

Phylogenetic signal and linear regression on species data

Liam J. Revell*

National Evolutionary Synthesis Center, Duke University, Durham, NC 27705, USA

Summary

1. A common procedure in the regression analysis of interspecies data is to first test the independent and dependent variables X and Y for phylogenetic signal, and then use the presence of signal in one or both traits to justify regression analysis using phylogenetic methods such as independent contrasts or phylogenetic generalized least squares.
2. This is incorrect, because phylogenetic regression assumes that the residual error in the regression model (not in the original traits) is distributed according to a multivariate normal distribution with variances and covariances proportional to the historical relations of the species in the sample.
3. Here, I examine the consequences of justifying and applying the phylogenetic regression incorrectly. I find that when used improperly the phylogenetic regression can have poor statistical performance, even under some circumstances in which the type I error rate of the method is not inflated over its nominal level.
4. I also find, however, that when tests of phylogenetic signal in phylogenetic regression are applied properly, and in particular when phylogenetic signal in the residual error is simultaneously estimated with the regression parameters, the phylogenetic regression outperforms equivalent non-phylogenetic procedures.

Key-words: comparative method, interspecific data, least squares, type I error

Introduction

Phylogenetic methods have become *de rigueur* in the analysis of interspecies data (Harvey & Pagel 1991). This is because species are non-independent for the purposes of statistical analysis due to their common history (Felsenstein 1985; Harvey & Pagel 1991). This problem of the statistical dependence of related species has been solved in various different ways for different types of data and scientific questions (e.g. Ridley 1983; Felsenstein 1985, 2005, 2008; Cheverud, Dow & Leutenegger 1985; Grafen 1989; Pagel & Harvey 1989; Maddison 1990; Garland *et al.* 1993; Hansen 1997; Pagel 1999; Garland & Ives 2000; Rohlf 2001; Butler & King 2004; Ives, Midford & Garland 2007; Revell *et al.* 2007; Hansen, Pienaar & Orzack 2008; Lajeunesse 2009; Revell & Collar 2009; reviewed in Harvey & Pagel 1991; Felsenstein 2004). Arguably, the most widely used statistical method for the analysis of interspecific data that accounts for the historical relationships of species is the phylogenetic regression (Felsenstein 1985; Grafen 1989).

Typical linear regression analysis is of the form: $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$, with the ordinary least squares (OLS) solution:

$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$, in which \mathbf{y} is an $n \times 1$ vector (for n species) containing values for the dependent variable, Y ; \mathbf{X} is an $n \times (m + 1)$ matrix containing 1·0s in the first column and the m independent (explanatory) variables of the model in columns two through $m + 1$; and $\boldsymbol{\beta}$ is a vector containing the parameter estimates (including intercept) of the fitted univariate or multivariate linear regression model (Rencher & Schaalje 2008). $\boldsymbol{\varepsilon}$ is an $n \times 1$ vector containing the residual error in the model, and under OLS it is assumed that $\boldsymbol{\varepsilon}$ is multivariate normally distributed with a variance–covariance matrix given by $\sigma_{\varepsilon}^2\mathbf{I}$. Here, \mathbf{I} is the identity matrix (an $n \times n$ matrix containing 1·0s on the diagonal and zeroes elsewhere), and σ_{ε}^2 is the residual variance of the model (i.e. the variability in Y not explained by the regressors).

If the residuals in $\boldsymbol{\varepsilon}$ are not distributed according to $\sigma_{\varepsilon}^2\mathbf{I}$, but instead according to $\sigma_{\varepsilon}^2\mathbf{C}$ in which \mathbf{C} is known and is not proportional to \mathbf{I} (i.e. $\mathbf{C} \neq k\mathbf{I}$ for $k \in \mathbf{R}$ and $k > 0.0$), then fitting the regression model becomes a generalized (instead of ordinary) least squares problem (Rohlf 2001; Kariya & Karuta 2004; Rencher & Schaalje 2008). For non-phylogenetic data, $\mathbf{C} \neq k\mathbf{I}$ might be true, for example, if the sampling variance of Y is uneven across data points (i.e. if our data for Y have been collected with varying amounts of error). In this situation, \mathbf{C} would be a diagonal matrix containing the n sampling

*Correspondence author. E-mail: lrevell@nescent.org
Correspondence site: <http://www.respond2articles.com/MEE/>

variances of each of the observations for Y . Here, the generalized least squares regression would be the same as a weighted regression in which the weights are proportional to the inverse of the sampling variances for each observation of Y . In the phylogenetic case, the problem is not usually that the diagonal of \mathbf{C} is uneven – all extant taxa in a phylogeny are temporally equidistant from the root of the tree (by definition) so they are frequently assumed to have equivalent variances (given that they are all extant and have been measured with comparable accuracy; but see Ives, Midford & Garland 2007). Rather, in the phylogenetic case, it is that the off-diagonals of \mathbf{C} are non-zero due to the correlated histories of related species (Butler, Schoener & Losos 2000; Garland & Ives 2000).

To solve this problem, we can find the minimum variance regression slope and intercept using the generalized least squares estimating equation (or Gauss–Markov estimator; Kariya & Karuta 2004):

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{C}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{C}^{-1}\mathbf{y}.$$

This approach to the regression of interspecies data was first suggested by Grafen (1989), and has since been showed to be exactly equivalent to regression estimated using the contrasts method of Felsenstein (1985; Garland & Ives 2000; Rohlf 2001). The generalized least squares estimating equation is similar to the OLS estimator (given above), except that now we have down-weighted each observation for Y (and corresponding row of \mathbf{X}) depending on the correlation of its residual error with the other observations in our set.

Under a simple Brownian motion model for evolutionary change in Y and the X s (Cavalli-Sforza & Edwards 1967; Felsenstein 1985, 2004), \mathbf{y} (or any column of \mathbf{X} , barring the first) is expected to be distributed as a multivariate normal with variance–covariance matrix given by $\sigma_y^2\mathbf{C}$ (or $\sigma_x^2\mathbf{C}$) in which \mathbf{C} contains the height of each of the n tips of the tree on its diagonal, as well as the heights of the most recent common ancestor of each species pair i and j in each i,j th off-diagonal position (Felsenstein 1973; O'Meara *et al.* 2006). σ_y^2 (or σ_x^2) gives the phylogenetic variance or 'evolutionary rate' for Y (or X ; O'Meara *et al.* 2006; Revell 2008). More importantly, however, $\boldsymbol{\varepsilon} = \mathbf{y} - \mathbf{X}\boldsymbol{\beta}$ will also be distributed according to a multivariate normal with variance–covariance matrix given by $\sigma_\varepsilon^2\mathbf{C}$ under this evolutionary scenario. Figure 1(b) shows the computation of \mathbf{C} from a simplified five taxon tree given in Fig. 1(a).

When data for our dependent and independent variables come from species it is a common procedure to estimate the degree to which each variable is distributed according to the variance–covariance matrices $\sigma_y^2\mathbf{C}$ and $\sigma_x^2\mathbf{C}$. This measurement, which can be taken in a variety of ways, is usually described as a measure of 'phylogenetic signal' for the characters in question (e.g. Blomberg & Garland 2002; Freckleton, Harvey & Pagel 2002; Blomberg, Garland & Ives 2003; Revell, Harmon & Collar 2008). If X and Y have been evolved by Brownian motion evolution, then their phylogenetic signal will be high (i.e. close to 1.0; Revell, Harmon & Collar 2008). Furthermore, if X and Y have evolved by Brownian motion then

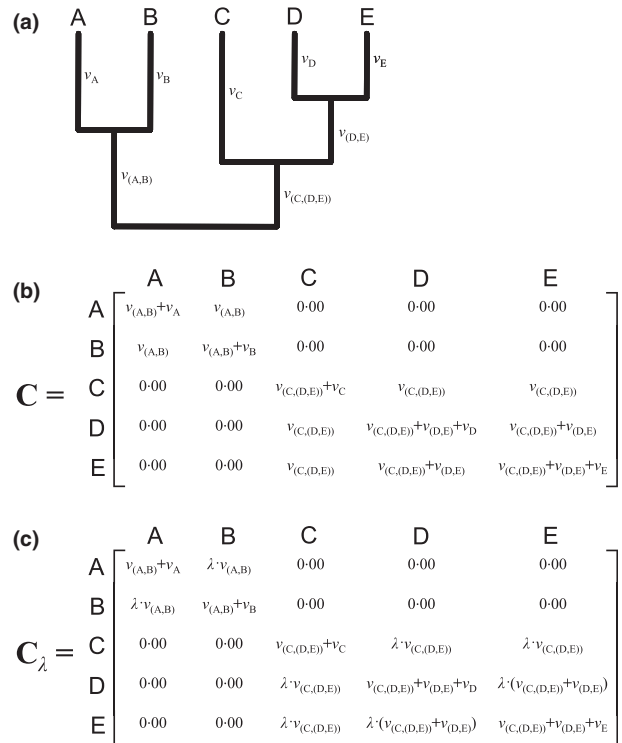


Fig. 1. (a) A simple, five taxon tree with branch lengths. (b) The matrix \mathbf{C} , which is proportional to the expected variance–covariance matrix of the residual error in a PGLS phylogenetic regression model. (c) A more flexible residual error matrix incorporating the parameter λ , which can be estimated using maximum likelihood.

$\boldsymbol{\varepsilon} = \mathbf{y} - \mathbf{X}\boldsymbol{\beta}$ will be distributed according to $\sigma_\varepsilon^2\mathbf{C}$ and the phylogenetic regression is an appropriate method to analyze the relationship between the independent variables contained in \mathbf{X} and the response variable of our model, Y . Thus, it is tempting to use high phylogenetic signal in the dependent and/or independent variables as a justification for the phylogenetic regression. This is, in fact, commonly done (e.g. Ashton 2002; Gustafsson & Lindenfors 2004; Rezende, Bozinovic & Garland 2004; Muñoz-García & Williams 2005; Collen *et al.* 2006; Ebensperger & Blumstein 2006; Ezenwa *et al.* 2006; Duminil *et al.* 2007; Hendrixson, Sterner & Kay 2007; Johnson, Isaac & Fisher 2007; Rönn, Katvala & Arnqvist 2007; Beaulieu *et al.* 2008; Capellini *et al.* 2008; Møller, Neilsen & Garamzegi 2008; Lovegrove 2009; Lindenfors, Revell & Nunn 2010).

However, it does not follow that if phylogenetic signal for X and/or Y is relatively high then $\boldsymbol{\varepsilon} = \mathbf{y} - \mathbf{X}\boldsymbol{\beta}$ will necessarily be distributed according to $\sigma_\varepsilon^2\mathbf{C}$. Furthermore, it is also possible that even if phylogenetic signal is very low, $\boldsymbol{\varepsilon} = \mathbf{y} - \mathbf{X}\boldsymbol{\beta}$ may still be distributed with variance–covariance matrix $\sigma_\varepsilon^2\mathbf{C}$. Thus, the appropriate test for phylogenetic signal is actually on the residual variability in Y given our regression model – a test which is relatively infrequently applied. In this study, I simulate scenarios in which X and/or Y have relatively high phylogenetic signal, but in which $\boldsymbol{\varepsilon} = \mathbf{y} - \mathbf{X}\boldsymbol{\beta}$ is non-phylogenetic and thus the phylogenetic regression is inappropriate. I show

that using a phylogenetic regression here will induce increased variance on the regression estimator. I also examine the possibility that X and/or Y are non-phylogenetic, but that $\boldsymbol{\varepsilon} = \mathbf{y} - \mathbf{X}\boldsymbol{\beta}$ is distributed according to $\sigma_e^2\mathbf{C}$. In this case, the phylogenetic regression is appropriate; however, standard diagnostic tests on X and Y might be taken to imply that ‘phylogenetic correction’ of the regression is unnecessary. I show that ignoring phylogeny in this case can lead to poor statistical performance of the regression. Finally, I repeat a maximum likelihood procedure using the λ statistic of Pagel (1999) in which we simultaneously estimate phylogenetic signal and the regression parameters (e.g. Revell 2009), thus obviating the need for a priori estimation of phylogenetic signal in the regression variables.

Materials and methods

To illustrate the case in point, I conducted four sets of numerical simulations, each under various conditions. First, I simulated X with phylogenetic signal, but in which the residual error in $\boldsymbol{\varepsilon} = \mathbf{y} - \mathbf{X}\boldsymbol{\beta}$ was distributed according to $\sigma_e^2\mathbf{I}$; i.e. it was non-phylogenetic. Depending on the relationship between X and Y (i.e. $\boldsymbol{\beta}$), as well as the size of the residual variance σ_e^2 , Y may or may not also show phylogenetic signal in this model. Second, I simulated X without phylogenetic signal, but in which the residual variation in Y ($\boldsymbol{\varepsilon} = \mathbf{y} - \mathbf{X}\boldsymbol{\beta}$) was distributed according to $\sigma_e^2\mathbf{C}$; i.e. it was phylogenetic. Again, depending on $\boldsymbol{\beta}$ and σ_e^2 , Y may or may not have signal in this model. Third, I simulated X and the residual variability in Y both with phylogenetic signal. This is the traditional Brownian motion model (Felsenstein 1985, 2004). Fourth, I simulated neither X nor Y with phylogenetic signal. I conducted all four of these simulation scenarios across several conditions for the magnitude of residual variability, σ_e^2 , and for the relationship between X and Y , $\boldsymbol{\beta}$. In particular, for fixed variance of X ($\sigma_x^2 = 1.0$), I simulated low, medium and high residual variation in Y ($\sigma_e^2 = 0.01, 0.1, \text{ and } 1.0$). For each value of residual variability in Y , I simulated a low, moderate and high regression relationship between X and Y (in which ‘low’ means close to zero in this case, not negative). I used the regression slopes of $\beta_1 = 0.00, 0.75$ and 0.90 for these conditions.

Generating data according to these models is quite easy. First, I simulated 1000 stochastic pure-birth phylogenetic trees, each containing a fixed number of species ($n = 100$). I arbitrarily rescaled these trees to have a total length from the root to any tip of 1.0. To simulate data on the trees in which X has signal, but in which the residual variation in Y given X is uncorrelated with the tree, I simply generated my data vector for X , as $\mathbf{x} = \text{chol}(\sigma_x^2\mathbf{C})'\mathbf{u}$, in which $\text{chol}(\sigma_x^2\mathbf{C})$ denotes the upper triangular Cholesky decomposite of \mathbf{C} times the rate of evolution in X (σ_x^2), and \mathbf{u} is a vector of values sampled randomly from the standard normal distribution. I then generated a similar vector for the residual error of my regression model, but in this case I simulated the error to be uncorrelated, i.e. $\boldsymbol{\varepsilon} = \sqrt{\sigma_e^2}\mathbf{v}$. \mathbf{v} is a vector of uncorrelated random standard normal deviates, as before. I then simply computed $\mathbf{y} = \mathbf{x}\beta_1 + \boldsymbol{\varepsilon}$, where β_1 is the desired slope of the regression relationship between X and Y and the intercept of the model is (arbitrarily) set to zero.

Generating data for X that is uncorrelated with the tree, but in which residual variability in the model $\mathbf{y} = \mathbf{x}\beta_1 + \boldsymbol{\varepsilon}$ is phylogenetic, was equally straightforward. Here, I just simulated X as $\mathbf{x} = \sqrt{\sigma_x^2}\mathbf{u}$, residual error $\boldsymbol{\varepsilon}$ as $\boldsymbol{\varepsilon} = \text{chol}(\sigma_e^2\mathbf{C})'\mathbf{v}$, and computed $\mathbf{y} = \mathbf{x}\beta_1 + \boldsymbol{\varepsilon}$, as before.

I also simulated data for X and residual error in Y that were both correlated with the tree. To accomplish this, I just calculated $\mathbf{x} = \text{chol}(\sigma_x^2\mathbf{C})'\mathbf{u}$, $\boldsymbol{\varepsilon} = \text{chol}(\sigma_e^2\mathbf{C})'\mathbf{v}$ and $\mathbf{y} = \mathbf{x}\beta_1 + \boldsymbol{\varepsilon}$, for vectors of random standard normal variates \mathbf{u} and \mathbf{v} .

Finally, I simulated data for X and residual error in Y that was uncorrelated with the phylogeny. In this case, I just generated the random vector for \mathbf{x} and residual error as $\mathbf{x} = \sqrt{\sigma_x^2}\mathbf{u}$ and $\boldsymbol{\varepsilon} = \sqrt{\sigma_e^2}\mathbf{v}$, respectively, for standard normal vectors \mathbf{u} and \mathbf{v} , as before. Then, I computed $\mathbf{y} = \mathbf{x}\beta_1 + \boldsymbol{\varepsilon}$.

The last two generating models are commonly assumed or discussed in the phylogenetic literature. For instance, the model in which both data for X and residual error in Y have phylogenetic signal is exactly equivalent to the common Brownian motion model (e.g. Felsenstein 1973, 1985; O’Meara *et al.* 2006). The model in which neither X nor Y have phylogenetic signal might be expected if Y evolves as a strong adaptive response to X , but X is uncorrelated with the phylogeny (e.g. Blomberg, Garland & Ives 2003; Butler & King 2004; Hansen, Pienaar & Orzack 2008; Lavin *et al.* 2008).

The first two generating conditions are much more rarely considered. I can suggest a couple of scenarios to which they might well apply; however, I am sure that biologically savvy readers of this article will come up with others. For example, phylogenetic signal in X but no phylogenetic signal in Y given X would be expected if X evolved by Brownian motion, and Y was determined completely by X , but was measured or phenotypically expressed with error. As long as the measurement or expression error was not phylogenetically correlated, this would represent an example of the first generating model. In the second generating model, there is no phylogenetic signal in X , but residual values for Y given X have signal. We might expect this pattern of interspecific variation if Y represented a phenotypically plastic response to a random (non-phylogenetic) environment X , but the magnitude of the plasticity was phylogenetically autocorrelated. Other biological processes that could result in generating conditions one or two are certainly possible.

For each simulation model and parameter conditions, I fit three different linear regression models for the relationship between X and Y . First, I fit an OLS regression, for which the estimating equation is given above, using $\mathbf{X} = [\mathbf{1} \ \mathbf{x}]$, where $\mathbf{1}$ is a column vector of 1.0s. To test the null hypothesis that the regression slope is $\beta_1 = 0.0$, we need to calculate the variance–covariance matrix of our estimator, $\hat{\boldsymbol{\beta}}$, which is given by:

$$\mathbf{V} = \sigma_e^2(\mathbf{X}'\mathbf{X})^{-1} \quad \text{or} \quad \mathbf{V} = (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})'(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})(n - 2)^{-1}(\mathbf{X}'\mathbf{X})^{-1}.$$

We can then compute $t(\text{d.f.} = n - 2) = \beta_1/\sqrt{V_{11}}$, which should be distributed as a t -statistic with $n - 2 = 98$ d.f. (Rencher & Schaalje 2008). With the term V_{11} , I am referring to the estimation variance component corresponding to β_1 (not β_0 , the intercept of the model), which is actually in the second row and column of \mathbf{V} .

Second, I fit the phylogenetic generalized least squares model (PGLS), in which the error structure of the residual vector, $\boldsymbol{\varepsilon} = \mathbf{y} - \mathbf{X}\boldsymbol{\beta}$, is assumed to be given by $\sigma_e^2\mathbf{C}$. As noted earlier, this model will yield exactly the same regression slope estimate as the procedure of independent contrasts followed by regression through the origin (Felsenstein 1985; Garland, Harvey & Ives 1992; Garland & Ives 2000; Rohlf 2001). Here, \mathbf{C} is a matrix with the tree height in every diagonal position, and the heights of the most recent common ancestor of species i and j in each i,j th off-diagonal position (Fig. 1b). The typical generalized least squares estimating equation, $\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{C}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{C}^{-1}\mathbf{y}$, is also given above. To conduct a

test of the hypothesis that $\beta_1 = 0.0$, we again compute $t(\text{d.f.} = n - 2) = \beta_1 / \sqrt{V_{11}}$, but this time using

$$\mathbf{V} = (\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}})' \mathbf{C}^{-1} (\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}) (n - 2)^{-1} (\mathbf{X}' \mathbf{C}^{-1} \mathbf{X})^{-1}.$$

Third, I simultaneously optimized the phylogenetic signal of the residuals of Y along with our statistical model (PGLS $_{\lambda}$). To do this, I used the parameter λ , which is a multiplier of the off-diagonal elements in \mathbf{C} (e.g. Pagel 1999; Freckleton, Harvey & Pagel 2002; Revell & Harrison 2008). Figure 1(c) gives the hypothetical computation of \mathbf{C} (herein, \mathbf{C}_{λ}) for a given λ on the five taxon tree in Fig. 1(a). In addition to use of the λ parameter of Pagel (1999), there are a variety of different ways in which \mathbf{C} can be transformed (Grafen 1989; Pagel 1999; Blomberg, Garland & Ives 2003; Hansen, Pienaar & Orzack 2008); and more or less the same effect that is accomplished by λ and related parameters can also be achieved by transforming the branch lengths of the tree prior to analysis (e.g. Garland, Harvey & Ives 1992; Clobert, Garland & Barbault 1998; Díaz-Uriarte & Garland 1998; Blomberg, Garland & Ives 2003; Lavin *et al.* 2008; Gartner *et al.* 2010). Here, I focus on λ due to its relative simplicity and because it can easily be simultaneously optimized with the regression slope and intercept of our model. To accomplish this, we optimize the following equation for the log-likelihood (L), based on the multivariate normal equation:

$$L(\hat{\boldsymbol{\beta}}, \hat{\sigma}_e^2, \hat{\lambda}) = -(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}})' (\hat{\sigma}_e^2 \mathbf{C}_{\lambda})^{-1} (\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}) / 2 - \log(|\hat{\sigma}_e^2 \mathbf{C}_{\lambda}|) / 2 - n \log(2\pi) / 2.$$

Here, \mathbf{C}_{λ} is the variance-covariance matrix, \mathbf{C} , to which the λ transformation has been applied (Fig. 1c; Pagel 1999; Freckleton, Harvey & Pagel 2002). We do not have an analytic solution for this equation, so it must be optimized numerically; however, the difficulty of this optimization is alleviated considerably by the fact that for any given value of λ (and thus \mathbf{C}_{λ}), our conditional maximum likelihood estimates for $\hat{\boldsymbol{\beta}}$ and $\hat{\sigma}_e^2$ can be obtained as follows (Rencher & Schaalje 2008):

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}' \mathbf{C}_{\lambda}^{-1} \mathbf{X})^{-1} \mathbf{X}' \mathbf{C}_{\lambda}^{-1} \mathbf{y} \quad \text{and} \quad \hat{\sigma}_e^2 = (1/n) (\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}})' \mathbf{C}_{\lambda}^{-1} (\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}).$$

By substituting \mathbf{C}_{λ} for \mathbf{C} , we can conduct our hypothesis test of $\beta_1 = 0.0$ using the same calculations as for PGLS, above; however, we should test our t -statistic against a t -distribution with $n - 3$ d.f. due to the one additional parameter (λ) estimated in the PGLS $_{\lambda}$ model. In this study, I limited estimation of λ to the interval $0.0 \leq \hat{\lambda} \leq 1.0$, because most values of λ outside this interval result in a likelihood equation that is not defined; however, in theory $\lambda > 1.0$ or $\lambda < 0.0$ are possible (Freckleton, Harvey & Pagel 2002).

For each simulated data set, I compared the performance of OLS and PGLS by determining which estimation procedure produced an estimated regression slope, β_1 , that was closest to its generating value. I also counted the number of significant regressions of each type to estimate the type I error (when the generating regression slope was $\beta_1 = 0.00$) or power (when $\beta_1 > 0.00$) of each procedure. Finally, I determined the bias of each estimating procedure by computing the mean parameter estimate across all the simulated data sets for each of the three estimators. Several studies have shown that OLS and GLS are unbiased even if the structure of the error term is specified incorrectly (e.g. Pagel 1993; Rohlf 2006; Rencher & Schaalje 2008; Revell 2009), so I did not expect bias to be significant for any of my estimation methods.

Finally, for each simulated data set I also computed a slew of phylogenetic diagnostics on the variables X and Y . I estimated λ (Pagel 1999; Freckleton, Harvey & Pagel 2002), now for each character

separately; I computed K , a measure of phylogenetic signal developed by Blomberg, Garland & Ives (2003); and, finally, I also calculated independent contrasts (Felsenstein 1985), and computed both the Pearson correlation (r) and Spearman rank correlation (ρ) between the absolute values of the standardized contrasts and their expected standard deviations prior to standardization (Garland, Harvey & Ives 1992).

For a single character, contained in (say) character vector \mathbf{y} , λ is optimized using likelihood by maximizing the following equation:

$$L(\hat{\lambda}, \hat{\sigma}_y^2, \hat{a}) = -(\mathbf{y} - \hat{a}\mathbf{1})' (\hat{\sigma}_y^2 \mathbf{C}_{\lambda})^{-1} (\mathbf{y} - \hat{a}\mathbf{1}) / 2 - \log(|\hat{\sigma}_y^2 \mathbf{C}_{\lambda}|) / 2 - n \log(2\pi) / 2,$$

in which $\mathbf{1}$ is a vector of 1.0s, as before; and the conditional maximum likelihood estimates of \hat{a} and $\hat{\sigma}_y^2$ are given by $\hat{a} = (\mathbf{1}' \mathbf{C}_{\lambda}^{-1} \mathbf{1})^{-1} \mathbf{1}' \mathbf{C}_{\lambda}^{-1} \mathbf{y}$ and $\hat{\sigma}_y^2 = (1/n) (\mathbf{y} - \hat{a}\mathbf{1})' \mathbf{C}_{\lambda}^{-1} (\mathbf{y} - \hat{a}\mathbf{1})$, respectively (Freckleton, Harvey & Pagel 2002). This equation is maximized using numerical methods. As before, we limited estimation of λ to the interval $0.0 \leq \hat{\lambda} \leq 1.0$.

Blomberg, Garland & Ives (2003) proposed an alternative measure of phylogenetic signal which is receiving wide utility. Their measure, K , can be computed as follows:

$$K = \frac{(\mathbf{y} - \hat{a}\mathbf{1})' (\mathbf{y} - \hat{a}\mathbf{1})}{(\mathbf{y} - \hat{a}\mathbf{1})' \mathbf{C}^{-1} (\mathbf{y} - \hat{a}\mathbf{1})} \bigg/ \frac{\text{tr}(\mathbf{C}) - n(\mathbf{1}' \mathbf{C}^{-1} \mathbf{1})^{-1}}{n - 1}.$$

Here, $\hat{a} = (\mathbf{1}' \mathbf{C}^{-1} \mathbf{1})^{-1} \mathbf{1}' \mathbf{C}^{-1} \mathbf{y}$; $\text{tr}(\mathbf{C})$ indicates that the trace of \mathbf{C} has been calculated; and all other terms have been previously defined (Revell, Harmon & Collar 2008).

As noted above, I computed the Pearson and Spearman correlations between the absolute values of the standardized independent contrasts of Felsenstein (1985) and the square roots of their expected variances (Garland, Harvey & Ives 1992). This method is intended to measure whether contrasts have been standardized appropriately – which is not expected to have been the case if Brownian motion is an inappropriate model for character evolution in our phylogeny. Thus, a non-significant relationship between the contrasts and their standard deviations is often taken as evidence that phylogenetic methods for regression are appropriate (Garland, Harvey & Ives 1992; Nunn & Barton 2000; Fisher, Blomberg & Owens 2003).

In a detailed Appendix, I have provided computer code written in the R programming language (R Development Core Team 2009) that implements the phylogenetic regression methods described above. I have also provided code to fit both Pagel's (1999) λ (Freckleton, Harvey & Pagel 2002) and Blomberg, Garland & Ives's (2003) K . Both are often used as measures of phylogenetic signal on individual traits.

Results

Table 1 shows the mean parameter estimates, type I error rates, and power for each estimating procedure (OLS, PGLS and PGLS $_{\lambda}$), for data generated with phylogenetic signal in the independent variable, X , but no phylogenetic signal in the model residuals. As expected, both OLS and PGLS were unbiased. The most obvious features of Table 1 are fourfold. First, estimation accuracy is substantially higher for OLS than PGLS. Averaged across simulation conditions, in 84.2% of simulations OLS produced a better (i.e. closer to its generating value) estimate of the regression slope than PGLS (Table 1; Fig. 2). Second, variance among estimated values of β_1 was

Table 1. Parameter estimation, type I error and power for OLS, PGLS and PGLS_λ for data generated with phylogenetic signal in *X*, but no phylogenetic signal for the model residuals

β_1	σ_ϵ^2	$\hat{\beta}_{OLS}$	$\text{var}(\hat{\beta}_{OLS})$	Error/ power	$\hat{\beta}_{PGLS}$	$\text{var}(\hat{\beta}_{PGLS})$	Error/ power	OLS : PGLS	$\hat{\beta}_\lambda$	$\text{var}(\hat{\beta}_\lambda)$	Error/ power
0.00	0.01	-9.62×10^{-4}	1.49×10^{-4}	0.055	-3.37×10^{-3}	3.49×10^{-3}	0.051	0.836 : 0.164	-9.26×10^{-4}	1.50×10^{-4}	0.050
	0.10	-8.86×10^{-4}	1.65×10^{-3}	0.060	3.17×10^{-3}	5.68×10^{-2}	0.053	0.825 : 0.175	-8.22×10^{-4}	1.70×10^{-3}	0.057
	1.00	-4.67×10^{-3}	1.46×10^{-2}	0.047	-1.06×10^{-2}	3.97×10^{-1}	0.045	0.829 : 0.171	-4.61×10^{-3}	1.51×10^{-2}	0.045
0.75	0.01	0.749	1.49×10^{-4}	1.000	0.749	4.41×10^{-3}	0.997	0.836 : 0.164	0.749	1.47×10^{-4}	1.000
	0.10	0.751	1.50×10^{-3}	1.000	0.764	7.67×10^{-2}	0.952	0.867 : 0.133	0.751	1.55×10^{-3}	1.000
	1.00	0.753	1.40×10^{-2}	1.000	0.765	4.79×10^{-1}	0.394	0.850 : 0.150	0.752	1.43×10^{-2}	1.000
0.90	0.01	0.901	1.63×10^{-4}	1.000	0.901	5.69×10^{-3}	0.998	0.831 : 0.169	0.901	1.67×10^{-4}	1.000
	0.10	0.901	1.62×10^{-3}	1.000	0.897	3.54×10^{-2}	0.966	0.848 : 0.152	0.901	1.64×10^{-3}	1.000
	1.00	0.906	1.45×10^{-2}	1.000	0.902	5.64×10^{-1}	0.511	0.853 : 0.147	0.907	1.48×10^{-2}	1.000

β_1 and σ_ϵ^2 indicate the generating regression slope and residual error, respectively. $\hat{\beta}_{OLS}$ indicates the mean parameter estimate by OLS and $\text{var}(\hat{\beta}_{OLS})$, the variation among simulations in the estimated slope. $\hat{\beta}_{PGLS}$ and $\text{var}(\hat{\beta}_{PGLS})$ are likewise interpreted. OLS : PGLS indicates the fraction of simulations for which OLS provided a better estimate of the regression slope (i.e. one closer to its generating value) compared with the fraction for which PGLS provided a better estimate. $\hat{\beta}_\lambda$ indicates the mean parameter estimate for the regression slope when the regression model was estimated simultaneously with λ and $\text{var}(\hat{\beta}_\lambda)$, the variation among simulations. Finally, error/power indicates the type I error rate (if $\beta_1 = 0.00$) or power of each estimation procedure.

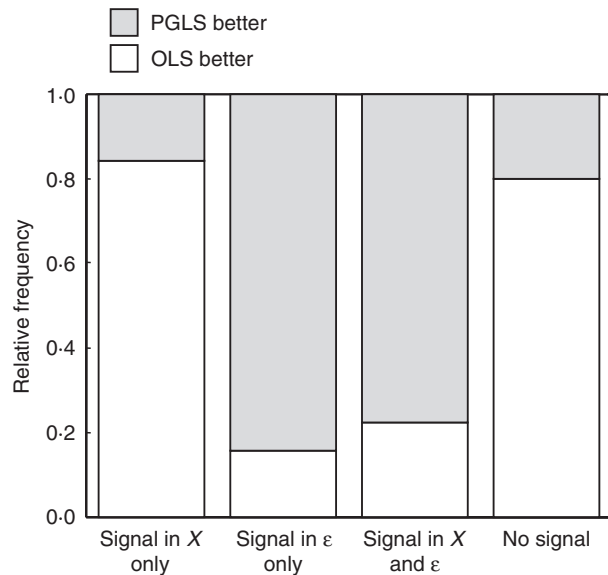


Fig. 2. A stacked bar graph showing the fraction of analyses in which OLS or PGLS yielded a better (i.e. closer to its generating value) parameter estimate of the regression slope, β_1 , for each of the four simulation models of this study. Results are averaged across simulation conditions for each model.

much higher (on average 32.9 times higher) for PGLS than OLS. Third (and in spite of points one and two), type I error was not elevated relative to its nominal level for PGLS. Finally, in simulations of $\beta_1 \neq 0.00$, power of the PGLS estimator was deflated for conditions of high residual error, σ_ϵ^2 .

Table 2 shows phylogenetic diagnostics of *X*, *Y*, and the bivariate regression model including λ . Estimated phylogenetic signal was invariably high for *X*; however, it was also quite high for *Y* when the generating value of β_1 was high and σ_ϵ^2 was relatively low. I also computed two different independent contrasts diagnostics: the Pearson and Spear-

man rank correlation coefficients for the correlation between the absolute values of the standardized contrasts and the square root of their expected variances prior to standardization. In general, these correlations were near zero for *X* (as expected), and negative for *Y*, although for increasing β_1 and sufficiently low σ_ϵ^2 the strength of the correlations decreased along with my power to detect a significant relationship (Table 2).

My results for data generated with no phylogenetic signal in the independent variable, *X*, but phylogenetically correlated residual variation in *Y* are given in Tables 3 and 4. Here, the PGLS estimator was better on average 84.5% of the time when compared with the estimated regression slope obtained using OLS (Fig. 2). However, OLS did not suffer from increased type I error when the generating regression slope was $\beta_1 = 0.00$ (Table 3). Phylogenetic signal was invariably low for *X* (Table 4). Phylogenetic signal was also generally quite low for *Y*, except when $\beta_1 = 0.00$. Diagnostic statistics on the contrasts for both *X* and *Y* overwhelming indicate inadequate standardization, again except when $\beta_1 = 0.00$, in which case they indicated that *Y*, but not *X*, has been adequately standardized in the computation of independent contrasts (Table 4).

When I generated data with phylogenetic signal in both the independent variable and in the model residuals, I found that PGLS vastly outperformed OLS. In this case, the PGLS estimator was better on average 77.8% of the time (Table 5; Fig. 2). I also found that type I error was substantially increased for OLS, unlike in Table 3 when only the residual error was simulated with phylogenetic signal and where OLS had type I error near the nominal level. All measures of phylogenetic signal indicated high phylogenetic signal for both *X* and *Y*; and no independent contrasts diagnostic suggested that the contrasts for *X* or *Y* had been inadequately standardized (Table 6).

Finally, when I generated data with no phylogenetic signal for *X* and no phylogenetic signal for the residual variability in

Table 2. Phylogenetic diagnostics for X , Y and the bivariate regression model, where the data have been simulated with phylogenetic signal in X , but no signal in the residual variability of Y given X

β_1	σ_ϵ^2	$\lambda(X)$	$\lambda(Y)$	$\lambda(Y \sim X)$	$K(X)$	$K(Y)$	$r(X)$	Error/ power	$r(Y)$	Error/ power	$\rho(X)$	Error/ power	$\rho(Y)$	Error/ power
0.00	0.01	0.998	0.014	0.012	0.991	0.066	-0.003	0.050	-0.521	1.000	-0.004	0.047	-0.538	1.000
	0.10	0.998	0.014	0.012	0.993	0.068	-0.001	0.055	-0.523	1.000	-0.002	0.052	-0.549	1.000
	1.00	0.998	0.016	0.013	1.010	0.068	0.005	0.038	-0.519	1.000	0.004	0.045	-0.545	1.000
0.75	0.01	0.998	0.979	0.010	0.995	0.714	0.001	0.061	-0.187	0.496	0.001	0.047	-0.121	0.245
	0.10	0.998	0.827	0.013	1.004	0.272	0.002	0.047	-0.413	0.993	6.43×10^{-5}	0.041	-0.340	0.919
	1.00	0.998	0.300	0.012	1.020	0.094	0.007	0.052	-0.503	1.000	0.006	0.060	-0.501	0.997
0.90	0.01	0.998	0.984	0.013	1.001	0.777	1.30×10^{-4}	0.057	-0.157	0.393	8.26×10^{-4}	0.052	-0.099	0.175
	0.10	0.998	0.872	0.010	0.996	0.337	0.003	0.052	-0.384	0.976	0.002	0.045	-0.299	0.838
	1.00	0.998	0.386	0.012	0.980	0.105	-0.001	0.052	-0.497	1.000	6.95×10^{-4}	0.054	-0.486	0.998

β_1 and σ_ϵ^2 are as in Table 1. $\lambda(X)$ and $\lambda(Y)$ indicate the mean value of phylogenetic signal estimated using the λ method for each character, X and Y , separately. $\lambda(Y \sim X)$ indicates the mean fitted λ , where λ was estimated simultaneously with the regression model. $K(X)$ and $K(Y)$ indicate the mean value of phylogenetic signal in X and Y , respectively, estimated using the K method. $r(X)$ and $r(Y)$ indicate the mean Pearson product-moment correlation between the absolute values of the independent contrasts and the square roots of their expected variances prior to standardization. $\rho(X)$ and $\rho(Y)$ are the mean values of the corresponding Spearman rank correlations. Error/power indicates the type I error or power of each contrasts-based diagnostic.

Table 3. Parameter estimation, type I error and power for OLS, PGLS and PGLS $_\lambda$ for data generated with no phylogenetic signal in X , but phylogenetic signal for the model residuals

β_1	σ_ϵ^2	$\hat{\beta}_{OLS}$	$\text{var}(\hat{\beta}_{OLS})$	Error/ power	$\hat{\beta}_{PGLS}$	$\text{var}(\hat{\beta}_{PGLS})$	Error/ power	OLS : PGLS	$\hat{\beta}_\lambda$	$\text{var}(\hat{\beta}_\lambda)$	Error/ power
0.00	0.01	-2.47×10^{-4}	8.04×10^{-5}	0.049	-1.27×10^{-6}	5.21×10^{-6}	0.055	0.151 : 0.849	-1.50×10^{-4}	5.47×10^{-6}	0.049
	0.10	-1.01×10^{-3}	7.90×10^{-4}	0.049	2.53×10^{-4}	5.84×10^{-5}	0.053	0.160 : 0.840	2.45×10^{-4}	6.29×10^{-5}	0.047
	1.00	-2.39×10^{-4}	8.17×10^{-3}	0.049	2.40×10^{-4}	5.76×10^{-4}	0.049	0.176 : 0.824	5.94×10^{-4}	6.04×10^{-4}	0.044
0.75	0.01	0.750	7.38×10^{-5}	1.000	0.750	5.14×10^{-6}	1.000	0.157 : 0.843	0.750	5.70×10^{-6}	1.000
	0.10	0.751	7.92×10^{-4}	1.000	0.750	5.44×10^{-5}	1.000	0.153 : 0.847	0.750	6.02×10^{-5}	1.000
	1.00	0.749	8.54×10^{-3}	1.000	0.750	5.43×10^{-4}	1.000	0.134 : 0.866	0.750	5.65×10^{-4}	1.000
0.90	0.01	0.900	8.02×10^{-5}	1.000	0.900	5.69×10^{-6}	1.000	0.152 : 0.848	0.900	6.08×10^{-6}	1.000
	0.10	0.900	7.53×10^{-4}	1.000	0.900	6.24×10^{-5}	1.000	0.153 : 0.847	0.900	6.55×10^{-5}	1.000
	1.00	0.902	8.00×10^{-3}	1.000	0.900	5.80×10^{-4}	1.000	0.163 : 0.837	0.900	6.18×10^{-4}	1.000

Column headers are as in Table 1.

Table 4. Phylogenetic diagnostics for X , Y and the bivariate regression model, where the data have been simulated with no phylogenetic signal in the independent variable, X , but signal in the residual error

β_1	σ_ϵ^2	$\lambda(X)$	$\lambda(Y)$	$\lambda(Y \sim X)$	$K(X)$	$K(Y)$	$r(X)$	Error/ power	$r(Y)$	Error/ power	$\rho(X)$	Error/ power	$\rho(Y)$	Error/ power
0.00	0.01	0.016	0.998	0.998	0.066	1.001	-0.520	1.000	-0.003	0.053	-0.549	1.000	-0.003	0.055
	0.10	0.016	0.998	0.998	0.067	1.002	-0.520	1.000	-2.02×10^{-4}	0.048	-0.550	1.000	-9.87×10^{-4}	0.042
	1.00	0.015	0.998	0.998	0.068	1.011	-0.521	1.000	-6.09×10^{-4}	0.044	-0.546	1.000	-0.002	0.058
0.75	0.01	0.014	0.022	0.998	0.067	0.068	-0.519	1.000	-0.518	1.000	-0.546	1.000	-0.544	1.000
	0.10	0.015	0.113	0.998	0.067	0.076	-0.520	1.000	-0.514	1.000	-0.546	1.000	-0.529	1.000
	1.00	0.015	0.593	0.998	0.066	0.146	-0.518	1.000	-0.473	1.000	-0.544	1.000	-0.436	0.990
0.90	0.01	0.014	0.017	0.998	0.066	0.067	-0.522	1.000	-0.521	1.000	-0.548	1.000	-0.546	1.000
	0.10	0.013	0.074	0.998	0.068	0.074	-0.521	1.000	-0.516	1.000	-0.547	0.999	-0.534	0.999
	1.00	0.013	0.497	0.998	0.066	0.125	-0.520	1.000	-0.485	1.000	-0.549	1.000	-0.458	0.994

Column headers are as in Table 2.

Y , OLS outperformed PGLS, as expected. Here, the OLS estimator was better on average 79.9% of the time (Table 7; Fig. 2). I also found that the type I error rate when $\beta_1 = 0.00$

was elevated for PGLS relative to its nominal level (Table 7), unlike the situation in Table 1. All measures of phylogenetic signal indicated low signal, and furthermore all

Table 5. Parameter estimation, type I error and power for OLS, PGLS and PGLS_λ for data generated with phylogenetic signal in *X* and the model residuals

β_1	σ_ϵ^2	$\hat{\beta}_{OLS}$	$\text{var}(\hat{\beta}_{OLS})$	Error/ power	$\hat{\beta}_{PGLS}$	$\text{var}(\hat{\beta}_{PGLS})$	Error/ power	OLS : PGLS	$\hat{\beta}_\lambda$	$\text{var}(\hat{\beta}_\lambda)$	Error/ power
0.00	0.01	-5.15×10^{-4}	6.72×10^{-4}	0.426	-8.18×10^{-4}	1.02×10^{-4}	0.044	0.224 : 0.776	-7.96×10^{-4}	1.02×10^{-4}	0.045
	0.10	-2.54×10^{-3}	7.27×10^{-3}	0.446	-8.25×10^{-4}	1.05×10^{-3}	0.055	0.213 : 0.787	-9.90×10^{-4}	1.07×10^{-3}	0.055
	1.00	-6.30×10^{-4}	7.61×10^{-2}	0.463	-1.71×10^{-3}	1.02×10^{-2}	0.045	0.222 : 0.778	-1.36×10^{-3}	1.02×10^{-2}	0.045
0.75	0.01	0.751	7.70×10^{-4}	1.000	0.750	9.89×10^{-5}	1.000	0.201 : 0.799	0.750	1.00×10^{-4}	1.000
	0.10	0.749	7.02×10^{-3}	1.000	0.751	1.06×10^{-3}	1.000	0.223 : 0.777	0.751	1.07×10^{-3}	1.000
	1.00	0.744	6.61×10^{-2}	0.981	0.752	1.07×10^{-2}	1.000	0.237 : 0.763	0.752	1.08×10^{-2}	1.000
0.90	0.01	0.899	7.73×10^{-4}	1.000	0.900	1.03×10^{-4}	1.000	0.239 : 0.761	0.900	1.03×10^{-4}	1.000
	0.10	0.901	6.74×10^{-3}	1.000	0.901	1.06×10^{-3}	1.000	0.217 : 0.783	0.901	1.06×10^{-3}	1.000
	1.00	0.895	7.32×10^{-2}	0.991	0.902	1.05×10^{-2}	1.000	0.223 : 0.777	0.902	1.04×10^{-2}	1.000

Column headers are as in Tables 1 and 3.

Table 6. Phylogenetic diagnostics for *X*, *Y* and the bivariate regression model, where the data have been simulated with phylogenetic signal in both *X* and the model residuals

β_1	σ_ϵ^2	$\lambda(X)$	$\lambda(Y)$	$\lambda(Y \sim X)$	$K(X)$	$K(Y)$	$r(X)$	Error/ power	$r(Y)$	Error/ power	$\rho(X)$	Error/ power	$\rho(Y)$	Error/ power
0.00	0.01	0.998	0.998	0.998	0.991	1.001	-0.003	0.050	-0.003	0.053	-0.004	0.047	-0.003	0.055
	0.10	0.998	0.998	0.998	0.993	1.002	-0.001	0.055	-2.02×10^{-4}	0.048	-0.002	0.052	-9.87×10^{-4}	0.042
	1.00	0.998	0.998	0.998	1.006	1.013	-2.89×10^{-4}	0.060	-3.17×10^{-4}	0.052	-6.59×10^{-5}	0.062	-1.67×10^{-4}	0.050
0.75	0.01	0.998	0.998	0.998	1.001	1.000	8.53×10^{-4}	0.061	5.36×10^{-4}	0.060	9.11×10^{-4}	0.047	2.05×10^{-4}	0.042
	0.10	0.998	0.998	0.998	1.004	0.997	0.002	0.047	1.44×10^{-4}	0.051	6.43×10^{-5}	0.041	-0.002	0.045
	1.00	0.998	0.998	0.998	1.011	0.990	6.91×10^{-4}	0.038	-0.001	0.038	3.85×10^{-4}	0.050	-0.002	0.044
0.90	0.01	0.998	0.998	0.998	1.001	1.000	1.30×10^{-4}	0.057	2.69×10^{-4}	0.056	8.26×10^{-4}	0.052	0.001	0.054
	0.10	0.998	0.998	0.998	0.996	0.998	0.003	0.052	2.75×10^{-4}	0.046	0.002	0.045	7.02×10^{-4}	0.048
	1.00	0.998	0.998	0.998	0.980	0.984	0.002	0.057	0.002	0.055	6.03×10^{-4}	0.054	2.00×10^{-4}	0.048

Column headers are as in Tables 2 and 4.

Table 7. Parameter estimation, type I error and power for OLS, PGLS and PGLS_λ for data generated with no phylogenetic signal in *X* nor in the model residuals

β_1	σ_ϵ^2	$\hat{\beta}_{OLS}$	$\text{var}(\hat{\beta}_{OLS})$	Error/ power	$\hat{\beta}_{PGLS}$	$\text{var}(\hat{\beta}_{PGLS})$	Error/ power	OLS : PGLS	$\hat{\beta}_\lambda$	$\text{var}(\hat{\beta}_\lambda)$	Error/ power
0.00	0.01	-8.37×10^{-4}	1.03×10^{-4}	0.043	-1.11×10^{-3}	1.71×10^{-3}	0.503	0.792 : 0.208	-8.18×10^{-4}	1.03×10^{-4}	0.041
	0.10	-8.54×10^{-4}	1.05×10^{-3}	0.055	4.54×10^{-3}	1.56×10^{-2}	0.519	0.799 : 0.201	-7.69×10^{-4}	1.06×10^{-3}	0.054
	1.00	-1.26×10^{-3}	1.02×10^{-2}	0.046	-9.34×10^{-3}	1.65×10^{-1}	0.520	0.797 : 0.203	-1.25×10^{-3}	1.02×10^{-2}	0.045
0.75	0.01	0.750	9.90×10^{-5}	1.000	0.749	1.71×10^{-3}	1.000	0.812 : 0.188	0.750	9.95×10^{-5}	1.000
	0.10	0.751	1.07×10^{-3}	1.000	0.758	1.94×10^{-2}	0.999	0.799 : 0.201	0.751	1.07×10^{-3}	1.000
	1.00	0.753	1.08×10^{-2}	1.000	0.744	2.23×10^{-1}	0.952	0.793 : 0.207	0.753	1.08×10^{-2}	1.000
0.90	0.01	0.900	1.03×10^{-4}	1.000	0.901	2.47×10^{-3}	1.000	0.782 : 0.218	0.900	1.03×10^{-4}	1.000
	0.10	0.901	1.06×10^{-3}	1.000	0.896	1.83×10^{-2}	0.999	0.809 : 0.191	0.901	1.07×10^{-3}	1.000
	1.00	0.902	1.06×10^{-2}	1.000	0.914	2.37×10^{-1}	0.966	0.807 : 0.193	0.902	1.07×10^{-2}	1.000

Column headers are as in Tables 1, 3 and 5.

contrasts-based diagnostics indicated inadequate standardization of *X* and *Y* (Table 8).

Notably, PGLS_λ, in which phylogenetic signal is estimated simultaneously with our regression model, effectively recovers the performance of the best model (either OLS or PGLS) under all of the simulation conditions of the study.

This is evidenced by the very low estimation variance of β_1 in the PGLS_λ model, regardless of the generating conditions (Tables 1, 3, 5, 7). Because fitting the PGLS_λ models requires the estimation of one additional parameter relative to OLS or PGLS, the accuracy of PGLS_λ is slightly decreased relative to OLS, when the assumptions of OLS

Table 8. Phylogenetic diagnostics for X , Y and the bivariate regression model. Here, neither X nor the model residuals have been simulated with phylogenetic signal.

β_1	σ_ϵ^2	$\lambda(X)$	$\lambda(Y)$	$\lambda(Y \sim X)$	$K(X)$	$K(Y)$	$r(X)$	Error/ power	$r(Y)$	Error/ power	$\rho(X)$	Error/ power	$\rho(Y)$	Error/ power
0.00	0.01	0.016	0.014	0.015	0.066	0.066	-0.520	1.000	-0.521	1.000	-0.549	1.000	-0.548	1.000
	0.10	0.016	0.014	0.015	0.067	0.068	-0.520	1.000	-0.523	1.000	-0.550	1.000	-0.549	1.000
	1.00	0.015	0.016	0.016	0.068	0.068	-0.521	1.000	-0.519	1.000	-0.546	1.000	-0.545	1.000
0.75	0.01	0.014	0.015	0.013	0.067	0.067	-0.519	1.000	-0.519	1.000	-0.546	1.000	-0.546	1.000
	0.10	0.015	0.014	0.016	0.067	0.067	-0.520	1.000	-0.519	1.000	-0.546	1.000	-0.547	1.000
	1.00	0.015	0.013	0.015	0.066	0.066	-0.518	1.000	-0.519	1.000	-0.544	1.000	-0.545	0.999
0.90	0.01	0.014	0.014	0.016	0.066	0.067	-0.522	1.000	-0.522	1.000	-0.548	1.000	-0.548	1.000
	0.10	0.013	0.015	0.013	0.068	0.068	-0.521	1.000	-0.521	1.000	-0.547	0.999	-0.548	1.000
	1.00	0.013	0.016	0.015	0.066	0.066	-0.520	1.000	-0.517	1.000	-0.549	1.000	-0.546	1.000

Column headers are as in Tables 2, 4 and 6.

hold, or PGLS, when the assumptions of PGLS hold. However, the advantage of PGLS $_\lambda$ is that it very nearly recovers the performance of the best model when no particular level of phylogenetic signal in the residual error can be safely assumed *a priori* (as will most often be the case for empirical studies).

Discussion

Over the past 20 years the phylogenetic regression has become among the most commonly applied methods in comparative biology (Felsenstein 1985; Grafen 1989). However, its assumptions are still not widely understood (Rohlf 2006). To examine how and when it should be applied I will review, in turn, each of the simulated scenarios for this study.

PHYLOGENETIC SIGNAL IN THE INDEPENDENT VARIABLE

When the generating model was one in which I simulated phylogenetic signal in the independent variable, but uncorrelated residual error in Y , a phylogenetic regression is not necessary. In this case, the assumption of independent errors holds and OLS is a perfectly appropriate method for fitting the regression model.

Consistent with this assertion, I found that OLS overwhelmingly provided a better regression slope estimate (84.2% of the time, averaged across simulation conditions; Fig. 2) than PGLS. Furthermore, the variance among simulations in the regression estimator was much higher for PGLS than OLS (on average 32.9 times higher; Table 1). However, I also found that type I error was not elevated for PGLS when the generating regression slope was $\beta_1 = 0.00$ (Table 1). This interesting result will be discussed further below. Performance of the phylogenetic regression was fully recovered by simultaneously estimating the λ parameter of Pagel (1999; Revell 2009), as discussed in the 'Materials and methods'. In general, the fitted value of λ for the regression model was very low (close to zero; Table 2) in this case, which makes PGLS $_\lambda$ nearly equivalent to OLS thus explaining its good statistical performance here.

NO PHYLOGENETIC SIGNAL IN THE INDEPENDENT VARIABLE

When I used a generating model for simulation with no phylogenetic signal in the independent variable, but phylogenetic signal in the residual error for the dependent variable, a phylogenetic regression is appropriate. Here, the standard OLS assumption of independent errors is not true, and thus OLS is not an appropriate method to fit our bivariate regression model and PGLS should be used. As such, I was not surprised to find that PGLS regression yielded a better estimate of the generating regression slope for the vast majority (84.5%; Fig. 2) of simulated data sets across all simulation conditions.

In these simulations, however, I also found that oftentimes phylogenetic signal for X and Y was low by both of our chosen metrics (the λ statistic of Pagel 1999; and the K statistic of Blomberg, Garland & Ives 2003; Table 4). Furthermore, for all generating conditions except $\beta_1 = 0.00$, Pearson and Spearman correlation-based diagnostic analysis of the independent contrasts suggested that contrasts had been inadequately standardized. Thus, for data generated under these conditions, standard diagnostics computed for the dependent and independent variables separately might be interpreted to suggest that standard PGLS or contrasts are inappropriate, even though they are called for in this case. As before, type I error of OLS was not inflated over its nominal level, a result which I will discuss at greater length below.

PHYLOGENETIC SIGNAL IN THE INDEPENDENT VARIABLE, AND THE RESIDUALS

This is the traditional Brownian motion model for character evolution in X and Y . In this case, it should be no surprise that PGLS outperforms OLS, producing a parameter estimate for the regression slope that is closer to its generating value in 77.8% of simulations, averaged across conditions (Table 5; Fig. 2). I also found that type I error was considerably inflated if OLS was used, which is also consistent with earlier studies (e.g. Rohlf 2006; Revell 2009). Under these simulation conditions, all phylogenetic diagnostics

(phylogenetic signal, Pearson and Spearman correlation diagnostics on the contrasts) suggest that the Brownian motion is an appropriate model for evolution and that the phylogenetic regression is indicated.

NON-PHYLOGENETIC INDEPENDENT VARIABLE AND ERROR

When I simulated data for X and residual error in Y that were entirely non-phylogenetic, I found that OLS was vastly superior to PGLS in both estimation and hypothesis testing, with PGLS showing severely inflated type I error for $\beta_1 = 0.00$ (Table 7). Unsurprisingly, estimated phylogenetic signal was invariably low and significant negative Pearson and Spearman rank correlations between the absolute values of the independent contrasts and the square roots of their expected variances suggested inadequate standardization in the contrasts procedure (Table 8).

DIAGNOSTICS AND THE PHYLOGENETIC REGRESSION

I found mixed results with regard to diagnostics and the phylogenetic regression. Certainly, as in Table 6, when all diagnostics indicated high phylogenetic signal for the independent and dependent variables considered separately, as well as no relation between the absolute value of independent contrasts and their expected variances, I found that the phylogenetic regression as traditionally applied performed extremely well. However, I also identified conditions in which phylogenetic signal for X and/or Y was high, and in which correlation-based diagnostics generally indicated that the phylogenetic regression was appropriate – and yet residual error was uncorrelated among species, and thus OLS (not PGLS) was called for (e.g. Table 2). In addition, in some cases, I found conditions in which phylogenetic signal was low, correlation-based contrasts analysis indicated inadequate standardization, and yet the phylogenetic regression was fully appropriate (e.g. Table 4).

These results suggest that the tenet of this article, that assessing phylogenetic signal in the original variables is generally insufficient to diagnose whether the phylogenetic regression is appropriate, is correct. Instead, researchers should consider simultaneously fitting a model for phylogenetic signal in the residual error along with their regression model, as I have done using the parameter λ of Pagel (1999), above.

SIMULTANEOUS ESTIMATION OF λ

It is possible to optimize the error structure of the residuals simultaneously with fitting the regression model via least squares (e.g. Grafen 1989). As noted in the preceding paragraphs, this approach seems preferable in general, and in this study, I found that the performance of the best model (OLS or PGLS, depending on the simulation conditions) could be fully or nearly fully recovered through simultaneous optimization of the λ parameter of Pagel (1999) (Tables 1, 3, 5, 7). However, this method has its own limitations. In particular, there are

many ways in which the true error structure of our residuals could differ from the hypothesized error structure given by the tree (e.g. Fig. 1a) other than that described by λ (e.g. Blomberg, Garland & Ives 2003; Hansen, Pienaar & Orzack 2008; Lavin *et al.* 2008). To the end of obtaining an even better fit, Pagel (1999) proposed several other parameters which can be simultaneously optimized. Furthermore, Garland, Harvey & Ives (1992) and others have shown that many analogous transformations can be achieved by manipulating the branch lengths of the tree in various ways.

TYPE I ERROR OF THE 'WRONG' REGRESSION ESTIMATOR

For some of the simulation conditions of this study, I found that the incorrect regression estimator still yielded appropriate type I error rates for the generating condition of $\beta_1 = 0.00$. This result is somewhat perplexing. In general, I found that the incorrect estimation procedures overwhelmingly lead to estimated regression coefficients that were farther from the generating values than estimates obtained using the correct procedure (e.g. Tables 1, 3, 5, 7; Fig. 2). Thus, one might naively suspect that for a generating slope of $\beta_1 = 0.00$, these coefficients might also more often be found to be significantly different than 0.00. This is not true because the estimated standard error for the incorrect model in these cases (in particular, phylogenetic signal in X , but not the model residuals; and no phylogenetic signal in X , but in the residuals) increases in direct proportion to the square root of the variance in the estimated regression slopes among simulations (Table 9). In fact, this is generally a property of regression tests that yielded appropriate type I error; but not for tests that produced error inflated over its nominal level (Table 9). This result is also somewhat encouraging, because it suggests that although type I error for

Table 9. The ratio of the square root of the variance of estimates of β_1 across simulation conditions by both OLS and PGLS; and the mean ratio of the estimated standard errors for β_1 by OLS and PGLS. If our estimate of the standard error captured the uncertainty in $\hat{\beta}_1$, then the two ratios should scale proportionally (as they do for the first two simulation conditions). When the data were generated with phylogenetic signal in X and the model residuals, our estimated standard errors for β_1 by OLS was too low; conversely, when the data were generated without signal, our estimated standard errors for β_1 by PGLS were too low. This is why type I errors are inflated when the incorrect estimation procedure is used in each case. Standard deviations of the mean ratios among simulation conditions are given in parentheses after each entry

Simulation model	SD($\hat{\beta}_{OLS}$)/SD($\hat{\beta}_{PGLS}$)	SE($\hat{\beta}_{OLS}$)/SE($\hat{\beta}_{PGLS}$)
Phylogenetic X ; non-phylogenetic ε	0.178 (0.023)	0.224 (0.005)
Non-phylogenetic X ; phylogenetic ε	3.768 (0.144)	3.882 (0.033)
Phylogenetic X and ε	2.631 (0.105)	1.000 (0.005)
Non-phylogenetic X and ε	0.234 (0.018)	1.001 (0.011)

the phylogenetic regression can be elevated under some simulation conditions (e.g. Table 7), these are not conditions in which diagnostic tests on the original variables or independent contrasts would generally suggest that the phylogenetic regression was appropriate.

GENERAL RECOMMENDATIONS

The general recommendations that can be derived from this study are straightforward. Firstly we cannot diagnose whether a phylogenetic regression is appropriate based on univariate measures of phylogenetic signal calculated on the individual variables in our analysis. Although such measures might be interesting for other reasons (e.g. Freckleton, Harvey & Pagel 2002; Blomberg, Garland & Ives 2003; but see Revell, Harmon & Collar 2008), they are not useful in assessing whether or not a phylogenetic regression is appropriate. Secondly, the suitability of a phylogenetic regression should actually be diagnosed by estimating phylogenetic signal in the residual deviations of Y given our predictors (X_1, X_2 , etc.). However, thirdly, an alternative approach, which I recommend, is the simultaneous estimation of phylogenetic signal and the regression model. One example (Pagel's λ) is provided herein; although many other potentially suitable transformations are also available (e.g. Garland, Harvey & Ives 1992; Pagel 1999; Blomberg, Garland & Ives 2003; Hansen, Pienaar & Orzack 2008; Lavin *et al.* 2008).

Conclusions

Ordinary least squares regression assumes that the residual error in our regression model is independent among observations. Commonly, this will be an incorrect assumption for various types of data, particularly for data from species related by a phylogenetic tree (Felsenstein 1985; Grafen 1989; Harvey & Pagel 1991). The phylogenetic regression (here, PGLS, but – equivalently – regression through the origin of independent contrasts; Felsenstein 1985; Garland & Ives 2000; Rohlf 2001) can be used for data from species in which the residual error is distributed with covariances between samples that are proportional to the amount of shared branch length from the root node of the tree to the common ancestor of each pair of species in the sample (Fig. 1b; Rohlf 2001). However, as a means of diagnosing *a priori* whether or not a phylogenetic regression is appropriate, it has become common practice to either: compute independent contrasts-based diagnostics, such as the correlation between the absolute values of standardized contrasts and the square roots of their expected variances; or to estimate phylogenetic signal in the independent and dependent variables of our model. In this study, I have shown that these measures can sometimes be inadequate, and even misleading, regarding whether a phylogenetic regression is called for. However, I have also shown that under conditions in which phylogenetic signal in the independent variable is high, but the phylogenetic regression is inappropriate (or *vice versa*), type I error is not inflated over its nominal level, even though the accuracy of parameter estimation is substantially decreased.

Finally, I showed that simultaneously optimizing the error structure of our generalized least squares model along with the parameters of the model can be a useful approach when the suitability of our data for phylogenetic regression is not known.

Acknowledgements

L. Harmon, M. Lajeunesse, J. Losos and two anonymous reviewers kindly provided comments on this article. The National Evolutionary Synthesis Center (NSF EF-0423641) supports the author in his current position.

References

- Ashton, K.G. (2002) Patterns of within-species body size variation of birds: strong evidence for Bergmann's rule. *Global Ecology and Biogeography*, **11**, 505–523.
- Beaulieu, J.M., Leitch, I.J., Patel, S., Pendharkar, A. & Knight, C.A. (2008) Genome size is a strong predictor of cell size and stomatal density in angiosperms. *New Phytologist*, **179**, 975–986.
- Blomberg, S.P. & Garland Jr, T. (2002) Tempo and mode in evolution: phylogenetic inertia, adaptation and comparative methods. *Journal of Evolutionary Biology*, **15**, 899–910.
- Blomberg, S.P., Garland Jr, T. & 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. *American Naturalist*, **164**, 683–695.
- Butler, M.A., Schoener, T.W. & Losos, J.B. (2000) The relationship between sexual size dimorphism and habitat use in Greater Antillean *Anolis* lizards. *Evolution*, **54**, 259–272.
- Capellini, I., Barton, R.A., McNamara, P., Preston, B.T. & Nunn, C.L. (2008) Phylogenetic analysis of the ecology and evolution of mammalian sleep. *Evolution*, **62**, 1764–1776.
- Cavalli-Sforza, L.L. & Edwards, A.W.F. (1967) Phylogenetic analysis: models and estimation procedures. *American Journal of Human Genetics*, **19**, 233–257.
- Cheverud, J.M., Dow, M.M. & Leutenegger, W. (1985) The quantitative assessment of phylogenetic constraints in comparative analyses: sexual dimorphism in body weight among primates. *Evolution*, **39**, 1335–1351.
- Clobert, J., Garland Jr, T. & 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.
- Collen, B., Bykova, E., Ling, S., Milner-Gulland, E.J. & Purvis, A. (2006) Extinction risk: a comparative analysis of Central Asian vertebrates. *Biodiversity and Conservation*, **15**, 1859–1871.
- Diaz-Uriarte, R. & Garland Jr, T. (1998) Effects of branch length errors on the performance of phylogenetically independent contrasts. *Systematic Biology*, **47**, 654–672.
- Duminil, J., Fineschi, S., Hampe, A., Jordano, P., Salvini, D., Vendramin, G.G. & Petit, R.J. (2007) Can population genetic structure be predicted from life-history traits. *American Naturalist*, **169**, 662–672.
- Ebensperger, L.A. & Blumstein, D.T. (2006) Sociality in New World hystricognath rodents is linked to predators and burrow digging. *Behavioral Ecology*, **17**, 410–418.
- Ezenwa, V.O., Price, S.A., Altizer, S., Vitone, N.D. & Cook, K.C. (2006) Host traits and parasite species richness in even and odd-toed hoofed mammals, Artiodactyla and Perissodactyla. *Oikos*, **115**, 526–536.
- Felsenstein, J. (1973) Maximum likelihood estimation of evolutionary trees from continuous characters. *American Journal of Human Genetics*, **25**, 471–492.
- Felsenstein, J. (1985) Phylogenies and the comparative method. *American Naturalist*, **125**, 1–15.
- Felsenstein, J. (2004) *Inferring Phylogenies*. Sinauer, Sunderland MA.
- Felsenstein, J. (2005) Using the quantitative genetic threshold model for inferences between and within species. *Philosophical Transactions of the Royal Society Series B*, **360**, 1427–1434.
- Felsenstein, J. (2008) Comparative methods with sampling error and within-species variation: contrasts revisited and revised. *American Naturalist*, **171**, 713–725.
- Fisher, D.O., Blomberg, S.P. & Owens, I.P.F. (2003) Extrinsic versus intrinsic factors in the decline and extinction of Australian marsupials. *Proceedings of the Royal Society London Series B*, **270**, 1801–1808.

- Freckleton, R.P., Harvey, P.H. & Pagel, M. (2002) Phylogenetic analysis and comparative data: a test and review of evidence. *American Naturalist*, **160**, 712–726.
- Garland Jr, T., Harvey, P.H. & Ives, A.R. (1992) Procedures for the analysis of comparative data using phylogenetically independent contrasts. *Systematic Biology*, **41**, 18–32.
- Garland Jr, T. & Ives, A.R. (2000) Using the past to predict the present: confidence intervals for regression equations in phylogenetic comparative methods. *American Naturalist*, **155**, 346–364.
- Garland Jr, T., Dickerman, A.W., Janis, C.M. & Jones, J.A. (1993) Phylogenetic analysis of covariance by computer simulation. *Systematic Biology*, **42**, 265–292.
- Gartner, G.E.A., Hicks, J.W., Manzani, P.R., Andrade, D.V., Abe, A.S., Wang, T., Secor, S.M. & Garland Jr, T. (2010) Phylogeny, ecology, and heart position in snakes. *Physiological and Biochemical Zoology*, **83**, 43–54.
- Grafen, A. (1989) The phylogenetic regression. *Philosophical Transactions of the Royal Society Series B*, **326**, 119–157.
- Gustafsson, A. & Lindenfors, P. (2004) Human size evolution: no evolutionary allometric relationship between male and female stature. *Journal of Human Evolution*, **47**, 253–266.
- 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.
- Harvey, P.H. & Pagel, M.D. (1991) *The Comparative Method in Evolutionary Biology*. Oxford University Press, Oxford UK.
- Hendrixson, H.A., Sterner, R.W. & Kay, A.D. (2007) Elemental stoichiometry of freshwater fishes in relation to phylogeny, allometry and ecology. *Journal of Fish Biology*, **70**, 121–140.
- Ives, A.R., Midford, P.E. & Garland Jr, T. (2007) Within-species variation and measurement error in phylogenetic comparative methods. *Systematic Biology*, **56**, 252–270.
- Johnson, C.N., Isaac, J.L. & Fisher, D.O. (2007) Rarity of a top predator triggers continent-wide collapse of mammal prey: dingoes and marsupials in Australia. *Proceedings of the Royal Society Series B*, **274**, 341–346.
- Kariya, T. & Karuta, H. (2004) *Generalized Least Squares*. John Wiley & Sons, West Sussex UK.
- Lajeunesse, M.J. (2009) Meta-analysis and the comparative phylogenetic method. *American Naturalist*, **174**, 369–381.
- Lavin, S.R., Karasov, W.H., Ives, A.R., Middleton, K.M. & Garland Jr, T. (2008) Morphometrics of the avian small intestine compared with that of nonflying mammals: a phylogenetic approach. *Physiological and Biochemical Zoology*, **81**, 526–550.
- Lindenfors, P., Revell, L.J. & Nunn, C.L. (2010) Sexual dimorphism in primate aerobic capacity: a phylogenetic test. *Journal of Evolutionary Biology*, **23**, 1183–1194.
- Lovegrove, B.G. (2009) Age at first reproduction and growth rate are independent of basal metabolic rate in mammals. *Journal of Comparative Physiology B*, **179**, 391–401.
- Maddison, W.P. (1990) A method for testing the correlated evolution of two binary characters: are gains or losses concentrated on certain branches of a phylogenetic tree? *Evolution*, **44**, 539–557.
- Møller, A.P., Neilsen, J.T. & Garamzegi, L.Z. (2008) Risk taking by singing males. *Behavioral Ecology*, **19**, 41–53.
- Muñoz-García, A. & Williams, J.B. (2005) Basal metabolic rate in carnivores is associated with diet after controlling for phylogeny. *Physiological and Biochemical Zoology*, **78**, 1039–1056.
- Nunn, C.L. & Barton, R.A. (2000) Allometric slopes and independent contrasts: a comparative test of Kleiber's law in primate ranging patterns. *American Naturalist*, **156**, 519–533.
- O'Meara, B.C., Ané, C., Sanderson, M.J. & Wainwright, P.C. (2006) Testing for different rates of continuous trait evolution using likelihood. *Evolution*, **60**, 922–933.
- Pagel, M. (1993) Seeking the evolutionary regression coefficient: an analysis of what comparative methods measure. *Journal of Theoretical Biology*, **164**, 191–205.
- Pagel, M. (1999) Inferring the historical patterns of biological evolution. *Nature*, **401**, 877–884.
- Pagel, M.D. & Harvey, P.H. (1989) Comparative methods for examining adaptation depend on evolutionary models. *Folia Primatologica*, **53**, 203–220.
- R Development Core Team (2009) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Rencher, A.C. & Schaalje, G.B. (2008) *Linear Models in Statistics*. John Wiley & Sons, Hoboken NJ.
- Revell, L.J. (2008) On the analysis of evolutionary change along single branches in a phylogeny. *American Naturalist*, **172**, 140–147.
- Revell, L.J. (2009) Size-correction and principal components for interspecific comparative studies. *Evolution*, **63**, 3258–3268.
- Revell, L.J. & Collar, D.C. (2009) Phylogenetic analysis of the evolutionary correlation using likelihood. *Evolution*, **63**, 1090–1100.
- 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.
- Revell, L.J., Johnson, M.A., Schulte II, J.A., Kolbe, J.J. & Losos, J.B. (2007) A phylogenetic test for adaptive convergence in rock-dwelling lizards. *Evolution*, **61**, 2898–2912.
- Rezende, E.L., Bozinovic, F. & Garland Jr, T. (2004) Climatic adaptation and the evolution of basal and maximum rates of metabolism in rodents. *Evolution*, **58**, 1361–1374.
- Ridley, M. (1983) *The Explanation of Organic Diversity: The Comparative Method and Adaptations for Mating*. Oxford University Press, Oxford, UK.
- Rohlf, F.J. (2001) Comparative methods for the analysis of continuous variables: geometric interpretations. *Evolution*, **55**, 2143–2160.
- Rohlf, F.J. (2006) A comment on the phylogenetic correction. *Evolution*, **60**, 1509–1515.
- Rönn, J., Katvala, M. & Arnqvist, G. (2007) Coevolution between harmful male genitalia and female resistance in seed beetles. *Proceedings of the National Academy of Sciences*, **104**, 10921–10925.

Received 11 March 2010; accepted 7 May 2010
 Handling Editor: Robert P. Freckleton

Supporting Information

Additional Supporting Information may be found in the online version of this article.

Appendix S1. R code for regression and calculation of phylogenetic signal.

As a service to our authors and readers, this journal provides supporting information supplied by the authors. Such materials may be re-organised for online delivery, but are not copy-edited or typeset. Technical support issues arising from supporting information (other than missing files) should be addressed to the authors.