

Correlated evolution and independent contrasts

TREVOR PRICE

Department of Biology 0116, University of California at San Diego, La Jolla, CA 92093, USA

CONTENTS

	PAGE
1. Introduction	519
2. Data	520
3. Independent contrasts	521
4. An alternative evolutionary model	524
5. Discussion	526
References	528

SUMMARY

The use of the independent contrast method in comparative tests is studied. It is assumed that: (i) the traits under investigation are subject to natural selection; (ii) closely related species are similar because they share many characteristics of their niche, inherited from a common ancestor; and (iii) the current adaptive significance of the traits is the focus of investigation. The main objection to the use of species values in this case is that third variables which are shared by closely related species confound the relationship between the focal traits. In this paper, I argue that third variables are largely not controlled by the contrast methods, which are designed to estimate correlated evolution. To the extent that third variables also show correlated evolution, the true relationship among the traits of interest will remain obscured. Although the independent contrast method does not resolve the influence of third traits it does, in principle, provide a greater resolution than the use of the species mean values. However, its validity depends on the applicability of an evolutionary model which has a substantial stochastic component. To illustrate the consequences of relaxing this assumption I consider an alternative model of an adaptive radiation, where species come to fill a fixed niche space. Under this model, the expected value for the contrast correlation differs from that for the species correlation. The two correlations differ because contrasts reflect the historical pattern of diversification among species, whereas the species values describe the present-day relationships among the species. If the latter is of interest, I suggest that assessing significance based on the species correlations can be justified, providing that attention is paid to the role of potentially confounding third traits. Often, differences between contrast and species correlations may be biologically informative, reflecting changes in correlations between traits as an adaptive radiation proceeds; contrasts may be particularly useful as a means of investigating past history, rather than current utility of traits.

1. INTRODUCTION

The use of phylogeny in comparative studies has become routine. The most widely employed method is that of independent contrasts, introduced by Felsenstein (1985). This method is designed to detect correlated evolution (the extent to which change in one trait is associated with change in another; Felsenstein 1985; Pagel 1993). Estimates of phylogenetic relationships are used to transform n species mean values into $n-1$ contrast values. Under a specified model of evolution (usually Brownian motion) these contrasts are independent, and can be used in regression and correlation studies (see Garland *et al.* 1992). With a fully resolved phylogeny no degrees of freedom are lost by this transformation, and apart from error introduced by uncertainties in phylogenetic relationships and the assumed evolutionary model, there is no loss in power.

This method was introduced as a means of testing the significance of a correlation or regression coefficient (Felsenstein 1985), without any assumptions about adaptation. However, it has almost invariably been used as a means of evaluating the adaptive significance of traits, and this will be the focus of this paper (but for a rare exception see Clark & Wang (1994)). The adaptive significance of a trait is assessed by questioning the strength of its relationship with other traits or environmental variables, when all traits are measured as species mean values. By adaptive significance I mean contemporary value rather than evolutionary history. According to the definition of Gould & Vrba (1982) this is therefore the study of aptation, rather than strict adaptation, which is concerned with the causes of origin of traits. I consider the traits studied here to be held by natural selection near to some optimum. Typically, closely related species will tend to be similar in those traits (as well as many

others) because they share a similar niche, inherited from a recent common ancestor. They are, therefore, subject to similar stabilizing selection pressures (Grafen 1989, p. 144; Harvey & Pagel 1991, p. 38; Harvey & Purvis 1991).

In this paper I evaluate the use of independent contrasts in this narrower framework, addressing three issues. First, I ask whether the independent contrasts method makes a substantial difference as compared to the use of species mean values. It does not appear to do so, apparently because the effect of common ancestry has to be severe, a result anticipated from simulation studies (Martins & Garland 1991; Losos 1994). Contrast methods measure correlated evolution, and in the second section I argue that the evolutionary correlation may be as uninformative (or as informative) as the correlation between the species' values. Finally, I consider some evolutionary models where this is definitely the case. These models lead to a changing relationship between traits at different taxonomic levels. In this case contrast correlations, which average over these changes in a different way from species correlations, will give different results. I conclude that one valuable approach is to compare the species and contrast correlations. Differences between them may be of biological significance.

2. DATA

Throughout this paper I use the term 'species correlation' to describe the correlation between species' mean values, and the term 'contrast correlation' to describe the correlation between the contrasts obtained by transforming the species values using a phylogeny. The evolutionary correlation refers to the correlation of evolutionary changes in two traits (Martins & Garland 1991). They distinguish between the input evolutionary correlation which describes the parametric correlation in a bivariate normal distribution from which evolutionary changes are stochastically drawn, and the realized evolutionary correlation which is the evolutionary correlation obtained from the sample drawn from this distribution. Here, I use the evolutionary correlation to refer to the input correlation, unless stated.

I extracted correlations for species and contrasts from 19 studies (see figure 1). The correlations are not independent, for example many of the bird studies use the Sibley & Ahlquist (1990) phylogeny, and most studies present multiple correlations, all of which are included in the analysis. Nevertheless, there is a high concordance between the species and contrast correlations (see figure 1). Few correlations are significantly different, if we use $p < 0.05$ as a criterion (see figure 1). The correlation between the species correlations and the contrast correlations is 0.86, indicating that the species correlation is usually a good guide to the contrast correlation. A similar analysis on a largely non-overlapping data set has been independently presented by Ricklefs & Starck (1996), with similar conclusions.

When the phylogeny is a 'starburst' there will be no difference between the two correlations. The reason for

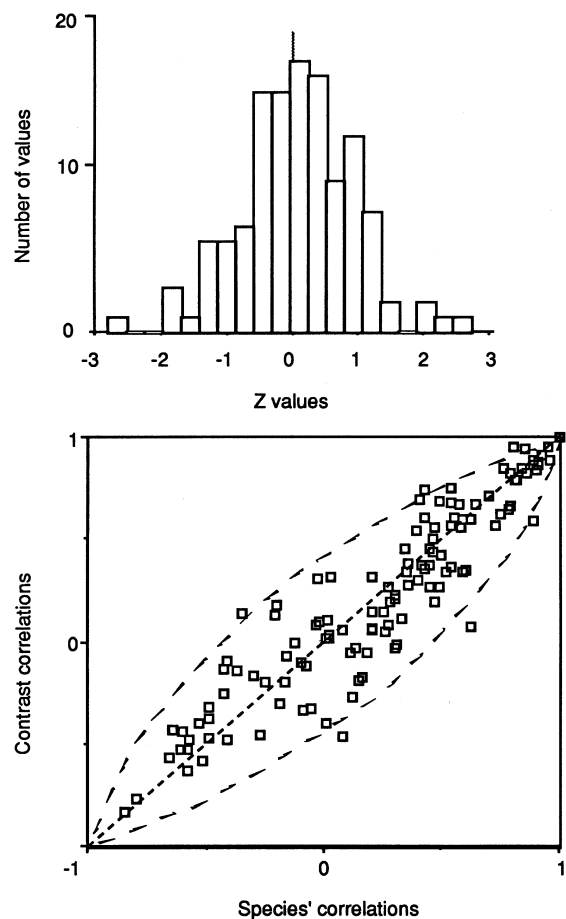


Figure 1. Lower panel. Scatter plot of species and contrast correlations, as taken from the following studies: Brandl *et al.* (1994), Briskie & Montgomerie (1992), Briskie *et al.* (1994), De Voogd *et al.* (1994), Garcia-Barros (1994), Garland *et al.* (1991), Garland & Janis (1993), Martin (1993), Martins (1993), Møller & Birkhead (1993), Poiani (1993), Promislow (1992), Promislow *et al.* (1993), Richman & Price (1992), Saether & Gordon (1994), Sessions & Larson (1987), Suhonen *et al.* (1994), Telleria & Carrascal (1994), and Walton (1993). Number of correlations from any one study varies from one to 47. Contrasts were calculated using Felsenstein's (1985) method or Pagel's (1992) modification for use when phylogenies are not resolved into a fully bifurcating tree. Dashed line is line of equality, and ellipse describes 95% confidence limits for pairs of correlation coefficients drawn from this line (constructed by simulation using the reduced major axis criterion, and assuming a sample size of 30). Upper panel. Distribution of z-values obtained from Fisher's transformation for the paired comparison of the contrast and species correlations (Zar 1996, p. 381), assuming a (usually conservative) sample size of 30 for each comparison. Z-values > 1.96 or < -1.96 would imply a significant difference between the correlation coefficients at $p < 0.05$ (ignoring problems of multiple tests and non-independence).

the similarity between the contrast and species correlations seems to be that the influence of phylogeny has to be large to have much of an effect (Martins & Garland 1991; Losos 1994) by which I mean that ecologically similar groups of species have to be closely related to each other, and distantly related from other groups. The empirical finding that the contrast and species correlations are similar implies that in nature,

phylogenies are often not constructed in this way. For illustrative purposes consider 20 analyses of trait associations, as might result from different studies of the type illustrated in figure 1. In this example, I assume that the underlying evolutionary correlation (defining the bivariate distribution from which evolutionary changes in two traits are drawn) differs for each analysis, i.e. 0 for one study, 0.05 for another, 0.1 for another, and so on up to 0.95. For each evolutionary correlation I calculated the species and contrast correlations by simulating evolution on a 16-species tree, and thus obtained 20 pairs of species and contrast correlations (simulation methods were as described by Martins & Garland 1991). For a perfectly symmetrical tree with equal branch lengths, the correlation between the species and contrast correlations is 0.74 (standard deviation of 20 repeat runs ± 0.18). Thus even when the phylogeny is quite hierarchically structured there is a good association between the species and contrast correlations. Lengthening the tips by a factor of ten lessens the effect of phylogeny, and the correlation in this case is 0.85 ± 0.08 . Lengthening the basal pair of branches by a factor of ten increases the effect of phylogeny and the correlation in this case is 0.51 ± 0.38 .

Both the data and the simulations imply that the contrast method gives correlations and regressions which are quite similar to those obtained when the species values are used instead. Correlation coefficients can be used as a guide to the p values for correlation and regression: the larger the correlation coefficient the lower the p value. Identical correlation coefficients with the same number of degrees of freedom imply that the regressions have the same p values. When a fully resolved tree is available there are the same number of degrees of freedom for both the species and contrast correlations, and so the significance of the association between two traits will be similar whichever method is used. This is surprising because closely related species are similar in many traits, and it is this similarity which usually provides the argument for including phylogeny in comparable studies (e.g. Harvey *et al.* 1995). I argue that correlations between independent contrasts do not control many of the similarities we worry about.

3. INDEPENDENT CONTRASTS

(a) *Neutral traits*

As a starting point, before considering adaptive traits, I discuss the evolution of traits subject solely to genetic drift. Under a null hypothesis the traits are genetically uncorrelated, and hence evolving in an uncorrelated manner. Chance coordinate changes will result in a correlation between the species mean values. Repeated runs of the same evolutionary process on the same phylogeny give different correlations, whose expected value is zero. The more hierarchically structured is the phylogeny, the greater the variance in the correlations obtained from repeated runs will be. This is because a coordinate change deep in the phylogeny (whether it be positive or negative) is preserved through later speciation events (Martins &

Garland 1991). Because we observe only one run of evolution, the correlation we measure may be extreme due to a large change early in the species' radiation. The contrast method corrects for this by giving equal weight to evolution above each node in the tree. The greater variability of the species correlation coefficients as compared to the contrast correlations from repeated evolutionary runs implies that the null hypothesis will more often be incorrectly rejected when the species correlations are used (Martins & Garland 1991; see also table 1).

Now consider that the null hypothesis is false because the two traits are genetically correlated and hence show correlated evolution. Evolution is still occurring as the sole result of genetic drift. In this case the principles are similar to the situation where the null hypothesis is true, but results are confounded by the asymmetric distribution of the correlation coefficient, which is bounded between +1.0 and -1.0. This asymmetry results in the contrast correlation being a slightly biased estimate of the evolutionary correlation (Martins & Garland 1991). Pagel (1993) studied the regression coefficient instead of the correlation coefficient and showed that, for the particular model of evolution he used, both the contrast and species regressions provide unbiased estimates of the evolutionary regression (the regression of the changes in one trait on the changes in the other). Note that this refers to the expected values of the regression over multiple evolutionary reruns (of which we actually observe just one), and the outcome of any one run of evolution could give a very different measure for the contrast and species values. Some results of simulations for different phylogenies are shown in table 1. The correlation between the species regressions and contrast regressions is usually low: thus even though they both estimate the evolutionary regression (and the contrast regression does so slightly better) the two estimates in any one study can differ. More detailed simulation studies are given by Martins & Garland (1991).

The reason for controlling for phylogeny reflects the stochastic nature of evolution by genetic drift, implying that the pattern would be different if we could play God and rerun evolution during a repeat, equivalent, species radiation (Pagel 1993). Rerunning evolution on the same tree is also the way simulation studies of the contrast method have been conducted (Martins & Garland 1991). It can be justified by the inherent assumption of the contrast method—that contrasts at each node in the phylogeny are identically independently distributed—and it is a matter of chance whether a large change occurs near the base of the phylogeny (inflating the ensuing correlation across the species) or near the tips (with little effect on the overall species correlation). Such a stochastic view (see figure 2) is crucial to the model of evolution that forms the basis of the independent contrast method. As Felsenstein (1985) noted, adaptive traits may evolve differently; for example, a large change lower in the tree might result in smaller changes higher up, or some clades in the tree might have higher rates of evolution than others. Under either of these two scenarios contrast values will not be independent of each other:

Table 1. Correlations and regressions between two variables (X and Y) produced by stochastic evolution on a symmetric 16-species tree (500 replicates). The tree either has all branch lengths equal (basal branches short), or the basal pair of branches ten times longer than the others. Simulation methods are as described by Martins and Garland (1991). The last two columns give the correlation between the contrast and species associations, as measured on each replicate

evolutionary correlation	basal branches	observed correlations		observed regressions		correlation between	
		contrasts	species	contrasts	species	contrasts	regression
		$\bar{x} \pm \text{SD}$	$\bar{x} \pm \text{SD}$	$\bar{x} \pm \text{SD}$	$\bar{x} \pm \text{SD}$		
0	short	0.01 ± 0.35	0.02 ± 0.36	0.00 ± 0.42	0.02 ± 0.40	0.264	0.225
0	long	0.00 ± 0.36	0.04 ± 0.53	0.00 ± 0.40	0.04 ± 0.75	0.098	0.106
0.5	short	0.48 ± 0.29	0.45 ± 0.30	0.52 ± 0.38	0.48 ± 0.37	0.278	0.273
0.5	long	0.47 ± 0.29	0.46 ± 0.43	0.50 ± 0.37	0.54 ± 0.61	0.071	0.035

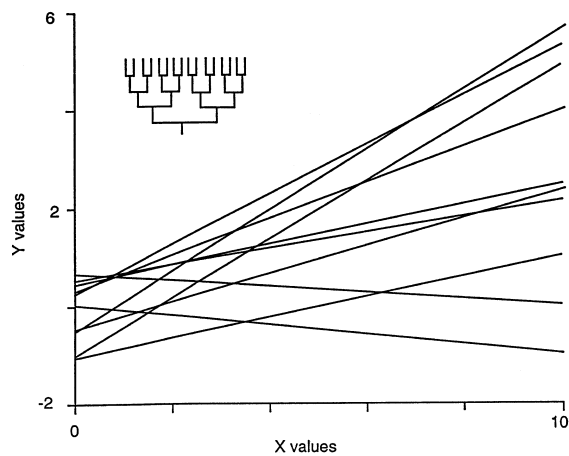


Figure 2. Ten regression lines obtained from the first ten runs of the second line of table 2. The 16-species symmetrical phylogeny with equal branch lengths used for the simulations is also illustrated.

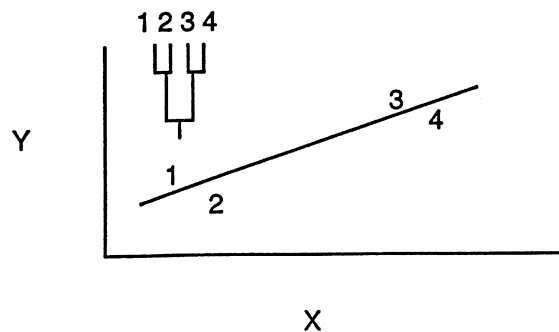


Figure 3. Regression of Y on X for four species, whose phylogenetic relations are also shown in the figure. This is a 'worse case' scenario for the problem of non-independence.

contrasts calculated from some parts of the tree will tend to be more similar to each other than contrasts from other parts of the tree. I now turn to the justification of contrast methods (and estimates of correlated evolution in general) when used with adaptive traits.

(b) Adaptive traits

Consider the allometric relationship between four species for two traits, Y and X , as illustrated in figure 3 (X may be a trait, for example, body size; or an

environmental variable, for example, climate). Two pairs of closely related species differ in their mean values. This is therefore a classic case of non-independence. However, assuming that the allometric relationship is linear, and that each species is being held at its mean value by natural selection, each species should provide an independent estimate of that relationship and could, therefore, be used as an independent datum. This is the argument that was often used until the mid-1980s—for example: 'common ancestry and phylogenetic inertia are unimportant when there exists additive genetic variance for the character in question' (Burt & Bell 1987, quoted in Ridley (1989)). The implication is that each species has been independently tested by selection. Otherwise it might have moved away from its close relative. In some cases this assumption has been tested. Lynch (1991) shows that stabilizing selection has almost certainly been involved in preventing diversification in several adaptive radiations, because genetic drift would lead to more variation among species than is actually observed. The assumption of stabilizing selection implies that the traits are useful in their current environment, and the regression of one trait on the other can be used as a guide to their adaptive significance.

Ridley (1989) and Grafen (1989) in their discussions of the comparative method accept that the species values could possibly be used in this case, but only under very restricted circumstances. If X is the sole variable affecting Y , each species can be considered an independent estimate of the allometric relationship between the two traits, and the adaptive significance of Y interpreted in terms of X . Under this assumption, if evolution is 'rerun' it would be entirely deterministic and always lead to exactly the same allometric relationship. The great variability in outcomes under the stochastic evolution model (see figure 2) would be replaced by a single outcome (see figure 3), which could be estimated by the species correlation and regression.

The difficulty with accepting this as justification for using the species regression is that other traits beside X will almost certainly affect the relationship between Y and X . According to Ridley (1989, p. 363): 'A group of species that share a particular character state by common descent probably also share many other character states and it could be that an association between two characters is truly due to an association with some unobserved third variable'. The com-

Table 2. Associations between Y and X when Y is determined by the relationship $Y = \beta_x X + \beta_z Z$, where β_x and β_z are partial regression coefficients (equation (1)). A symmetric 16-species tree is used, with the basal pair of branches $10 \times$ the length of the other branches. Simulations were replicated 500 times. For all runs the variance in X and the variance in Z is set to be 1.0, and $\beta_z = 1.0$. The input is the true effect of X on Y (β_x , the partial regression coefficient of Y on X , holding Z constant) and the correlation between X and Z (r_{xz}). The parametric correlation between Y and X (r_{yx}) and the parametric regression of Y on X (b_{yx}) is calculated from these values. The estimated partial regression coefficient (last column) is the regression of Y on X , holding Z fixed: results are the same whether the species or contrast values are used

β_x	r_{xz}	r_{yx}	b_{yx}	associations between X and Y				
				correlation		regression		partial regression
				contrast	species	contrast	species	
				$\bar{x} \pm \text{SD}$	$\bar{x} \pm \text{SD}$	$\bar{x} \pm \text{SD}$	$\bar{x} \pm \text{SD}$	$\bar{x} \pm \text{SD}$
0	0	0	0	0.01 ± 0.36	0.02 ± 0.51	0.01 ± 0.40	0.02 ± 0.72	0.0 ± 0.0
0	0.5	0.5	0.5	0.47 ± 0.28	0.44 ± 0.44	0.51 ± 0.36	0.51 ± 0.61	0.0 ± 0.0
0.5	0.5	0.76	1.0	0.72 ± 0.19	0.68 ± 0.34	0.98 ± 0.36	0.98 ± 0.66	0.5 ± 0.0

parative method uses phylogeny to control for unobserved variables, but I will argue that it rarely controls for them completely.

Following Grafen (1989) we can formalize Ridley's statement. Write

$$Y = \beta_x X + \beta_z Z, \quad (1)$$

where β_x and β_z are regression coefficients. Here the mean value for a species in trait Y is attributable to a measured trait, X , and an unmeasured trait, Z . In fact there may be several measured traits, and there will be many unmeasured variables which affect Y (Grafen 1989, pp. 143–144), but for illustrative purposes we will consider one measured trait and one unmeasured trait. The philosophy behind multiple traits is identical. Apart from the influence of measurement and sampling error on the estimate of the species mean values (which, following Grafen (1989) we assume to be small) it is probable that the trait Y can be written completely in terms of its determination by other traits and environmental variables. To illustrate, suppose Y is brain size, and X is body size. One unmeasured variable (Z) might be intelligence. We will assume that from a knowledge of intelligence and body size, brain size can be accurately predicted (and ignore other variables such as climate, diet, etc.).

Unmeasured variables affect the relationship between measured traits in two ways. First, they cause deviations around the regression line (Grafen 1989; Pagel & Harvey 1991, p. 184; Riska 1991). For example, intelligent species show positive deviations of brain size from a brain–body regression. Often these errors will be phylogenetically correlated. Thus primates are related and show positive deviations in a regression of mammals' brain sizes on body sizes.

Second, unmeasured variables confound the relationship between two traits, even in the absence of any correlation between the deviations from a regression line. Felsenstein (1985, p. 5) gives a hypothetical example where there is a strong influence of phylogeny, but the errors are not phylogenetically distributed. A similar example is shown in figure 3: the average deviation from the regression line within each pair of closely related species is zero, and so the deviations are not phylogenetically distributed. In this case, unmeasured variables are still confounding the relationship between Y and X by affecting the strength

of the correlation or regression. For example the two clades (of two species each) may differ in diet. Diet affects Y , and is associated with X , and therefore causes the observed association between Y and X . It is the effect of third variables on the strength of the association between Y and X which is of most concern.

The influence of unmeasured variables on the correlation and regression can be described formally from equation (1). The covariance between X and Y ($\text{Cov}(X, Y)$) is written as:

$$\text{Cov}(X, Y) = \beta_x \text{Var}(X) + \beta_z \text{Cov}(X, Z), \quad (2)$$

where β_x and β_z are partial regression coefficients describing the effect of each trait, X and Z , on Y when the other is held fixed. Equation (2) shows that an observed covariance between X and Y can be attributed to two components. The first term on the right-hand side describes the direct effect of X on Y , and is the relationship which is being studied. The second term describes an association between X and Y which arises indirectly, because an unknown trait Z affects Y . This term is zero if X and Z are uncorrelated.

From equation (2) it can be seen that a positive covariance between Z and X increases the observed covariance between X and Y (and hence correlation and regression), and a negative covariance between Z and X decreases the covariance between X and Y . Even if Z and X are evolving in an uncorrelated manner they will often show a too-high correlation across the species when the phylogeny is hierarchically structured, in exactly the same manner as was demonstrated for the case of two traits evolving solely as a result of genetic drift (see table 1; Martins & Garland 1991). This inflated sampling variance of the correlation coefficient between Z and X is reflected in an inflated sampling variance of the correlation between X and Y (equation (2)). Some simulation results are given in the first row of table 2, and see also Grafen (1989). Under this model of adaptive evolution there is a fair degree of stochasticity in the outcome, because traits Z and X are assumed to evolve independently of each other. This is appropriately controlled for by the contrast method. The significance of the contrast correlation provides a better guide to the adaptive significance of any observed relationship between X and Y , because the influence of the unknown trait Z has been accounted for.

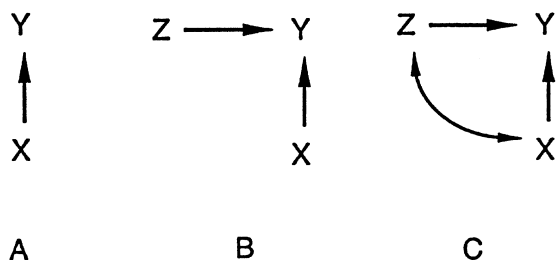


Figure 4. Three path models describing possible the causal relationships between Y and X , and unmeasured variables (Z).

The crux of the argument of this paper is that unmeasured variables rarely evolve independently of measured traits, and are thus unlikely to be completely controlled for by the contrast method. For example, intelligence may be more likely to arise in large rather than small animals, and thus intelligence and body size evolve in a correlated manner.

Results of some simulations are shown in table 2. When there is no evolutionary correlation between Z and X (i.e. Z and X are evolving truly independently), and the null hypothesis that there is no association between Y and X is true, the familiar effect of phylogeny in increasing the variance of correlation and regression coefficients is apparent. This is true also when Z and X are correlated, but the true association between Y and X is masked by the correlation between Z and X ; this bias often has more influence than the occasional large coefficient due to random associations between Z and X . There is no inherent difference between the confounding of the relationship between Y and X due to associations with unmeasured variables, whether they be a 'random association' or a 'correlated association', and therefore contrasts may control for a very small part of confounding third variables. One way to get at these effects is to use a partial regression or correlation, holding identified third variables fixed, as this always gives the correct result (see table 2).

Causal diagrams (see figure 4) can be used to illustrate the three different models discussed here. In figure 4a the only trait affecting Y is X , and the species regression is fully justified. In figure 4b all unmeasured traits affecting Y are not inherently correlated with X . However, inflated correlations due to sampling arise through the hierarchical structure of the phylogeny, affecting the observed relationship between Y and X . Here the contrast method is more appropriate. In figure 4c unmeasured traits are correlated with X . Only figure 4c is likely to be realistic. Before considering this further, I will discuss an alternative evolutionary model, which results in the contrast and species values estimating different things.

4. AN ALTERNATIVE EVOLUTIONARY MODEL

It has been stressed repeatedly that any comparative study based on phylogeny makes assumptions about the evolutionary model (e.g. Felsenstein 1985; Harvey & Pagel 1991; Harvey & Purvis 1991; Pagel 1993). The model that has been adopted generally for applications of independent contrasts is that of sto-

chastic evolution, usually Brownian motion (Diaz-Uriarte & Garland 1996). Under the Brownian motion model, the independent contrast method controls at least partly for confounding third traits, as shown here (see table 2) and elsewhere (e.g. Grafen 1989; Martins & Garland 1991). It might be argued that the partial resolution of unmeasured traits provided by the contrast method is better than nothing, thereby justifying its application. In this section I consider an alternative evolutionary model, which results in the expected value of the contrast correlation being lower than that of the species correlation. The contrast correlation in this case could be construed as providing a poor guide to the significance and strength of allometric relationship across species, but much is to be gained from a comparison of the species and contrast correlations.

I term this an 'adaptive radiation' model. It is based on Grafen's (1989, p. 143) discussion of why closely related species are similar. I define an ecological niche space as the parameters of a bivariate normal distribution, and then model an adaptive radiation from a single ancestor randomly positioned within this space. New species are added one after the other at regular intervals. Each species is dropped randomly on to the niche space, subject to the proviso that it cannot be too close to any other species. New species always evolve from ancestral forms to which they are ecologically closest. Once formed, a species shows no further evolution, and there is no extinction. A single run showing the build-up of a ten-species phylogeny is depicted in figure 5. Contrast correlations from the simulated phylogenies were estimated assuming both a punctuational model of evolution (i.e. what was actually simulated), as well as one in which branch-lengths were set proportional to time. In all test cases the punctuational and gradual models gave very similar results. Some results are shown in table 3.

This model ties adaptation to speciation. For the phylogenies which are generated, and if the parametric correlation between Y and X which forms the niche space is equal to zero, the standard deviations of the species and contrast correlations are similar (see the first line of table 3). This means that one is about as equally likely to reject the null hypothesis if the species or contrast correlations are used. If alternate hypotheses about the parametric correlation between Y and X are true, the species regressions and correlations provide a good estimate of the parameters of the bivariate normal distribution describing niche space; but the contrast regressions and correlations are often low. The reason for this is that most of the correlation between species is due to early speciation events: the first few species are dispersed through the niche space and other species radiate in no particular direction from one or other of those species. Thus a few evolutionary changes early in the history of the group are accounting for the correlation across species. In cases where the contrast correlation is lower than species correlation, this reflects a change in the evolutionary correlation at different taxonomic levels (high early, essentially zero later): a real feature of the biology of the radiation.

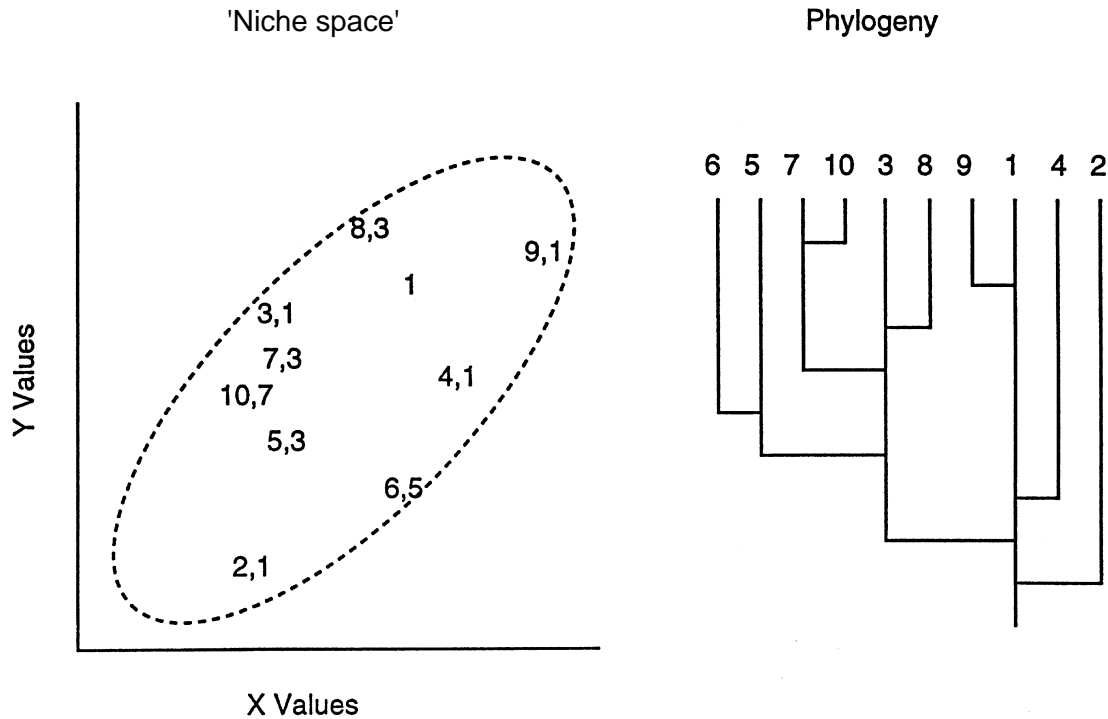


Figure 5. Illustration of an alternative model of evolution, as used to generate the results in table 3. In this example a bivariate normal distribution with correlation $r_{XY} = 0.7$, and variances of X and Y equal to 1.0 describes the niche space (95% ellipse is shown). Species are randomly dropped on to the niche space in order from 1–10; in this example the minimum distance permitted between species was 0.01. Once dropped on, the species is connected to its ecologically nearest neighbour, assumed to be its ancestor, to produce the phylogeny on the right. The model is supposed to reflect new niches uniformly appearing in time, and then being occupied by a descendant of the extant species ecologically most similar to the newly available niche. Position of a species in niche space is indicated by the first number in each pair, and the second number indicates the ancestor from which it evolved. Contrast and species correlations are obtained from the positions in niche space and the phylogeny.

Table 3. Some results of the adaptive radiation model (figure 5). The input is the variance of X and of Y (both set to be unity) and the bivariate correlation describing niche space (r_{XY}), which is equal to the regression of Y on X in this case (500 replicates). For all runs 16 species were randomly and sequentially dropped on to the space, with a requirement that there be a minimum distance between them of 0.25 units. Contrasts were calculated assuming a punctuational model of evolution

	correlations		regressions	
	contrast	species	contrast	species
r_{XY}	$\bar{x} \pm \text{SD}$	$\bar{x} \pm \text{SD}$	$\bar{x} \pm \text{SD}$	$\bar{x} \pm \text{SD}$
0	0.01 ± 0.24	0.00 ± 0.24	-0.01 ± 0.25	-0.01 ± 0.25
0.5	0.28 ± 0.23	0.49 ± 0.18	0.29 ± 0.26	0.50 ± 0.22
0.9	0.72 ± 0.13	0.90 ± 0.05	0.73 ± 0.16	0.91 ± 0.10

Some examples do exist where the contrast correlation is appreciably lower than the species correlations (Ricklefs & Starck 1996). Marchetti *et al.* (1995) studied the association between wing-length and migration distance in leaf warblers. The correlation across the species is $r = 0.77$, $p < 0.01$ ($N = 14$), but the contrast correlation is $r = 0.41$ ($p > 0.1$). The study combined species in Japan with long wings and long migration distances with species in the Himalayas with generally shorter wings and short migration distances. Species with long wings are closely related to each other (Richman & Price 1992), so the correlation across the species can be largely attributed to ancestral differences between Japan and the Himalayas, with species differences within these

locations being less strongly associated (the weaker association within regions is at least partly attributable to uncertainties in the estimates of migration distance).

When results from a contrast and species analysis differ, the interpretation depends on the assumed model. Under a Brownian motion model, provided migration distance evolves independently of unmeasured traits which might affect wing-length (for example, climate), such traits are controlled for by the contrast method. What is seen happening within the Himalayan and Japanese clades is used as a guide to assuming similar processes between other clades. The reason that the contrast correlation is lower than the species correlation is because a large coordinate change in migration distance and wing-length occurred

through a chance association with an unmeasured trait early in the species' radiation.

Under the adaptive radiation model, the low observed evolutionary correlation during the later stages of the adaptive radiation is contingent on there being a high correlation earlier. The large coordinate change in migration distance and wing-length early in the species' radiation is an important component of the species correlation, even if it occurred only once, and even if it is due to an associated third trait. The different evolutionary correlations reflect predictable differences in the influence of unmeasured traits at different taxonomic levels, and there is no statistical justification for combining the levels.

However, whatever the model, the species' values represent the outcome of a unique adaptive radiation. In the Discussion I will consider the meaning of a p -value attached to an analysis based on the species values.

5. DISCUSSION

The main assumption of this paper is that the traits of interest are of adaptive significance in the current environment, and that this significance forms the focus of study. There are at least two difficulties attached to such a study (e.g. Leroi *et al.* 1994). The first is that a measured association between traits may be attributable to other evolutionary processes besides direct selection on those traits. This criticism is not addressed here, but it has been considered in depth elsewhere (e.g. Gould & Lewontin 1979; Leroi *et al.* 1994).

The second difficulty is that an association between two traits is potentially confounded by other variables (e.g. Lauder 1981; Harvey & Pagel 1991; Leroi *et al.* 1994). For this reason it has been argued that we need to look at independently evolving groups: 'An association that has been established for many independently evolved groups is less likely to be due to a random association with a confounded third variable' (Ridley 1989, p. 363). This argument may have force for discrete traits, because the number of positions in the phylogeny where one or the other trait has evolved (i.e. changed state) will be typically far fewer than the number of species, and may be dispersed throughout a phylogeny. The argument is less compelling for continuously varying traits.

The mean values for continuously varying traits will usually differ between any pair of species (e.g. Garland & Adolph 1994), and, for a fully resolved phylogeny, the number of evolutionary changes in each trait is taken to be only one less than the number of species. I suggest that, generally, confounding third variables will show some tendency to evolve in an associated direction (either positive or negative) with the traits of interest, and therefore an association that has been established for the many evolving groups does not circumvent the problem of confounding variables. The clearest example of this is the extreme case, when the correlation between several traits is very high, so that they all evolve together (the species, contrast and evolutionary correlations are all close to 1.0; see Losos 1994). In this case, if closely related species are similar

in one trait, they will be similar in all of them, but these similarities are not addressed by associations between contrasts. However, the association between confounding traits and measured traits does not need to be strong, or even present in all parts of the tree, before it is not controlled by methods which appeal to measures of correlated evolution.

Evolutionary correlations between traits can arise in a very general way, such as a common response to some environmental gradient, with no implied causation between the traits themselves (Björklund 1994). For example, species typically differ in body size and hence in the size of many other morphological traits, as well as in physiological and behavioural variables. Some of these traits are likely to influence the traits being studied; chains of causation can be very long and tenuous. Suppose we wish to investigate an association between prey type and habitat among leaf warblers. It is known that habitat is associated with colour pattern in leaf warblers (Marchetti 1993), and that the feeding habits of the different species are also related to their coloration (Marchetti & Price 1997). Therefore coloration may be related to both prey type and habitat through various causal pathways. Coloration is a confounding variable which needs to be examined, and its effects will not be controlled for fully by comparative method techniques.

It might be argued that, as the contrast method provides a partial resolution of the problem of confounding variables, it provides a better guide to the significance of a relationship than the species association could. However, the potentially small added resolution provided by contrasts comes at the expense of additional assumptions. These assumptions are: (i) that a tree of phylogenetic relationships is accurately known (Felsenstein 1985; Björklund 1996); and (ii) that evolution occurs according to a specified model of evolution (Felsenstein 1985; Pagel 1993). In our research we have found that striking differences between contrast and species correlations are sometimes attributable to uncertainties in the tree, which place two very different species as sisters—a result also commented on by Losos (1994) in his investigations of random trees.

The question of an evolutionary model is deeper. At one extreme, one could make the assumption that evolution is entirely deterministic, and that an adaptive radiation from the same ancestor in the same environment would always produce exactly the same constellation of species. If one is prepared to make this assumption, then the observed species need not be considered a sample, but a population, and the standard error of the regression line simply needs to account for measurement and sampling error in estimating each species mean (Riska 1991). In this case the influence of confounding traits can only be assessed by multiple regression techniques. Few would make the assumption of entirely deterministic evolution, but it is not clear that the stochastic approach embodied in Brownian motion models (see figure 2) is appropriate for adaptive traits either. The difficulty is that one can imagine an infinite set of alternative evolutionary models. Some assumptions underlying particular evol-

utionary models, such as a normal distribution of contrast values, can be evaluated with the data (Garland 1992; Garland *et al.* 1992; Losos & Miles 1994, pp. 87–88; Diaz-Uriarte & Garland 1996). However, the inherent assumption of contrast methods, that evolution has a considerable stochastic component, cannot be tested.

Large coordinate changes in X and Y which occur early in an adaptive radiation give rise to the situation where incorporation of phylogeny makes most difference to comparative methods. Such large changes may well reflect associations of X and Y with a third trait, whose association is less strong later in evolution. One example is the migration distance: wing size correlation discussed previously; another might be brain–body regression, where diet differences lead to small and large taxa (Pagel & Harvey 1989). For the purposes of discussion I will consider a third, hypothetical, example where two distantly related clades differ in a discrete trait. For example, one clade might contain green species and the other clade red species. Colour affects Y directly, and is also associated with X . The contrast correlation between X and Y will be low, and third traits such as colour will have been partly controlled for. In this case it appears that the contrast correlation is a better guide to the significance of the relationship. Note that a strict Brownian motion model will be violated because colour affects trait Y . As there is no evolution of colour within clades, the rate of evolution of Y will probably change. However, the method may be robust to such changes (Martins & Garland 1991; Diaz-Uriarte & Garland 1996). There are still two reasons why the species' correlation might be used.

First, the true model of evolution may differ more strongly from the Brownian motion assumptions. In particular, under the adaptive radiation model described in this paper the contrast regression and correlation is a biased estimator of the expected value of the species regression and correlation. The usual statement regarding the need to control for phylogeny is that species values are not independent (Harvey & Pagel 1991; Harvey & Purvis 1991). Non-independence does not affect bias, but rather reduces standard errors, making one more likely to reject the null hypothesis when it is true. Thus under stochastic evolutionary models the contrast and species regressions are both unbiased estimates of the 'true' species regression, but the estimate based on the contrast regression has a more acceptable standard error (Pagel 1993). In the model under discussion, the contrast regressions do not provide an unbiased estimate of the 'true' species regression. On the other hand, the species regression and correlation are unbiased estimators of the niche dimensions, as described by the parameters of the bivariate normal distribution. If one decides to use species values because they give an unbiased estimate of niche dimensions, it is reasonable to use the significance levels associated with those species values.

The second reason why the species correlations and regressions may be preferred is that some confounded third variables are likely to show some correlated evolution, and thus will not be completely controlled

for by contrasts, even if the evolutionary model is adequately described by Brownian motion. The contrast correlation partly controls for confounding variables, but how much is usually unknown. The correlation between species values does not control for third variables at all, but does indicate something meaningful: it describes the relationship among the traits of interest for the species being studied (without any reference to causation).

The arguments of this paper provide some justification for studies which compare only two species. These studies use individuals as replicates to demonstrate that species differ in two traits, and ask why the differences occur in the direction they do (Garland & Adolph 1994). It is true that a single instance of coordinate change could be attributed to many other traits or environments other than those under investigation (Lauder 1981; Garland & Adolph 1994). It is also true that divergence between any pair of species is likely to be affected by a different constellation of confounding traits than divergence between any other pair (Garland & Adolph 1994). Indeed, if this were not generally true (given a standard linear model of relationships among traits), the correlation across a group of species would be 1.0. The two-species approach therefore suffers from the problem of confounding traits, but so does the contrast method. Many two-species studies qualitatively, but often convincingly, assess and dismiss the possibility that other traits cause the observed relationship (e.g. Schluter 1993; Davies *et al.* 1996). The virtue of the two-species or few-species study is that it can be much more exhaustive, enabling in-depth examination of potential confounding variables. A second way to justify these studies would be to argue for an evolutionary model which always produces the same pattern of divergence of the two species from a common ancestor (i.e. is completely deterministic), so that the measured association is repeatable. This is less satisfactory than scrutinizing possible third variables because it does not separate causation from correlation. However, it might also be argued that this is as reasonable assumption to make as a model of stochastic evolution, which cannot be used without more species and estimates of phylogeny.

I have argued that confounding unmeasured variables cause the main difficulties in comparative studies, and that comparative methods which incorporate phylogeny do not control for such confounding variables particularly well. If we use the species values we do not control for them at all, but are describing the relationship among traits in their current environment. In this case we are treating species values as fixed effects, and the p -value could be considered a conservative guide to the effects of sampling and measurement error on estimates of the correlation and regression. The p -value is conservative because species deviate from a linear regression line due to the influences of third traits as well (increasing the error), but in such an analysis the effects of third traits are considered part of the fixed species values.

One mechanism to control for third variables is by multiple regression, using either the species or the

contrast values. The species analysis will generally be more powerful, because it requires fewer assumptions, and multiple regression will remove all effects of third variables included in the analysis, whether or not they are 'randomly' associated or correlated (see table 2). However, the use of multiple regression may not always be possible, because the variables cannot be identified, or more likely, because of collinearity and degrees of freedom problems. Typically, closely related species share many traits, which are correlated across species, and the inclusion of more than a few of them in multiple regression will be impossible.

Contrast correlations should be presented wherever possible, but agreement between contrast and species correlations in no way obviates the need to consider potentially confounding traits. Perhaps the greatest value of contrasts is that they provide one means of investigating the history of the group under investigation (e.g. Harvey *et al.* 1991; Richman & Price 1992). Differences between species and contrast correlations are likely to be of biological significance, suggesting predictable patterns of change in the evolutionary correlation during a species radiation. Similar arguments have been made by Ricklefs & Starck (1996), who refer to some alternative methods to detect patterns of evolution.

I thank S. Bensch, T. Case, D. Irwin, J. Kohn, E. Martins, A. Purvis, A. Richman, K. Roy, D. Schluter, A. Suarez, M. Turelli and especially K. Marchetti (Pondicherry beer) for much discussion and comments. I thank J. Losos for the initial discussions which led to this paper, and for a careful review of the manuscript.

REFERENCES

- Björklund, M. 1994 The independent contrast method in comparative biology. *Cladistics* **10**, 425–433.
- Björklund, M. 1997 Mortality vs brightness: comments on recent comparative analyses. *Evolution*. (In the press.)
- Brandl, R., Kristin, A. & Leisler, B. 1994 Dietary niche breadth in a local community of passerine birds: an analysis using phylogenetically independent contrasts. *Oecologia* **98**, 109–116.
- Briskie, J. V. & Montgomerie, R. 1992 Sperm size and sperm competition in birds. *Proc. R. Soc. Lond. B* **247**, 89–95.
- Briskie, J. V., Naugler, C. T. & Leech, S. M. 1994 Begging intensity of nestling birds varies with sibling relatedness. *Proc. R. Soc. Lond. B* **258**, 73–78.
- Burt, A. & Bell, G. 1987 Red queen versus tangled bank models. *Nature, Lond.* **330**, 118.
- Clark, A. & Wang, L. 1994 Comparative evolutionary analysis of metabolism in nine *Drosophila* species. *Evolution* **48**, 1230–1243.
- Davies, N. B., Hartley, I. R., Hatchwell, B. J. & Langmore, N. E. 1996 Female control of copulations to maximize male help—a comparison of the polygynandrous alpine accentor, *Prunella collaris*, and dunnock, *P. modularis*. *Anim. Behav.* **51**, 27–47.
- De Voogd, T. J., Krebs, J. R., Healy, S. D. & Purvis, A. 1993 Relations between song repertoire size and the volume of brain nuclei related to song—comparative evolutionary analyses amongst oscine birds. *Proc. R. Soc. Lond. B* **254**, 75–82.
- Diaz-Uriarte, R. & Garland, T. 1996 Testing hypotheses of correlated evolution using phylogenetically independent contrasts: sensitivity to deviations from Brownian motion. *Syst. Biol.* **45**, 27–47.
- Felsenstein, J. 1985 Phylogenies and the comparative method. *Am. Nat.* **125**, 1–15.
- García-Barros, E. 1994 Egg size variation in European Satyrine butterflies (Nymphalidae, Satyrinae). *Biol. J. Linn. Soc.* **51**, 309–324.
- Garland, T. 1992 Rate tests for phenotypic evolution using phylogenetically independent contrasts. *Am. Nat.* **140**, 509–519.
- Garland, T. & Adolph, S. C. 1994 Why not to do two-species comparative tests—limitations on inferring adaptations. *Phys. Zool.* **67**, 797–828.
- Garland, T. & Janis, C. M. 1993 Does metatarsal femur ratio predict maximal running speed in cursorial mammals? *J. Zool.* **229**, 133–151.
- Garland, T., Huey, R. B. & Bennett, A. F. 1991 Phylogeny and coadaptation of thermal physiology in lizards—a reanalysis. *Evolution* **45**, 1969–1975.
- Garland, T., Harvey, P. H. & Ives, A. R. 1992 Procedures for the analysis of comparative data using phylogenetically independent contrasts. *Syst. Biol.* **41**, 18–32.
- Gould, S. J. & Lewontin, R. C. 1979 The spandrels of San Marco and the Panglossian paradigm: a critique of the adaptationist programme. *Proc. R. Soc. Lond. B* **205**, 581–598.
- Gould, S. J. & Vrba, E. S. 1982 Exaptation—a missing term in the science of form. *Paleobiology* **8**, 4–15.
- Grafen, A. 1989 The phylogenetic regression. *Phil. Trans. R. Soc. Lond. B* **326**, 119–157.
- Harvey, P. H. & Pagel, M. D. 1991 *The comparative method in evolutionary biology*. New York: Oxford University Press.
- Harvey, P. H. & Purvis, A. 1991 Comparative methods for explaining adaptations. *Nature, Lond.* **351**, 619–624.
- Harvey, P. H., Pagel, M. D. & Rees, J. A. 1991 Mammalian metabolism and life histories. *Am. Nat.* **137**, 556–566.
- Harvey, P. H., Read, A. F. & Nee, S. 1995 Why ecologists need to be phylogenetically challenged. *J. Ecology* **83**, 535–536.
- Lauder, G. V. 1981 Form and function: structural analysis in evolutionary morphology. *Paleobiology* **7**, 430–442.
- Leroi, A. M., Rose, M. R. & Lauder, G. V. 1994 What does the comparative method reveal about adaptation? *Am. Nat.* **143**, 381–402.
- Losos, J. B. 1990a Ecomorphology, performance capability, and scaling of West Indian *Anolis* lizards—an evolutionary analysis. *Ecol. Monogr.* **60**, 369–388.
- Losos, J. B. 1990b Concordant evolution of locomotor behaviour, display rate and morphology in *Anolis* lizards. *Anim. Behav.* **39**, 879–890.
- Losos, J. 1994 An approach to the analysis of comparative data when a phylogeny is unavailable or incomplete. *Syst. Biol.* **43**, 117–123.
- Losos, J. & Miles, D. 1994 Adaptation, constraint and the comparative method: phylogenetic issues and methods. In *Ecological morphology: integrative organismal biology* (eds P. Wainwright and S. Reilly), pp. 60–98. University of Chicago Press.
- Lynch, M. 1991 The rate of morphological evolution in mammals from the standpoint of the neutral theory. *Am. Nat.* **136**, 727–741.
- Marchetti, K. 1993 Dark habitats and bright birds illustrate the role of the environment in species divergence. *Nature, Lond.* **362**, 149–152.
- Marchetti, K., Price, T. & Richman, A. 1995 Associations of wing morphology with ecology in *Phylloscopus* species. *J. Avian Biol.* **26**, 177–181.
- Marchetti, K. & Price, T. 1997 The adaptive significance of colour patterns in warblers. *Oikos*. (In the press.)

- Martin, T. E. 1993 Evolutionary determinants of clutch size in cavity-nesting birds—nest predation or limited breeding opportunities. *Am. Nat.* **142**, 937–946.
- Martins, E. P. 1993 A comparative study of the evolution of *Sceloporus* push-up displays. *Am. Nat.* **142**, 994–1018.
- Martins, E. P. & Garland, T. 1991 Phylogenetic analyses of the correlated evolution of continuous characters—a simulation study. *Evolution* **45**, 534–557.
- Møller, A. P. & Birkhead, T. R. 1993 Certainty of paternity covaries with paternal care in birds. *Behav. Ecol. Sociobiol.* **33**, 261–268.
- Owens, I. P. F. & Bennett, P. M. 1994 Mortality costs of parental care and sexual dimorphism in birds. *Proc. R. Soc. Lond. B* **257**, 1–8.
- Pagel, M. 1992 A method for the analysis of comparative data. *J. Theor. Biol.* **156**, 431–442.
- Pagel, M. 1993 Seeking the evolutionary regression coefficient—an analysis of what comparative methods measure. *J. Theor. Biol.* **164**, 191–205.
- Pagel, M. D. & Harvey, P. H. 1989 Taxonomic differences in the scaling of brain on body weight among mammals. *Science, Wash.* **244**, 1589–1593.
- Poiani, A. 1993 Small clutch sizes as a possible adaptation against ectoparasitism—a comparative analysis. *Oikos* **68**, 455–462.
- Promislow, D. 1992 Costs of sexual selection in natural populations of mammals. *Proc. R. Soc. Lond. B* **247**, 203–210.
- Promislow, D., Montgomerie, R. & Martin, T. 1992 Mortality costs of sexual dimorphism in birds. *Proc. R. Soc. Lond. B* **250**, 143–150.
- Richman, A. D. & Price, T. 1992 Evolution of ecological differences in the Old World leaf warblers. *Nature, Lond.* **355**, 817–821.
- Ricklefs, R. E. & Starck, J. M. 1996 Applications of phylogenetically independent contrasts: a mixed progress report. *Oikos* **77**, 167–172.
- Ridley, M. 1989 Why not to use species in comparative tests. *J. Theor. Biol.* **136**, 361–364.
- Riska, B. 1991 Regression models in evolutionary allometry. *Am. Nat.* **138**, 283–299.
- Saether, B. & Gordon, I. 1994 The adaptive significance of reproductive strategies in ungulates. *Proc. R. Soc. Lond. B* **256**, 263–268.
- Sessions, S. K. & Larson, A. 1987 Developmental correlates of genome size in Plethodontid salamanders and their implications for genome evolution. *Evolution* **41**, 1239–1251.
- Schluter, D. 1993 Adaptive radiation in sticklebacks—size, shape and habitat use efficiency. *Ecology* **74**, 699–709.
- Sibley, C. & Ahlquist, J. E. 1990 *Phylogeny and classification of birds: a study in molecular evolution*. New Haven: Yale University Press.
- Suhonen, J., Alatalo, R. V. & Gustafsson, L. 1994 Evolution of foraging ecology in Fennoscandian tits (*Parus* spp.). *Proc. R. Soc. Lond. B* **258**, 127–131.
- Telleria, J. & Carrascal, L. 1994 Weight–density relationships between and within bird communities—implications of niche space and vegetation structure. *Am. Nat.* **143**, 1083–1092.
- Walton, B. M. 1993 Physiology and phylogeny: the evolution of locomotor energetics in the Hylid frogs. *Am. Nat.* **141**, 26–50.
- Zar, J. H. 1996 *Biostatistical analysis*, 3rd. edn. New Jersey, USA: Prentice Hall.

Received 3 September 1996; accepted 7 October 1996