Evolutionary Ecology Research, 2008, **10**: 787–796

On comparative analyses involving non-heritable traits: why half a loaf is sometimes worse than none

William E. Kunin

Earth & Biosphere Institute and Institute for Integrative and Comparative Biology, University of Leeds, Leeds, UK

ABSTRACT

Question: Should phylogenetically informed (PI) analyses be used when not all variables in an analysis are heritable?

Methods: I simulated phylogenetic trees, with randomized traits (1) inherited with variation, (2) assigned directly to extant species, or (3) made partially dependent on a heritable variable, testing the frequency of 'significant' correlations between variables using conventional and two different PI techniques.

Results: 'Significance' was inflated in analyses of heritable variables, and this was corrected by both PI methods. However, where one variable was heritable and the other not, conventional analyses provided unbiased probability estimates. Modelled correlations between heritable and non-heritable traits were more readily detected by conventional analyses, but analyses involving 'incorrect' heritable traits sometimes showed spurious correlations.

Conclusions: The results suggest that PI analyses are inappropriate when only one of a pair of variables displays phylogenetic pattern. Where intrinsically non-heritable traits display phylogenetic pattern, conventional analyses are appropriate as an initial approach, but residuals should be tested for phylogenetic patterning.

Keywords: abundance, comparative method, correlated traits, distribution, heritability, independent contrasts, phylogenetic analyses.

INTRODUCTION

The use of interspecific comparisons is central to the study of ecology and evolutionary biology. The practice is particularly important in addressing issues that operate on temporal or spatial scales too vast to allow experimental manipulation, in particular the evolution of particular species characteristics. Thus if species with, for example, polygamous mating systems also tended to have pronounced sexual dimorphism, this might be taken as evidence that the two variables were linked in some way: perhaps the dimorphism evolved as a result of the increased sexual selection resulting from polygamy (e.g. Dunn *et al.*, 2001). Until

* Correspondence: W.E. Kunin, Earth & Biosphere Institute and Institute for Integrative and Comparative Biology, University of Leeds, Leeds LS2 9JT, UK. e-mail: w.e.kunin@leeds.ac.uk

fairly recently, it was common practice to perform such comparative studies by treating each species as an independent data point. Felsenstein (1985), however, noted that such analyses can be seriously misleading, as traits may be associated with one another by inheritance rather than due to adaptive evolution. To treat each species as an independent event in such cases clearly exaggerates the statistical power of the analysis (Harvey and Pagel, 1991).

The growing appreciation of these matters has revolutionized the study of ecology, comparative morphology, and evolutionary biology. Although there have been some dissenting voices (e.g. Westoby *et al.*, 1995; Price 1997), most ecologists and evolutionary biologists have either adopted phylogenetically informed (PI) methods willingly or have grudgingly accepted them at the behest of journal editors and statistical referees. As the consensus for using such methods has grown, an increasingly diverse armoury of techniques has been developed (e.g. Grafen, 1989; Garland *et al.*, 1993; Martins and Hansen, 1997), allowing PI analyses of diverse data types under a range of statistical and evolutionary models. Many of these methods make use of 'independent contrasts' – examining the *differences* in trait values between related species rather than the raw trait values themselves. These contrast values can be interpreted as measures of the evolutionary change that has occurred since species diverged from a common ancestor, so that each such contrast represents in principle a statistically independent event. In well-resolved phylogenies, this gain in statistical independence comes at little cost in statistical power, as N species can provide $N - 1$ contrasts for analysis.

However, it is far less clear what should be done when only a subset of variables displays such phylogenetic structure. In a recent paper, Rheindt and colleagues (2004) examined the relationship between habitat – a highly phylogenetically conserved trait in the set of West African birds they studied – and various song characteristics, which evolve so rapidly that little if any phylogenetic signal remains (Abouheif, 1999). This is, to my knowledge, the first paper to consider how best to test for relationships between variables that differ in their phylogenetic properties. In such circumstances, one might be tempted to make use of phylogenetically independent contrasts in the variable that is highly conserved (e.g. habitat), but to use raw trait values for the other trait (e.g. song frequency) that is otherwise uncoupled from the phylogeny. This solution, however, creates a second problem. As noted above, even in a fully resolved phylogeny containing N species, there are only $N - 1$ independent contrasts – each is a difference between two species (or rather two nodes) rather than a species value *per se*. How can one test for correlations between N species values of one variable and $N - 1$ interspecific contrast values of the other? To resolve this, Rheindt *et al.* (2004) make use of a 'star phylogeny' in analyses of the non-phylogenetic traits, retaining the branching structure of the original phylogeny but reducing all branch lengths except the terminal ones to zero. This method treats each species as if it had evolved separately and simultaneously from a single common ancestor, and indeed Purvis and Garland (1993) showed that it provides equivalent results to standard non-phylogenetic methods. They then have both heritable and non-heritable traits in the common currency of contrasts (albeit from phylogenies with different relative branch lengths), allowing them to compare both phylogenetically conserved and non-conserved variables in a single analysis. Interestingly, in their analyses, Rheindt *et al.* (2004) find a strong association between habitat and one aspect of bird song (maximum frequency) in both conventional and fully phylogenetic analyses, but find that the relationship disappears when analysed with what they believe to be an appropriately semi-phylogenetic method.

The issue of relationships between variables with differing degrees of statistical independence, however, is not unique to phylogenetics. In many ways, an analogous

problem arises in spatial statistics, when considering the effects of spatial autocorrelation. Just as with phylogenetic comparisons, the problem of spatial autocorrelation is essentially one of lack of statistical independence: two samples collected in close physical proximity to each other are not fully independent, just as two species closely related to each other are not. The similarity between issues of phylogenetic and spatial non-independence has been noted before (e.g. Cheverud *et al.*, 1985), although the direct application of spatial methods to phylogenies is not straightforward (Rohlf, 2001). As it happens, the issue of how best to analyse relationships between variables with different degrees of such non-independence has already been considered in spatial analyses. The answer may surprise many evolutionary biologists: when analysing a relationship between two variables, one of which is strongly autocorrelated in space but the other of which is not, there is *no need* to correct for autocorrelation (Clifford *et al.*, 1989). By analogy, if we are analysing a relationship between two species traits, one of which is phylogenetically conserved and the other of which is not, it seems possible that no phylogenetic correction may be required.

In this paper, I examine whether phylogenetic correction, either by standard PI methodologies or by that proposed by Rheindt *et al.* (2004), is appropriate in cases of this sort, by testing them on simple simulated phylogenies. I then discuss the issue of how best to analyse the relationship between pairs of variables where only one displays phylogenetic autocorrelation. Finally, I address the rather trickier question of how best to deal with variables that may show some phylogenetic correlation but which are intrinsically incapable of being inherited.

METHODS

To test these ideas, a set of simple phylogenetic trees was constructed. Each tree contained 32 species, arranged for simplicity as a set of symmetrical bifurcations with equal time intervals for all internal and terminal branches. Traits were mapped onto this phylogeny in several ways.

Fully *independent phylogenetic* (IP) traits were simulated by beginning with a random ancestral value for the basal species in the tree, and then varying the trait value randomly at each step in the phylogeny from its ancestral value. Both additive and multiplicative models were tested: in the former, a uniform random variable with mean = 0 and range (max–min) of r was added at each step; in the latter, the ancestral trait value was multiplied by a uniform random variable with geometric mean = 1 and range r. Note that the latter case converges on the former when results are plotted logarithmically. For simplicity, only the additive results will be reported here (with $r = 5$), although the multiplicative model or different r values leads to identical conclusions on the substantive points discussed below.

By contrast, *independent non-phylogenetic* (IN) traits were assigned randomly to extant species at the tips of the phylogenetic tree. This would be appropriate for rapidly evolving traits (Rheindt *et al.*, 2004), or for traits that were not heritable at all. Clearly for such traits, ancestral values could not actually be inferred from current trait values. Nonetheless, ancestral values within the phylogeny were impugned using standard phylogenetic methods (Felsenstein, 1985; Harvey and Pagel, 1991), as an analyst would do if applying standard PI techniques. In addition, a 'star' phylogeny was used, in which all but terminal branches in the phylogeny were reduced to zero length, as recommended by Rheindt and colleagues (2004).

A final category of variables was simulated as *correlated non-phylogenetic* (CN) traits. In this case, even though the trait itself was not directly heritable, its values were partially

dependent on another (IP) trait that *was* heritable. Thus, for example, species abundance cannot be directly inherited from an ancestor, but it may be highly correlated with other traits (e.g. body size) that are directly heritable. Values were assigned in CN traits by assigning only a proportion $1 - p$ of the value to extant species at random (as in IN traits). The remaining fraction (p) was assigned as a linear function ($y = a + bx$) of an IP trait's value. For simplicity, the simulations reported here will be for the simplest linear function, with a set to 0 and b set to 1 so that $y = x$, although simulations with other linear functions were explored as well. The resulting trait may show phylogenetic pattern, but does so because it has been affected by a heritable trait, even though much of its variation ($1 - p$) is non-heritable. As in IN traits, putative 'ancestral' values were then assigned using both standard phylogenetic techniques and through the use of a star phylogeny.

Analyses were performed on four different combinations of variable types: (a) two uncorrelated phylogenetic variables (IP × IP′); (b) two uncorrelated variables of which only one was heritable (IN × IP); (c) a correlated non-heritable trait and its associated phylogenetic trait (CN × IP); and (d) a correlated non-heritable trait and an independent phylogenetic trait other than the one to which it is related (CN × IP′). A spreadsheet was developed (in Microsoft Excel; available for download from the author) to conduct all four comparisons simultaneously. Uncorrelated analyses (cases (a) and (b), above) were simulated 400 times each. For CN × IP analyses (case (c), above), 200 sets of simulations each were performed for $p = 0.3$, $p = 0.2$, and $p = 0.1$, thus allowing the effect of correlation strength to be assessed. Analyses involving a CN variable and the 'incorrect' IP variable (case (d)) were done only for a single correlation strength ($p = 0.2$) and again replicated 200 times. For each simulation, the Pearson's product–moment correlation coefficient (r) was calculated, and used to assess 'significance' relative to the appropriate $P = 0.05$ critical value (0.349 for the raw data with $n = 32$; 0.355 for the phylogenetic analyses with $n = 31$). R^2 values were computed to measure the strength of each association.

RESULTS

The relative value of the different methods employed varied strikingly, depending on context. In classical phylogenetic contexts (case (a): IP × IP′), phylogenetic analyses clearly proved their worth, as expected. Here there was no intrinsic correlation between the variables, and so any appearance of a relationship between variables is due to chance or methodological artefact. Where the analyses are not corrected for phylogeny, they reveal substantially inflated R^2 values, and show substantially higher numbers of nominally 'significant' correlations than would be expected by chance (Table 1): 85 instead of the expected 20 of 400 runs ($\chi^2 = 222.37$, $P < 0.0001$). Phylogenetically corrected analyses, by contrast, produced roughly the number of significant runs expected: 25 out of the 400 runs ($\chi^2 = 1.316$, $P = 0.251$).

The situation becomes rather different, however, when we examine the relationship between a phylogenetic trait and a second trait that is not heritable. Where the two traits are uncorrelated (case (b): IN × IP) we would expect to get a 'significant' result in only 5% of simulations. That is precisely what we find (Table 1), whether we analyse the relationship using phylogenetic correction or not, and indeed, whether the phylogeny employed is bifurcating or effectively 'star' shaped, following the advice of Rheindt *et al.* (2004). Thus of 400 runs, 22 showed 'significant' uncorrected analyses ($\chi^2 = 0.211$, $P = 0.646$), compared with 31 for independent contrast methods ($\chi^2 = 6.368$, $P = 0.012$) and 26 when a star

Table 1. Pearson product–moment correlations between sets of independently simulated traits

	Uncorrected analysis		Conventional phylogenetic		Star phylogentic	
	'Significant' runs (of 400)	Mean R^2 ± S.D.	'Significant' runs (of 400)	Mean R^2 ± S.D.	'Significant' runs (of 400)	Mean R^2 ± S.D.
Case (a)	85 ***	0.0740 [a]	25	0.0356 [b]	N.A.	
IP × IP′		± 0.079		± 0.047		
Case (b)	22	0.0328	31*	0.0399	26	0.0360
IN × IP		± 0.043		± 0.055		± 0.047

Note: Comparing two uncorrelated phylogenetic traits (case (a): IP × IP′) shows spuriously high levels of significant effects, whereas when only one trait is heritable (case (b): IN × IP) no such inflation of significance occurs. Asterisks indicate significant differences from the expected 5% of cases (20), as tested by χ^2 analysis: *$P < 0.05$, **$P < 0.001$, ***$P < 0.0001$. Superscript letters indicate significant differences between R^2 values.

phylogeny was used ($\chi^2 = 1.895$, $P = 0.169$). Oddly, the standard PI independent contrast method produced significantly more than the expected number of 'significant' results here, although a single such anomalous finding might be expected among the large number of tests reported here. As suggested in the Introduction, it is clear that phylogenetic correction is unnecessary in cases where one of the variables is independent of phylogeny.

This point is underscored when we consider a similar case, but now build in a correlation between the phylogenetic and non-phylogenetic traits (case (c): CN × IP). Here the use of a phylogenetically informed analysis (either conventional or star) weakens the measured relationship. Where weak correlations were imposed between the variables ($p = 0.1$ and $p = 0.2$), the relationship was substantially less apparent to the two phylogenetically informed techniques than it was to conventional (uncorrected) analyses (Table 2). Thus of 200 runs with $p = 0.2$, 164 showed significant correlations in conventional analyses, while only 103 and 124 were significant when analysed by the two phylogenetic techniques (both significantly lower than conventional; $\chi^2 = 74.49$ and 33.95, both $P < 0.00001$). This difference largely disappears when the strength of the effect is made stronger ($p = 0.3$), as all three techniques uncovered significant correlations in almost all runs at this level, but even in these runs the measured strength of the correlation was affected. While neither phylogenetic technique performed as well as the uncorrected analysis, the star phylogeny method proposed by Rheindt and colleagues (2004) generally performed less poorly than did conventional phylogenetic techniques.

The final case considered involved comparing a correlated non-phylogenetic trait with the 'wrong' phylogenetically heritable trait (case (d): CN × IP′) – that is, to a trait other than the one to which the CN trait was related. In this case, the statistical bias reappears: more significant correlations are detected than the 5% that would be expected by chance alone (Table 2). Even with the relatively modest phylogenetic effect modelled here ($p = 0.2$), more than twice the expected number of 'significant' correlations were displayed ($\chi^2 = 15.16$, $P < 0.001$).

Table 2. Results of simulations of 'correlated non-phylogenetic' (CN) traits, and their relationships with independent phylogenetic (IP) traits upon which they depend (case (c)) and with other IP traits with no intrinsic correlation (case (d))

	Uncorrected analysis		Conventional phylogenetic		Star phylogentic	
	'Significant' runs (of 200)	Mean R^2 ± S.D.	'Significant' runs (of 200)	Mean R^2 ± S.D.	'Significant' runs (of 200)	Mean R^2 ± S.D.
Case (c): Correlated non-phylogenetic × 'correct' phylogenetic trait						
$p = 0.1$	42 [a]	0.074 [a] ± 0.069	21 [b]	0.051 [b] ± 0.059	22 [a,b]	0.056 [b] ± 0.059
$p = 0.2$	164 [a]	0.243 [a] ± 0.125	103 [b]	0.038 [c] ± 0.050	124 [b]	0.180 [b] ± 0.114
$p = 0.3$	200	0.493 [a] ± 0.108	193	0.328 [b] ± 0.128	197	0.362 [b] ± 0.107
Case (d): Correlated non-phylogenetic × 'incorrect' phylogenetic trait						
$p = 0.2$	22**	0.043 ± 0.054	14	0.038 ± 0.050	7	0.033 ± 0.039

Note: For case (c), superscripts represent significant differences between methods. For case (d), one should expect 5% of trials (10) to appear significant by chance, and deviations from this expectation are indicated by asterisks, as in Table 1.

DISCUSSION

The aim of this paper is to make two main points, one of which is simple to demonstrate, but the second of which may require some subtlety.

The first point is to demonstrate that *no phylogenetic correction is required in any bivariate analysis where at least one of the two variables considered is unrelated to phylogeny*. I argued in the Introduction that full statistical independence could be assumed in such cases. The results of my simulations support this assertion. Despite the fact that spurious 'significance' appears when relationships between two phylogenetically heritable variables are analysed, no such inflation of significance is seen when one variable is heritable but the other is not. The method proposed by Rheindt and colleagues (2004) is clever, but it is unnecessary. Making use of independent contrast methods in cases like this unnecessarily reduces the power of comparative analyses, and obfuscates otherwise clear trait comparisons by converting them from directly observable values to contrast scores on divergent phylogenies. Worse, in the case cited in the original paper (Rheindt *et al.*, 2004), the technique successfully masks a biologically meaningful pattern uncovered in the (correct) non-phylogenetic analysis (and indeed in the fully phylogenetic one).

This result matches closely the established result for spatial statistics (Clifford *et al.*, 1989), where no correction is recommended where one variable in a bivariate analysis is spatially autocorrelated and the other is not. In that case, so long as one of the two variables is 'free', then the relationship between them is unconstrained by space and no loss of statistical independence is experienced; the same applies here to phylogenetic correlation. To take the specific case posed by Rheindt *et al.* (2004), the association between (rapidly evolving) bird

song characteristics and (phylogenetically conserved) habitat need not have been subjected to the specialized analyses used, because each bird species *independently* evolves its song characteristics in response to its habitat, and so the original (uncorrected) finding of a relationship between the two should stand.

That far I can state with confidence. The second point I expect to be more contentious, but I nonetheless put it forward for discussion. At issue is how we should deal with analyses of intrinsically non-heritable traits that may nonetheless display phylogenetic correlation.

There are some widely studied ecological traits that are not in themselves heritable, but which are likely to be affected by other traits that are. The best examples tend to be emergent 'population-level' traits rather than traits of individuals: traits such as local abundance, spatial patterning within or between populations, geographical range extent, and other aspects of species commonness or rarity are not in themselves heritable in the way that morphological traits such as feather colour or leaf shape are. Depending on the mode of speciation, one might argue that there may be some degree to which these properties in a daughter species might be affected by the parent species' traits immediately after speciation (e.g. in a vicariant speciation event), but the relationship will not be one of 'inheritance' in the usual sense, and should fade with time.

However, the fact that these traits are not directly heritable does not necessarily mean that they are uncorrelated to phylogeny. Population density, for instance, may be affected by body size (Blackburn and Gaston, 1999); spatial pattern may be affected by dispersal ability (e.g. Pocock *et al.*, 2006; Rundle *et al.*, 2007); range size may be affected by temperature tolerance (Duncan *et al.*, 2001; Pither, 2003) or diet breadth (Beck and Kitching, 2007) – all traits that are themselves heritable. Indeed, there is evidence that some population distributional characteristics *do* show phylogenetic patterns (Jablonski, 1987; but see Webb and Gaston, 2003, 2005; Hunt *et al.*, 2005), presumably due to the effects of heritable traits of this sort. 'Correcting' for phylogeny in situations of this sort risks discarding the ecological baby along with the methodological bathwater.

How, then, should one analyse such cases? On the face of it, the simulations reported here appear to give clear advice: standard, non-phylogenetic analyses are substantially more likely to detect such a pattern when it is present (case (c)), but do not inflate significance to falsely indicate such a relationship when it is not present (case (b)). The two phylogenetically informed techniques considered here increased false-negatives (type II errors) without reducing false-positives (type I errors), and so seem to have little justification.

The final case considered (case (d)), however, provides some grounds for caution. In analysing the relationship between an intrinsically non-heritable (but phylogenetically correlated) variable and a genuinely heritable one, we run the risk of being fooled by the effects of some other heritable variable not considered in the analysis. To take the example of the body size–abundance relationship, it might be that abundance was influenced not by body size but rather by some other heritable trait (e.g. anti-predator defences or habitat preference), and that this other factor was responsible for the phylogenetic pattern displayed in the (non-heritable) abundance trait. If this other factor was not included in our analyses, we might spuriously detect a body size–abundance relationship due to the potential phylogenetic correlation between the (incorrect) heritable trait considered (body size) and the correct heritable trait that was not included in the analysis. Of course, problems caused by the effects of excluded but correlated variables are not restricted to phylogenetic analyses; one must always be careful to consider this possibility in any analysis. However, if the problem is not unique to heritable variables, it is uniquely powerful in such cases. Among heritable variables, there will not only be correlations between functionally linked

traits (e.g. body size and metabolic rate), there will also often be correlations due to common descent, even where traits are functionally independent. Thus the risk of effects being caused by correlated variables not included in an analysis are increased, as the number of candidate correlated variables is itself increased. Phylogenetically informed analyses will greatly reduce the risk of being fooled by variables correlated by common descent – although they will not allay concerns about correlations due to functional links.

The situation is explored graphically in Fig. 1. In a classical phylogenetic analysis (Fig. 1a), two heritable traits (e.g. body size and metabolic rate) are considered, both of which may carry direct phylogenetic signals, allowing spurious correlations to arise. On the other hand, when a non-heritable trait (e.g. abundance) is involved (Fig. 1b), a phylogenetic pattern in it can only occur if it is driven by the heritable trait under consideration (here, body size), or by some other heritable trait not considered in the analysis. This latter possibility would be reflected as a phylogenetically correlated error term.

This then suggests a practical strategy. On balance, the best approach may be to use both methods. When considering the relationship between a heritable trait (such as body size) and an intrinsically non-heritable trait that nonetheless shows phylogenetic pattern (such as abundance), the first step should be to conduct conventional, non-phylogenetic analyses. Residuals from this analysis should then be examined for phylogenetic pattern [using, for

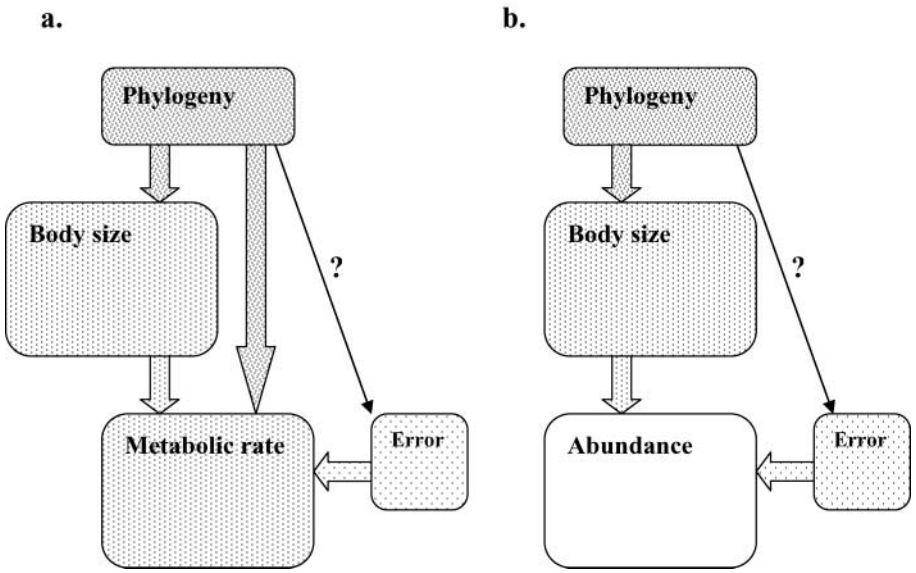

Fig. 1. Diagrammatic representation of two contrasting phylogenetic issues discussed in the text. Panel (a) represents the typical situation considered in phylogenetically informed analyses, where two heritable traits (here, body size and metabolic rate) are considered. If each is correlated with phylogeny, spurious relationships between them may emerge from traditional analyses. Panel (b), on the other hand, represents a case where one heritable variable (again, body size) is considered in relation to a second variable (here, abundance) that is intrinsically non-heritable. If a phylogenetic pattern is seen in this variable, it must either be due to the relationship under study (in which case it is of direct interest), or due to the influence of some other (phylogenetically inherited) variable not considered in the analysis. This latter possibility would be reflected in phylogenetic patterns in the error term.

example, the methods proposed by Abouheif (1999)]. If the phylogenetic pattern of the non-heritable variable was caused by its relationship with the heritable trait(s) in the analysis, that pattern should be removed once the relationship has been statistically accounted for, leaving no phylogenetic signal in the errors. If, on the other hand, the residuals continue to show significant phylogenetic pattern, it suggests that some other heritable variable that was not included in the analysis may be responsible. This then raises serious concerns about spurious correlations, and so a phylogenetically informed analysis should be considered. On the evidence of this paper, the semi-phylogenetic approach proposed by Rheindt *et al.* (2004) may be the preferred option in this case.

ACKNOWLEDGEMENTS

The author's work is supported by the Natural Environment Research Council (Grants GR3/11916 and NE/C507837/1), and by the EU FP6 ALARM project. Helpful comments were provided by Jack Lennon, Stephen Hartley, Chris Thomas, Ted Garland, and Wally Gilks.

REFERENCES

Abouheif, E. 1999. A method for testing the assumption of phylogenetic independence in comparative data. *Evol. Ecol. Res.*, **52**: 1197–1204.

Beck, J. and Kitching, I.J. 2007. Correlates of range size and dispersal ability: a comparative analysis of sphingid moths from the Indo-Australian tropics. *Global Ecol. Biogeogr.*, **16**: 341–349.

Blackburn, T.M. and Gaston, K.J. 1999. The relationship between animal abundance and body size: a review of the mechanisms. *Adv. Ecol. Res.*, **28**: 181–210.

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

Clifford, P., Richardson, S. and Hemon, D. 1989. Assessing the significance of the correlation between two spatial processes. *Biometrics*, **45**: 123–134.

Duncan, R.P., Bomford, M., Forsyth, D.M. and Conibear, L. 2001. High predictability in introduction outcomes and the geographical range size of introduced Australian birds: a role for climate. *J. Anim. Ecol.*, **70**: 621–632.

Dunn, P.O., Whittingham, L.A. and Pitcher, T.E. 2001. Mating systems, sperm competition and the evolution of sexual dimorphism in birds. *Evolution*, **55**: 161–175.

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

Garland, T., Jr., Dickerman, A.W., Janis, C.M. and Jones, J.A. 1993. Phylogenetic analysis of covariance by computer simulation. *Syst. Biol.*, **42**: 265–292.

Grafen, A. 1989. The phylogenetic regression. *Phil. Trans. R. Soc. Lond. B*, **326**: 119–156.

Harvey, P.H. and Pagel, M.D. 1991. *The Comparative Method in Evolutionary Biology*. Oxford: Oxford University Press.

Hunt, G., Roy, K. and Jablonski, D. 2005. Species-level heritability reaffirmed: a comment on 'On the heritability of geographic range sizes'. *Am. Nat.*, **166**: 129–135.

Jablonski, D. 1987. Heritability at the species level: analysis of the species ranges of Cretaceous molluscs. *Science*, **238**: 360–363.

Martins, E.P. and Hansen, T.F. 1997. Phylogenies and the comparative method: a general approach for incorporating phylogenetic information into analysis of interspecific data. *Am. Nat.*, **149**: 646–667.

Pither, J. 2003. Climate tolerance and interspecific variation in geographic range size. *Proc. R. Soc. Lond. B*, **270**: 475–481.

Pocock, M.J.O., Hartley, S., Telfer, M.G., Preston, C.D. and Kunin, W.E. 2006. Ecological correlates of range structure in rare and scarce British plants. *J. Ecol.*, **94**: 581–596.

Price, T. 1997. Correlated evolution and independent contrasts. *Phil. Trans. R. Soc. Lond. B*, **352**: 519–529.

Purvis, A. and Garland, T., Jr. 1993. Polytomies in comparative analyses of continuous variables. *Syst. Biol.*, **42**: 569–575.

Rheindt, F.E., Grafe, T.U. and Abouheif, E. 2004. Rapidly evolving traits and the comparative method: how important is testing for phylogenetic signal? *Evol. Ecol. Res.*, **6**: 377–396.

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

Rundle, S.D., Bilton, D.T., Abbott, J.C. and Foggo, A. 2007. Range size in North American Enallagma damselflies correlates with wing size. *Freshw. Biol.*, **52**: 471–477.

Webb, T.J. and Gaston, K.J. 2003. On the heritability of geographic range sizes. *Am. Nat.*, **161**: 553–566.

Webb, T.J. and Gaston, K.J. 2005. Heritability of geographic range sizes revisited: a reply to Hunt *et al. Am. Nat.*, **166**: 136–143.

Westoby, M., Leishman, M.R. and Lord, J. 1995. On misinterpreting the 'phylogenetic correction'. *J. Ecol.*, **83**: 531–534.