Received: 14 October 2017 | Accepted: 13 January 2018

DOI: 10.1111/2041-210X.12977

RESEARCH ARTICLE

Methods in Ecology and Evolution | BRITISH ECOLOGICAL SOCIETY

Comparing evolutionary rates between trees, clades and traits

Liam J. Revell[1,2] | Laura E. González-Valenzuela[3] | Alejandro Alfonso[3] | Luisa A. Castellanos-García[3] | Carlos E. Guarnizo[3] | Andrew J. Crawford[3]

[1]Programa de Biología, Facultad de Ciencias Naturales y Matemáticas, Universidad del Rosario, Bogotá, Colombia

[2]Department of Biology, University of Massachusetts Boston, Boston, MA, USA

[3]Departamento de Ciencias Biológicas, Universidad de los Andes, Bogotá, Colombia

Correspondence
Liam J. Revell
Email: liam.revell@phytools.org

Funding information
Departamento Administrativo de Ciencia, Tecnología e Innovación, Grant/Award Number: 120456934310; United States National Science Foundation, Division of Environmental Biology, Grant/Award Number: 1350474

Handling Editor: Daniele Silvestro

Abstract

1. The tempo of evolutionary change through time is among the most heavily studied dimensions of macroevolution using phylogenies.
2. Here, we present a simple, likelihood-based method for comparing the rate of phenotypic evolution for continuous characters between trees. Our method is derived from a previous approach published by Brian O'Meara and colleagues in 2006.
3. We examine the statistical performance of the method and find that it suffers from the typical downward bias expected for maximum likelihood estimates of the variance, but only for very small trees. We find that evolutionary rates are estimated with minimal bias for trees of even relatively modest size.
4. We also find that type I error rates based on a likelihood-ratio test are minimally elevated above the nominal level, even for small phylogenies. The type I error rate can be reduced to a level at or below its nominal value by substituting a test-statistic distribution obtained via simulation under the null hypothesis of no difference in evolutionary rate among trees.
5. We discuss the consequences of failing to account for uncertainty in the estimation of species means or in the phylogeny, and describe strategies for taking this uncertainty into consideration during estimation. We also identify how our approach is related to previous methods for comparing the rate of evolution among different clades of a single tree or between different phenotypic traits. Finally, we describe how the method can be applied to different evolutionary models and to discrete characters—options that are already implemented in software.
6. Evolutionary biologists continue to be intrigued by changes in the tempo of phenotypic evolution across the tree of life. The method described herein should be useful for contexts in which changes in the evolutionary rate or process between two or more clades of distant or unknown relationship are of interest.

KEYWORDS
comparative methods, maximum likelihood, phylogeny

1 | INTRODUCTION

At least since Simpson (1944), if not before, evolutionary biologists have been fascinated by the tempo or rate of evolutionary change through time. Studies of changes in evolutionary rate have grown into a major component of research on macroevolution using phylogenies (e.g. Collar, O'Meara, Wainwright, & Near, 2009; Mahler, Revell, Glor, & Losos, 2010). A wide variety of hypotheses have been

© 2018 The Authors. Methods in Ecology and Evolution © 2018 British Ecological Society

posited for why evolutionary rates may differ between time periods or among lineages. For instance ecological opportunity offered by a new environment has been asserted as a mechanism that can spur an increase in the evolutionary rate through time (e.g. Ehrlich & Raven, 1964; reviewed in Yoder et al., 2010). A lineage colonizing an area absent of competitors might be expected to diversify, in species number as well as in ecologically relevant traits, more rapidly than related clades found in more species-rich, ecologically saturated communities (e.g. Gavrilets and Losos 2009; Schluter, 2000). For instance Mahler et al. (2010) presented evidence suggesting that colonization of a new island by lizards of the genus *Anolis* tended to result in an increase in the rate of phenotypic diversification in inverse relation to the number of related lineages already present on each island.

The evolution of a new trait or organ, sometimes referred to as a "key innovation," may also stimulate species diversification or accelerate phenotypic change (Hunter & Jernvall, 1995; Liem, 1973; Simpson, 1944; Van Valen, 1971). One example of such an innovation might be the origin of the novel pharyngeal jaw form found in cichlid fishes (Liem, 1973). The specialization of the pharyngeal jaw on food processing is hypothesized to have liberated the oral jaw to differentiate in form and function, thus permitting cichlids to diversify remarkably in their feeding niche (Liem, 1973). Finally, changes in genomic architecture or genome size may also be responsible for increases in the rate of phenotypic evolution through time (e.g. Brunet et al., 2006; Hansen, 2006). For instance a whole-genome duplication event creates the opportunity for different gene copies to assume different functional roles, potentially accelerating the rate of phenotypic change in a clade.

Early methods for studying phenotypic evolution on phylogenies used Felsenstein's (1985) important method of independent contrasts. For instance Garland (1992) proposed a method in which the absolute values of independent contrasts are statistically compared between clades. The clade with the larger standardized contrasts is deemed to have evolved under a faster tempo of evolution than the clade in which the absolute values of the contrasts are smaller. Subsequently, O'Meara, Ané, Sanderson, and Wainwright (2006; also Thomas, Freckleton, & Székely, 2006) introduced an elegant method based on the statistical estimation procedure of maximum likelihood. This method, like the contrasts algorithm of Felsenstein (1985) that preceded it, explicitly assumes a Brownian process of evolutionary change.

Brownian motion is perhaps the most commonly used model for studying phenotypic evolution of continuously valued (i.e. metric) character traits on a phylogenetic tree (Felsenstein, 1985). Brownian motion is merely a continuous-time stochastic diffusion process in which the expected value is constant through time (and thus the process is directionless), and in which the anticipated variance between any pair of lineages increases as a direct function of the product of the time elapsed since the two lineages shared a common ancestor and the rate parameter of the model, normally denoted as σ^2 (O'Meara et al., 2006). Typically, σ^2 is given as the instantaneous variance parameter of the Brownian diffusion process, or the evolutionary rate. Thus, most contemporary phylogenetic studies in which the evolutionary rates between clades have been compared are in fact comparing the value of the rate parameter, σ^2 (O'Meara et al., 2006; Revell, 2008; Revell, Harmon, & Collar, 2008).

In 2006, Brian O'Meara and colleagues introduced a simple and ingenious statistical approach using likelihood to compare the evolutionary rate among lineages. The underlying idea is that the process of Brownian motion should result in a covariance structure among the observations for a given trait at the tips of the tree that is multivariate normal and in which the variance for each terminal taxon is equal to the rate of evolution (σ^2) multiplied by the total time elapsed from the global root of the tree to that species (Felsenstein, 1973; see also Revell, 2008). For an ultrametric phylogeny in which the rate of evolution is constant through time and all species are extant, this should be a constant value for all species equivalent to the total height of the tree multiplied by σ^2. The covariances between species, then, are equal to the amount of shared history between each pair of taxa multiplied by the rate of evolution through time, σ^2. The amount of shared history is merely the time from the global root of the tree to the common ancestor of each pair of terminal taxa. The likelihood is then computed from the multivariate normal density given the correlation structure implied by the phylogeny, the character values ($\mathbf{x}$) at the tips of the tree, the value of σ^2, to be estimated, and the value of the phenotypic trait at the global root of the tree, a, which must also be estimated.

O'Meara et al. (2006) showed that fitting a model in which the evolutionary rate has changed through time requires only that one accumulate the variances and covariances for each species and pair of taxa as the sum of the products of each branch segment under each rate regime, multiplied by the different rates for these different regimes: σ^2_1, σ^2_2, σ^2_3, and so on. The values at the tips of the tree under this heterogeneous rate process retain a multivariate normal density, but the process becomes one in which the variances for species and the expected covariances between them are no longer directly proportional to the elapsed time from the root to a common ancestor for each tip or pair of taxa. Rather, the variances and covariances can be computed by summing the different amounts of variance that would accumulate along each edge in each rate regime from the root to each terminal taxon or internal node.

The above procedure requires a number of parameters to be estimated that is equal to the number of postulated rate regimes, k, plus one parameter for the global root value of the trait. These parameters can be estimated by maximizing the likelihood, and then this fitted model can be compared to other simpler or more complex models using the standard machinery of likelihoods, such as the likelihood-ratio test or information theory (O'Meara et al., 2006).

In O'Meara et al. (2006), the authors presented two alternative models for heterogeneity in the evolutionary rate through time. One is a model that they referred to as the "non-censored" model in which a rate regime is postulated for every edge (or partial edge) of the tree. This method has seen wide application in empirical research (e.g. Collar, Near, & Wainwright, 2005; Collar et al., 2009; Edwards & Smith, 2010; Kozak & Wiens, 2010; Price, Holzman, Near, &

Wainwright, 2011). Their second model they denominated the "censored" model. According to this model, different rate regimes are postulated for different subtrees or bipartitions of the phylogeny, but no hypothesis is presented for the process along the intervening edge separating these two bipartitions. (It would not be correct to refer to these parts of the tree as "clades," since one subtree might be a clade that renders the ancestral regime paraphyletic, as in figure 3 of O'Meara et al., 2006.) This model has not $k + 1$ but $2k$ parameters, since one must separately estimate the ancestral value of each subtree or bipartition of the phylogeny. The term "censored" refers to the procedure of censoring (removing) the edge connecting the two or more subtrees with different hypothesized rate regimes.

Although not explicitly designed for this purpose, it recently occurred to us that the censored approach of O'Meara et al. (2006) could be applied equally well to compare the rate of evolution of a given trait among independent phylogenetic trees, rather than merely among subtrees of a particular phylogeny. In this article, we describe this approach, which is effectively a simple adaptation of the censored model of O'Meara et al. (2006). We then proceed to examine its statistical performance over a range of values for the difference in rate between clades, and over a number of different sizes of phylogeny. Although we focus herein on comparing the evolutionary rate between phylogenetic trees, afterwards we will discuss how the method is related to existing approaches for comparing the rate of evolution among traits (Adams, 2013). Finally, we consider the issues of error in the estimation of species means and error in the phylogeny, as well as the application of this general approach to different evolutionary models and discrete characters.

This method has already been implemented in software and is publicly available in the form of the function *ratebytree* which forms part of lead author Revell's PHYTOOLS R package (Revell, 2012). PHYTOOLS in turn depends on the R statistical computing environment (R Core Team, 2017), and on the important R phylogenetics packages APE (Paradis, Claude, & Strimmer, 2004) and PHANGORN (Schliep, 2011), along with several other contributed R function libraries (Azzalini & Genz, 2016; Becker, Wilks, Brownrigg, Minka, & Deckmyn, 2016; Chasalow, 2012; Gilbert & Varadhan, 2016; Harmon, Weir, Brock, Glor, & Challenger, 2008; Jackson, 2011; Lemon, 2006; Ligges & Mächler, 2003; Pinheiro, Bates, DebRoy, Sarkar, & R Core Team, 2017; Plummer, Best, Cowles, & Vines, 2006; Qiu & Joe, 2015; Venables & Ripley, 2002; Xie, 2013).

2 | MODEL, METHODS AND RESULTS

2.1 | The model

The basic method presented here is a straight-forward adaptation of O'Meara et al.'s (2006) censored approach for comparing rates among subtrees or phylogenetic bipartitions. We envision fitting two models. The more parameter-rich of these is a model in which each phylogenetic tree in our study is permitted to have a different evolutionary rate for the trait of interest. The log-likelihood of this model is merely the sum of the log-likelihoods of each phylogeny with its own tree-specific rate, given the tip data for our continuous character and the topology and branch lengths of our trees. This log-likelihood is consequently the sum of a series of logarithm-transformed multivariate normal densities and can be written as follows:

$$L=\sum_{i=1}^{k}\log\left(\frac{\exp\left[-\frac{1}{2}\left(\mathbf{x}_i-\mathbf{1}a_i\right)'\left(\sigma_i^2\mathbf{C}_i\right)^{-1}\left(\mathbf{x}_i-\mathbf{1}a_i\right)\right]}{\sqrt{(2\pi)^{N_i}\left|\sigma_i^2\mathbf{C}_i\right|}}\right)$$

Here, $\mathbf{x}_i$ is a vector of trait values for each tip of the ith tree; $\mathbf{1}$ is a conforming vector of 1.0 s; a_i is the ancestral state at the root of the ith tree; σ_i^2 is the rate of evolution for the ith tree; $\mathbf{C}_i$ is a matrix containing the heights above the global root of the most recent common ancestor of each pair of tips in the ith tree (e.g. Revell, 2008); and N_i is the number of tips in the ith tree. This model has a total of $2k$ parameters to be estimate for k trees. These parameters consist of k different rates for our different phylogenies plus k values for the root states (a_1, a_2, and so on) of each tree.

We can compare this parameter-rich model to a simpler model containing $k - 1$ fewer parameters in which we have assumed that there is but a single rate of evolution, σ^2, of our continuous character for all trees. The log-likelihood of this model is also the sum of logarithm-transformed multivariate normal densities and can be written down as follows:

$$L=\sum_{i=1}^{k}\log\left(\frac{\exp\left[-\frac{1}{2}\left(\mathbf{x}_i-\mathbf{1}a_i\right)'\left(\sigma^2\mathbf{C}_i\right)^{-1}\left(\mathbf{x}_i-\mathbf{1}a_i\right)\right]}{\sqrt{(2\pi)^{N_i}\left|\sigma^2\mathbf{C}_i\right|}}\right)$$

Note that we have permitted each tree to have a different ancestral state, which makes sense because even in our model in which the rate is constant among trees we have no *a priori* reason to presume that evolution in each of our phylogenies began from the same initial value.

An important caveat to note is that for this method to genuinely involve the comparison of rates between trees, the edge lengths of the trees should be represented in identical units. These need not necessarily be units of time, but the edge length units must match for the rate comparisons to be meaningful. For instance two or more trees in which the units of branch length were organismal generations would consequently permit a comparison of rates between phylogenies in units of evolutionary change per generation.

2.2 | Simulation tests of the method and results

We undertook the following tests of this method for comparing the evolutionary rates among trees. First, we simulated under a constant rate of evolution between two or three trees containing $N_1 = N_2 = 10$, $N_1 = N_2 = 50$, $N_1 = 10$ and $N_2 = 50$, or $N_1 = N_2 = N_3 = 30$ taxa. Trees were simulated under a constant-rate pure-birth model in which the birth rate, λ, was set equal to 1.0 and each simulation was ended just prior to the moment of the $(N+1)$th event. (That is, when the tree contained N taxa and the waiting time to the next speciation event had elapsed,

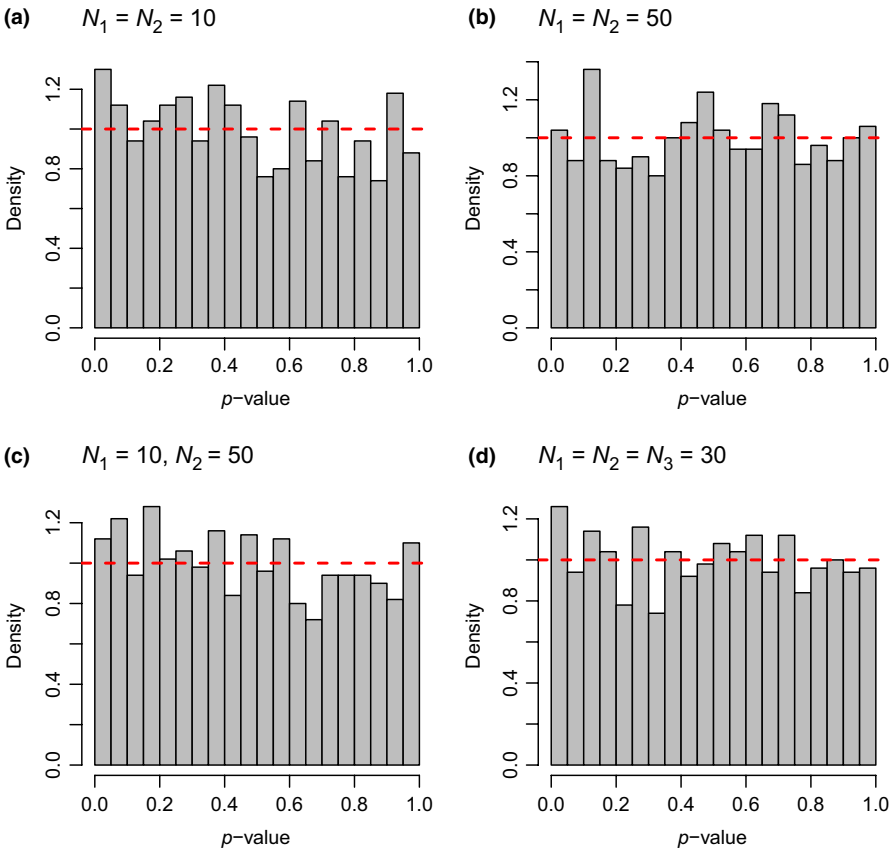

FIGURE 1 Distribution of *p*-values obtained from hypothesis tests when data were simulated under the null hypothesis of no difference in rate between trees. The expected distribution is uniform on the interval of [0, 1]. Panels (a)–(d) show the number of species in the sets of two or three trees simulated for each condition

TABLE 1 Type I error rates of a likelihood method of comparing evolutionary rates between trees when the null distribution for the likelihood-ratio test statistic was either: a χ^2 distribution with degrees of freedom equal to the number of trees, *k*, minus 1; or a distribution obtained via simulation under the null hypothesis of no difference in rate. *p*-values were obtained by comparison of the observed type I error rate to a binomial distribution with probability set to the level of α employed in the statistical test (0.05)

N_i	Type I error (χ^2)	*p*(binomial test)	Type I error (simulation)	*p*(binomial test)
$N_1 = N_2 = 10$	0.065	.015	0.045	.739
$N_1 = N_2 = 50$	0.052	.351	0.052	.351
$N_1 = 10, N_2 = 50$	0.056	.172	0.043	.827
$N_1 = N_2 = N_3 = 30$	0.063	.028	0.047	.634

but before the next speciation event was permitted to occur.) As a consequence, for a particular set of simulations, the expected total tree length is $[\log(N_i+1) - \log(2)]/\lambda$; however, trees simulated in this way will vary stochastically in their total length.

The tests of type I error involved a total of 1,000 simulations of each condition. We employed both null hypothesis testing using a likelihood-ratio test in which we compared our test-statistic to a χ^2 distribution with $k - 1$ degrees freedom, as well as via a procedure in which we generated our null distributions via simulation.

In general, we found that that the method had good statistical performance with respect to type I error. Figure 1 gives the distribution of P-values obtained under various conditions when simulating under the null. All panels fairly closely approximate a uniform distribution on the interval [0, 1], which is what we would expect under the null hypothesis if the statistical test is performing correctly. Although all simulation conditions resulted in a type I error rate that was elevated over the nominal level of 0.05, this increase was small in all cases and non-significant based on comparison to a binomial distribution for cases in which one of the input trees contained 50 taxa (Table 1). It is important to emphasize, however, that for the smallest trees of this study type I error was indeed significantly elevated above its nominal level (although this elevation was not particularly large). Fortunately, we found that substituting a null distribution of the likelihood-ratio test statistic obtained via simulation for the χ^2 distribution with $k - 1$ degrees of freedom invariably resulted in a rate of type I error statistically indistinguishable from the nominal level (Table 1). This option is already available in the *ratebytree* function of PHYTOOLS.

Next, to examine the questions of power and parameter estimation, we simulated a genuine rate difference between trees. For these simulations we invariably fixed one of the two rates at $\sigma_1^2 = 1.0$ and varied the difference between rates from no

difference ($\sigma_2^2/\sigma_1^2 = 1.0$) through $\sigma_2^2/\sigma_1^2 = 4.0$ in intervals of $\Delta(\sigma_2^2/\sigma_1^2)$ of 0.2. We used the same combinations of tree sizes simulated in the type I error analysis described above; however, for the case of $N_1 = N_2 = N_3 = 30$, we set both σ_1^2 and σ_3^2 equal to 1.0 and varied only σ_2^2/σ_1^2. We employed a smaller number of simulations in this analysis: a total of 100 for each set of tree sizes and rate ratios. For these simulations, we evaluated power to reject the null hypothesis of equal rates, along with proximity of the estimated rates to the generating values for each simulation. Figure 2 shows an example simulated dataset for $N_1 = N_2 = N_3 = 30$ in which $\sigma_1^2 = \sigma_3^2 = 1.0$ and $\sigma_2^2 = 2.0$, illustrated with a "traitgram" (i.e. a projection of the phylogeny into phenotype space; Evans, Smith, Flynn, & Donoghue, 2009) generated using the PHYTOOLS function *phenogram* (Revell, 2012, 2014).

We found that power to detect a significant difference in rate among trees was generally modest for very small trees. In the case of $N_1 = N_2 = 10$, for example, power was only about 50% even for the highest rate difference simulated ($\sigma_2^2/\sigma_1^2 = 4.0$; Figure 3). For trees of intermediate size, power was much better, and a 2- to 3-fold difference in rate was significant in over 50% of cases. For $N_1 = N_2 = 50$ and $N_1 = N_2 = N_3 = 30$, for example, over 90% of simulations were significant for rate ratios (σ_2^2/σ_1^2) equal to or greater than about 2.6 and 3.2 respectively (Figure 3). Parameter estimation was also reasonably unbiased. For the smallest simulated trees ($N_1 = N_2 = 10$) rate estimates were downwardly biased; however, this bias was minimal for larger trees (Figure 4).

2.3 | Empirical example: Rates of body size and shape evolution in two lizard clades

To illustrate the method we obtained body size and shape data for two lizard clades: the iguanian tribe Liolaemini, which includes the diverse South American genus *Liolaemus*; and the primarily North American subfamily Phrynosomatinae. The data we used come from Harmon et al. (2010a,b). We used the root node ages of 75 and 74 million years (Myr) for the two clades, respectively; however, we rescaled the trees to be in units of 100 Myr (thus having total lengths of 0.75 and 0.74) to improve estimation and to ensure that the estimated rates did not take very small values (which makes some optimization routines work less efficiently). Tree rescaling does not affect the likelihood of the different models. To obtain the values for overall body size for each species we simply log-transformed snout-to-vent length (SVL in units of mm). This follows Harmon et al. (2010a). To obtain shape, we first computed phylogenetic principal components (Revell, 2009) on each log-transformed dataset consisting of measurements for overall body size, tail length, forelimb length and hindlimb length. Next, we computed the mean eigenstructure across both phylogenies. Finally, we rotated each of the original datasets into the average eigenspace of the two different clades. Following Harmon et al. (2010a), we used scores from the second principal component as a measure of "shape" as the first principal component is dominated by overall size.

Figure 5 shows a projection of each of the two clades into a phenotype space defined by body size (log-scaled SVL) on the vertical axis and time since the root (in intervals of 100 Myr) in the horizontal dimension of each subplot. Results are summarized in Table 2. In general, we found that Liolaemini showed an approximately 1.7-fold higher rate of evolution of overall body size in spite of the fact that the overall range of body size exhibited by the clade is actually slightly lower than that seen in phrynosomatines (Figure 5). We also discovered that there was no evidence to support a difference in the rate of body shape evolution between the two lizard clades (Table 2).

3 | DISCUSSION

3.1 | Relationship with other methods

In this article, we describe a method that is not entirely novel—in the sense that it is mathematically identical to the censored rate test of O'Meara et al. (2006). Nonetheless, O'Meara et al. (2006) did not describe the model as applying to the comparison of rates among trees, nor have we encountered this application of the censored method to this particular problem. Consequently, we believe that the application of this model to the problem of comparing rates among phylogenies is relatively novel. Also in 2006, Gavin Thomas and colleagues independently published a method highly similar to O'Meara et al.'s (2006) non-censored model. The method of the current article is most appropriately viewed as an adaptation of O'Meara et al.'s (2006) censored model; however, all three approaches are nonetheless closely related.

Our approach is also related to other phylogenetic methods, as well. In particular, Adams (2013) developed a method permitting the comparison of Brownian rate for different characters on a single tree. The approach presented here could reasonably be viewed as a version of the method of Adams (2013), but where the covariance between characters is zero, and in which the underlying phylogenetic structure is permitted to differ between "traits." (In our case, we envision the traits as the same trait on different phylogenies—but the idea is similar.) In fact, our method can already be used to compare the rate of evolution of different characters under this assumption if the user merely supplies the function *ratebytree* with a list of duplicate phylogenies and a list of character vectors in which each vector contains the species values for a different trait. Much earlier, Garland (1992) presented a method in which the absolute values of Felsenstein's (1985) contrasts are compared. Had Garland (1992) elected to compare the squared values of the contrasts, as opposed to their absolute values, we believe that this would correspond to an REML (restricted maximum likelihood) version of the method herein described.

3.2 | Error in the estimation of species' means

A consideration worth noting is the effect of sampling error in the species' means on the estimation of evolutionary rates. In general,

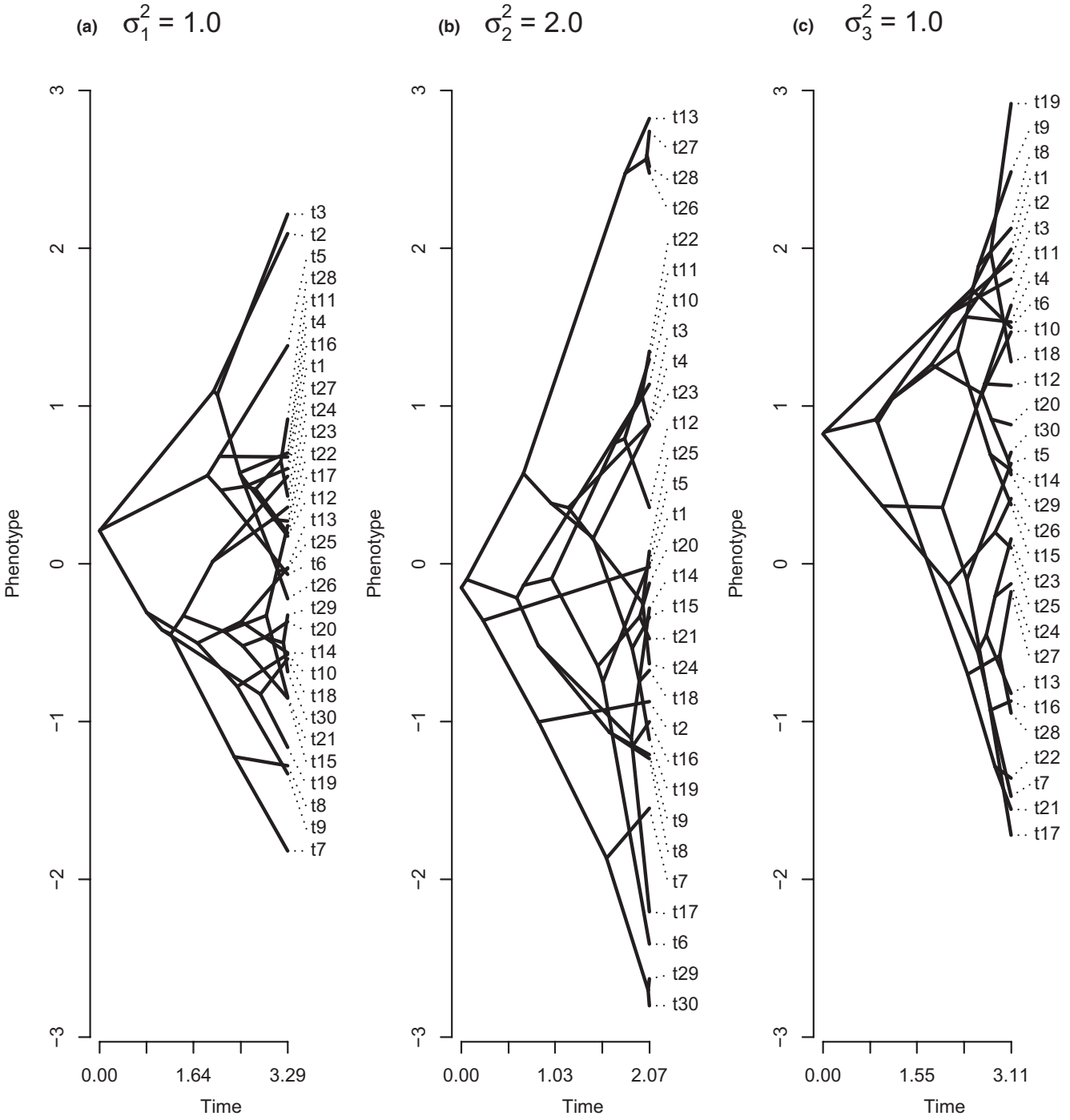

FIGURE 2 Example simulated trees and data for $N_1 = N_2 = N_3 = 30$ species in which Brownian rate parameters, $\sigma_1^2 = \sigma_3^2 = 1.0$ and $\sigma_2^2 = 2.0$. Each panel shows a traitgram, i.e. a projection of the phylogeny into phenotypic space. Note that the trees were simulated under a scenario of pure-birth with a constant rate of speciation and a taxon-number stopping condition, thus resulting in different total depths

uncertainty in the estimation of species' means will tend to result in an overestimation of the evolutionary rate for a clade. This is because the model treats uncertainty in species means as "extra" evolution at the tips of the trees (Felsenstein, 2008; Ives, Midford, & Garland, 2007). This extra evolution can only be explained by fitting a model with a higher evolutionary rate, σ^2. Consequently, we hypothesized that a consistent difference in the sampling error between clades might tend to result in an estimated Brownian rate that is higher in one tree than the other, even in the absence of a genuine difference in the rate of evolution between trees.

To explore the importance of this effect, we simulated data on two trees with $N_1 = N_2 = 50$. We simulated the tip data with varying degrees of sampling error in the estimation of species' means and then we analysed our simulated data using the method of this article. We simulated sampling error in which the sampling variance for each species mean was drawn from an exponential distribution.

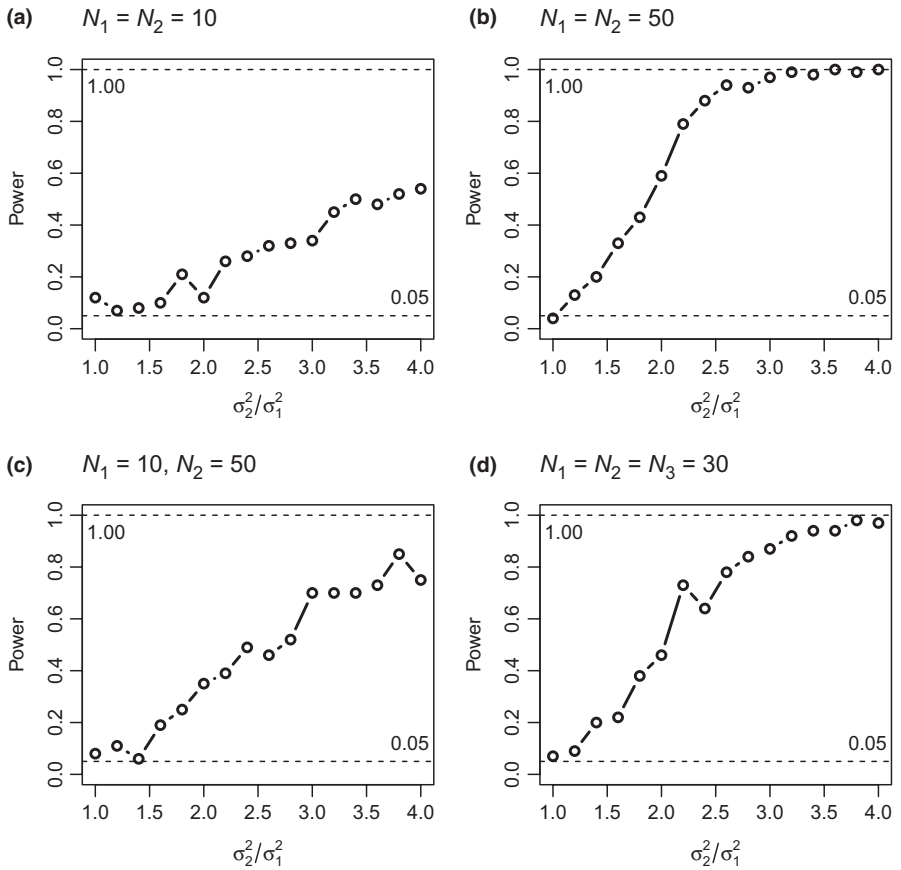

FIGURE 3 Power to reject the null hypothesis of equal rates under different simulation conditions of the Brownian rate parameters, σ_1^2 and σ_2^2 (and σ_2^2, for panel d). Invariably, the generating value of σ_1^2 (and σ_3^2, if applicable) was fixed at 1.0, whereas σ_2^2 was varied between 1.0 and 4.0 in intervals of 0.2. Panels (a)–(d) are as in Figure 1

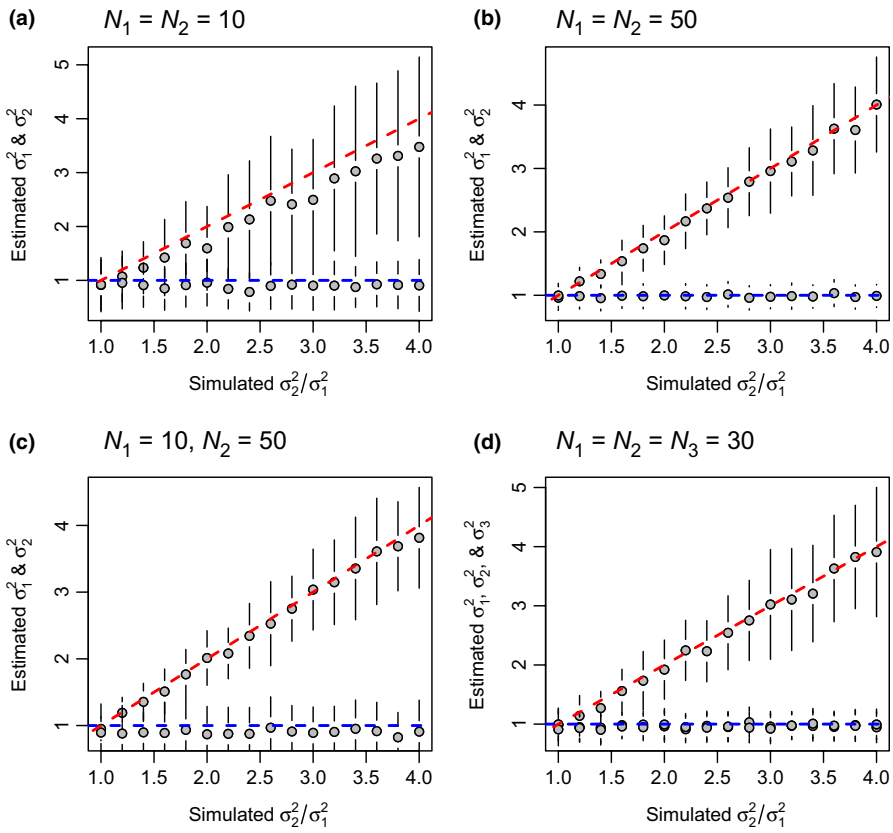

FIGURE 4 Parameter estimation for simulation conditions involving a genuine difference in rate between trees. As in Figure 2, the generating value of the Brownian rate parameter σ_1^2 (and σ_3^2, if applicable) was fixed at 1.0 while σ_2^2 varied between 1.0 and 4.0. For all panels the blue line represents the generating values of σ_1^2 (and σ_3^2, if applicable) and the red line represents the generating values of σ_2^2. Panels (a)–(d) are as in Figures 1 and 2

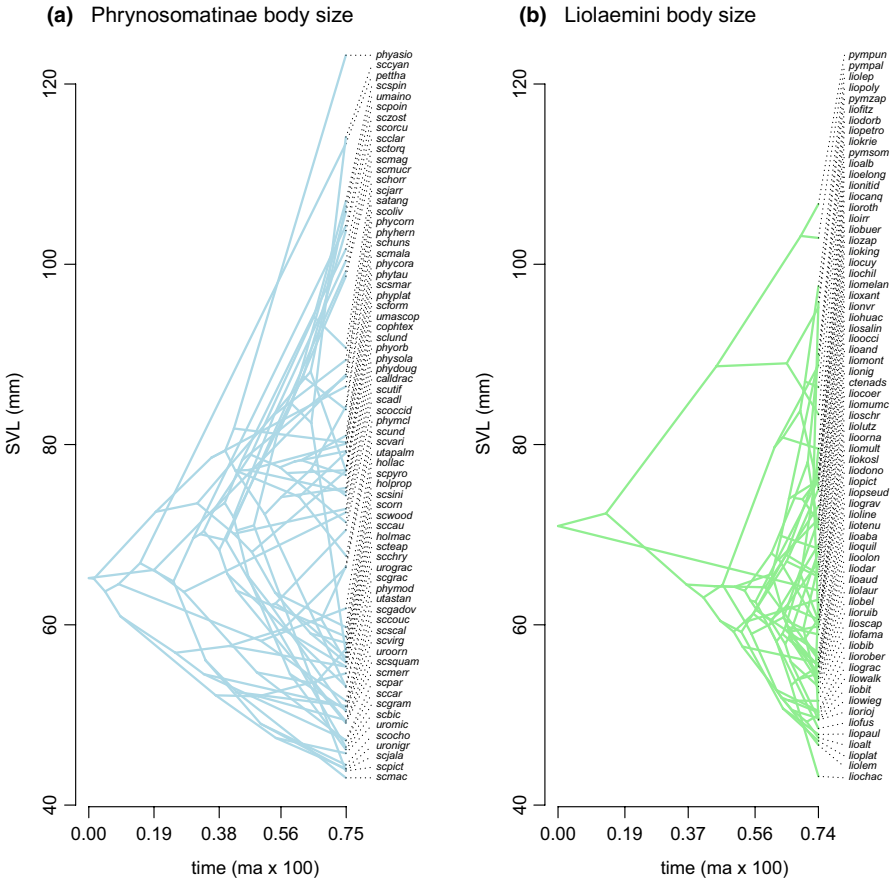

FIGURE 5 Projection of the tree into phenotype space (traitgram) for body size evolution in two lizard clades: the North American phrynosomatines; and the South American tribe, Liolaemini. Rate estimates and ancestral states for the projection were obtained on a natural logarithm scale and then back translated onto a linear scale for the purposes of visualization only

We elected to use this distribution because it is a flexible, positively valued distribution that can result in large differences in sampling error among species—or small ones, depending on the value of the rate parameter, λ. For a given value of λ, the mean sampling variance has an expected value equal to $(1/\lambda)$, whereas the expected variance among species in sampling variance is equal to $(1/\lambda)^2$. For one of the two trees, we used exponentially distributed sampling variances with a rate parameter, λ_1, of 3.0. For the second tree, we simulated random sampling variances drawn from an exponential distribution with $\lambda_2 = 1.0$ through $\lambda_2 = 10$, incrementing in intervals of 1.0. We conducted 100 simulations for each condition. Finally, we computed the type I error of the method.

Fortunately, it is straightforward to explicitly account for (known) sampling error as an additional component of our model, following the method presented by Anthony Ives and colleagues in 2007. We have employed this approach in our implementation of the method in the PHYTOOLS R package (Revell, 2012). Consequently, in analysing the results of the simulations described above we fit the model of this study, first ignoring and then accounting for known sampling error (using the method of Ives et al., 2007) in the estimation of the species' means of our simulated datasets. In the latter analysis, we assumed that within-species sampling errors were known precisely, rather than estimated from the data (as would typically be done in empirical studies); however, we suspect that our general result is unchanged as a consequence of this simplification.

TABLE 2 Body size (SVL) and shape (common phylogenetic PC2) for two lizard clades: (1) the North American iguanian subfamily Phrynosomatinae; and (2) the South American lizard tribe Liolaemini

	σ^2_1	σ^2_2	a_1	a_2	k	log(L)
Body size (SVL)						
ML common-rate model:						
Value	0.26	–	4.18	4.26	3	−4.85
SE	0.03	–	0.15	0.24		
ML multi-rate model:						
Value	0.19	0.33	4.18	4.26	4	−2.19
SE	0.03	0.06	0.13	0.27		
Likelihood ratio: 5.32; *p*-value (based on χ^2, *df* = 1): 0.021						
Body shape (common PC2)						
ML common-rate model:						
Value	0.11	–	0.00	0.00	3	47.4
SE	0.01	–	0.10	0.16		
ML multi-rate model:						
Value	0.12	0.11	0.00	0.00	4	47.5
SE	0.02	0.02	0.10	0.16		
Likelihood ratio: 0.057; *p*-value (based on χ^2, *df* = 1): 0.81						

Our results are given in Figure 6. We found that the type I error rate was substantially elevated when sampling error was not taken into account—regardless of the relative values of the rate

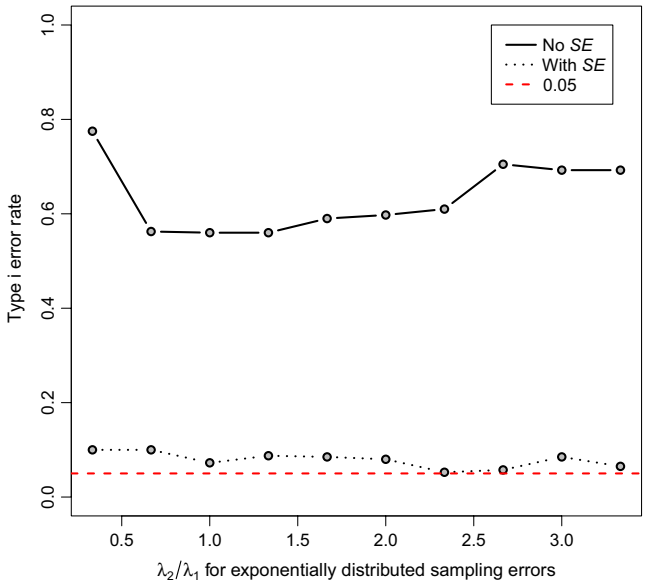

FIGURE 6 Type I error rate estimated using simulated data with sampling error in which the error is either ignored (solid line) or explicitly modelled, following the method of Ives et al. (2007; dotted line). Simulations were conducted by drawing sampling variances for each species from an exponential distribution with different values of the rate parameter, λ, for different trees. In all simulations λ_1 was set to 1.0 while λ_2 was varied. The red horizontal dashed line shows the nominal α level of 0.05

parameters, λ_1 and λ_2, of the simulated sampling variances in each tree. Conversely, when sampling error was considered explicitly in the model, the type I error rate was recovered to approximately its nominal level (Figure 6). This serves to emphasize that ignoring sampling error in the estimation of species means will tend to inflate the type I error rate of the method of this article, even if the degree of sampling error is relatively similar among our two or more phylogenies.

3.3 | Error in the phylogeny

In the simulations of this project we invariably assumed that the phylogeny was known without error. Of course, in empirical studies phylogenetic error can be considerable. Prior research has shown that phylogenetic error can introduce bias, rather than simply random error, in the estimation of certain evolutionary parameters, including the rate parameter of Brownian motion evolution, σ^2. In particular, phylogenetic error tends to inflate the estimated value of σ^2 relative to its true value (e.g. Rabosky, 2015). Consequently, we advise fitting evolutionary models to the best point estimate of the tree (obtained, for instance using maximum likelihood, or by computing a maximum clade credibility or a consensus tree from a Bayesian posterior sample), rather than simply averaging over a set of bootstrapped trees or sample of trees from a posterior distribution and assuming that the mean parameter estimate across trees is unbiased. A reasonable strategy instead might be to first compute our parameter estimates using only the best tree, and then proceed to measure uncertainty (due to phylogeny reconstruction) around these estimated values by fitting our model across a sample of trees and calculating dispersion across the set by some measure. The resultant distribution of estimated values of σ^2 is likely to be right-skewed, particularly if phylogenetic uncertainty is substantial. If this is the case than it may be advisable to use quantiles from the posterior set to measure uncertainty in σ^2 due to the phylogeny, rather than the simple variance of estimated values among phylogenies in our sample, although we believe this should also be the subject of further research. As a set of trees can easily be read into R, or generated within R during phylogeny inference using packages such as PHANGORN (Schliep, 2011), it is already straightforward to repeat estimation over a sample of trees from bootstrapping or Bayesian Markov chain Monte Carlo.

3.4 | Trait scale

We think it will probably be most common to compare rates among relatively closely related organisms, such as between different genera of frogs, or between subfamilies of lizards, such as in this study. However, there is no theoretical difficulty in comparing the rates of evolution of like traits between relatively distantly related clades—such as the rate of body size evolution in rodents vs. turtles. (Not that we suppose this comparison would be particularly interesting—we merely identify it as a case of broadly disparate organisms that *could* be compared by the method). In this case, and perhaps under all circumstances, we would advocate transforming the data to a logarithmical scale in advance of analysis. In this way, evolutionary changes of equal proportion are equivalent, regardless of scale (e.g. Adams, 2013).

3.5 | Comparison among sets of trees

A natural extension of this model would be to compare the rate (or process) of evolution among sets of trees. That is if we *a priori* hypothesize that trees 1, 2 and 3 have been evolving with one rate, and tree 4 under another, then we ought to be able to test that hypothesis. In fact, this is fairly straightforward. We merely need to think of this as fitting multiple common-rate models and then summing the log-likehoods. This model can be compared to one in which the rate is held constant among all trees, or, alternatively, a more complex model in which all trees are permitted to evolve with different rates. All of these options are already available in the function *ratebytree* of the PHYTOOLS package, along with various helper methods (such as *AIC*; Akaike, 1974) to facilitate the comparison of alternative models. Obviously, if users elect to fit and compare a wide variety of alternative models then issues of multiple testing should be taken into consideration.

3.6 | Other models of trait evolution

Herein we have focused on the comparison of the rate of evolution of a continuously valued trait on the tree assuming a model of

Brownian evolution through time. The same general approach could be employed to compare different trees under alternative models of evolutionary change, such the Ornstein-Uhlenbeck model (Butler & King, 2004; Hansen, 1997) or the Early-Burst model (Blomberg, Garland, & Ives, 2003; Harmon, Schulte, Larson, & Losos, 2003; Harmon et al., 2010a). In this case, the method would be the same. We would merely accumulate the likelihood across multiple trees under a flexible scenario in which the process of evolutionary change, the position of the adaptive optima, or the rate of change in the tempo of evolution through time is permitted to differ between trees, and then we would compare this likelihood to a model in which some or all of the parameters of the modelled process have been fixed to be constant among phylogenies. Similarly, the approach we have identified could be extended in a relatively straightforward way to identify differences in the rate of discrete character evolution among trees (Lewis, 2001; Pagel, 1994). In this case, we would merely fit two models of discrete character evolution: one in which the transition rates between character states are permitted to be different in different phylogenies, and a second, simpler model in which they are constrained to be equal. Again, we would compare between these alternative models using the standard machinery of likelihoods, just as we have done in this study.

Model comparison for continuous character evolution under alternative models of evolution (specifically, the Ornstein-Uhlenbeck and Early-Burst models; Hansen, 1997; Blomberg et al., 2003), and comparison of both rate and model of evolution for discretely valued traits (Lewis, 2001; Pagel, 1994) have already been implemented as options in the PHYTOOLS package. We feel that exploring the statistical performance of inference under these models is beyond the scope of this article; however, we have conducted some preliminary exploration and they seem reasonable.

4 | CONCLUSION

The rate of evolutionary change through time has long been a popular theme of macroevolutionary research using comparative methods. In 2006, Brian O'Meara and colleagues presented a novel method based on likelihood, one version of which the authors denominated the censored model. This model can be employed to compare the rate of evolutionary change between arbitrary bipartitions of a phylogenetic tree, ignoring the edge or edges separating the two or more partitions identified. Here we show that this method can easily be adapted to the problem of comparing the rate of evolution among trees in which the relationship between the two or various phylogenies is unknown or unmeasured. In this article we have described this approach and presented a brief simulation study examining the statistical properties of the method. We also describe relatively simple extensions of this general approach to the comparison of evolutionary process between trees and to discrete character data. These extensions have already been applied in software. As interest in the heterogeneous nature of evolutionary change in the tree of life continues to grow, we hope this method will be useful for investigators concerned with comparing the tempo of the evolutionary process between different phylogenies. We also hope that this article will serve as a useful guide in this endeavour.

ACKNOWLEDGEMENTS

The authors thank the United States National Science Foundation (DEB 1350474 to L.J.R.) and Colombia's *Departamento Administrativo de Ciencia, Tecnología e Innovación* (*Colciencias*), *Programa Nacional en Ciencias Básicas* (award number 120456934310 to A.J.C.) for supporting portions of this research. The majority of this research was undertaken by L.J.R. during a sabbatical at the *Universidad de los Andes* in Bogotá, Colombia.

AUTHORS' CONTRIBUTIONS

All authors conceived the project and contributed to writing. L.J.R. and A.J.C. developed the method. L.J.R. implemented the method in software and conducted the simulation analyses. L.J.R. wrote the first manuscript draft.

DATA ACCESSIBILITY

Data have been deposited in the Dryad Digital Repository https://doi.org/10.5061/dryad.5st8m (Revell et al., 2018).

ORCID

Liam J. Revell http://orcid.org/0000-0003-0767-4713

REFERENCES

Adams, D. C. (2013). Comparing evolutionary rates for different phenotypic traits on a phylogeny using likelihood. *Systematic Biology*, *62*, 181–192. https://doi.org/10.1093/sysbio/sys083

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, *19*, 716–723. https://doi.org/10.1109/TAC.1974.1100705

Azzalini, A., & Genz, A. (2016). The R package "MNORMT": The multivariate normal and 't' distributions (version 1.5-5). http://azzalini.stat.unipd.it/SW/Pkg-mnormt.

Becker, R. A., Wilks, A. R., Brownrigg, R., Minka, T. P., & Deckmyn, A. (2016). MAPS: Draw geographical maps. R package version 3.1.1. https://CRAN.R-project.org/package=maps.

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. https://doi.org/10.1111/j.0014-3820.2003.tb00285.x

Brunet, F. G., Crollius, H. R., Paris, M., Aury, J.-M., Gilbert, P., Jaillon, O., … Robinson-Rechavi, M. (2006). Gene loss and evolutionary rates following whole-genome duplication in teleost fishes. *Molecular Biology and Evolution*, *23*, 1808–1816. https://doi.org/10.1093/molbev/msl049

Butler, M. A., & King, A. A. (2004). Phylogenetic comparative analysis: A modeling approach for adaptive evolution. *The American Naturalist*, *164*, 683–695. https://doi.org/10.1086/426002

Chasalow, S. (2012). COMBINAT: Combinatorics utilities. R package version 0.0-8. https://CRAN.R-project.org/package=combinat.

Collar, D. C., Near, T. J., & Wainwright, P. C. (2005). Comparative analysis of morphological diversity: Does disparity accumulate at the same rate in two lineages of centrarchid fishes? *Evolution*, *59*, 1783–1794. https://doi.org/10.1111/j.0014-3820.2005.tb01826.x

Collar, D. C., O'Meara, B. C., Wainwright, P. C., & Near, T. J. (2009). Piscivory limits diversification of feeding morphology in centrarchid fishes. *Evolution*, *63*, 1557–1573. https://doi.org/10.1111/j.1558-5646.2009.00626.x

Edwards, E. J., & Smith, S. A. (2010). Phylogenetic analyses reveal the shady history of C_4 grasses. *Proceedings of the National Academy of Sciences of the United States of America*, *107*, 2532–2537. https://doi.org/10.1073/pnas.0909672107

Ehrlich, P. R., & Raven, P. H. (1964). Butterflies and plants: A study in coevolution. *Evolution*, *18*, 586–608. https://doi.org/10.1111/j.1558-5646.1964.tb01674.x

Evans, M. E. K., Smith, S. A., Flynn, R. S., & Donoghue, M. J. (2009). Climate, niche evolution, and diversification of the "bird-cage" evening primroses (Oenothera, sections Anogra and Kleinia). *The American Naturalist*, *173*, 225–240. https://doi.org/10.1086/595757

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. *The American Naturalist*, *125*, 1–15. https://doi.org/10.1086/284325

Felsenstein, J. (2008). Comparative methods with sampling error and within-species variation: Contrasts revisited and revised. *The American Naturalist*, *171*, 713–725. https://doi.org/10.1086/587525

Garland, T. G., Jr. (1992). Rate tests for phenotypic evolution using phylogenetically independent contrasts. *The American Naturalist*, *140*, 509–519. https://doi.org/10.1086/285424

Gavrilets, S., & Losos, J. B. (2009). Adaptive radiation: Contrasting theory with data. *Science*, *323*, 732–737. https://doi.org/10.1126/science.1157966

Gilbert, P., & Varadhan, R. (2016). NUMDERIV: Accurate numerical derivatives. R package version 2016.8-1. https://CRAN.R-project.org/package=numDeriv.

Hansen, T. F. (1997). Stabilizing selection and the comparative analysis of adaptation. *Evolution*, *51*, 1341–1351. https://doi.org/10.1111/j.1558-5646.1997.tb01457.x

Hansen, T. F. (2006). The evolution of genetic architecture. *Annual Review of Ecology, Evolution, and Systematics*, *37*, 123–157. https://doi.org/10.1146/annurev.ecolsys.37.091305.110224

Harmon, L. J., Losos, J. B., Davies, J., Gillespie, R. G., Gittleman, J. L., Jennings, W. B., … Mooers, A. Ø. (2010a). Early bursts of body size and shape evolution are rare in comparative data. *Evolution*, *64*, 2385–2396. https://doi.org/10.1111/j.1558-5646.2010.01025.x

Harmon, L. J., Losos, J. B., Davies, J., Gillespie, R. G., Gittleman, J. L., Jennings, W. B., … Mooers, A. Ø. (2010b). Data from: Early bursts of body size and shape evolution are rare in comparative data. *Dryad Digital Repository*, https://doi.org/10.5061/dryad.f660p

Harmon, L. J., Schulte, J. A., II, Larson, A., & Losos, J. B. (2003). Tempo and mode of evolutionary radiation in iguanian lizards. *Science*, *301*, 961–964. https://doi.org/10.1126/science.1084786

Harmon, L. J., Weir, J. T., Brock, C. D., Glor, R. E., & Challenger, W. (2008). GEIGER: Investigating evolutionary radiations. *Bioinformatics*, *24*, 129–131. https://doi.org/10.1093/bioinformatics/btm538

Hunter, J. P., & Jernvall, J. (1995). The hypocone as a key innovation in mammalian evolution. *Proceedings of the National Academy of Sciences of the United States of America*, *92*, 10718–10722.

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. https://doi.org/10.1080/10635150701313830

Jackson, C. H. (2011). Multi-state models for panel data: The MSM package for R. *Journal of Statistical Software*, *38*, 1–29. https://doi.org/10.18637/jss.v038.i08

Kozak, K. H., & Wiens, J. J. (2010). Accelerated rates of climatic-niche evolution underlie rapid species diversification. *Ecology Letters*, *13*, 1378–1389. https://doi.org/10.1111/j.1461-0248.2010.01530.x

Lemon, J. (2006). PLOTRIX: A package in the red light district of R. *R-News*, *6*, 8–12.

Lewis, P. O. (2001). A likelihood approach to estimating phylogeny from discrete morphological character data. *Systematic Biology*, *50*, 913–925. https://doi.org/10.1080/106351501753462876

Liem, K. F. (1973). Evolutionary strategies and morphological innovations: Cichlid pharyngeal jaws. *Systematic Zoology*, *22*, 425–441. https://doi.org/10.2307/2412950

Ligges, U., & Mächler, M. (2003). SCATTERPLOT3D – An R package for visualizing multivariate data. *Journal of Statistical Software*, *8*, 1–20. https://doi.org/10.18637/jss.v008.i11

Mahler, D. L., Revell, L. J., Glor, R. E., & Losos, J. B. (2010). Ecological opportunity and the rate of morphological evolution in the diversification of Greater Antillean anoles. *Evolution*, *64*, 2731–2745. https://doi.org/10.1111/j.1558-5646.2010.01026.x

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. https://doi.org/10.1111/j.0014-3820.2006.tb01171.x

Pagel, M. (1994). Detecting correlated evolution on phylogenies: A general method for the comparative analysis of discrete characters. *Proceedings of the Royal Society London B*, *255*, 37–45. https://doi.org/10.1098/rspb.1994.0006

Paradis, E., Claude, J., & Strimmer, K. (2004). APE: Analyses of phylogenetics and evolution in R language. *Bioinformatics*, *20*, 289–290. https://doi.org/10.1093/bioinformatics/btg412

Pinheiro, J., Bates, D., DebRoy, S., & Sarkar, D., & R Core Team. (2017). NLME: Linear and nonlinear mixed effects models. R package version 3.1-131. https://CRAN.R-project.org/package=nlme.

Plummer, M., Best, N., Cowles, K., & Vines, K. (2006). CODA: Convergence diagnosis and output analysis for MCMC. *R-News*, *6*, 7–11.

Price, S. A., Holzman, R., Near, T. J., & Wainwright, P. C. (2011). Coral reefs promote the evolution of morphological diversity and ecological novelty in labrid fishes. *Ecology Letters*, *14*, 462–469. https://doi.org/10.1111/j.1461-0248.2011.01607.x

Qiu, W., & Joe, H. (2015). CLUSTERGENERATION: Random cluster generation (with specified degree of separation). R package version 1.3.4. https://CRAN.R-project.org/package=clusterGeneration.

R Core Team. (2017). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Rabosky, D. L. (2015). No substitute for real data: A cautionary note on the use of phylogenies from birth-death polytomy resolvers for downstream comparative analyses. *Evolution*, *69*, 3207–3216. https://doi.org/10.1111/evo.12817

Revell, L. J. (2008). On the analysis of evolutionary change along single branches in a phylogeny. *The American Naturalist*, *172*, 140–147. https://doi.org/10.1086/588078

Revell, L. J. (2009). Size-correction and principal components for interspecific comparative studies. *Evolution*, *63*, 3258–3268. https://doi.org/10.1111/j.1558-5646.2009.00804.x

Revell, L. J. (2012). PHYTOOLS: An R package for phylogenetic comparative biology (and other things). *Methods in Ecology and Evolution*, *3*, 217–223. https://doi.org/10.1111/j.2041-210X.2011.00169.x

Revell, L. J. (2014). Graphical methods for visualizing comparative data on phylogenies. In L. Z. Garamszegi (Ed.), *Modern phylogenetic comparative methods and their application in evolutionary biology: Concepts and practice* (pp. 77–103). Berlin, Heidelberg: Springer-Verlag.

Revell, L. J., González-Valenzuela, L. E., Alfonso, A., Castellanos-García, L. A., Guarnizo, C. E., & Crawford, A. J. (2018). Data from: Comparing evolutionary rates between trees, clades and traits. *Dryad Digital Repository*, https://doi.org/10.5061/dryad.5st8m

Revell, L. J., Harmon, L. J., & Collar, D. C. (2008). Phylogenetic signal, evolutionary process, and rate. *Systematic Biology*, *57*, 591–601. https://doi.org/10.1080/10635150802302427

Schliep, K. P. (2011). PHANGORN: Phylogenetic analysis in R. *Bioinformatics*, *27*, 592–593. https://doi.org/10.1093/bioinformatics/btq706

Schluter, D. (2000). *The ecology of adaptive radiation*. Oxford, UK: Oxford University Press.

Simpson, G. G. (1944). *Tempo and mode in evolution*. New York, NY: Columbia University Press.

Thomas, G. H., Freckleton, R. P., & Székely, T. (2006). Comparative analyses of the influence of developmental mode on phenotypic diversification rates in shorebirds. *Proceedings of the Royal Society B*, *273*, 1619–1624. https://doi.org/10.1098/rspb.2006.3488

Van Valen, L. (1971). Adaptive zones and the orders of mammals. *Evolution*, *25*, 420–428. https://doi.org/10.2307/2406935

Venables, W. N., & Ripley, B. D. (2002). *Modern applied statistics with S* (4th ed.). New York, NY: Springer. https://doi.org/10.1007/978-0-387-21706-2

Xie, Y. (2013). ANIMATION: An R package for creating animations and demonstrating statistical methods. *Journal of Statistical Software*, *53*, 1–27. https://doi.org/10.18637/jss.v053.i01

Yoder, J. B., Clancey, E., Des Roches, S., Eastman, J. M., Gentry, L., Godsoe, W., … Harmon, L. J. (2010). Ecological opportunity and the origin of adaptive radiations. *Journal of Evolutionary Biology*, *23*, 1581–1596. https://doi.org/10.1111/j.1420-9101.2010.02029.x

How to cite this article: Revell LJ, González-Valenzuela LE, Alfonso A, Castellanos-García LA, Guarnizo CE, Crawford AJ. Comparing evolutionary rates between trees, clades and traits. *Methods Ecol Evol*. 2018;00:1–12. https://doi.org/10.1111/2041-210X.12977