Functional Ecology

British Ecological Society

Functional Ecology 2009, **23**, 1059–1069 doi: 10.1111/j.1365-2435.2009.01596.x

New multivariate tests for phylogenetic signal and trait correlations applied to ecophysiological phenotypes of nine *Manglietia* species

Li Zheng[1], Anthony R. Ives*[,2], Theodore Garland Jr[3], Bret R. Larget[4], Yang Yu[5] and Kunfang Cao*[,1]

[1]*Xishuangbanna Tropical Botanical Garden, Chinese Academy of Sciences, Kunming 650223, China;* [2]*Department of Zoology, UW-Madison, Madison, Wisconsin 53706, USA;* [3]*Department of Biology, University of California, Riverside, Riverside, California 92521, USA;* [4]*Departments of Botany and Statistics, UW-Madison, Madison, Wisconsin 53706, USA; and* [5]*Yunnan Provincial Environmental Protection Department, Appraisal Center for Environment and Engineering, 27 Xiyuannan Road, Kunming 650032, China*

Summary

1. Phylogenetic signal – the similarity in trait values among phylogenetically related species – is pervasive for most types of traits in most organisms. Traits can often be categorized *a priori* into groups based on the level of biological organization, functional relations, developmental origins, or genetic underpinnings. Traits within such groups are often expected to be correlated and hence show similar levels of phylogenetic signal.

2. We developed multivariate statistical methods to test for phylogenetic signal in groups of traits while also incorporating estimates of trait measurement error (including within-species variation) that can obscure phylogenetic signal. Simultaneously, these methods produce estimates of correlations between traits that are corrected for phylogenetic relationships among species.

3. We applied these methods to data for 13 morphological and physiological traits gathered in a common-garden study of nine species of *Manglietia* (Magnoliaceae). The 13 traits fell into four groups: three traits involved photosynthesis [maximum net photosynthesis (A_{max}), light saturation point (LSP), light compensation point]; three described leaf morphology (thickness of leaves, palisade tissue, sponge tissue); four related to plant growth (basal stem diameter, crown volume, leaf area, relative growth rate); and three measured thermal tolerance [critical temperature (T_{ch}), peak temperature (T_{max}), temperature of half-inactivation (T_{50})]. We also constructed a molecular phylogeny for these species from 219 AFLP markers via maximum likelihood estimation under the assumption of sequential binary changes in DNA sequences.

4. Of the 13 traits, only two photosynthesis traits (A_{max} and LSP) exhibited statistically detectable phylogenetic signal ($P < 0{\cdot}05$) when analysed separately, whether using previously published univariate tests or our new univariate tests that incorporate measurement error. In contrast, multivariate analyses of the four trait groups, estimating simultaneously the phylogenetic signal for all traits and the correlations between traits, revealed a statistically significant phylogenetic signal for two of the four groups (photosynthesis and plant growth), comprising seven traits in total.

5. Our results demonstrate that even when the number of species in a comparative study is small, resulting in low power for univariate tests, phylogenetic signal can nonetheless be detected with multivariate tests that incorporate measurement error. Furthermore, our simulations show that the joint estimation of phylogenetic signal and trait correlations can lead to better (less biased and more precise) estimates of both.

Key-words: character syndromes, comparative methods, Magnoliaceae, phylogenetic inertia, phylogenetic signal, shade tolerance, strategy

*Correspondence authors. E-mail: arives@wisc.edu, caok@xtbg.en

Introduction

Statistical methods that incorporate phylogenetic information are now common in comparative analyses of trait variation and covariation (e.g. Clobert, Garland & Barbault 1998; Housworth, Martins & Lynch 2004; Hansen, Pienaar & Orzack 2008; Lavin *et al*. 2008), because these methods enhance both biological insight and statistical validity. Biologically, shared ancestry should cause related species to resemble each other for many traits, a pattern referred to as phylogenetic signal (Blomberg & Garland 2002). For example, in a survey of traits taken from many studies, Blomberg, Garland & Ives (2003) demonstrated that phylogenetic signal was pervasive. Nonetheless, they found that phylogenetic signal was weaker on average for behavioural traits than for body size or size-adjusted morpholometric traits. This finding may indicate that behavioural traits experience selection that breaks down phylogenetic patterns (Revell, Harmon & Collar 2008), giving inferential information about evolution of behavioural traits (although the possibility that they exhibit greater measurement error cannot be discounted; Ives, Midford & Garland 2007).

Statistically, analyses that require interspecific comparisons are challenging, because species cannot be assumed to be independent data points, violating the foremost assumption of many standard statistical analyses. When standard statistical methods are applied to phylogenetically related data, type I errors (rejecting the null hypothesis when in fact it is true) are often inflated, and coefficients estimated from statistical models (such as regression slopes) may not be minimum variance (e.g. Grafen 1989; Diaz-Uriarte & Garland 1996; Garland & Díaz-Uriarte 1999; Rohlf 2006). To address both biological and statistical issues, a large number of phylogenetically informed approaches have been developed for a wide range of analyses (reviewed in Garland, Bennett & Rezende 2005; Lavin *et al*. 2008).

Here, we extend this programme by developing new methods that estimate the correlation in trait values among species while simultaneously estimating the strength of phylogenetic signal and accounting for measurement error. This overcomes a limitation of the most commonly used phylogenetic statistical methods; they require an *a priori* assumption about the strength of phylogenetic signal in the data to be analysed. For example, Felsenstein's independent contrasts method (Felsenstein 1985) assumes that evolution follows a 'Brownian motion' process in which trait values increase or decrease randomly as evolution proceeds incrementally up a hierarchical phylogenetic tree. The assumption would be invalid, however, if trait evolution did not proceed in a Brownian motion fashion (e.g. Diaz-Uriarte & Garland 1996). In contrast, our methods estimate phylogenetic signal simultaneously with trait correlations, rather than making *a priori* assumptions about its strength, and so they should improve the estimation of trait correlations (Ackerly 1999) and related inferences, such as determining whether correlations differ statistically from zero.

Our methods are related to recently developed approaches used for detecting phylogenetic signal (Hansen 1997; Ackerly & Reich 1999; Freckleton, Harvey & Pagel 2002; Blomberg, Garland & Ives 2003; Butler & King 2004, 2005; Hansen, Pienaar & Orzack 2008; Scales, King & Butler 2009) and for performing regression analyses while simultaneously estimating the strength of phylogenetic signal (Lavin *et al*. 2008). Our methods use a conceptually related approach, but applied to correlation analyses. The approach is similar to that proposed by Freckleton, Harvey & Pagel (2002) and generalized by Revell & Harrison (2008), although it is based on an explicit model of evolution (see Blomberg, Garland & Ives 2003).

In addition to eliminating the need to make *a priori* assumptions about the strength of phylogenetic signal, our methods have two advantages. First, they allow for the joint estimation of phylogenetic signal in multiple traits (as does Freckleton, Harvey & Pagel 2002). Based on simulations of single continuous-valued traits, Blomberg, Garland & Ives (2003) found that *c*. 20 species are required to detect phylogenetic signal with a statistical power of about 0·8. However, statistical power could be improved if several traits are analysed simultaneously. For example, phylogeny aside, a multivariate analysis of variance (MANOVA) of several (correlated) traits typically yields much higher power to detect group differences compared with an ANOVA of a single trait (Willig & Owen 1987; Schmitz, Cherny & Fulker 1998). Our statistical approach allows for the joint estimation of phylogenetic signal for any group of traits, but rather than create groups arbitrarily, it makes most sense to combine traits based on level of biological organization (e.g. behaviour or biochemistry), functional relations (e.g. traits involved in locomotor or feeding performance), developmental origins (e.g. Dohm & Garland 1993) or genetic underpinnings (e.g. Arnold *et al*. 2008), because traits within such groups are expected to be correlated.

Second, our methods make it possible to incorporate within-species variation (or measurement error *sensu* Ives, Midford & Garland 2007). Real data sets contain within-species variation, and Ives, Midford & Garland (2007) and Felsenstein (2008) have shown that incorporating measurement error improves the accuracy of statistical tests and increases the statistical power to detect phylogenetic signal, group differences and so forth. By incorporating independent estimates of measurement error, our methods should give better estimates of trait correlations and phylogenetic signal.

We illustrate these methods by analysis of 13 traits measured in a common-garden study of nine species of *Manglietia* (Magnoliaceae). Below, we first focus on phylogenetic signal, asking which of 13 physiological and morphological traits show interspecific variation that to some extent reflects the relationships of the underlying phylogenetic tree. We then address the estimation of correlations among traits. For both phylogenetic signal and trait correlations, we present simulation studies to explore the statistical properties of the methods.

Materials and methods

PLANT TRAITS AND PHYLOGENETIC RECONSTRUCTION

We performed a common-garden experiment using nine *Manglietia* species: *M. grandis*, *M. hookeri*, *M. insignis*, *M. szechuanica*, *M. megaphylla*, *M. kwangtungensis*, *M. fordiana*, *M. chingii* and *M. pachyphylla* (for details, see Appendix S1, Supporting Information). After growing four seedlings of each species for 11 months in a greenhouse at the Xishuangbanna Tropical Botanical Garden, we performed 10 physiological and morphological measurements that divide into three categories: photosynthesis (maximum net photosynthesis A_{max}, light saturation point LSP and light compensation point LCP), plant growth (basal stem diameter, crown volume, plant height, leaf area, relative growth rate) and thermal tolerance (critical temperature T_{ch}, peak temperature T_{max} and temperature half-inactivation T_{50}). Data for the fourth group of traits, leaf morphology (thickness of leaves, palisade tissue, sponge tissue), were obtained for each species from the literature (Xie & Zhenghai 2000). Because all traits except the measurements of leaf morphology obtained from the literature were measured on multiple individuals, we used the standard error of the mean values of the traits for each species to estimate within-species variability, or measurement error (Ives, Midford & Garland 2007). Because multiple trait values were obtained from the same individuals, within-species variability could be correlated among different traits; for example, one individual that has a thick pallisade layer might also have a thin sponge layer relative to other individuals in the same species. Therefore, we calculated the correlations in within-species variability among individuals and incorporated these into the statistical models along with the within-species variability.

Inferring phylogenetic relationships among closely related plant species is often difficult due to the lack of molecular markers exhibiting enough nucleotide variability. Therefore, we used many amplified fragment length polymorphisms (AFLPs) distributed throughout the whole genome, which proved capable of generating a hypothesis for the phylogenetic relationships among species (Appendix S1). For AFLP analyses, individual plants for each of the nine species (n = 2–11 individuals per species) were collected in the field or obtained from two botanical gardens, located in Kunming and Wenshan, Yunnan, China. We also included *Liriodendron chinense* (n = 3 individuals) as an outgroup to root the tree. A phylogenetic tree was constructed from the AFLP data using maximum likelihood with PAUP 4.0 (Swofford 1999). The analysis did not assume a molecular clock and resulted in a single maximum likelihood tree for the nine *Manglietia* species where the root position was determined by the location of the outgroup *Liriodendron*. The phylogenetic tree for the nine *Manglietia* species obtained using AFLP markers is presented in Fig. 1, and the AFLP data and the measurements of the 13 traits analysed in this article are given in Appendices S2 and S3.

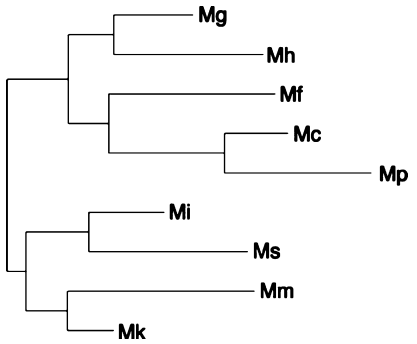

Fig. 1. Hypothesized phylogenetic tree for nine *Manglietia* species based on AFLP markers. Species codes are Mg, *M. grandis*; Mh, *M. hookeri*; Mf, *M. fordiana*; Mc, *M. chingii*; Mp, *M. pachyphylla*; Mi, *M. insignis*; Ms, *M. szechuanica*; Mm, *M. megaphylla*; and Mk, *M. kwangtungensis*. The tree is given in Nexus format with branch lengths in Appendix S2.

PERMUTATION TEST FOR PHYLOGENETIC SIGNAL

A simple permutation test for phylogenetic signal is given by Blomberg & Garland (2002) and Blomberg, Garland & Ives (2003), and here we modify it to test for phylogenetic signal in multiple traits simultaneously. The permutation test is based on the null hypothesis that phylogenetic signal is absent, so that under the null hypothesis trait values can be permuted among species without changing the statistical characteristics of the data. For each permutation data set, the mean squared error (MSE) is calculated under the assumption that phylogenetic signal exists. Specifically, the MSE is calculated assuming that evolution proceeds as a Brownian motion process using the specified hierarchical phylogenetic tree, so that the covariance in trait values between species is proportional to the amount of their shared phylogenetic ancestry (i.e. the branch-length distance from the root of the tree to their last common ancestor); more closely related species have a longer shared ancestry and hence greater predicted covariances in trait values. (The MSE calculated in this way is equivalent to the variance of standardized phylogenetically independent contrasts; Blomberg & Garland 2002; Blomberg, Garland & Ives 2003.) The distribution of MSE values calculated for the permutation data sets is compared with the MSE calculated for the observed data. If phylogenetic signal exists, then the MSE of the observed data will be distinctly low relative to the distribution of permutation MSEs, because incorporating phylogenetic structure (i.e. the assumption of Brownian motion evolution) leads to a better explanation (lower MSE) of the observed data.

Although Blomberg, Garland & Ives (2003) apply this test only to single traits taken separately, it can also be performed on multiple traits using the joint MSE for all traits. In this test, we want to weight each trait equally, so we standardized trait values x to have mean zero and variance 1, $z = \frac{x-\bar{x}}{\mathrm{var}\{x\}^{1/2}}$. In other applications researchers might want to transform some traits, such as log-transform body size, in which case this should be done on the values of x before they are standardized.

PHYLOGENETIC SIGNAL AND TRAIT CORRELATIONS

Here, we describe the estimation of trait correlations with the simultaneous estimation of phylogenetic signal. The utility of this approach is that it leads to a full statistical model of both correlations among traits and phylogenetic correlations among species that can be used to test multiple hypotheses about trait correlations and phylogenetic signal. All of our computations for joint trait analyses were performed by the program CORRELATIONv2.m written in Matlab (Mathworks, Inc. 1996), available on request from T.G.

Our approach for the joint estimation of phylogenetic signal and trait correlations is based on a specific model of evolution under an Ornstein–Uhlenbeck (OU) process (Uhlenbeck & Ornstein 1930; Felsenstein 1988; Hansen 1997; Martins & Hansen 1997; Butler & King 2004; Scales, King & Butler 2009), although other models of evolution could be used (Grafen 1989; Martins & Hansen 1997; Pagel 1997;

Freckleton, Harvey & Pagel 2002; Housworth, Martins & Lynch 2004); for a review, see Lavin *et al*. (2008 Appendix A) and Revell, Harmon & Collar (2008). Our OU model is a multi-trait extension of the single-trait OU process as modelled by Blomberg, Garland & Ives (2003) in which evolution of the traits is correlated. This formulation differs slightly from that proposed by Martins & Hansen (1997), which assumes that the basal trait values are selected from the stationary distribution of an OU process. Instead, we assume that the variances in trait values at the base of the tree are zero, which has the advantage that our model becomes a Brownian motion model of evolution (which is a non-stationary process) as a special case.

To illustrate this model for two traits, suppose that over some arbitrarily small time step Δt, the evolution of trait values $x(t)$ and $y(t)$ for a given species is described by

$$X(t+\Delta t) = d_x X(t) + \gamma_x(t)$$
$$Y(t+\Delta t) = d_y Y(t) + \gamma_y(t) \qquad \text{eqn 1}$$

where d_x and d_y measure the strength of stabilizing selection for traits x and y, and $\gamma_x(t)$ and $\gamma_y(t)$ are random variables with means zero, variances σ_x^2 and σ_y^2, and correlation r. Mapping this model onto a phylogenetic tree (under the assumption that the variance in trait values at the base of the tree is zero), the covariance in trait x between species i and j is

$$\mathrm{cov}\{X_i, X_j\} = d_x^{\tau_i+\tau_j}\frac{1-d_x^{2\tau_{ij}}}{1-d_x^2}\sigma_x^2 \qquad \text{eqn 2}$$

where X_i and X_j are the values of trait x for species i and j, τ_i and τ_j are the distances from the node representing their most recent common ancestor to the tip of the tree for species i and j, and τ_{ij} is the branch length from the base to the most recent common ancestor (Blomberg, Garland & Ives 2003). An identical expression holds for the covariance in values of trait y among species. Using a similar approach, it can be shown that the covariance between trait x for species i and trait y for species j is given by

$$\mathrm{cov}\{X_i, Y_j\} = d_x^{\tau_i} d_y^{\tau_j}\frac{1-(d_x d_y)^{\tau_{ij}}}{1-d_x d_y} r\sigma_x\sigma_y \qquad \text{eqn 3}$$

Thus, from the model of trait evolution (eqn 1) and a phylogenetic tree with N species, we can construct the variances and covariances in trait values among all species that incorporate both correlations between traits r and the strength of phylogenetic signal d. The stronger the stabilizing selection towards an optimum for a trait (the larger the phylogenetic signal parameter d), the weaker the phylogenetic correlation between species for that trait, because phylogenetic history is erased by selection towards the optimum (Felsenstein 1985; Blomberg, Garland & Ives 2003); a value of $d = 0$ corresponds to no phylogenetic signal (trait values among species are independent), and a value of $d = 1$ corresponds to Brownian motion evolution.

For use in statistical analyses, these variances (eqn 2) and covariances (eqn 3) must be combined into a covariance matrix. For illustrative purposes, consider hypothetical example of three species and two traits (Fig. 2). This leads to a 3×3 matrix $\mathbf{C}$ whose diagonal elements τ_i contain the distance from the base to the tip for species i, and whose off-diagonal elements τ_{ij} give the shared distance on the phylogenetic tree between species i and j. In our example, we have standardized the branch lengths so that $\tau_i = 1$ (although the methods do not require that phylogenetic tree be ultra-metric, that is, have contemporaneous tips), and the shared branch between species 1 and 2 has length $\tau_{ij} = 0{\cdot}7$ (Fig. 2a). In the statistical analysis, both the trait correlations r and the strength of phylogenetic signal d are estimated; for our example, there is a single value of r and two values of d, one for each trait. The effect of d on the strength of phylogenetic correlations can be visualized using a phylogenetic tree, but now with branch lengths proportional to the estimated phylogenetic signal. The value of phylogenetic signal for trait x, d_x, is 2, leading to a covariance of 0·8 between species, while for trait y, $d_y = 0{\cdot}25$ and the covariance between species is 0·4; these covariances are contained within the covariance matrices $\mathbf{C}_x(d_x)$ and $\mathbf{C}_y(d_y)$ that are visualized by either extending (trait x) or contracting (trait y) the shared branch lengths between species (Fig. 2b). With $d_x = 2$ and $d_y = 0{\cdot}25$, the correlation between traits is $r = 0{\cdot}5$ and the variances for traits 1 and 2 are $\sigma_x^2 = 1{\cdot}4$ and $\sigma_y^2 = 18{\cdot}9$, resulting in the joint covariance matrix $\sigma^2\mathbf{\Psi}$ (Fig. 2c). Here, we use the standard convention of scaling the covariance matrix by σ^2 that gives the overall variance of the data (e.g. Judge *et al*. 1985).

The two blocks of the covariance matrix $\sigma^2\mathbf{\Psi}$ along the diagonal (shaded) correspond to the covariance matrices for traits x and y; each matrix gives the variances for each of the three species along the diagonal and the phylogenetic covariances in the off-diagonals. The two blocks of matrix $\sigma^2\mathbf{\Psi}$ on the off-diagonals (not shaded) give the covariances among species between traits x and y. For example, the value 2·7 is the covariance between traits x and y expressed by species

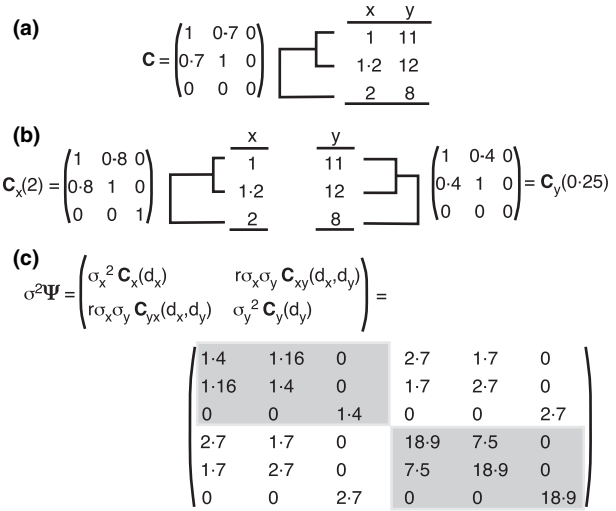

Fig. 2. Construction of the joint trait-phylogeny covariance matrix that incorporates correlations among traits and correlations among phylogenetically related species. (a) The base information provided for a three-species data set with two traits x and y. The covariance matrix $\mathbf{C}$ describes the phylogeny; the covariance in trait values between species is proportional to the shared branch lengths. (b) Strengths of phylogenetic signal when the values of $d_x = 2$ and $d_y = 0{\cdot}25$ are estimated from the data. (Note that these values of d are hypothetical and in fact values of d cannot be estimated for data sets with fewer than four species.) The transformed matrices $\mathbf{C}_x$ and $\mathbf{C}_y$ incorporate the strengths of phylogeny signal which are depicted by the phylogenetic trees with transformed branch lengths. (c) The overall trait–species covariance matrix, $\sigma^2\mathbf{\Psi}$. The 3×3 submatrices along the diagonal of $\sigma^2\mathbf{\Psi}$ (shaded) give the covariance matrices containing correlations in values of a single trait among species. The 3×3 submatrices on the off-diagonal of $\sigma^2\mathbf{\Psi}$ give the covariance matrices containing correlations between traits x and y. The terms σ_x^2 and σ_y^2 give the variances in traits x and y, and r is the correlation between them.

1, while the value 1·7 is the covariance between trait x for species 1 and trait y for species 2. Thus, the matrix $\sigma^2\mathbf{\Psi}$ gives covariances between all trait–species combinations.

The statistical method we developed estimates the components of the joint covariance matrix, specifically the parameters r, d and σ^2, using Restricted Maximum Likelihood estimation; the details are presented in Appendix S4. The covariance matrix $\sigma^2\mathbf{\Psi}$ can also be modified to incorporate measurement error. We treat as measurement error any within-species variation that might include error in the measurement of the trait or variation among individuals of the same species. (Population differences would also contribute to this error if they had been pooled to obtain a single sample to represent a species.) This measurement error is captured by the standard error of the within-species estimates of trait values (Ives, Midford & Garland 2007). To determine whether the estimates of d differ from zero or one, we used likelihood ratio tests. Specifically, we calculated the maximum restricted log likelihoods both including (LL) and excluding (LL_0) the parameters of interest, and tested the significance of the inclusion of the parameters using the result that $-2(LL_0 - LL)$ is asymptotically χ^2-distributed with degrees of freedom equal to the number of parameters differing between models; for example, the test for $d \neq 0$ for three traits involves comparing the models with and without the three parameters d, and the resulting chi-squared distribution for $-2(LL_0 - LL)$ has 3 degrees of freedom.

In addition to estimating trait and phylogenetic correlations simultaneously for groups of traits, giving estimates we refer to as r_{joint} and d_{joint}, we modified the methods to calculate the value of d assuming that all species within the group have the same value d_{common} and estimate the corresponding trait correlations r_{comon}. This provides a test for dissimilarity in phylogenetic signal among traits by comparing the fits of the model using d_{joint} and d_{common}. To compare with the jointly estimated values d_{joint}, we also computed the values of d estimated for each trait separately, d_{sep}; we do not calculate the trait correlations r along with d_{sep}, because the correlations are components of the covariance matrix $\sigma^2\mathbf{\Psi}$ for which values of d for each species should be estimated simultaneously (i.e. d_{joint}).

STATISTICAL POWER TO IDENTIFY PHYLOGENETIC SIGNAL

To compare the statistical power of the two methods (joint permutation test, likelihood ratio test) to detect phylogenetic signal in multiple traits, we performed a simulation study for three traits using the phylogenetic tree of the nine *Manglietia* species (Fig. 1). We simulated data sets under the assumption that evolution of the three traits followed an OU process with the same 'true' value of d for all traits. We assumed traits were not correlated ($r = 0$), although simulations with $r \neq 0$ gave similar results (data not presented). For each value of d between 0 and 1 in increments of 0·1, we simulated 2000 data sets, and for each simulated data set we performed the joint permutation test and the likelihood ratio test (obtained from the estimation of d_{joint}); we used an $\alpha = 0{\cdot}05$ level to identify statistically significant phylogenetic signal.

STATISTICAL PROPERTIES OF ESTIMATORS OF *D* AND *R*

We investigated the statistical properties of the estimators of phylogenetic signal, d, and trait correlations, r, using simulations assuming that there are three traits under study for nine species with the same phylogeny as the nine *Manglietia* species (Fig. 1). To address the issue of sample size, we also considered the case of 49 species using the phylogenetic tree provided by Garland *et al.* (1993). For both simulations, we assumed values of d are 0·2, 1 and 1·5, and values of r are 0·5, −0·5 and −0·25. Finally, we simulated data both without and with measurement error; measurement error for the nine-species simulations were the same as for the nine *Manglietia* species, while for the 49-species simulations we assumed that the measurement error for each trait and the covariances between measurement errors were the same for all species and equal to the mean measurement errors for the corresponding traits in the nine *Manglietia* species data set. For the simulations with measurement error, we estimated d_{joint} and d_{sep} both incorporating and not incorporating measurement error.

Results

PHYLOGENETIC SIGNAL

The four different groups of traits (photosynthesis, leaf morphology, plant growth and thermal tolerance) demonstrated contrasting patterns of phylogenetic signal. The photosynthesis traits showed very strong phylogenetic signal (Table 1). For two of the three traits (A_{max} and LSP), the single-trait permutation tests were statistically significant and the estimates of d_{sep} for each trait analysed separately were greater than 1. The estimates of d_{joint} gave a similar pattern, and the joint permutation test was statistically significant. Finally, under the assumption that all three traits have the same value of d, the estimate of d_{common} was greater than 1 both with and without the inclusion of measurement error.

We used likelihood ratios to test three null hypotheses regarding the estimation of d_{joint} (Table 2): (i) whether there is phylogenetic signal in the group of traits (H_0:no signal), (ii) whether the species covariances differ from those anticipated under Brownian motion evolution (H_0:Brownian), and (iii) whether there is heterogeneity in values of d_{joint} among species (H_0:$d_{joint} = d_{common}$). For the three photosynthesis traits, only the first of these hypotheses was rejected. We conclude that the phylogenetic signal is statistically indistinguishable from Brownian motion evolution, and that all three traits show indistinguishable degrees of phylogenetic signal.

For the three leaf morphology traits, the estimated values of d_{sep} and d_{joint} are large, yet they nonetheless cannot be statistically distinguished from zero; the separate likelihood ratio tests on d_{sep} (Table 1) and the joint likelihood ratio test for d_{joint} (Table 2) are not statistically significant. Nonetheless, the joint permutation test showed statistically significant phylogenetic signal (Table 1, $P < 0{\cdot}03$). This illustrates that the tests using d and the permutation test do not necessarily give the same results for any given data set (see also Blomberg, Garland & Ives 2003). Below we demonstrate that the joint permutation test has greater statistical power than the likelihood ratio test for d_{joint}. Thus, it is not surprising that the permutation test identifies statistically significant phylogenetic signal while the likelihood ratio test does not. Despite the fact that the null hypothesis of no phylogenetic signal was rejected by the permutation test, we nonetheless rely solely

Table 1. Measures of phylogenetic signal for 13 plant traits. The permutation test gives a nonparametric assessment (*P*-value) for statistical departure from the case in which species are phylogenetically independent. Estimates of phylogenetic signal are given by the parameter d of an Ornstein–Uhlenbeck model of stabilizing selection, which is estimated for traits separately, d_{sep}, simultaneously for all traits within a group, d_{joint}, and simultaneously for all traits in a group under the assumption that all traits share the same value, d_{common}. Likelihood ratio tests were used to test the significant departure of d_{sep} from zero (†$P < 0{\cdot}1$; *$P < 0{\cdot}05$; **$P < 0{\cdot}02$), using likelihoods LL_{sep} and LL_0 calculated from the model including and excluding d_{sep}. The estimates of leaf morphology traits were taken from the literature and do not have associated measurement errors, and therefore values of d were not estimated with measurement error (shown by dashes in the table)

	Permutation test (*P*-values)		d without ME					d with ME				
Trait	Single	Joint	d_{sep}	LL_{sep}	LL_0	d_{joint}	d_{common}	d_{sep}	LL_{sep}	LL_0	d_{joint}	d_{common}
Photosynthesis												
A_{max}	0·03	0·03	4·55**	−3·69	−6·44	2·30	1·44	4·44†	−4·90	−6·67	1·82	1·67
LSP	0·02		2·66*	−38·29	−40·55	2·34		3·22*	−43·12	−45·25	3·93	
LCP	0·52		0·16	−8·54	−8·56	0·12		0·12	−9·59	−9·59	0·01	
Leaf morphology												
Thickness of leaves	0·21	0·04	2·65	−32·65	−33·01	0·74	0·26	–	–	–	–	–
Thickness of palisade tissue	0·14		1·98	−31·04	−31·45	1·36		–	–	–	–	
Thickness of sponge tissue	0·14		1·85	−34·00	−34·30	0·23		–	–	–	–	
Plant growth												
Basal stem diameter	0·36	0·59	0·24	−0·47	−0·48	0·14	1·52	0·47	−0·95	−0·97	1·51	0·83
Crown volume	0·66		0·11	−86·89	−86·89	1·39		0·11	−96·80	−96·80	0	
Leaf area	0·86		0·10	−48·29	−48·29	1·95		0·00	−52·92	−52·92	0·01	
Relative growth rate	0·28		0·44	−3·15	−3·24	1·97		0·46	−3·31	−3·37	0	
Thermal tolerance												
T_{ch}	0·74	0·88	0·08	−7·32	−7·32	0	0·00	0·00	−7·48	−7·48	0·30	0
T_{max}	0·96		0·41	−3·59	−3·59	0		0·00	−3·26	−3·26	0·04	
T_{50}	0·70		0·05	−5·25	−5·25	0·10		0·00	−5·25	−5·25	0·16	

Table 2. Joint tests for statistical significance of phylogenetic signal for four categories of traits. H_0:no signal is the null hypothesis that there is no phylogenetic signal ($d = 0$ for all traits). H_0:Brownian is the null hypothesis that the phylogenetic signal can be described by a Brownian motion model of evolution ($d = 1$ for all traits). H_0:$d = d_{common}$ is the null hypothesis that the strength of phylogenetic signal is the same for all traits within a category. ME denotes measurement error and only *P*-values $< 0{\cdot}10$ are shown. Degrees of freedom (d.f.) for the chi-squared tests equal the number of parameters that differ between models; for H_0:no signal and H_0:Brownian, d.f. = the number of traits per group, and for H_0:$d = d_{common}$, d.f. = the number of traits − 1. Leaf morphology traits do not have associated measurement errors, and therefore models incorporating measurement error were not used

	Without ME						With ME					
	H_0:no signal		H_0:Brownian		H_0: $d = d_{common}$		H_0:no signal		H_0:Brownian		H_0: $d = d_{common}$	
Trait category	χ^2	*P*	χ^2	*P*	χ^2	*P*	χ^2	*P*	χ^2	*P*	χ^2	*P*
Photosynthesis	11·2	0·01	5·8		5·3		12·7	0·005	8·1	0·04	5·3	
Leaf morphology	2·34		2·57		2·17		–		–		–	
Plant growth	12·6	0·01	13·5	0·009	8·3	0·08	15·3	0·004	14·3	0·006	9·0	0·06
Thermal tolerance	0·14		6·9	0·015	0·14		3·3		6·0		3·3	

on the likelihood ratio test for reasons described below, concluding that there is no phylogenetic signal in leaf morphology traits.

The four traits for plant growth give an example in which the analyses of single traits revealed little evidence of phylogenetic signal, yet phylogenetic signal was found for the group of traits analysed jointly. Specifically, the null hypothesis H_0:no signal was rejected whether measurement error was not ($P < 0{\cdot}01$) or was ($P < 0{\cdot}004$, Table 2) included. Moreover, the null hypothesis H_0:Brownian is also rejected ($P < 0{\cdot}009$ and $< 0{\cdot}006$ with and without measurement error), indicating that even though there is evidence for phylogenetic signal, the Brown motion evolutionary model is also not supported. The nearly significant rejection of the null hypothesis H_0:$d = d_{common}$ ($P < 0{\cdot}08$ and $< 0{\cdot}06$ with and without measurement error) suggests that traits differ in whether or not they show phylogenetic signal, which is supported by the very different estimates among species of both d_{sep} and d_{joint} (Table 1). Note that in contrast to the leaf morphology traits, although the likelihood ratio tests identify joint phylogenetic signal, the permutation test does not. We return to this in the Discussion.

For the fourth category of traits (thermal tolerance), there is little indication of phylogenetic signal in any analysis.

STATISTICAL POWER TO IDENTIFY PHYLOGENETIC SIGNAL

In our simulation study, the joint permutation test had greater statistical power than the likelihood ratio test (Fig. 3), showing greater ability to reject the null hypothesis of no phylogenetic signal than the likelihood ratio test for $d > 0$. Furthermore, the likelihood ratio test had a slightly inflated type I error rate (probability of rejecting the null hypothesis when it is true), with 6·6% of the data sets rejected at the $\alpha = 0{\cdot}05$ level when $d = 0$; this compared to a type I error rate of 5·6% for the permutation test. Importantly, only 2·0% of the simulated data sets were rejected by both the joint permutation test and the likelihood ratio test. Therefore, if one were to apply both tests to the same data set and conclude that the data set had phylogenetic signal if either one or the other of the tests rejected the null hypothesis, then the type I error rate would be 10·2%; the type I error rates for the permutation and likelihood ratio tests were 5·6% and 6·6%, so the probability of one or the other test rejecting the null hypothesis is 5·6% + 6·6% − 2·0% = 10·2%. This suggests that it is necessary to make an *a priori* decision about which statistical test to use – and use only that one – in order to obtain the correct type I error rates.

TRAIT CORRELATIONS

We computed trait correlations within each of the four groups. Table 3 gives four types of correlation coefficients: (i) Pearson correlation coefficients obtained under the assumption that there is no phylogenetic signal; (ii) coefficients calculated under the assumption of Brownian motion evolution; (iii) coefficients calculated at the same time as d_{joint} for all three traits, r_{joint}; and (iv) r_{joint} calculated while incorporating measurement error. For brevity, we do not report correlations r_{common} obtained when assuming all species have a common value of d.

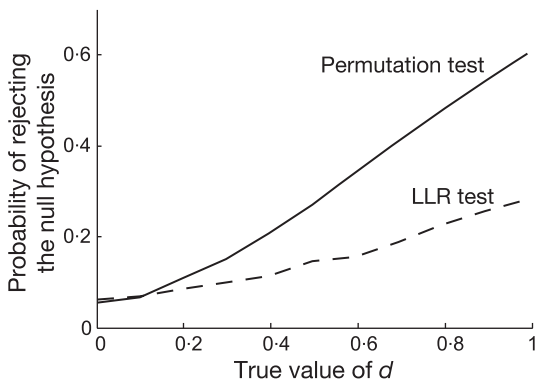

Fig. 3. Power analysis for the joint permutation and likelihood ratio tests for three traits simulated using the phylogenetic tree of nine *Manglietia* species (Fig. 1). Each trait was simulated using an Ornstein–Uhlenbeck evolutionary process with a common value of d, and traits were assumed to be independent. For values of d between zero and one in increments of 0·1, 2000 simulations were performed, and the figure gives the fraction for which the null hypothesis of $d = 0$ was rejected using the joint permutation test and likelihood ratio test at an $\alpha = 0{\cdot}05$ level.

For all groups of traits, the values of the correlation coefficients computed using all four methods exhibit general correspondence (Table 3); this empirical correspondence between standard and phylogenetic correlations has been shown previously (Ricklefs & Starck 1996; Ackerly & Reich 1999). Furthermore, there was statistically significant correlation between traits within all four groups; all likelihood ratio tests of the hypothesis that all correlations are zero were rejected using all four of the approaches for measuring correlation coefficients (Table 3). Comparing statistical tests among the four approaches used for estimating correlation coefficients, the most striking result is the stronger rejections (lower P-values) of the null hypotheses when the analyses incorporated measurement error; for the three groups for which measurement error was available (photosynthesis, leaf morphology and thermal tolerance), the null hypotheses of zero correlations were rejected at the $P < 0{\cdot}03$, 0·01 and 0·001 levels using Pearson correlations (assuming no measurement error and no phylogenetic signal), whereas all these hypotheses were rejected at the $P < 0{\cdot}001$ level when estimating r_{joint} with measurement error (Table 3).

STATISTICAL PROPERTIES OF *d* AND *r*

To explore the statistical properties of the joint estimation of d and r, we performed a simulation study using phylogenetic trees for nine species (Fig. 1) and 49 species (Garland *et al.* 1993). For simulations without measurement error, the estimates of d_{joint} were less biased and more precise (i.e. less variable) than the estimates calculated for each species separately, d_{sep}, when there were nine species (Fig. 4a). Increasing the sample size to 49 removed any bias for both methods, and the estimates of d_{joint} are only very slightly more precise (Fig. 4b). Thus, for small sample sizes the estimates of d_{joint} are preferred over d_{sep}, even without the added advantage of d_{joint} that it provides statistical tests of phylogenetic signal for groups of traits. (We should of course note that this is true only for the specific simulations with the specific trait correlations r that we investigated, and that researchers might want to perform simulations based on their own data to confirm this. Nonetheless, we suspect this pattern will hold broadly.) In simulations including measurement error (Fig. 4c,d), the estimates of d that ignored the simulated measurement error were more biased than those incorporating measurement error, especially for the case with 49 species (Fig. 4d). Furthermore, the estimates of d_{joint} were more precise than those for d_{sep} for small sample sizes (Fig. 4c). Thus, for data sets with substantial measurement error, estimating d_{joint} with measurement error incorporated has the best statistical properties.

In the same simulations we used to estimate values of d (Fig. 4), we also obtained estimates of r (Fig. 5); we did not

	P-value	Trait pairs					
		1–2	1–3	2–3	1–4	2–4	3–4
(a) Photosynthesis traits							
No signal	0·04	0·63	−0·63	−0·30			
Brownian	0·005	0·71	−0·49	−0·18			
r_{joint}	0·03	0·62	−0·56	−0·26			
r_{joint} with ME	0·001	0·71	−0·52	0			
Ave. d_{joint} with ME		2·82	0·61	1·97			
Ave. d_{joint}		2·32	1·21	1·23			
(b) Leaf morphology traits							
No signal	0·002	0·34	0·61	−0·42			
Brownian	0·002	0·30	0·56	−0·47			
r_{joint}	0·003	0·25	0·52	−0·55			
Ave. d_{joint}		1·04	0·47	0·80			
(c) Growth traits							
No signal	0·01	0·30	0·29	0·37	0·84	0·65	0·75
Brownian	0·01	0·32	0·24	0·40	0·92	0·60	0·74
r_{joint}	0·001	0·35	0·29	0·92	0·45	0·63	0·83
r_{joint} with ME	0·001	0·26	−0·02	0·36	0·32	0·83	0·54
Ave. d_{joint} with ME		0·75	0·76	0·01	0·75	0·00	0·01
Ave. d_{joint}		0·77	1·05	1·67	1·06	1·68	1·96
(d) Thermal tolerance traits							
No signal	0·001	0·80	0·85	0·64			
Brownian	0·01	0·82	0·86	0·68			
r_{joint}	0·001	0·80	0·85	0·64			
r_{joint} with ME	0·001	0·50	0·99	0·52			
Ave. d_{joint}		0	0·05	0·05			
Ave. d_{joint} with ME		0·17	0·23	0·10			

Table 3. Correlation coefficients calculated using four methods. The 'no signal' and 'Brownian motion' calculations assume that there is no phylogenetic signal ($d = 0$) and signal given by Brownian motion evolution along the original specified phylogenetic tree ($d = 1$). Coefficients were also calculated using phylogenetic signal estimated simultaneously for all traits, d_{joint}, with and without incorporating measurement error. When d was estimated, the average of the estimates of d is provided for each pair of species. The traits are listed in order: (a) photosynthesis traits (A_{max}, LSP, LCP), (b) leaf morphology traits (thickness of leaves, thickness of palisade tissue, and thickness of sponge tissue), (c) growth traits (basal stem diameter, crown volume, leaf area, relative growth rate) and (d) thermal tolerance traits (critical temperature T_{ch}, peak temperature T_{max}, temperature half-inactivation T_{50}). *P*-values correspond to the likelihood ratio test for the null hypothesis H_0: $r = 0$ for all species pairs

include estimates of r when estimating d_{sep} for reasons described in Material and methods. In simulations not including measurement error (Fig. 5a,b), the estimates of r under the assumption of no phylogenetic signal or Brownian motion evolution were both biased for the larger sample size of 49 species, whereas the estimates of r_{joint} were not. In simulations with measurement error (Fig. 5c,d), with 49 species the estimates using all methods were biased except for the joint estimate of r_{joint} incorporating measurement error. These results illustrate the relatively good statistical performance of estimators of d_{joint} and r_{joint} that simultaneously include phylogenetic signal and trait correlations.

Discussion

Our statistical methods to measure phylogenetic signal in multiple traits simultaneously revealed statistically significant signal in the group of three traits involved in photosynthesis (A_{max}, LSP and LCP) and the group of four traits involving plant growth (basal stem diameter, crown volume, leaf area and relative growth rate). A third group of three traits involving leaf morphology showed phylogenetic signal in the permutation test but not in the likelihood ratio tests; estimates of measurement error were not available, so our method for incorporating measurement error could not be applied. Finally, the fourth group of thermal tolerance traits exhibited no phylogenetic signal. These results from analyses that consider traits jointly within the same group contrast the conclusions obtained when analysing the 13 traits separately. Analysed separately, we found strong phylogenetic signal in only two traits, A_{max} and LSP, both of which involve photosynthesis. Failure to detect phylogenetic signal in traits treated separately is not surprising, however, because statistical tests of phylogenetic signal often lack power when sample sizes are small (Blomberg, Garland & Ives 2003).

For the groups of traits showing statistically significant phylogenetic signal, all we conclude is that phylogenetically related species are more likely to share the same values of the group of traits. Although our methods rely on a specific model of evolution, assuming trait values evolve according to an OU process of stabilizing selection across the phylogeny (eqn 1), it would be an over-interpretation of the results to conclude that the data were produced under an OU process. A formal statistical test of the specific model of evolution would consist of constructing competing models derived under different evolutionary processes and then selecting the best amongst them; although this procedure is straightforward, we suspect that large amounts of high-quality data would be required to statistically distinguish among models. Our goal, however, was only to identify the existence of phylogenetic signal, not to test any specific mechanisms that underlie phylogenetic patterns (cf. Blomberg & Garland 2002; Blomberg, Garland & Ives 2003).

Previous studies of plants have documented cases of both strong phylogenetic signal and weak or absent signal. Ackerly & Reich (1999) assembled a data set of 108 tree species for which eight leaf traits were available. All of the traits showed phylogenetic signal, although this was largely due to the data set spanning angiosperms and gymnosperms, two clades with distinct leaf morphology (broad leaves vs. needles); in similar

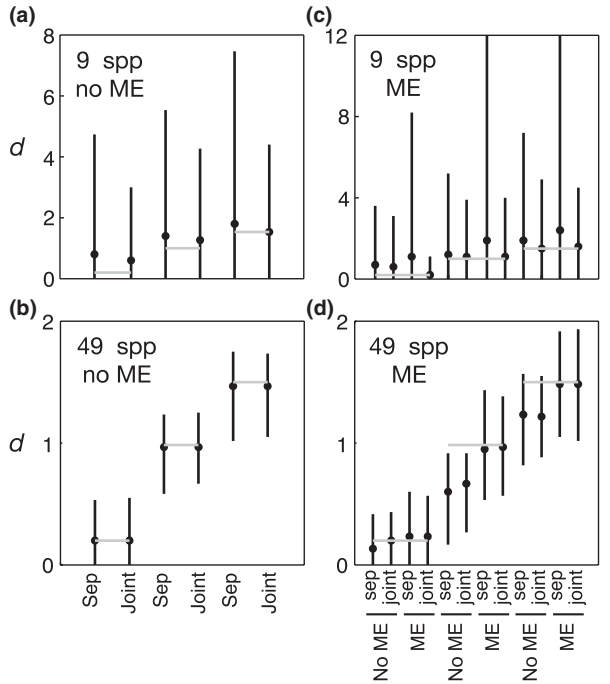

Fig. 4. Estimates of phylogenetic signal, d, for simulated data for three traits. In (a) and (b) data were simulated without measurement error, and d_{sep} and d_{joint} were estimated excluding measurement error. In (c) and (d) measurement error was simulated, and both d_{sep} and d_{joint} were estimated including and not including measurement error. In (a) and (c) the phylogenetic tree for the nine *Manglietia* species was used for simulations, and in (b) and (d) the tree for 49 species from Garland *et al.* (1993) was used. Points give the mean of the estimates from simulations, vertical lines give the 95% inclusion intervals and horizontal lines give the true values of d for the three traits used in the simulations: 0·2, 1 and 1·5. Values of $r = 0{\cdot}5$, $-0{\cdot}5$ and $-0{\cdot}25$. For (a) and (c) 1000 data sets were simulated, and for (b) and (d) 200 data sets were simulated.

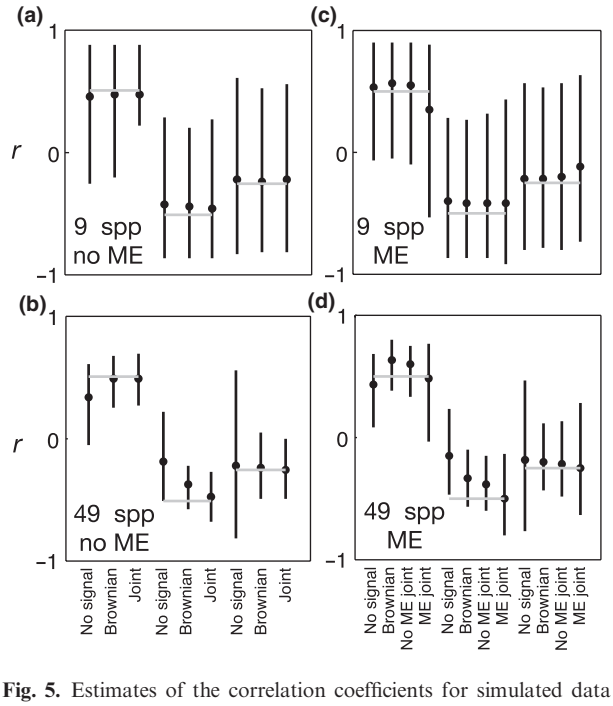

Fig. 5. Estimates of the correlation coefficients for simulated data sets for three traits. The same data sets were used as those producing corresponding panels in Fig. 4. In (a) and (b) data were simulated without measurement error, and d_{sep} and d_{joint} were estimated excluding measurement error. The correlation coefficients were then estimated assuming no phylogenetic signal, Brownian motion evolution and phylogenetic signals given by d_{joint}. In (c) and (d) measurement error was simulated, and d_{joint} was estimated including and not including measurement error. In (a) and (c) the phylogenetic tree for the nine *Manglietia* species was used for simulations, and in (b) and (d) the tree for 49 species from Garland *et al.* (1993) was used. Points give the mean of the estimates from simulations, vertical lines give the 95% inclusion intervals and horizontal lines give the true values of r for the three traits (0·5, $-0{\cdot}5$ and $-0{\cdot}25$). For (a) and (c) 1000 data sets were simulated, and for (b) and (d) 200 data sets were simulated.

analyses confined to angiosperms, they found little phylogenetic signal in the same traits. Zanne, Chapman & Kitajima (2005) found strong phylogenetic signal in cotyledon type but not seed mass among 70 species of trees and shrubs occupying the same habitat in Uganda. In a survey of pollination traits among 288 species, Ornelas *et al.* (2007) found phylogenetic signal in all traits involving nectar production, as well as corolla length, suggesting strong evolutionary conservatism in plant reproductive biology.

These studies were broad taxonomic surveys, however, involving diverse plant lineages. In contrast, our study compared a small group of congeners. At a similar taxonomic scale, Blomberg, Garland & Ives (2003) analysed two plant data sets consisting of morphological traits for maple trees (Ackerly & Donoghue 1998) and *Tithonia* (Asteraceae) (Morales 2000). Despite the larger size of these data sets (17 and 32 species vs. our 9), phylogenetic signal was found for only about one-third of traits (4 of 12 and 5 of 14 traits respectively) based on the same permutation test that we employed here (with P-values not corrected for multiple comparisons). In a common-garden experiment on drought tolerance of eight species from five genera, Valladares & Sanchez-Gomez (2006) documented phylogenetic signal in several traits (as determined by clustering of traits in a Principle Components Analysis), with similar drought tolerance exhibited by members of the same genera. In comparison with the taxonomically broad surveys, these studies suggest that phylogenetic signal in plant traits is harder to identify among taxonomically similar species. However, this may also be the result of sample size, as confining analyses to closely related species often limits the number of species that can be analysed, thus reducing statistical power.

Our statistical methods provide a means for estimating simultaneously the correlations of multiple traits and the phylogenetic correlations in values of traits among species. In a simulation study, we showed that correlations calculated when either ignoring phylogenetic signal or assuming phylogenetic signal generated a Brownian motion model of evolution often produced biased estimates of the true correlation (see also Rohlf 2006). In contrast, estimating trait correlations while simultaneously estimating phylogenetic signal, r_{joint}, gave less-biased estimates. This demonstrates that correctly accounting for phylogenetic signal, rather than making an *a priori* assumption that it is zero or given

by Brownian motion evolution, leads to better estimates of correlations.

Although we present two tests for phylogenetic signal in suites of traits, the joint permutation test and the likelihood ratio test using estimates of d_{joint}, only one of these tests should be selected *a priori* and used to determine statistical significance. This is because the two tests do not necessarily identify the same data sets as showing phylogenetic signal in suites of traits (see also Blomberg, Garland & Ives 2003). To give an extreme example, if both tests had rejection rates of the null hypothesis of 5%, yet data sets rejected by one test were never rejected by the other, then 10% of the data sets would be rejected by one or the other test. Therefore, if a researcher were to assume that a data set showed phylogenetic signal for a suite of traits if one or the other statistical test gave statistical significance, then in this example the type I error rate (the rate of rejecting a null hypothesis that is true) would be inflated to 10%. In our simulations, only 2·0% of simulated data sets had null hypotheses that were rejected by both tests, indicating potentially serious type I errors if both tests were applied. The underlying explanation for this statistical complexity is that the permutation and likelihood ratio tests are rejecting the null hypothesis based on different characteristics of the data sets; they are employing different alternative (H_1) hypotheses.

Given that either the permutation test or the likelihood ratio test should be selected *a priori* as the group test for phylogenetic signal, which one should it be? The joint permutation test has the advantage of greater power, and therefore it might be preferred if detecting phylogenetic signal is the only objective. The likelihood ratio test has the advantage of being a model-based test that gives more information about the data set, including an estimate of the strength of phylogenetic signal and phylogenetically correct correlations among traits. In our study, we were interested in not only phylogenetic signal but also trait correlations, and therefore we would pick the likelihood ratio test over the permutation test. Thus, although the permutation test shows phylogenetic signal for the group of leaf morphology traits, we do not conclude that they in fact exhibit phylogenetic signal.

The product of the estimation of d_{joint} and r_{joint} is the joint covariance matrix $\sigma^2\mathbf{\Psi}$ that incorporates correlations due to phylogenetic relationships among species, associations between traits and measurement error (Fig. 1, Appendix S1). This covariance matrix can be best explained by considering each value of a trait for each species as a trait–species datum. The elements of the covariance matrix are then the covariances between each pair of trait–species values. In the structure assumed for the covariance matrix $\sigma^2\mathbf{\Psi}$, trait–species values might be correlated if species are phylogenetically related. They might also be correlated if the two traits involved in the trait–species values have experienced correlated evolution along the phylogenetic tree. Although the structure we assume for the covariance matrix is flexible to include different strengths of phylogenetic signal and different trait correlations, it nonetheless imposes constraints on the covariances. For example, it assumes that closely related species are at least as likely to share similar trait values as more distantly related species. This assumption might be violated, however, if there is convergent evolution, in which distantly related species have become similar in multiple traits. Although the structure of the covariance matrix $\sigma^2\mathbf{\Psi}$ cannot accommodate convergent evolution, it can nonetheless be used to provide a null hypothesis that no convergence exists. This could be done, for example, by fitting the covariance matrix $\sigma^2\mathbf{\Psi}$ to a data set, using the fitted model to simulate data, and then looking for patterns of convergence in the observed data that are not found in the simulated data; methods to do this, however, will require further development. In general, by fitting an explicit model to the covariance structure of a data set, our approach provides an avenue to investigate complex evolutionary hypotheses about the variation in multiple traits among phylogenetically related species.

Acknowledgements

We thank especially Stacey Smith for comments and help with this project. It was supported in part by NSF grants DEB-0816613 to ARI and DEB-0416085 to D.N. Reznick, M.S. Springer and T.G.

References

Ackerly, D.D. (1999) Comparative plant ecology and the role of phylogenetic information. *Physiological Plant Ecology* (eds M. Press, J. Scholes & M. Barker), pp. 391–413. Blackwell Science, Oxford.

Ackerly, D.D. & Donoghue, M.J. (1998) Leaf size, sapling allometry, and Corner's rules: phylogeny and correlated evolution in maples (Acer). *The American Naturalist*, **152**, 767–791.

Ackerly, D.D. & Reich, P.B. (1999) Convergence and correlations among leaf size and function in seed plants: a comparative test using independent contrasts. *American Journal of Botany*, **86**, 1272–1281.

Arnold, S.J., Burger, R., Hohenlohe, P.A., Ajie, B.C. & Jones, A.G. (2008) Understanding the evolution and stability of the G-matrix. *Evolution*, **62**, 2451–2461.

Blomberg, S.P. & Garland, T., Jr (2002) Tempo and mode in evolution: phylogenetic inertia, adaptation and comparative methods. *Journal of Evolutionary Biology*, **15**, 899–910.

Blomberg, S.P., Garland, T., Jr & Ives, A.R. (2003) Testing for phylogenetic signal in comparative data: behavioral traits are more labile. *Evolution*, **57**, 717–745.

Butler, M.A. & King, A.A. (2004) Phylogenetic comparative analysis: a modeling approach for adaptive evolution. *The American Naturalist*, **164**, 683–695.

Butler, M.A. & King, A.A. (2005) Ouch! An improved method for phylogenetic tests of adaptive evolution. *Integrative and Comparative Biology*, **45**, 973–973.

Clobert, J., Garland, T., Jr & Barbault, R. (1998) The evolution of demographic tactics in lizards: a test of some hypotheses concerning life history evolution. *Journal of Evolutionary Biology*, **11**, 329–364.

Diaz-Uriarte, R. & Garland, T., Jr (1996) Testing hypotheses of correlated evolution using phylogenetically independent contrasts: sensitivity to deviations from Brownian motion. *Systematic Biology*, **45**, 27–47.

Dohm, M.R. & Garland, T., Jr (1993) Quantitative genetics of scale counts in the garter snake *Thamnophis sirtalis*. *Copeia*, **4**, 987–1002.

Felsenstein, J. (1985) Phylogenies and the comparative method. *The American Naturalist*, **125**, 1–15.

Felsenstein, J. (1988) Phylogenies and quantitative characters. *Annual Review of Ecology and Systematics*, **19**, 445–471.

Freckleton, R.P., Harvey, P.H. & Pagel, M. (2002) Phylogenetic analysis and comparative data: a test and review of evidence. *The American Naturalist*, **160**, 712–726.

Garland, T., Jr, Bennett, A.F. & Rezende, E.L. (2005) Phylogenetic approaches in comparative physiology. *Journal of Experimental Biology*, **208**, 3015–3035.

Garland, T., Jr & Díaz-Uriarte, R. (1999) Polytomies and phylogenetically independent contrasts: an examination of the bounded degrees of freedom approach. *Systematic Biology*, **48**, 547–558.
Garland, T., Jr, Dickerman, A.W., Janis, C.M. & Jones, J.A. (1993) Phylogenetic analysis of covariance by computer-simulation. *Systematic Biology*, **42**, 265–292.
Grafen, A. (1989) The phylogenetic regression. *Transactions of the Royal Society of London B, Biological Sciences*, **326**, 119–157.
Hansen, T.F. (1997) Stabilizing selection and the comparative analysis of adaptation. *Evolution*, **51**, 1341–1351.
Hansen, T.F., Pienaar, J. & Orzack, S.H. (2008) A comparative method for studying adaptation to a randomly evolving environment. *Evolution*, **62**, 1965–1977.
Housworth, E.A., Martins, E.P. & Lynch, M. (2004) The phylogenetic mixed model. *The American Naturalist*, **163**, 84–96.
Ives, A.R., Midford, P.E. & Garland, T., Jr (2007) Within-species variation and measurement error in phylogenetic comparative methods. *Systematic Biology*, **56**, 252–270.
Judge, G.G., Griffiths, W.E., Hill, R.C., Lutkepohl, H. & Lee, T.-C. (1985) *The Theory and Practice of Econometrics*. John Wiley and Sons, New York.
Lavin, S.R., Karasov, W.H., Ives, A.R., Middleton, K.M. & Garland, T., Jr (2008) Morphometrics of the avian small intestine, compared with non-flying mammals: a phylogenetic approach. *Physiological and Biochemical Zoology*, **81**, 526–550.
Martins, E.P. & Hansen, T.F. (1997) Phylogenies and the comparative method: a general approach to incorporating phylogenetic information into the analysis of interspecific data. *The American Naturalist* **149**, 646–667. Erratum 153:448.
Mathworks, Inc. (1996) *MATLAB*. The MathWorks, Inc., Natick, MA.
Morales, E. (2000) Estimating phylogenetic inertia in Tithonia (Asteracea): a comparative approach. *Evolution*, **54**, 475–484.
Ornelas, J.F., Ordano, M., De-Nova, A.J., Quintero, M.E. & Garland, T., Jr (2007) Phylogenetic analysis of interspecific variation in nectar of hummingbird-visited plants. *Journal of Evolutionary Biology*, **20**, 1904–1917.
Pagel, M. (1997) Inferring evolutionary processes from phylogenies. *Zoologica Scripta*, **26**, 331–348.
Revell, L.J., Harmon, L.J. & Collar, D.C. (2008) Phylogenetic signal, evolutionary process, and rate. *Systematic Biology*, **57**, 591–601.
Revell, L.J. & Harrison, A.S. (2008) PCCA: a program for phylogenetic canonical correlation analysis. *Bioinformatics*, **24**, 1018–1020.
Ricklefs, R.E. & Starck, J.M. (1996) Applications of phylogenetically independent contrasts: a mixed progress report. *Oikos*, **77**, 167–172.
Rohlf, F.J. (2006) A comment on phylogenetic correction. *Evolution*, **60**, 1509–1515.
Scales, J.A., King, A.A. & Butler, A.M. (2009) Running for your life or running for your dinner: what drives fiber-type evolution in lizard locomotor muscles? *The American Naturalist*, **173**, 543–553.
Schmitz, S., Cherny, S.S. & Fulker, D.W. (1998) Increase in power through multivariate analyses. *Behavior Genetics*, **28**, 357–363.
Swofford, D.L. (1999) *PAUP: Phylogenetic Analysis Using Parsimony (*and Other Methods)*. Sinauer Associates, Inc., Sunderland, MA.
Uhlenbeck, G.E. & Ornstein, L.S. (1930) On the theory of Brownian motion. *Physical Review*, **36**, 823–841.
Valladares, F. & Sanchez-Gomez, D. (2006) Ecophysiological traits associated with drought in Mediterranean tree seedlings: Individual responses versus interspecific trends in eleven species. *Plant Biology* **8**, 688–697.
Willig, M.R. & Owen, R.D. (1987) Univariate analyses of morphometric variation do not emulate the results of multivariate analyses. *Systematic Zoology*, **36**, 398–400.
Xia, C. & Zhenghai, H. (2000) Comparative studies on leaf structure and oil cells of the Magnaliaeae in China. *Acta Phytotaxonomica Sinica*, **38**, 218–230.
Zanne, A.E., Chapman, C.A. & Kitajima, K. (2005) Evolutionary and ecological correlates of early seedling morphology in East African trees and shrubs. *American Journal of Botany*, **92**, 972–978.

Received 20 January 2009; accepted 19 May 2009
Handling Editor: James Cresswell

Supporting information

Additional supporting information may be found in the online version of this article:

Appendix S1. Methodological details of the common-garden experiments and AFLP analysis for phylogenetic tree construction for nine *Manglietia* species.

Appendix S2. AFLP data for nine *Manglietia* species and the outgroup *Liriodendron chinense* used to construct the phylogenetic tree (Fig. 2), including the phylogenetic tree in Nexus form.

Appendix S3. Data for 13 traits for the nine *Manglietia* species.

Appendix S4. Derivations for the statistical methods used in the analyses.

Please note: Wiley-Blackwell is not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.