted topology, the characters themselves might be used to obtain branch lengths. For example, a distance matrix could be constructed and used in the KITSCH or FITCH procedures of J. Felsenstein's PHYLIP package. Grafen (1989: 123) suggested an arbitrary but consistent way of assigning branch lengths if no branch length information is available (Fig. 2b) (see also Pagel, 1992).

Regardless of what "starter" branch lengths are employed (see also Cheverud et al., 1985; Gittleman and Kot, 1990; Garland et al., 1991; Harvey and Pagel, 1991; Lynch, 1991; Martins and Garland, 1991), independent contrasts must be adequately standardized so that they will receive equal weighting in subsequent correlation or regression analyses. As argued elsewhere (Garland et al., 1991; Garland, 1992; Garland and Janis, 1992), one approach to verification is to plot the absolute value of each standardized independent contrast versus its standard deviation (i.e., the square root of the sum of its branch lengths). (The contrasts and their standard deviations can be obtained easily from the "CMSINGLE" program of Martins and Garland [1991].) Any significant linear or nonlinear trend in the plot indicates that the contrasts are not adequately standardized. This approach is equivalent to plotting any ratio versus its denominator to determine whether scaling effects of the denominator are effectively removed.

Figures 3a and 4a depict such plots for body mass and for home range area, respectively, using the branch lengths of Figure 1. Figure 3a shows no clear relationship, and a Pearson product-moment correlation is not significant ($r = -0.070$,

$n$ = 42 contrasts, two-tailed $P$ = 0.658). Contrasts at lower standard deviations appear perhaps more variable, but there are many more of them, so this is difficult to judge. Figure 4a, for home range area, shows a negative trend ($r$ = −0.299, $P$ = 0.055), indicating that contrasts involving long branch lengths (i.e., large standard deviations) tend to be overstandardized relative to contrasts involving shorter branch lengths.

Given that a plot such as Figure 4a shows some trend, what options are available? First, the phenotypic data for the tip species can be transformed, new contrasts computed and standardized, and a new plot examined. Because home range area had already been log transformed, given expectations for its allometric scaling with body mass (e.g., Janis, in press), this approach would not seem the best choice for our present example.

Second, the branch lengths themselves can be transformed (Garland, 1992; Garland and Janis, 1992). For an initial plot showing a negative trend (as in Fig. 4a), a log transformation of branch lengths may be appropriate because it will shrink long branches (and so standard deviations) relative to short ones (if some branch lengths were ≤1, then all branches would need to be multiplied by a constant prior to log transformation). An initial plot showing a positive trend (indicating that contrasts involving large branch lengths are understandardized) might be corrected by squaring branch lengths. In general, any transformation of possible use for tip data (e.g., Sokal and Rohlf, 1981) might also be tried for branch lengths.

Figure 2a depicts the phylogeny of Figure 1 but with its branches log transformed. These $\log_{10}$-transformed branch lengths yield more even standardization for home range area contrasts ($r$ = 0.021, $P$ = 0.895) (Fig. 4b). However, log-transformed branch lengths do not work so well for body mass, as they result in a significant positive trend (Fig. 3b; $r$ = 0.387, $P$ = 0.011).

Three general points should be noted. First, branch lengths, either before or after any transformation, are not constrained to indicate contemporaneous tips. For example, if fossil taxa were included in the analysis and branch lengths were in units of time, then they would appear as "pruned" branches terminating before the present.

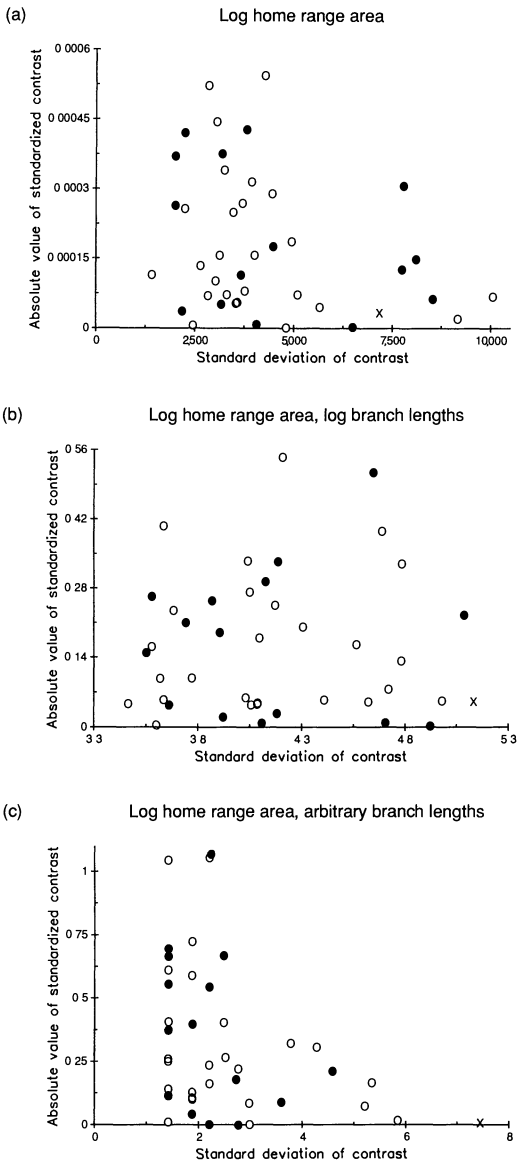


FIGURE 4. Absolute values of standardized independent contrasts in $\log_{10}$ home range area plotted versus their standard deviations. Branch lengths for (a) are from Figure 1 (raw), for (b) from Figure 2a ($\log_{10}$ transformed), and for (c) from Figure 2b (arbitrary). Only (b) indicates adequate standardization. Symbols as in Figure 3.

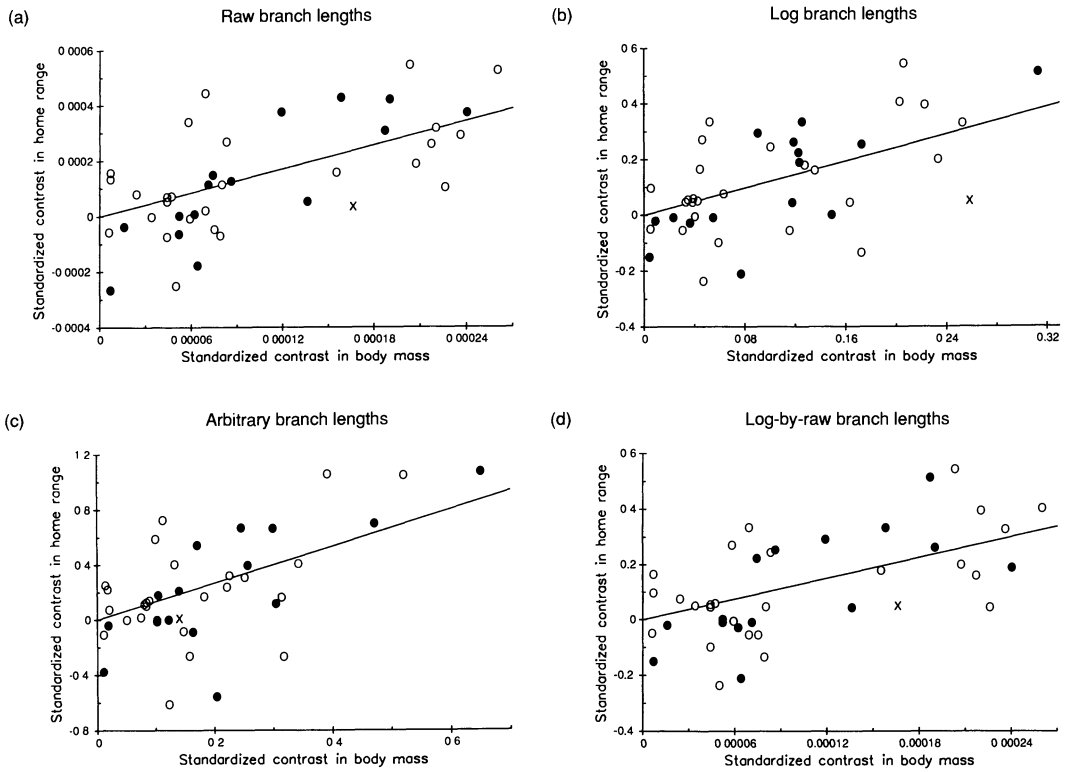


FIGURE 5. Bivariate scatterplots of relationship between standardized independent contrasts in $\log_{10}$ home range area and $\log_{10}$ body mass. Branch lengths for (a) are from Figure 1 (raw), for (b) from Figure 2a ($\log_{10}$ transformed), and for (c) from Figure 2b (arbitrary) (Grafen, 1989). Raw branch lengths for body mass and log branch lengths for home range area—the "best" combination of branch lengths, as suggested by Figures 3 and 4—are used for (d). Symbols as in Figure 3. Correlation coefficients and slopes are presented in Table 1.

Second, the statistical adequacy of any proposed branch lengths should be viewed as an empirical issue (cf. Grafen, 1989, 1992; Harvey and Pagel, 1991; Pagel and Harvey, 1992). Third, different transformations—and indeed entirely different branch lengths, based on different information—can be used for different characters. The use of different branch lengths for different characters would be analogous to using different transformations for the two variables in an ordinary correlation or regression. In a regression analysis, use of different branch lengths for the dependent and independent variables would, of course, complicate interpretation of the slope for a set of standardized independent contrasts (e.g., in an allometric analysis; see Table 1).

Grafen's (1989) "phylogenetic regression" is an application of the independent contrasts approach. Its perspective is that of multiple regression, as opposed to the correlation perspective of Felsenstein's (1985) original presentation. Grafen used maximum-likelihood techniques to estimate relationships of sets of standardized independent contrasts and simultaneously to (in effect) transform branch lengths. More specifically, Grafen's (1989) method first scales all branch lengths such that the basal node is at a height of 1 and the tip species have a height of zero (if the tree had noncontemporaneous tips, then only the highest tip would be at zero height). His implementation using the Generalised Linear Interactive Modelling (GLIM) system then estimates a parameter rho, which

is the positive power to which the heights (as opposed to the branch lengths themselves) should be raised, while simultaneously estimating regression slopes.

Three differences between Grafen's (1989) method and our suggestions should be noted. First, if the starter branch lengths indicated contemporaneous tip species (as in Fig. 1), then they would remain contemporaneous after transformation by rho (as shown in Grafen's [1989] Figs. 2 and 3). This effect differs from that of direct transformation of the branch lengths themselves, which will yield noncontemporaneous tips (cf. Figs. 1, 2a). Second, Grafen's implementation uses the same set of branch lengths for all characters in the data set. We suggest that different branch lengths, or different transformations of the same branch lengths (e.g., Fig. 5d), can be used for different characters. Third, although Grafen's (1989) original presentation used initial branch lengths derived as in our Figure 2b, this is not necessary. One could, for example, start with the branch lengths in Figure 1 and then implement Grafen's phylogenetic regression with GLIM.

From the perspective of regression, the residuals of a regression or multiple regression involving standardized independent contrasts must also be examined (Grafen, 1989). If residuals show patterns such as nonlinearity or heteroscedasticity, then (1) different transformations of the character data or of the branch lengths can be tried or (2) weighted regression can be used (Harvey and Pagel, 1991:151–152).

Returning to our empirical example, Figure 5 shows a series of bivariate scatterplots of standardized independent contrasts, using different branch lengths; Table 1 lists the correlations and slopes estimated for these different scatterplots. Figure 5a uses raw branch lengths in units of time (from Fig. 1), Figure 5b uses $\log_{10}$ transformations of these branch lengths (as shown in Fig. 2a), Figure 5c uses the arbitrary branch lengths of Figure 2b, and Figure 5d uses a combination of raw branch lengths for body mass (as suggested by Fig. 3a) and log-transformed branch lengths for home range area (as suggested by Fig. 4b). Comparison of these plots and their residuals (not shown) indicates that Figure 4b comes closest to meeting the assumptions of regression. Nevertheless, the correlations estimated for the four alternative sets of branch lengths are quite similar (Table 1). Thus, for this example data set, the branch lengths used have essentially no effect on the final conclusions reached.

TABLE 1. Correlations and least-squares linear regression slopes (both estimated through the origin) between standardized independent contrasts in $\log_{10}$ home range area and $\log_{10}$ body mass when using alternate branch lengths (see Fig. 5). All correlations are significant at $P < 0.0001$.

| Figure | Branch lengths | Correlation | Slope |
|---|---|---|---|
| 5a | Figure 1: raw time | 0.767 | 1.4230 |
| 5b | Figure 2a: $\log_{10}$ time | 0.722 | 1.2133 |
| 5c | Figure 2b: arbitrary | 0.702 | 1.3308 |
| 5d | Figure 2a: $\log_{10}$ time for home range<br>Figure 1: raw time for body mass | 0.736 | 1,259.5 |

Slopes, however, can vary widely (see Table 1) because they are dependent on the units of measurement both of the characters (always $\log_{10}$ transformed in our examples) and the branch lengths. The least squares slopes from conventional log-log allometric plots are 1.2384 for the 16 Carnivora ($r = 0.750$), 0.9518 for the 27 ungulates ($r = 0.686$), 0.6696 for the combined sample of 43 species ($r = 0.524$), and 1.0613 for the pooled within-groups slope from the analysis of covariance. These differences reflect the fact that ungulates have smaller home ranges than do Carnivora (e.g., see Janis, in press) and are somewhat larger in body mass.

Although Figures 3–5 might suggest that branch lengths in units of time are superior to arbitrary branch lengths, this comparison is unfair. Arbitrary branch lengths can also be transformed (Grafen, 1989), just as we have done for the time branch lengths of Figure 1 to produce Figure 2a. Transformation could be done either as we have suggested or through Grafen's (1989, 1992) GLIM implementation of the phylogenetic regression.

### The Possibility of Not Standardizing Contrasts

The main reason for standardizing contrasts is to achieve equal weighting and hence allow the use of ordinary probability tables for statistical tests (e.g., Zar's [1984] Table B.16 or Rohlf and Sokal's [1981] Table 25 for Pearson product-moment correlation coefficients). As discussed elsewhere (Harvey and Pagel, 1991; Martins and Garland, 1991), however, certain types of evolutionary correlations are better estimated by nonstandardized contrasts (e.g., the computations termed "FL2G" and "FL2P" by Martins and Garland [1991]). If these correlations are desired, then significance tests are available through reference to computer-simulated empirical null distributions. Creation of these distributions requires information on branch lengths (Martins and Garland, 1991), and so it does not obviate the need for them.

Alternatively, if assuming a punctuational (Martins and Garland, 1991) or "speciational" (Rohlf et al., 1990) model of character change when applying Felsenstein's (1985) method, then all branch lengths are set equal to any arbitrary value. (Note that any set of branch lengths transformed by raising to the zero power yields a new set of branches that are all equal to unity.) However, these branch lengths are still used in the computation of nodes and hence contrasts, and they are also used to standardize the contrasts. To employ standard statistical analyses, these branch lengths still must be taken as representing expected variance of change. Under a speciational model, character change can only occur in association with a speciation event, and so the branch lengths must represent the number of speciation events that have occurred along them. The assumption is that all speciation events in the clades being studied have been accounted for, including those leading to both extant and extinct species, regardless of whether those species are represented in the phenotypic data set being analyzed (see Martins and Garland, 1991). In other words, any speciation event that has occurred anywhere in the lineages leading to the species being studied must be accounted for when determining branch lengths. Given our uncertainty as to the number of extinct species in most clades, and indeed our uncertainty as to what exactly constitutes a speciation event, the number of these events will be difficult to know for real data. In the limit, if speciation occurred each and every generation, then models of punctuational or speciational change would converge on models of gradual change (see also Friday, 1987:71–74, and references therein; Felsenstein, 1988:466). Strictly speaking, if speciation occurred each generation, the variances of contrasts of resulting tip species would not be linear with time nor precisely normal but rather something like a mixture of Poisson and normal (J. Felsenstein, pers. comm.). Even purely gradual Brownian motion models of evolutionary change will sometimes yield periods of rapid change (e.g., cf. Bookstein's [1988] Fig. 2 of a random walk process). In any case, whatever branch lengths are employed, their adequacy for standardizing the independent contrasts must be verified.

### Population Versus Species Comparisons

In most cases, we study phylogenies and phenotypic data for species, but populations can also be studied with the independent contrasts approach. To the extent that populations experience gene flow from other populations, whereas different "species" generally do not, complications may arise in estimating branch lengths. In principle, gene flow between two diverging sister populations (i.e., the members of one independent contrast) can be modeled simply by shortening branch lengths since their divergence. Thus, a comparative study involving two populations from each of several species provides no unusual analytical problems and could provide more information than a purely interspecific comparison. For example, a comparative study could be designed in which pairs of populations from each of a series of fairly distantly related species were measured.

The contrasts involving pairs of conspecific populations would then provide information on microevolutionary (within-species) phenomena, whereas the contrasts involving their nodes and those deeper in the tree would reflect macroevolutionary phenomena. Using more than two populations per species would complicate analyses because gene flow between nonsister populations would result in nonindependence of otherwise independent contrasts (cf. Felsenstein, 1982).

### Regression and Correlation Using Standardized Independent Contrasts

To calculate independent contrasts between two nodes, the values of the variables in question must be subtracted. The direction of subtraction is entirely arbitrary, and consequently there is ambiguity in assigning signs to the contrasts. For example, if $\Delta x$ and $\Delta y$ represent the values of two contrasts between two specified tips or nodes, the pair of contrasts can be given equally well by either $(\Delta x, \Delta y)$ or $(-\Delta x, -\Delta y)$, depending on the direction of subtraction from the nodes.

The ambiguity associated with determining the sign of independent contrasts places restrictions on the methods used to regress one independent contrast against another. Specifically, regression through the origin is necessary (Appendix 1; see also Grafen, 1992). As $n - 1$ contrasts are available, and only one degree of freedom is lost in regression (or correlation) through the origin (as opposed to two in conventional regression, one for the slope and one for the $y$-intercept), significance is tested with the usual $n - 2$ degrees of freedom.

Regression through the origin can be computed in most commercially available statistical packages. In SPSS/PC+ (Norusis, 1988), the "/ORIGIN" option is used somewhere before indicating the dependent variable. In SPSS, the correlation coefficients listed will be the correct ones, as will the slope. The necessary formulas to do computations by hand are presented in Appendix 2.

Regression through the origin is not limited to a single independent variable, and standard multiple regression (through the origin) techniques are applicable. Testing for clade differences in the slopes of relationships can be accomplished with dummy variables (e.g., ANCOVA). The clades are coded as 0 or 1 in dummy variables; the number of dummy variables required is one fewer than the number of clades to be compared. Contrasts not contained exclusively within one or another clade are excluded from the analysis (Garland and Janis, 1992). Crossproduct terms are then computed as each dummy variable multiplied by the independent variable (e.g., body mass). Finally, the crossproduct term(s) is (are) entered into a multiple regression with the independent variable, and its (their) significance is tested in the usual way with partial $F$ statistics. The dummy variables themselves are not tested because all regression lines must pass through the origin.

### Graphical Analyses and Nonlinearities

The indeterminate signs of the independent contrasts mean that no representation of a scatterplot of one set of contrasts versus another is unique. Because of this ambiguity, we suggest the convention of giving a positive sign to the independent contrast graphed on the horizontal axis and simultaneously switching the sign of the other independent contrast as needed (Harvey and Pagel, 1991; Garland and Janis, 1992). A difficulty arises, however, if any of the contrasts to be plotted on the horizontal axis are exactly zero because the sign of the other independent contrast becomes arbitrary. Contrasts that equal zero do not compromise the computations of slopes and correlations discussed in the appendices, although they might affect application of some nonparametric methods (see Felsenstein, 1985, 1988; Grafen, 1989; Harvey and Pagel, 1991).

Inspection and/or fitting of a line to a bivariate plot of standardized independent contrasts, with the horizontal axis "positivized" as just described, may suggest a nonzero $y$-intercept. However, nonzero $y$-intercepts are artifactual and cannot be

interpreted in any meaningful way (see also Grafen, 1992). In some cases, apparent nonzero $y$-intercepts may indicate significant nonlinearity in the relationship between sets of independent contrasts (Harvey and Pagel, 1991). Existence of significant nonlinearities may be tested either by examining residuals from the regression through the origin (e.g., using a Durbin-Watson test), by fitting polynomials and identifying significant nonlinear terms, or by fitting other types of nonlinear equations. In all cases, however, the regression equation $\Delta y(\Delta x)$ must pass through and be symmetrical about the origin, having the property that $\Delta y(\Delta x) = -\Delta y(-\Delta x)$; Appendix 1 gives further restrictions on the loss function used to fit the regression line.

The existence of significant nonlinearities in the relationship between two sets of standardized independent contrasts suggests that the evolutionary correlation is itself changing, although significant nonlinearity does not give any information regarding the precise nature of such changes. At least four explanations are possible.

1. The strength or even the sign of the evolutionary correlation may be different within different portions of the phylogeny. In this case, multiple regression can be used to compare monophyletic groups (clades) indicated by dummy variables (Garland and Janis, 1992). In our empirical example, we see no indication that the slope of the relationship between home range area and body mass differs between Carnivora and ungulates (see Fig. 5).
2. The evolutionary correlation may be different for contrasts involving relatively long versus short branch lengths (the depth of the split between two nodes). In this case, branch length and/or (branch length)$^2$ can be used as a covariate(s) in a multiple regression through the origin.
3. The evolutionary correlation may be influenced by the actual values of the traits in question (e.g., for large versus small animals). Thus, the correlation between contrasts might depend on the magnitude of the traits at the two tips or nodes used to calculate contrasts. To investigate this, multiple regression can be used in which either or both values of the trait at the tips or nodes, or perhaps their mean value, are used as covariates. However, this may be just a more complicated way of performing a nonlinear regression.
4. A nonlinear relationship between contrasts might be an artifact caused by improper standardization of the contrasts, although the adequacy of branch lengths should have been investigated previously. In our empirical example, Figure 5d seems the most clearly linear for both clades.

Related to these four points, Grafen (1989: 146) suggested plotting residuals from a regression involving standardized independent contrasts versus the heights of their corresponding nodes. All of these procedures may also be of use when the relationship is linear because they may provide additional information about the relationship between $\Delta x$ and $\Delta y$ (e.g., does it vary among clades or in relation to depth in the phylogeny?).

Standardized independent contrasts essentially indicate minimum average "rates" (or, in Bookstein's [1988] terminology, "reduced speeds") of character change occurring along two branches deriving from a common ancestor (Garland, 1992). Because they are standardized, they do not indicate absolute amounts of change. For example, a large contrast (absolute difference) between two species at the tips of a phylogeny may yield a small standardized contrast if the branch lengths of that contrast are long.

## Allometry and Removing the Effects of Body Size

A common problem in comparative studies is that the characters of interest (e.g., brain size, basal metabolic rate, home range area) are highly correlated with animal size (Clutton-Brock and Harvey, 1979; Garland

and Huey, 1987; Pagel and Harvey, 1988, 1989; Harvey and Pagel, 1991). A common solution is to compute residuals from regressions of each character on body size (these residuals will by definition be uncorrelated with body size) and then correlate the residuals. This solves only one problem, however. We still must deal with the possible confounding effects of phylogeny, which is easily done with independent contrasts (Losos, 1990; Harvey and Pagel, 1991; Martins and Garland, 1991; Garland and Janis, 1992).

First, standardized independent contrasts are computed both for the traits of interest and for the measure of body size (e.g., body mass, snout-vent length). (Log transformation, if appropriate, must be done before computing contrasts.) Second, the contrasts for body size are "positivized," as described above. Third, regressions through the origin of contrasts are computed for each independent variable on the "positivized" body size contrasts. Fourth, residuals from these regressions through the origin are computed. Such residuals will not necessarily sum to zero because that is not a necessary property of regression through the origin. Nonetheless, when tested by regression through the origin, these residuals will be perfectly uncorrelated with the body size contrasts (regression through the origin is necessary by the arguments presented above and in Appendix 1). Also, a choice must be made among least-squares, model II regression, reduced major axis, and major axis slopes to compute residuals. In any case, once these size- and phylogeny-corrected residuals are computed, they can be analyzed for correlation, again using regression through the origin (any nonlinear regression line must also be symmetrical through the origin). Graphs of the relationships between pairs of residual traits need not be "positivized" because a unique representation is already guaranteed by the second step.

Whether some line other than least squares (e.g., reduced major axis, major axis) should be employed to describe the relationship between two variables is a complicated problem (Sokal and Rohlf, 1981; Rayner, 1985; Pagel and Harvey, 1988, 1989; LaBarbera, 1989; Harvey and Pagel, 1991; Riska, 1991). Principal components analysis can be performed on correlation coefficients (computed through the origin) of standardized independent contrasts. Reduced major axis slopes can be computed in the usual way as the ratio of standard deviations or as the least-squares regression (through the origin) slope divided by the correlation coefficient. All lines must go through the origin, and none of them are restricted to go through the point (mean $x$, mean $y$). Formulas for computing reduced major axis and major axis slopes are provided in Appendix 3. All of the foregoing slopes are special cases of the general structural model. Whatever method is used to estimate a slope, its interpretation as a conventional allometric scaling exponent will depend on (1) log transformation of both characters prior to analyses and (2) use of the same branch lengths for both; these two criteria are met by each of the first three slopes listed in Table 1.

## What Kinds of Traits Can Be Analyzed?

The independent contrasts approach is designed to investigate the correlated evolution of traits that are inherited from ancestors, whatever the cause of that heritability. Thus, the phenotypic data for tip species are generally assumed to reflect underlying genetic differences among species, as could be verified through common-garden experiments (Garland and Adolph, 1991). The tip species for one set of contrasts can even be different from those for the other set, as long as the phylogenetic trees are isomorphic. This realization makes it possible to use independent contrasts to examine coevolution of phenotypic traits such as body size and life span in coevolved host/parasite systems (see Harvey and Keymer, 1991).

In addition to the usual phenotypic traits (e.g., body size, metabolic rate), cultural (see Cavalli-Sforza and Feldman, 1981), environmental (e.g., soil or water pH, mean annual temperature), and other traits that

are difficult to categorize (e.g., home range area) can be studied as long as they are passed on from ancestral to descendent species (or populations) and have a continuous distribution. For example, many environmental properties, such as latitude or mean annual rainfall, are not inherited in the conventional (genetic) sense. Nevertheless, they are inherited in the sense that organisms are born into environmental conditions and locations experienced by their parents at the time of birth. Thus, the ancestor of two species living in a desert may also have lived in a desert (cf. Huey, 1987), or the ancestor of one high-latitude and one equatorial species may have lived at midlatitude. Similarly, if an environmental characteristic is determined solely (without externally imposed constraints) through a process of habitat selection, and if species differences in habitat selection are genetically based, then species differences in the environmental trait will be genetically based as well. Alternatively, if variation in some (genetically based) phenotypic trait can be used as a precise indicator of some environmental characteristic, then that phenotypic trait may be used as a surrogate for the environmental characteristic. For example, toe fringes in lizards might indicate occupancy of sandy habitats. Unfortunately, this is not unfailingly the case; some species that glide through the air or that run across water also possess toe fringes (Luke, 1986). Finally, paleoclimatological and historical biogeographical data might be used in conjunction to indicate environmental characteristics of hypothetical ancestral (as opposed to tip) species, but this takes us into the realm of other comparative methods, such as those based on minimum evolution reconstructions of ancestors (Huey, 1987; Harvey and Pagel, 1991; Maddison, 1991; Martins and Garland, 1991). In any case, techniques for correlating phenotypes with environmental characteristics require further study.

## ACKNOWLEDGMENTS

We thank C. M. Janis for very kindly providing access to unpublished information and allowing us to use it for the example, A. Grafen for extensive discussions on branch lengths, A. J. Purvis for motivation, J. A. Jones for assistance with computer programming, and A. W. Dickerman, J. Felsenstein, D. M. Hillis, R. E. Jung, A. Sparti, and an anonymous reviewer for comments on the manuscript. This work was supported by U.S. National Science Foundation grants BSR-9006083, BSR-9146864 (Research Experiences for Undergraduates), and BSR-9157268 (Presidential Young Investigator Award) to T.G. and by grant number SCI*0814.C from the Commission for the European Communities to P.H.H.

## REFERENCES

BELL, G. 1989. A comparative method. Am. Nat. 133: 553–571.

BOOKSTEIN, F. L. 1988. Random walk and the biometrics of morphological characters. Evol. Biol. 23: 369–398.

BURT, A. 1989. Comparative methods using phylogenetically independent contrasts. Oxf. Surv. Evol. Biol. 6:33–53.

CAVALLI-SFORZA, L. L., AND M. W. FELDMAN. 1981. Cultural transmission and evolution. Princeton Univ. Press, Princeton, New Jersey.

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.

CLUTTON-BROCK, T. H., AND P. H. HARVEY. 1979. Comparison and adaptation. Proc. R. Soc. Lond. B 205:547–565.

DONOGHUE, M. J. 1989. Phylogenies and the analysis of evolutionary sequences, with examples from seed plants. Evolution 43:1137–1156.

FELSENSTEIN, J. 1982. How can we infer geography and history from gene frequencies? J. Theor. Biol. 96:9–20.

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

FELSENSTEIN, J. 1988. Phylogenies and quantitative characters. Annu. Rev. Ecol. Syst. 19:445–471.

FRIDAY, A. 1987. Models of evolutionary change and the estimation of evolutionary trees. Oxf. Surv. Evol. Biol. 4:61–88.

GARLAND, T., JR. 1992. Rate tests for phenotypic evolution using phylogenetically independent contrasts. Am. Nat. (in press).

GARLAND, T., JR., AND S. C. ADOLPH. 1991. Physiological differentiation of vertebrate populations. Annu. Rev. Ecol. Syst. 22:193–228.

GARLAND, T., JR., AND R. B. HUEY. 1987. Testing symmorphosis: Does structure match functional requirements? Evolution 41:1404–1409.

GARLAND, T., JR., R. B. HUEY, AND A. F. BENNETT. 1991. Phylogeny and thermal physiology in lizards: A reanalysis. Evolution 45:1969–1975.

GARLAND, T., JR., AND C. M. JANIS. 1992. Does metatarsal/femur ratio predict maximal running speed in cursorial mammals? J. Zool. (Lond.) (in press).

GITTLEMAN, J. L., AND M. KOT. 1990. Adaptation:

Systematics and a null model for estimating phylogenetic effects. Syst. Zool. 39:227-241.

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. (in press).

HARVEY, P. H., AND A. E. KEYMER. 1991. Comparing life histories using phylogenies. Philos. Trans. R. Soc. Lond. B 332:31-39.

HARVEY, P. H., AND M. D. PAGEL. 1991. The comparative method in evolutionary biology. Oxford Univ. Press, Oxford, England.

HARVEY, P. H., AND A. PURVIS. 1991. Comparative methods for explaining adaptations. Nature 351: 619-624.

HUEY, R. B. 1987. Phylogeny, history, and the comparative method. Pages 76-98 *in* New directions in ecological physiology (M. E. Feder, A. F. Bennett, W. W. Burggren, and R. B. Huey, eds.). Cambridge Univ. Press, New York.

JANIS, C. M. In press. Do legs support the arms race hypothesis in mammalian predator/prey relationships? *In* Vertebrate behaviour as derived from the fossil record (J. R. Horner and L. Ellis, eds.). Columbia Univ. Press, New York.

LABARBERA, M. 1989. Analyzing body size as a factor in ecology and evolution. Annu. Rev. Ecol. Syst. 20:97-117.

LOSOS, J. B. 1990. The evolution of form and function: Morphology and locomotor performance in West Indian *Anolis* lizards. Evolution 44:1189-1203.

LUKE, C. 1986. Convergent evolution of lizard toe fringes. Biol. J. Linn. Soc. 27:1-16.

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

MADDISON, W. P. 1990. A method for testing the correlated evolution of two binary characters: Are gains or losses concentrated on certain branches of a phylogenetic tree? Evolution 44:539-557.

MADDISON, W. P. 1991. Squared-change parsimony reconstructions of ancestral states for continuous-valued characters on a phylogenetic tree. Syst. Zool. 40:304-314.

MARTINS, E. P., AND T. GARLAND, JR. 1991. Phylogenetic analyses of the correlated evolution of continuous characters: A simulation study. Evolution 45:534-557.

NETER, J., W. WASSERMAN, AND M. K. KUTNER. 1989. Applied linear regression models, 2nd edition. Richard D. Irwin, Homewood, Illinois.

NORUSIS, M. J. 1988. SPSS/PC+ V2.0 base manual. SPSS, Chicago.

PAGEL, M. D. 1992. A method for the analysis of comparative data. J. Theor. Biol. (in press).

PAGEL, M. D., AND P. H. HARVEY. 1988. The taxon-level problem in the evolution of mammalian brain size: Facts and artifacts. Am. Nat. 132:344-359.

PAGEL, M. D., AND P. H. HARVEY. 1989. Taxonomic differences in the scaling of brain on body weight among mammals. Science 244:1589-1593.

PAGEL, M. D., AND P. H. HARVEY. 1992. On solving the correct problem: Wishing does not make it so. J. Theor. Biol. (in press).

RAYNER, J. M. V. 1985. Linear relations in biomechanics: The statistics of scaling functions. J. Zool. Ser. A 206:415-439.

RIDLEY, M. 1983. The explanation of organic diversity: The comparative method and adaptations for mating. Clarendon, Oxford, England.

RISKA, B. 1991. Regression models in evolutionary allometry. Am. Nat. 138:283-299.

ROHLF, F. J., W. S. CHANG, R. R. SOKAL, AND J. KIM. 1990. Accuracy of estimated phylogenies: Effects of tree topology and evolutionary model. Evolution 44:1671-1684.

ROHLF, F. J., AND R. R. SOKAL. 1981. Statistical tables, 2nd edition. W. H. Freeman, San Francisco.

SIBLEY, C. G., AND J. E. AHLQUIST. 1991. Phylogeny and classification of birds. Yale Univ. Press, New Haven, Connecticut.

SOKAL, R. R., AND F. J. ROHLF. 1981. Biometry, 2nd edition. W. H. Freeman, San Francisco.

ZAR, J. H. 1984. Biostatistical analysis, 2nd edition. Prentice-Hall, Englewood Cliffs, New Jersey.

*Received 3 May 1991; accepted 17 December 1991*

## APPENDIX 1

In this appendix, we demonstrate why regression through the origin must be used for standardized independent contrasts. For the sake of this explanation, consider two nodes, $a$ and $b$, and two variables, $x$ and $y$, that have values $x_a$ and $y_a$ at node $a$ and $x_b$ and $y_b$ at node $b$. The value of the independent contrast in $x$ can be either $\Delta x = x_a - x_b$, or $-\Delta x = x_b - x_a$, where $\Delta x$ may either be positive or negative. In other words, the choice of the direction of subtraction of $x$ is arbitrary and consequently so is the sign of $\Delta x$. If $\Delta y = y_a - y_b$, then the pair of independent contrasts can be written as either $(\Delta x, \Delta y)$ or $(-\Delta x, -\Delta y)$; the direction of subtraction for the variable $y$ must be the same as that for the variable $x$ once the direction of subtraction of variable $x$ is chosen. For the special case when $\Delta x = 0$, the choice of the sign of $\Delta y$ becomes arbitrary because it does not matter whether $x_a$ is subtracted from $x_b$ or vice versa. Thus, because the sign of $\Delta y$ is arbitrary, the expected value of $\Delta y$ must be zero when $\Delta x = 0$. Note, however, that degrees of freedom for hypothesis testing are still $n - 1$, where $n$ = the number of independent contrasts, because points with $\Delta x = 0$ still provide information on the variance about the regression line.

Because $\Delta x$ has an arbitrary sign, further restrictions are placed on the type of regression analysis that can be used to compare the independent contrasts. Let $\Delta x$ and $\Delta y$ denote the random variables for the independent contrasts, and let $\Delta\hat{y}(\Delta x)$ denote the estimator of $\Delta y$ given $\Delta x$. This estimator may be a linear function of $\Delta x$ (linear regression) or any other function of $\Delta x$ (nonlinear regression). When $\Delta y$ is plotted as a function of $\Delta x$, the arbitrary choice of the sign of $\Delta x$ will not affect the fitted line provided:

1. $\Delta\hat{y}(\Delta x) = -\Delta\hat{y}(-\Delta x)$

2. The loss function, $L(\Delta x, \Delta y)$, used in regression (e.g., $L(\Delta x, \Delta y) = [\Delta y - \Delta\hat{y}(\Delta x)]^2$ for least squares) has the property $L(\Delta x, \Delta y) = L(-\Delta x, -\Delta y)$. This property clearly holds for least squares regression, and it will also hold for any loss function that weights points above and below the regression line equally.

As long as these two properties hold, a unique line is determined for $\Delta\hat{y}(\Delta x)$, regardless of the sign of the values of $\Delta x$.

## Appendix 2

Here we present formulas for regression through the origin using standardized independent contrasts (Neter et al., 1989:167–170; Martins and Garland, 1991).

Correlation coefficient:

$$r = \frac{\sum (\Delta x_i \cdot \Delta y_i)}{\{[\sum (\Delta x_i)^2][\sum (\Delta y_i)^2]\}^{0.5}}.$$

Regression equation:

$$\Delta y_i = \beta\, \Delta x_i + e_i,$$

where $e_i$ is independent and normally distributed with mean 0 and variance $\sigma^2$.

Least squares estimator of $\beta$:

$$\beta = \sum (\Delta x_i \cdot \Delta y_i) / \sum (\Delta x_i)^2$$

(equivalent to maximum likelihood estimator).

Residuals:

$$e_i = \Delta y_i - \beta\, \Delta x_i.$$

Unbiased estimator of $\sigma^2$:

$$\text{Mean squared error [MSE]} = \sum e_i^2/(n-1),$$

where $n$ = number of independent contrasts.

Variance of the estimator $\beta$:

$$s^2(\beta) = \text{MSE}/\sum (\Delta x_i)^2.$$

Confidence interval for $\beta$:

$$\beta \pm t \cdot s(b),$$

where $t$ is taken from Student's $t$ distribution with $n - 1$ df and $n$ = number of independent contrasts.

## Appendix 3

Here we present formulas for reduced major axis and major axis slopes computed through the origin using standardized independent contrasts (Sokal and Rohlf, 1981:550–552, 595).

Reduced major axis:

$$\text{slope} = \pm[\sum (\Delta y_i)^2 / \sum (\Delta x_i)^2]^{0.5} = \pm$$

where $\beta$ is the regression coefficient and $r$ is the product-moment correlation coefficient, both computed as given in Appendix 2.

Major axis (1st principal component):

$$\text{slope} = \text{Cov}/(\lambda x - S_x^2),$$

where Cov is the sample covariance

$$\sum (\Delta x_i \cdot \Delta y_i)/(n-1);$$

$S_x^2$ is the sample variance

$$\sum (\Delta x_i)^2/(n-1);$$

and

$$\lambda_x = \tfrac{1}{2}\{S_x^2 + S_y^2 + [(S_x^2 + S_y^2)^2 - 4(S_x^2 \cdot S_y^2 - \text{Cov}^2)]^{0.5}\}.$$