DR DAVID ALEJANDRO DUCHENE (Orcid ID : 0000-0002-5479-1974)

Article type : Research Papers

Phylogenetic estimates of diversification rate are affected by molecular rate variation

David A. Duchêne[1,2,*], Xia Hua[1], Lindell Bromham[1]

[1]*Macroevolution & Macroecology, Research School of Biology, Australian National University, Canberra ACT 0200, Australia*

[2]*School of Life and Environmental Sciences, University of Sydney, Sydney NSW 2006, Australia*

Running title: DNA rates affect diversification estimates

* Corresponding author
David A. Duchêne
School of Life and Environmental Sciences
Edgeworth David Building A11
University of Sydney
Sydney, NSW 2006
Australia
Telephone: +61 4 1202 6379
Email: david.duchene@sydney.edu.au

Abstract

Molecular phylogenies are increasingly being used to investigate the patterns and mechanisms of macroevolution. In particular, node heights in a phylogeny can be used to detect changes in rates of diversification over time. Such analyses rest on the assumption that node heights in a phylogeny represent the timing of diversification events, which in turn rests

on the assumption that evolutionary time can be accurately predicted from DNA sequence divergence. But there are many influences on the rate of molecular evolution, which might also influence node heights in molecular phylogenies, and thus affect estimates of diversification rate. In particular, a growing number of studies have revealed an association between the net diversification rate estimated from phylogenies and the rate of molecular evolution. Such an association might, by influencing the relative position of node heights, systematically bias estimates of diversification time. We simulated the evolution of DNA sequences under several scenarios where rates of diversification and molecular evolution vary through time, including models where diversification and molecular evolutionary rates are linked. We show that commonly-used methods, including metric-based, likelihood and Bayesian approaches, can have a low power to identify changes in diversification rate when molecular substitution rates vary. Furthermore, the association between the rates of speciation and molecular evolution rate can cause the signature of a slowdown or speedup in speciation rates to be lost or misidentified. These results suggest that the multiple sources of variation in molecular evolutionary rates need to be considered when inferring macroevolutionary processes from phylogenies.

Keywords

macroevolution, diversification speedup, diversification slowdown, LASER, gamma statistic, TESS, molecular evolutionary rates

Introduction

Macroevolutionary studies have undergone a revitalization, stimulated by the availability of large molecular phylogenies for a wide range of different biological groups. Molecular phylogenies are particularly useful for investigating macroevolution because they represent not only the relationships between living species, but a record of the diversification events that produced the current diversity. For example, molecular phylogenies have been used to show that adaptive radiation can result in the repeated evolution of similar ecomorphs independently on different islands (Gillespie, 2004; Mahler *et al.*, 2013). These studies of adaptive radiation relied on inferring the relationships between taxa among and between islands. But many other macroevolutionary analyses rely explicitly or implicitly on being able to infer the timing of evolutionary events from molecular data. In some cases, molecular dates are used to put specific events into a geological or biological context, for example linking diversification of lineages to periods of climate change (Crisp & Cook, 2013) or

changing ocean currents (Crottini *et al.*, 2012). Increasingly macroevolutionary studies are using molecular phylogenies to infer the timing of all nodes in the tree to infer patterns of diversification over time (Rabosky, 2006a; Harmon *et al.*, 2008). Researchers now routinely use these kinds of analyses to compare rates of diversification among groups, between areas, or over time (Ricklefs, 2007; Morlon, 2014).

Macroevolutionary analyses of molecular phylogenies have reinvigorated key debates in macroevolution and macroecology concerning the mechanisms of diversification, in particular fuelling a discussion on whether biodiversity continues to increase over time or whether it slows down as a "carrying capacity" for species is reached. Analyses that show a decreasing rate of phylogenetic branching events towards the tips of the phylogeny have been used to argue for the occurrence of slowdown in evolutionary rates over time (Phillimore & Price, 2008; Machac *et al.*, 2013). This result has been used to suggest that speciation rates are limited by the environmental capacity of the region to support more species: in other words, the process of diversification is density dependent, and slows as more species are added.

The generality and significance of slowdown in diversification rates has been vigorously debated (Moen & Morlon, 2014). Possible measurement artefacts that could lead to reduction in inferred speciation rate towards the tips of the phylogeny have been discussed. Many critiques have focused on the artificial generation of a pattern of slowdown through incomplete sampling. This incompleteness includes cases when extinct lineages result in missing history of diversification (Rabosky, 2009), when recent lineage divergence events are unrepresented due to lack of sampling of sub-specific lineages (Etienne & Rosindell, 2012), or if sampling is non-random across the phylogeny (Welch *et al.*, 2005; Cusimano & Renner, 2010). Tests of the veracity of phylogenetic tests of diversification dynamics have predominantly involved simulating the evolution of phylogenies under different macroevolutionary scenarios and asking whether the reconstructed history is a fair or biased estimate of the true history of diversification (Cusimano & Renner, 2010; Pigot *et al.*, 2010; Etienne & Rosindell, 2012; Quental & Marshall, 2013). But these tests implicitly assume that the phylogeny of the sampled species can be estimated accurately, and that any error is stochastic variation in the pattern of diversification from that phylogeny. This is because the simulations used to test macroevolutionary methods are based on birth-death models of lineage diversification. While the test may involve sampling simulated trees, it does not

model the error associated with the reconstruction of the tree itself. Even in cases where empirical phylogenies are used (rather than simulated birth-death trees), the error in phylogeny estimation is not modeled (Cusimano & Renner, 2010).

Surprisingly, there have been few studies that examine the reliability of inference of diversification rate from molecular phylogenies in light of potential biases in the reconstruction of time-scaled phylogenies from DNA data. Macroevolutionary inference from molecular phylogeny relies on the ability to correctly infer node heights as timing of speciation events. Estimating absolute or relative timescales from phylogenies relies on being able to correctly model the relationship between amount of accumulated molecular change and time since divergence. Molecular changes and divergence times are related by the substitution rate over the phylogeny, which must be modeled in phylogenetic inference. But substitution rates can be influenced by species traits, demographic history, and environmental variables, which themselves may evolve over the phylogeny. Therefore, we expect substitution rates to show complex patterns of variation across many phylogenies.

While recognition that substitution rates vary across lineages has led to some skepticism about the reliability of molecular dates of divergence, these fears have been largely allayed by the development of rate-variable molecular dating methods, which allow branches on the phylogeny to have different rates of molecular evolution. However, rate-variable dating methods may not provide a universal panacea to the problem of inferring dates from DNA sequences. These methods rely on a large number of assumptions about the evolutionary process and the nature of the dataset, and when these assumptions are not met, the results can be misleading (Ho *et al.*, n.d.; Welch & Bromham, 2005; Ho & Duchêne, 2014; Duchêne *et al.*, 2015a; b). Often we do not know which (if any) assumptions are realistic for a given dataset, so the choice made is somewhat arbitrary, yet these choices can have a large impact on the estimates of the timing of speciation events (Lepage *et al.*, 2007; Linder *et al.*, 2011; Bellot & Renner, 2014; Crisp *et al.*, 2014; Duchêne *et al.*, 2015a).

Despite the central importance of accurate inference of node heights for phylogenetic studies of macroevolution, studies of diversification rate give little attention to the problem of molecular rate variation. This is particularly a concern given the potential for macroevolutionary processes themselves to be associated with changing rates of molecular evolution. A significant correlation between inferred substitution rates and net diversification

rate has been noted for a range of datasets using a variety of methods (Barraclough *et al.*, 1996; Barraclough & Savolainen, 2001; Webster *et al.*, 2003; Pagel *et al.*, 2006; Eo & DeWoody, 2010; Lanfear *et al.*, 2010; Duchene & Bromham, 2013; Bromham *et al.*, 2015). These studies suggest that even if the link between diversification and molecular evolution is not universal (Goldie *et al.*, 2011), it is common enough to deserve our attention. The cause of the correlation between rates of molecular evolution and net diversification has been debated. It has been suggested that speciation accelerates the rate of molecular evolution (Venditti & Pagel, 2010), or that greater rates of molecular evolution drive higher diversification rates (Lanfear *et al.*, 2010; Hua & Wiens, 2013; Bromham *et al.*, 2015), or that the two may be indirectly linked through environmental factors (Davies *et al.*, 2004). Whatever the underlying cause of the correlation, if substitution rate is accelerated in lineages that have high net diversification rate then the relative distance between nodes might be increased, changing the estimates of diversification rate. Existing models of the substitution rate across lineages do not explicitly account for a link between diversification and molecular evolution. Instead, existing methods assume that variation can be modeled without describing the mechanisms driving molecular evolution (Lepage *et al.*, 2007; Ho & Duchêne, 2014).

Here we ask whether variation in the rate of molecular evolution can influence the accuracy of reconstruction of patterns of diversification rate over time, particularly the inference of diversification changes (speedup or slowdown). We focus on a particular pattern of rate variation that might be of concern to researchers using phylogenies to reconstruct diversification rate: the positive association between rate of diversification and rate of molecular evolution. Therefore, we use simulations to explore whether commonly used molecular phylogenetic estimates of diversification rate are influenced by an association between substitution rate and diversification rate. We simulate the evolution of DNA sequences under a number of models that allow both speciation rate and substitution rate to vary. We use the resulting DNA sequence to reconstruct the phylogenetic divergence times, which we then used in a number of common methods for analyzing diversification rates through time. Then we ask whether these commonly used phylogenetic methods can correctly recover temporal patterns in the relative rate of diversification.

METHODS

Our aim in this study is to ask whether variation in rate of molecular evolution might influence the detection of patterns of diversification rates over time from molecular phylogenetic analysis. We restrict our investigation to a defined set of macroevolutionary models that describe the relationship between rates of molecular evolution and patterns of diversification. In this section, we first describe these models, and then explain how we parameterized those models using data from the literature in order to provide realistic simulations. Then we describe how we simulated the evolution of DNA sequences under different models of rate variation and temporal patterns of diversification. We reconstructed the evolutionary history of these simulated sequences using commonly employed phylogenetic methods, and compare the reconstructions to the known history of the sequences. Because we want to find out whether these models that link diversification rates to molecular evolution rates could change our view of macroevolutionary processes, we use a number of popular methods for detecting changes in diversification rates over time, and ask whether the reconstructed data gives an accurate picture of the true history of the sequences.

Macroevolutionary models

We simulated phylogenies under six evolutionary birth-death models, which differ in the pattern of change in diversification rate over time, and in the relationship between speciation rate and rate of molecular evolution. In each model, both speciation rate and the rate of molecular evolution vary over time, but extinction rate is constant. The models differ in two respects: whether the change in speciation rate is stochastic or directional (biased toward increases or decreases), and whether changes in the rate of molecular evolution are linked to the speciation rate or independent from it (Table 1).

1) ***Stochastic-Unlinked (STU):*** Speciation rate and molecular evolution rate both change stochastically, independently of each other
2) ***Stochastic-Linked (STL)***: Speciation rate changes stochastically, and changes in molecular evolution rate are positively linked to speciation rate.
3) ***Slowdown-Unlinked (SLU):*** Speciation rates have a tendency to decrease over time; molecular evolution rate varies stochastically and independently of speciation rate.
4) ***Slowdown-Linked (SLL):*** Speciation rates have a tendency to decrease over time, and changes in molecular evolution rate are positively linked to speciation rate.

5) ***Speedup-Unlinked (SPU):*** Speciation rates have a tendency to increase over time; molecular evolution rate varies stochastically and independently of speciation rate.
6) ***Speedup-Linked (SPL)***: Speciation rates have a tendency to increase over time, and changes in molecular evolution rate are positively linked to speciation rate.

Selecting realistic parameters for simulations

In order to make our simulations representative of typical macroevolutionary studies, we need to select parameters that are reasonable representations of published studies. To determine an appropriate number of taxa per phylogeny, sequence length, and reconstruction methods, we sampled one hundred studies published over a ten-year period (2005-2015) that used phylogenetic analyses to detect changes in diversification rate (see Supplementary Information for details, Table S1). We accessed the relevant literature by searching in the freely available scholarly literature database Google Scholar. The search had the term *phylogeny "diversification rates"*, and was used to extract the first 100 articles within the selected time-period that used molecular phylogenies to make estimates of diversification rates or diversification rate dynamics. Based on the median values from this sample (see Table S2), we set the size of our simulated dataset to be 150 sequences of length 4000 bases. We also selected two of the most commonly used phylogeny reconstruction methods from the studies included in our survey: nonparametric rate smoothing, NPRS (Sanderson, 1997) and Bayesian "relaxed clock" phylogenies, as implemented in BEAST (Drummond *et al.*, 2006).

Our simulations require a set value for extinction rate and a starting value for speciation rate and molecular evolution rate. In order to condition our simulations to reasonable values, we used published estimates of these parameters for birds, as an example of a diverse group for which a link between diversification rates and molecular evolution rates has been demonstrated (Eo & DeWoody, 2010; Lanfear *et al.*, 2010). In all simulations we included a constant background extinction rate of 0.01 per time step per lineage, which is approximately the mean extinction rate reported in the past for birds (Jetz *et al.*, 2012). Speciation rates are drawn from a distribution with a maximum of 4.64 per million years per lineage, which represents the highest rate of diversification estimated for family-level bird lineages (Jetz *et al.*, 2012). The minimum speciation rate was 0.05. The maximum rate of molecular evolution was set to 0.01 substitutions per site per million years, based on estimates for avian nuclear genes, and the minimum rate of molecular evolution was set to 1×10^{-7} substitutions per site per million years (van Tuinen & Hedges, 2001).

All simulations began with the same initial rate of molecular evolution, 2.47×10^{-3} substitutions per site per million years, which is typical substitution rate estimates for avian nuclear genes (van Tuinen & Hedges, 2001), and is similar to some mitochondrial gene estimates for birds (Pereira & Baker, 2006). We chose a starting speciation rate of 0.10, which has previously been estimated as the median speciation rate for the birds (Jetz *et al.*, 2012). Each of our simulations follows a stochastic process forwards in time, with discrete time steps of 0.1 million years. At each time step, sequentially, we used a probability distribution to determine whether a speciation, extinction, or substitution event occurred. Speciation and substitution rates were sampled from a multivariate distribution with variables on a log-log scale, and a covariance of 2.5×10^{-7}, which has been used to describe the relationship between rates of diversification and rates of molecular evolution (Lanfear *et al.*, 2010). With these parameters, our final alignments had between 200 and 1000 variable sites.

Simulating evolution of DNA sequences under macroevolutionary models

Our aim in these simulations is to produce sequences under a biologically reasonable model of evolution, then to reconstruct the phylogenetic history of these sequences using commonly employed methods. Because variations in rate of molecular evolution and speciation rate are common, our model allows both speciation rate and molecular evolution rate to vary.

For substitutions in the DNA sequence, we used a general time reversible model (Tavaré, 1986) to determine which base a given position in the sequence would change to (R-matrix = 1.3472, 4.8145, 0.9304, 1.2491, 5.5587, 1.0000; base frequencies: A=0.2628, C=0.2605, G=0.2436, T=0.2331)(Murphy *et al.*, 2001). The simulation does not allow insertions or deletions, only changes to single bases in the sequence. If a speciation event occurs, the sequence is duplicated, and each copy then evolves independently of the other.

To determine the probability of each event occurring at a time step, we used the rate of events relative to each other. The following equations describe the total rate of an event of any kind, ρ, and the probability for each of the events occurring at a given time step of size dt, where λ is the speciation rate, ε is the extinction rate, s is the substitution rate, and l is the sequence length:

$$\rho = \lambda + \varepsilon + sl$$

$$P(event) = \rho e^{-\rho dt}$$

$$P(speciation \mid event) = \frac{\lambda}{\rho}$$

$$P(extinction \mid event) = \frac{\varepsilon}{\rho}$$

$$P(substitution \mid event) = \frac{sl}{\rho}$$

The simulation was run, sampling events at each time step, until the required number of extant taxa was reached, giving a phylogeny with 150 tips. Each run of this procedure generates an alignment (a set of contemporaneous sequences, which are by default aligned as they have no changes in sequence length or arrangement) and a phylogeny (the series of diversification events that generated the alignment, which represents the "true history" of evolutionary events that produced those sequences). The true history encapsulated in the simulated phylogeny includes lineages that terminated in an extinction event, and did not give rise to a sequence in the alignment. However, for the purposes of this study we will consider the "true tree" to be the series of bifurcation events that gave rise to the sequences in the alignment, not included extinct lineages. The code for running the simulations is available from GitHub (github.com/duchene/moldivlink).

Phylogeny reconstruction

For each of the six macroevolutionary models (Table 1), we produced 100 simulated phylogenies, using the procedures and parameter values described above. Each of these simulations produces a sequence alignment and a "true" phylogeny, which is the known history of diversification events that produced the observed sequences. We then estimated the phylogeny using only the sequence alignment, to give a reconstructed phylogeny.

For all phylogeny reconstruction analyses, we used a general time-reversible substitution model, with parameters estimated from the data as part of the inference procedure. Because there are no insertions, deletions or rearrangements, sequences are already perfectly aligned so we did not need to use alignment algorithms. In order to ask whether commonly used methods to detect changes in diversification rate could correctly infer the true history of these sequences, we reconstructed the phylogeny of the sequences using two commonly used phylogenetic methods.

One common reconstruction method, employed in just over half the of studies in our literature sample (Figure S1), was an autocorrelated rates model, which uses a maximum likelihood rate-smoothing procedure to assign similar rates to related branches (Sanderson, 1997). We implemented this method using non-parametric rate smoothing (NPRS) in the R package "ape" (Paradis *et al.*, 2004). In this method, an estimate of the optimal distribution of rates is obtained from a sequence-based cross-validation procedure. For each analysis, we selected a value of the smoothing parameter between minus one and six on a $\log_{10}$ scale in increments of 1 with the lowest cross-validation score (Sanderson, 2002). It was also necessary to estimate branch lengths in substitutions per site independently beforehand, which was done in the R package "phangorn" (Schliep, 2011). We used a single root calibration with a confidence interval of ±5% of the true age of the root node, derived from the true phylogenetic history for each simulated dataset.

The other common phylogeny reconstruction method, used in over a third of the sampled studies, was an uncorrelated rates model, where rates across branches are independent and identically distributed using a lognormal distribution (Drummond *et al.*, 2006), implemented in a Bayesian framework using the software BEAST 2.2 (Bouckaert *et al.*, 2014). We used a birth-death tree prior and a normally distributed root calibration prior with a standard deviation of 5% of the true age of the root node. All Bayesian analyses had a chain length of 20 million steps, sampled every two thousand steps. We discarded the burn-in after 2 million steps and visually checked for convergence of the likelihood. We also calculated effective sample sizes using the R package "CODA" (Plummer *et al.*, 2006): we considered only those phylogenies where the parameters had values above 200, otherwise we discarded the results of the analysis and reran the analysis. As per common practise in published studies (Table S1), we used the maximum clade credibility tree from each posterior distribution for subsequent analyses (Rannala & Yang, 1996).

Estimating diversification rates

We now have a true tree (the known history of the sequences) and a reconstructed tree (the inferred history of the sequences given the phylogenetic reconstruction method) for one hundred simulated datasets for each of the six models. Now we wish to ask whether commonly used macroevolutionary analysis methods can accurately detect the pattern of diversification that gave rise to those sequences. Although a wide variety of methods are used

to detect changes in diversification rate from phylogenies, we chose two representative methods from our sample of publications, plus one more recent method for comparison.

The first method is a metric-based method. The gamma statistic (used in one third of the sampled studies) describes the relative distribution of nodes in a phylogeny. This statistic is usually compared to the expected value for a phylogeny with a constant diversification process and with no extinction (Yule process), so this comparison is often referred to as the constant rates (CR) test (Pybus & Harvey, 2000). Under a constant speciation rate with no extinction, nodes will accumulate exponentially from the root to the tip, which gives gamma values distributed as a standard normal around zero. A slowdown in speciation rate over time is expected to produce a phylogeny with an excess of nodes in the early part of the history, which gives negative gamma values. When the gamma value is smaller than -1.645, we reject constant speciation rate, in support of a slowdown in speciation rate. A speedup in speciation rate will result in more nodes near the tips, giving positive gamma values. When the gamma value is larger than 1.645, we reject constant speciation rate, in support of a speedup in speciation rate. The constant rates test, based on these values, has been used in a large number of studies to determine the underlying diversification process (Figure S3). We calculated gamma for both the true tree (the expected value of gamma, γ_{exp}) and the reconstructed tree (the observed value of gamma, γ_{obs}) using the R package "ape" (Paradis *et al.*, 2004).

The second method is a maximum likelihood model-fitting procedure. There are a range of available methods but we selected the most commonly used implementation for these tests, from the R package 'LASER' (Rabosky, 2006a), which was used in over a quarter of the sampled studies (Figure S3). This approach starts by fitting alternative diversification models: a birth-death model with constant rates of speciation and extinction over the phylogeny, a slowdown model with decreasing rates of speciation over time, and a speedup model with increasing rates of speciation over time. Then, the method tests whether the model with constant rates can be rejected in favour of either of the other two models. This is done by approximating the distribution of the Akaike Information Criterion (AIC; Akaike, 1992) under the model with constant rates. The constant rates model is rejected if the AIC for a variable rates model falls outside the 95th percentile of the approximated distribution (Rabosky, 2006b).

The third method is a Bayesian model-fitting procedure. We selected a recently proposed approach implemented in the R package 'TESS' (Höhna *et al.*, 2015), although it is too recent to have been included in any of the papers in the literature survey (Table S1). We used this method to assess the same three models as in the 'LASER' package (see above). We used stepping stone sampling to estimate marginal likelihoods for each model, which is an estimate of model fit (Xie *et al.*, 2011). This model fitting procedure requires the user to provide priors for the parameters, which can be informative about the expected parameter values. We aimed to use minimally informative priors so that the data has relatively greater power to override the priors, so we used broad priors that provide little information about each parameter. The birth-death model was used with two parameters: diversification rate (prior exp[rate = 10]) and extinction rate (prior exp[rate = 10]). The other models (dec and inc) were used with three parameters: the extinction rate (prior exp[rate = 0.1]), the initial rate of speciation (prior exp[rate = 10]), and the rate of decay of speciation rates through time (prior exp[rate = 0.1]). Following standard practice in Bayesian statistics, we rejected the birth-death model if the ratio between its marginal likelihood and the competing model (the Bayes factor) was > 2 (Kass & Raftery, 1995; Lartillot & Philippe, 2006).

Comparing reconstructed diversification rates to true history

Because we wish to know whether the phylogenetic estimates of the pattern of diversification accurately capture the true history of the sequences, we apply the three methods for estimating diversification process to both the true tree (produced by the simulation) and the reconstructed tree (produced by analyzing the simulated sequences) for each simulated dataset. For the reconstructed tree, we used the maximum credibility tree. Uncertainty in the node age estimates is likely to have an impact on diversification rate estimates. We found our datasets to be highly informative, i.e., there is low overlap in node age estimates under different macroevolutionary models (Figure S2), such that this form of uncertainty is unlikely to change our results significantly. In analyses of empirical data with a small number of loci, it might be a better practice to consider the uncertainty in estimates of node ages, rather than the standard procedure of considering only the maximum clade credibility tree.

We then ask whether the analyses applied to the reconstructed phylogeny accurately reflect the underlying diversification process (the model under which the sequences were evolved). More specifically, for each method we asked: (i) does it correctly identify the underlying diversification process from the simulated (true) phylogenies? (ii) does it correctly identify

the diversification process from the reconstructed phylogenies (from NPRS and BEAST)? And (iii) can it identify the underlying diversification process when the rates of molecular substitution and rates of diversification are linked (models SLL and SPL)?

In models STU and STL, the speciation rate varies randomly over the tree without any tendency to speed up or slow down (Table 1). Under this condition, we expect a gamma value that does not indicate either speedup or slowdown, so that it falls approximately between -1.645 and 1.645, and we expect that model selection using LASER or TESS will not reject the birth-death process. For the slowdown models (SLU and SLL), speciation rates have a trend of decreasing over time (Table 1), so we expect a gamma value less than -1.645, and that model selection with LASER and TESS will reject the birth-death model in favour of speciation rate slowdown. For the speedup models (SPU and SPL), speciation rates have a trend of increasing over time, so we expect a value for gamma above 1.645, and that model selection with LASER and TESS will reject the birth-death model in favour of speciation rate speedup. For each of the models in Table 1, we have 100 simulated datasets, each of which lead to three phylogenies (true, reconstructed with NPRS, and reconstructed with BEAST). Each of these phylogenies was examined using the three macroevolutionary methods, gamma, LASER and TESS. For each set of 100 phylogenies and each macroevolutionary method, we report the proportion of replicates in which the corresponding diversification process is identified. To compare this proportion across schemes, we calculated the confidence interval around each sample proportion using a normal approximation. Our analyses were made using standard thresholds for rejecting the null model as used in previous research. In addition, we compared the power of the three macroevolutionary methods in terms of their sensitivity and specificity across a range of thresholds, using ROC curve analysis as implemented in the R package 'pROC' (Robin *et al.*, 2011).

RESULTS

Our first step was to investigate how often each of the three macroevolutionary methods correctly identified the pattern of diversification under which phylogenies were generated for the "true" simulated phylogenies. This tests the ability of each method to detect diversification rate trends when the phylogeny is known without error. When speciation rates vary stochastically with no directional change (models STU and STL), gamma and TESS identified the model with no directional change in diversification rate in over 80% of the simulated phylogenies (Table 2), and LASER identified the constant rates model in 65% of

the simulated phylogenies. When speciation rates varied with a tendency of slowdown over time (models SLU and SLL), gamma and LASER identified slowdown in over 70% of the simulated phylogenies, while TESS identified slowdown in 55% of the simulated phylogenies. When speciation rates varied with a tendency of speedup over time (models SPU and SPL), gamma and LASER identified speedup in over 70% of the simulated phylogenies and TESS identified speed up in all the simulated phylogenies. These results suggest that, within the parameter space explored in this simulation study, the commonly used macroevolutionary methods have reasonable power to detect directional change in speciation rates when speciation rates vary stochastically, although LASER is less reliable when speciation rates vary stochastically with no directional change and TESS is less reliable when speciation rates slow down.

Next, we investigate how often the macroevolutionary methods can identify the temporal pattern of diversification rates, when applied to phylogenies reconstructed from the simulated sequences. When speciation rates and molecular substitution rates vary independently of each other (models STU, SLU and SPU), the power of gamma to detect directional changes in speciation rate is significantly reduced (compared to its power on simulated (true) phylogenies), largely due to a much wider spread of gamma values in the reconstructed trees (Figure A). Similarly, LASER has significantly reduced power to identify directional changes in speciation rates from reconstructed phylogenies than from simulated (true) phylogenies (Figure B). This suggests that both gamma and LASER are less able to detect speciation rate slowdown or speedup when there is variation in rate of molecular evolution, even when such variation is modelled in the phylogenetic inference method. In contrast, the power of TESS to detect directional changes in speciation rate is similar for both the true and reconstructed trees (Figure C), suggesting that it is less affected by uncertainty in branch length reconstruction under rate variation than gamma and LASER are. However, TESS is relatively poor at detecting speciation rate slowdown in both the true trees and the reconstructed trees (Table 2). Analyses of ROC curves show that diversification rate estimates using TESS have a slightly higher true positive rate and lower false positive rate, suggesting that this approach has better statistical performance than the other methods when making inferences on diversification patterns (Figure S4).

In terms of the methods to reconstruct phylogenies, speedup was correctly identified significantly more often in NPRS reconstructions than in those from BEAST (Table 2). While slowdown was correctly identified more often in BEAST reconstructions than in those from NPRS, the difference was not significant in all comparisons. These results suggest that the phylogeny reconstruction method does have an impact on the ability to detect macroevolutionary dynamics from phylogenies, presumably due to differences in the ways that NPRS and BEAST infer branch lengths under variable substitution rates. As shown by ROC curve analysis, analyses using BEAST have a slightly higher true positive rate and lower false positive rates, suggesting these analyses might have better accuracy in phylogeny reconstruction (Figure S4).

Lastly, we examine whether a link between speciation rates and molecular substitution rates affect our ability to detect the underlying diversification dynamics, given that no commonly used methods explicitly account for such an association. When there is no directional change in speciation rates (models STU and STL), all three macroevolutionary methods tested have similar power to correctly identify that there has been no directional trend in speciation rates, whether or not speciation rates are linked to molecular rates. This suggests that the association between speciation rates and molecular substitution rates does not affect our ability to infer the underlying diversification process, as long as speciation rates do not have a consistent directional trend (speedup or slowdown).

However, when there is a directional change in speciation rates, and molecular rates vary either stochastically or in association with speciation rates, all the three macroevolutionary methods have lower power to identify either slowdown or speedup in speciation rates. We find this difference to be significant in several cases where the confidence intervals for the proportions of time that each process is identified do not overlap (Table 2). For example, when molecular dates are linked to speciation rates, gamma and TESS show a notable reduction in the power to detect speedup in diversification, and LASER show reduction in its power to detect slowdown. These results suggest that the power to detect changes in speciation rate is reduced when molecular rates and speciation rates are linked. In fact, the association between speciation rate and molecular substitution rates can sometimes lead to the false inference of the opposite pattern to the model under which the data was generated, such that a slowdown in speciation rates is detected as a speedup and vice versa (Figures A, B, C).

Discussion

Molecular phylogenies have become a valuable tool in macroevolutionary study, but they are typically employed with relatively little attention given to the underlying evolutionary processes that produced the DNA sequence data. The performance of these macroevolutionary methods is usually tested on simulated phylogenies, where the height of nodes is known with certainty. But for real molecular phylogenies, accurate inference of node heights is challenging due to the complicated and dynamic nature of molecular evolution (Ho *et al.*, n.d.).

While methods that allow for variation in rate of molecular evolution have been developed, these methods generally employ stochastic models of rate variation, where increases and decreases in rate are equally likely (Ho *et al.*, n.d.; dos Reis *et al.*, 2015). But we know that some variation in substitution rates is associated with lineage-specific traits, such as life-history, lifestyle and, potentially, environment. Since these properties may evolve in directional trends along lineages, it is possible that the stochastic models of rate variation do not fully capture this variation, and the relative position of nodes might therefore be affected by a failure to fully capture variation in substitution rates. Of particular concern for phylogenetic studies of diversification rate is the growing number of studies reporting an association between rate of diversification and rate of molecular evolution (reviewed in Hua & Bromham, 2017), as this has the potential to confound estimates of diversification rate from molecular phylogenies.

For this study, we simulated datasets under a biologically reasonable model of evolution to produce the kind of datasets that might typically be analyzed for macroevolutionary patterns, and we parameterized these simulations using values from empirical studies of molecular evolution and diversification rates. To make sure our results are relevant to common practice, we based the size of the dataset (150 sequences, 4000 bases long) and the methods of phylogeny reconstruction and macroevolutionary analysis on a sample of 100 published studies (Table S2). But we also included a newer method to allow for the fact that published studies may tend to represent previously popular methods and may not yet reflect recent suggested improvements.

By first analyzing the "true" simulated trees, we can test how well the methods perform when the phylogenies are known without error. In this case, the only source of uncertainty arises

from the stochastic nature of the birth-death process under which the trees were simulated. We show that for datasets typical of macroevolutionary studies, these methods can correctly identify diversification rate speedup or slowdown in the majority of cases. However, the misidentification rate is relatively high. For example, for the simulated trees, LASER fails to detect speedup in nearly a quarter of phylogenies, but falsely infers speedup in over one fifth of phylogenies generated under a model where speciation rates vary stochastically (Figure B). In contrast, TESS fails to detect diversification rate slowdown in nearly half of the simulated phylogenies: when speciation rates slowdown over time, TESS selects the non-directional diversification rate model almost as often as the slowdown model (Figure C).

Given that the models used in LASER and TESS are mathematically equivalent, the differences in the performance of the two methods are likely due to the differences between the model selection methods used. To select among candidate models LASER uses AIC, calculated using the maximum likelihood point estimate. Meanwhile, TESS uses Bayes factors, calculated by estimating the marginal likelihood of the model, which marginalizes over all possible parameter values. The marginal likelihood and AIC are not necessarily proportional, particularly when there is a large range of parameter values that give similar likelihoods to the maximum likelihood, and there is high uncertainty in parameter estimations. Depending on which model has high uncertainty, AIC can reject constant rates more or less often than Bayes factors. For example, cases where both LASER and TESS correctly identify slowdown have less uncertainty in the estimation of extinction rate in the constant rates model (Figure S5a), compared with cases when only LASER correctly identifies slowdown (Figure S5b). Similarly, cases where both LASER and TESS correctly identify speedup have less uncertainty in the speedup model (Figure S5c), compared with cases where only TESS correctly identifies speedup (Figure S5d). This suggests that no one model selection criterion performs best under all conditions. AIC may be more powerful to test slowdown, because uncertainty in the constant rates model can overwrite signals of slowdown.

We have demonstrated that all three of these methods perform less well on reconstructed phylogenies than they do on the “true” simulated phylogenies. This results warns us that testing macroevolutionary methods on simulated phylogenies without allowing for the uncertainty in phylogenetic inference from molecular data risks painting an overly optimistic picture of the power of these methods to reveal the underlying macroevolutionary processes.

Both of the reconstruction methods tested here, NPRS and BEAST, allow for variation in rates of molecular evolution, but reconstruction of node heights is subject to uncertainty. This error is most notable when rates of molecular evolution are linked to speciation rates: both reconstruction methods move the excess of nodes in the early part of the trees simulated under diversification rate slowdown towards the tips, and move the excess of nodes near the tips of the trees simulated under diversification rate speedup towards the root (Figure D). In other words, the reconstruction methods tend to erase signals on directional change in diversification rate when rates of molecular evolution are linked to diversification rate. This explains the decrease in the power of macroevolutionary methods to detect directional change in diversification rate.

Our data were not simulated under the models used in the macroevolutionary methods, but under biologically reasonable models and parameters, so as to mimic potential real datasets. For example, we do not simulate under a constant speciation rate, as in the birth-death model, but under a model where speciation rate can vary stochastically (but non-directionally) over time. So we are not testing whether the methods can recover the birth-death model from the data that was simulated under the birth-death model. Instead, we want to know whether these methods can reveal the underlying macroevolutionary processes from real datasets, and in this study we take the first step to answer this question by applying the methods to simulated datasets under more biologically reasonable models. Since these macroevolutionary methods are applied to a wide range of empirical datasets, it is important that their performance is evaluated not just on the specific models incorporated in the programs (e.g. the birth-death, diversification speedup, and diversification slowdown models in LASER and TESS) but under a much wider range of plausible macroevolutionary scenarios. In particular, our tests of macroevolutionary methods must incorporate possible confounding elements, such as variation in rates of both speciation and molecular evolution over time, and allowing for the interplay between rates of molecular evolution and diversification.

The association between diversification rates and rates of molecular evolution detected in the analysis of molecular phylogenies is emerging as a general pattern, reported in a wide range of taxa including a range of vertebrate and angiosperm taxa (Barraclough *et al.*, 1996; Barraclough & Savolainen, 2001; Webster *et al.*, 2003; Pagel *et al.*, 2006; Eo & DeWoody, 2010; Lanfear *et al.*, 2010; Duchene & Bromham, 2013; Ezard *et al.*, 2013; Bromham *et al.*, 2015). It seems unlikely that the association between rates of molecular evolution and

diversification rate is due to measurement artefact, as it has been detected in a wide variety of taxa, and also by a range of methods, including root-to-tip distances, tip-branch lengths, DNA distances and model-based estimates of synonymous substitution rate. Because the relationship has been detected for comparisons of synonymous substitution rates, which reflect differences in mutation rate, it seems likely that lineages that have higher mutation rates have higher speciation rates, possibly because they more rapidly develop genomic incompatibility between recently diverged populations (Hua & Bromham, 2017). For the purposes of this study, the cause of the link between diversification and molecular rate is immaterial, all that matters is that the link has been detected in a wide range of phylogenies (Webster *et al.*, 2003; Pagel *et al.*, 2006), so is likely to be a relatively common problem in the analysis of macroevolutionary patterns from molecular phylogenies.

Given that this relationship between diversification rate and molecular rates is common, and that we have shown it may reduce the power of phylogenetic methods to detect temporal patterns in diversification rate, and may indeed lead to false inference of temporal patterns, how can we increase the reliability of macroevolutionary inference from molecular phylogenies? Experience shows that changing the assumptions of molecular phylogenetic methods, including the substitution model and placement of calibrations, can have a substantial effect on the relative heights of nodes within the tree. Such variation will generally not be captured in confidence intervals from a single analysis, nor by using a sample of trees from the posterior, because these only represent the uncertainty in an analysis with a single model and set of assumptions. Therefore, studies are needed that examine the effect of uncertainty in phylogenetic inference on macroevolutionary studies. Future research might also focus on developing molecular rate models for inference of node ages that consider the link between the rate of diversification and the rate of molecular evolution. One approach to do this is by using data from morphological or life history traits that link the two processes. These approaches could be augmented further by including life history data from fossil samples for informing the strength and nature of the link.

While our investigation is specifically aimed at the inference of temporal patterns in speciation rate, it points to a more general need to critically examine the performance of phylogeny reconstruction methods under a range of models of rate variation. Testing “relaxed clock” methods only under stochastic changes in substitution rates will not tell us whether these methods perform well when there are directional changes in rates. Directional changes

in rate can be expected in many different macroevolutionary scenarios. For example, faster rates are associated with life history traits such as smaller body size, shorter generations, shorter lifespans and/or higher fecundity in many lineages (Bromham, 2009), including mammals (Welch *et al.*, 2008), fish (Hua *et al.*, 2015), invertebrates (Thomas *et al.*, 2010), plants (Bromham *et al.*, 2015), foraminifera (Ezard *et al.*, 2013) and bacteria (Weller & Wu, 2015). Our results show that if these life history traits evolve over a phylogeny in stochastic manner, then we can expect that this rate variation will introduce some error into diversification rate estimates performed on phylogenies, even when estimated under "relaxed clocks". However, if these life history traits evolve in a directional manner, then the error in phylogenetic diversification rates will be greater. For example, it has been suggested that studies of the timing of the radiation of placental mammal orders have been influenced by the association between body size and molecular evolution, in two ways. Firstly, average body sizes increased in most ordinal lineages during the diversification, potentially causing a directional slow down in rates, which may result in underestimation of diversification dates. Secondly, preferential calibration on larger bodied lineages may underestimate rates of molecular evolution across the phylogeny, potentially resulting in overestimation of node depths, pushing dates of the radiation back in time.

These examples, in addition to the results from this study, illustrate the potential for interaction between patterns of variation in rate of molecular evolution and macroevolutionary inference of diversification rates and patterns from molecular phylogenies. We need more studies on the influence of non-random patterns of rate of molecular evolution on the accuracy of branch length estimation on molecular phylogenies. We also need to test macroevolutionary methods not just on perfectly-known simulated phylogenies but on imperfectly-estimated molecular phylogenies, so that we can investigate the influence of rates of molecular evolution on our ability to infer macroevolutionary patterns from molecular data.

References

Akaike, H. 1992. Information Theory and an Extension of the Maximum Likelihood Principle. In: *Breakthroughs in Statistics*, pp. 610–624. Springer, New York.

Barraclough, T. & Savolainen, V. 2001. Evolutionary rates and species diversity in flowering plants. *Evolution* **55**: 677–683.

Barraclough, T.G., Harvey, P.H. & Nee, S. 1996. Rate of rbcL gene sequence evolution and

species diversification in flowering plants (Angiosperms). *Proc. R. Soc. B Biol. Sci.* **263**: 589–591.

Bellot, S. & Renner, S.S. 2014. Exploring new dating approaches for parasites: The worldwide Apodanthaceae (Cucurbitales) as an example. *Mol. Phylogenet. Evol.* **80**: 1–10.

Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C.-H., Xie, D., *et al.* 2014. BEAST 2: a software platform for Bayesian evolutionary analysis. *PLOS Comput. Biol.* **10**: e1003537. Public Library of Science.

Bromham, L. 2009. Why do species vary in their rate of molecular evolution? *Biol. Lett.* **5**: 401–404.

Bromham, L., Hua, X., Lanfear, R. & Cowman, P.F. 2015. Exploring the relationships between mutation rates, life history, genome size, environment, and species richness in flowering plants. *Am. Nat.* **185**: 507–524.

Crisp, M.D. & Cook, L.G. 2013. How Was the Australian Flora Assembled Over the Last 65 Million Years? A Molecular Phylogenetic Perspective. *Annu. Rev. Ecol. Evol. Syst.* **44**: 303–324. Annual Reviews .

Crisp, M.D., Hardy, N.B. & Cook, L.G. 2014. Clock model makes a large difference to age estimates of long-stemmed clades with no internal calibration: a test using Australian grasstrees. *BMC Evol. Biol.* **14**: 263.

Crottini, A., Madsen, O., Poux, C., Strauss, A., Vieites, D.R. & Vences, M. 2012. Vertebrate time-tree elucidates the biogeographic pattern of a major biotic change around the K-T boundary in Madagascar. *Proc. Natl. Acad. Sci. U. S. A.* **109**: 5358–63. National Academy of Sciences.

Cusimano, N. & Renner, S.S. 2010. Slowdowns in diversification rates from real phylogenies may not be real. *Syst. Biol.* **59**: 458–464.

Davies, T.J., Savolainen, V., Chase, M.W., Moat, J. & Barraclough, T.G. 2004. Environmental energy and evolutionary rates in flowering plants. *Proc. R. Soc. B Biol. Sci.* **271**: 2195–2200.

dos Reis, M., Donoghue, P.C.J. & Yang, Z. 2015. Bayesian molecular clock dating of species divergences in the genomics era. *Nat. Rev. Genet.* **17**: 71–80. Nature Research.

Drummond, A.J., Ho, S.Y.W., Phillips, M.J. & Rambaut, A. 2006. Relaxed phylogenetics and dating with confidence. *PLoS Biol.* **4**: e88.

Duchêne, D.A., Duchêne, S., Holmes, E.C. & Ho, S.Y.W. 2015a. Evaluating the adequacy of molecular clock models using posterior predictive simulations. *Mol. Biol. Evol.* **32**:

2986–2995.

Duchene, D. & Bromham, L. 2013. Rates of molecular evolution and diversification in plants: chloroplast substitution rates correlate with species-richness in the Proteaceae. *BMC Evol. Biol.* **13**: 65. BMC Evolutionary Biology.

Duchêne, D., Duchêne, S. & Ho, S.Y.W. 2015b. Tree imbalance causes a bias in phylogenetic estimation of evolutionary timescales using heterochronous sequences. *Mol. Ecol. Resour.* **15**: 785–794. Wiley Online Library.

Eo, S.H. & DeWoody, J.A. 2010. Evolutionary rates of mitochondrial genomes correspond to diversification rates and to contemporary species richness in birds and reptiles. *Proc. R. Soc. B Biol. Sci.* **277**: 3587–3592.

Etienne, R.S. & Rosindell, J. 2012. Prolonging the past counteracts the pull of the present: protracted speciation can explain observed slowdowns in diversification. *Syst. Biol.* **61**: 204–213.

Ezard, T.H.G., Thomas, G.H. & Purvis, A. 2013. Inclusion of a near-complete fossil record reveals speciation-related molecular evolution. *Methods Ecol. Evol.* **4**: 745–753.

Gillespie, R. 2004. Community assembly through adaptive radiation in Hawaiian spiders. *Science* **303**: 356–359. American Association for the Advancement of Science.

Goldie, X., Lanfear, R. & Bromham, L. 2011. Diversification and the rate of molecular evolution: no evidence of a link in mammals. *BMC Evol. Biol.* **11**: 286. BioMed Central Ltd.

Harmon, L.J., Weir, J.T., Brock, C.D., Glor, R.E. & Challenger, W. 2008. GEIGER: investigating evolutionary radiations. *Bioinformatics* **24**: 129–131.

Ho, S.Y.W. & Duchêne, S. 2014. Molecular-clock methods for estimating evolutionary rates and time scales. *Mol. Ecol.* **23**: 5947–5975.

Ho, S.Y.W., Duchene, S., Hua, X., Ritchie, A.M., Duchene, D.A. & Bromham, L.D. n.d. Bayesian molecular dating: opening up the black box. *Biol. Rev.*

Höhna, S., May, M.R. & Moore, B.R. 2015. TESS: an R package for efficiently simulating phylogenetic trees and performing Bayesian inference of lineage diversification rates. *Bioinformatics* **32**: 789–791. Oxford Univ Press.

Hua, X. & Bromham, L. 2017. Darwinism for the Genomic Age: Connecting Mutation to Diversification. *Front. Genet.* **8**: 12. Frontiers Media SA.

Hua, X., Cowman, P., Warren, D. & Bromham, L. 2015. Longevity is linked to mitochondrial mutation rates in rockfish: A test using Poisson regression. *Mol. Biol. Evol.* **32**: 2633–2645. Oxford University Press.

Hua, X. & Wiens, J.J. 2013. How does climate influence speciation? *Am. Nat.* **182**: 1–12.

Jetz, W., Thomas, G.H., Joy, J.B., Hartmann, K. & Mooers, A.O. 2012. The global diversity of birds in space and time. *Nature* **491**: 444–448.

Kass, R.E. & Raftery, A.E. 1995. Bayes Factors. *J. Am. Stat. Assoc.* **90**: 773–795. Taylor & Francis.

Lanfear, R., Ho, S.Y.W., Love, D. & Bromham, L. 2010. Mutation rate is linked to diversification in birds. *Proc. Natl. Acad. Sci. U. S. A.* **107**: 20423–20428.

Lartillot, N. & Philippe, H. 2006. Computing Bayes factors using thermodynamic integration. *Syst. Biol.* **55**: 195–207.

Lepage, T., Bryant, D., Philippe, H. & Lartillot, N. 2007. A general comparison of relaxed molecular clock models. *Mol. Biol. Evol.* **24**: 2669–2680.

Linder, M., Britton, T. & Sennblad, B. 2011. Evaluation of Bayesian models of substitution rate evolution--parental guidance versus mutual independence. *Syst. Biol.* **60**: 329–342.

Machac, A., Storch, D. & Wiens, J.J. 2013. Ecological causes of decelerating diversification in carnivoran mammals. *Evolution* **67**: 2423–33.

Mahler, D.L., Ingram, T., Revell, L.J. & Losos, J.B. 2013. Exceptional convergence on the macroevolutionary landscape in island lizard radiations. *Science* **341**: 292–295. American Association for the Advancement of Science.

Moen, D. & Morlon, H. 2014. Why does diversification slow down? *Trends Ecol. Evol.* **29**: 190–7.

Morlon, H. 2014. Phylogenetic approaches for studying diversification. *Ecol. Lett.* **17**: 508–525.

Murphy, W.J., Eizirik, E., Johnson, W.E., Zhang, Y.P., Ryder, O.A. & O'Brien, S.J. 2001. Molecular phylogenetics and the origins of placental mammals. *Nature* **409**: 614–8. Nature Publishing Group.

Pagel, M., Venditti, C. & Meade, A. 2006. Large punctuational contribution of speciation to evolutionary divergence at the molecular level. *Science* **314**: 119–121.

Paradis, E., Claude, J. & Strimmer, K. 2004. APE: Analyses of Phylogenetics and Evolution in R language. *Bioinformatics* **20**: 289–290.

Pereira, S.L. & Baker, A.J. 2006. A mitogenomic timescale for birds detects variable phylogenetic rates of molecular evolution and refutes the standard molecular clock. *Mol. Biol. Evol.* **23**: 1731–1740.

Phillimore, A.B. & Price, T.D. 2008. Density-dependent cladogenesis in birds. *PLoS Biol.* **6**: e71.

Pigot, A.L., Phillimore, A.B., Owens, I.P.F. & Orme, C.D.L. 2010. The shape and temporal dynamics of phylogenetic trees arising from geographic speciation. *Syst. Biol.* **59**: 660–673.

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

Pybus, O.G. & Harvey, P.H. 2000. Testing macro-evolutionary models using incomplete molecular phylogenies. *Proc. R. Soc. B Biol. Sci.* **267**: 2267–2272.

Quental, T.B. & Marshall, C.R. 2013. How the Red Queen drives terrestrial mammals to extinction. *Science* **341**: 290–2.

Rabosky, D.L. 2009. Heritability of extinction rates links diversification patterns in molecular phylogenies and fossils. *Syst. Biol.* **58**: 629–6240.

Rabosky, D.L. 2006a. LASER: a maximum likelihood toolkit for detecting temporal shifts in diversification rates from molecular phylogenies. *Evol. Bioinform. Online* **2**: 273–276.

Rabosky, D.L. 2006b. Likelihood methods for detecting temporal shifts in diversification rates. *Evolution* **60**: 1152–1164. Wiley Online Library.

Rannala, B. & Yang, Z. 1996. Probability distribution of molecular evolutionary trees: A new method of phylogenetic inference. *J. Mol. Evol.* **43**: 304–311. Springer-Verlag.

Ricklefs, R.E. 2007. Estimating diversification rates from phylogenetic information. *Trends Ecol. Evol.* **22**: 601–610.

Robin, X., Turck, N., Hainard, A., Tiberti, N., Lisacek, F., Sanchez, J.-C., *et al.* 2011. pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics* **12**: 77.

Sanderson, M.J. 1997. A nonparametric approach to estimating divergence times in the absence of rate constancy. *Mol. Biol. Evol.* **14**: 1218–1231. Chicago, Ill.: University of Chicago Press, c1983-.

Sanderson, M.J. 2002. Estimating Absolute Rates of Molecular Evolution and Divergence Times: A Penalized Likelihood Approach. *Mol. Biol. Evol.* **19**: 101–109.

Schliep, K.P. 2011. PHANGORN: phylogenetic analysis in R. *Bioinformatics* **27**: 592–593.

Tavaré, S. 1986. Some probabilistic and statistical problems in the analysis of DNA sequences. *Lect. Math. Life Sci.* **17**: 57–86.

Thomas, J. a, Welch, J.J., Lanfear, R. & Bromham, L. 2010. A generation time effect on the rate of molecular evolution in invertebrates. *Mol. Biol. Evol.* **27**: 1173–80.

van Tuinen, M. & Hedges, S.B. 2001. Calibration of avian molecular clocks. *Mol. Biol. Evol.* **18**: 206–213.

Venditti, C. & Pagel, M. 2010. Speciation as an active force in promoting genetic evolution. *Trends Ecol. Evol.* **25**: 14–20.

Webster, A.J., Payne, R.J.H. & Pagel, M. 2003. Molecular phylogenies link rates of evolution and speciation. *Science* **301**: 478.

Welch, J.J., Bininda-Emonds, O.R.P. & Bromham, L. 2008. Correlates of substitution rate variation in mammalian protein-coding sequences. *BMC Evol. Biol.* **8**: 53.

Welch, J.J. & Bromham, L. 2005. Molecular dating when rates vary. *Trends Ecol. Evol.* **20**: 320–327.

Welch, J.J., Fontanillas, E. & Bromham, L. 2005. Molecular dates for the "cambrian explosion": the influence of prior assumptions. *Syst. Biol.* **54**: 672–8.

Weller, C. & Wu, M. 2015. A generation-time effect on the rate of molecular evolution in bacteria. *Evolution* **69**: 643–652.

Xie, W., Lewis, P.O., Fan, Y., Kuo, L. & Chen, M.-H. 2011. Improving marginal likelihood estimation for Bayesian phylogenetic model selection. *Syst. Biol.* **60**: 150–160.

Table 1: Macroevolutionary models used to generate simulated datasets.

Model	Description	Speciation rates	Molecular rate
STU	Stochastic-Unlinked	Stochastic variation	Stochastic variation
STL	Stochastic-Linked	Stochastic variation	Linked to speciation rate
SLU	Slowdown-Unlinked	Decrease over time	Stochastic variation
SLL	Slowdown-Linked	Decrease over time	Linked to speciation rate
SPU	Speedup-Unlinked	Increase over time	Stochastic variation
SPL	Speedup-Linked	Increase over time	Linked to speciation rate

Table 2. Proportion of simulated datasets for which each macroevolutionary method correctly identifies the model of diversification rates under which the data was simulated for each of the macroevolutionary methods (gamma, LASER, and TESS), for each of the six models of simulation of the link between the rates of speciation and molecular substitutions (Table 1), and each of the three sources of phylogeny: simulated (true) trees, NPRS, and BEAST. Each proportion is given as range, corresponding to confidence intervals calculated using a normal approximation to proportion values. Bold values show the ranges from inferences that did not overlap with the values from simulated data.

Model		Simulated	NPRS	BEAST
Gamma				
Stochastic-Unlinked	STU	0.804 - 0.936	**0.525 - 0.715**	**0.452 - 0.648**
Stochastic-Linked	STL	0.829 - 0.951	**0.494 - 0.686**	**0.546 - 0.734**
Slowdown-Unlinked	SLU	0.632 - 0.808	**0.432 - 0.628**	0.442 - 0.638
Slowdown-Linked	SLL	0.578 - 0.762	**0.181 - 0.359**	**0.072 - 0.208**
Speedup-Unlinked	SPU	0.610 - 0.790	0.557 - 0.743	**0.362 - 0.558**
Speedup-Linked	SPL	0.510 - 0.690	**0.238 - 0.422**	**0.113 - 0.267**
LASER				
Stochastic-Unlinked	STU	0.557 - 0.743	0.578 - 0.762	0.535 - 0.725
Stochastic-Linked	STL	0.535 - 0.725	0.589 - 0.771	0.504 - 0.696
Slowdown-Unlinked	SLU	0.699 - 0.861	0.546 - 0.734	0.578 - 0.762
Slowdown-Linked	SLL	0.722 - 0.878	**0.333 - 0.527**	**0.238 - 0.422**
Speedup-Unlinked	SPU	0.643 - 0.817	0.589 - 0.771	**0.372 - 0.568**
Speedup-Linked	SPL	0.780 - 0.920	**0.165 - 0.335**	**0.156 - 0.324**
TESS				
Stochastic-Unlinked	STU	0.804 - 0.936	0.756 - 0.904	0.780 - 0.920
Stochastic-Linked	STL	0.756 - 0.904	0.816 - 0.944	0.841 - 0.959
Slowdown-Unlinked	SLU	0.452 - 0.648	0.422 - 0.618	0.452 - 0.648
Slowdown-Linked	SLL	0.422 - 0.618	**0.148 - 0.312**	**0.00 - 1.200**
Speedup-Unlinked	SPU	1.000 - 1.000	1.000 - 1.000	1.000 - 1.000
Speedup-Linked	SPL	1.000 - 1.000	**0.535 - 0.725**	**0.412 - 0.608**

Figure A. Gamma statistics from simulated (true) phylogenies (a) and phylogenies reconstructed using NPRS (b) and BEAST (c). Simulation models (coloured lines) are listed in Table 1. The data have been smoothed such that the y-axis indicates the density of samples and each curve integrates to 1.

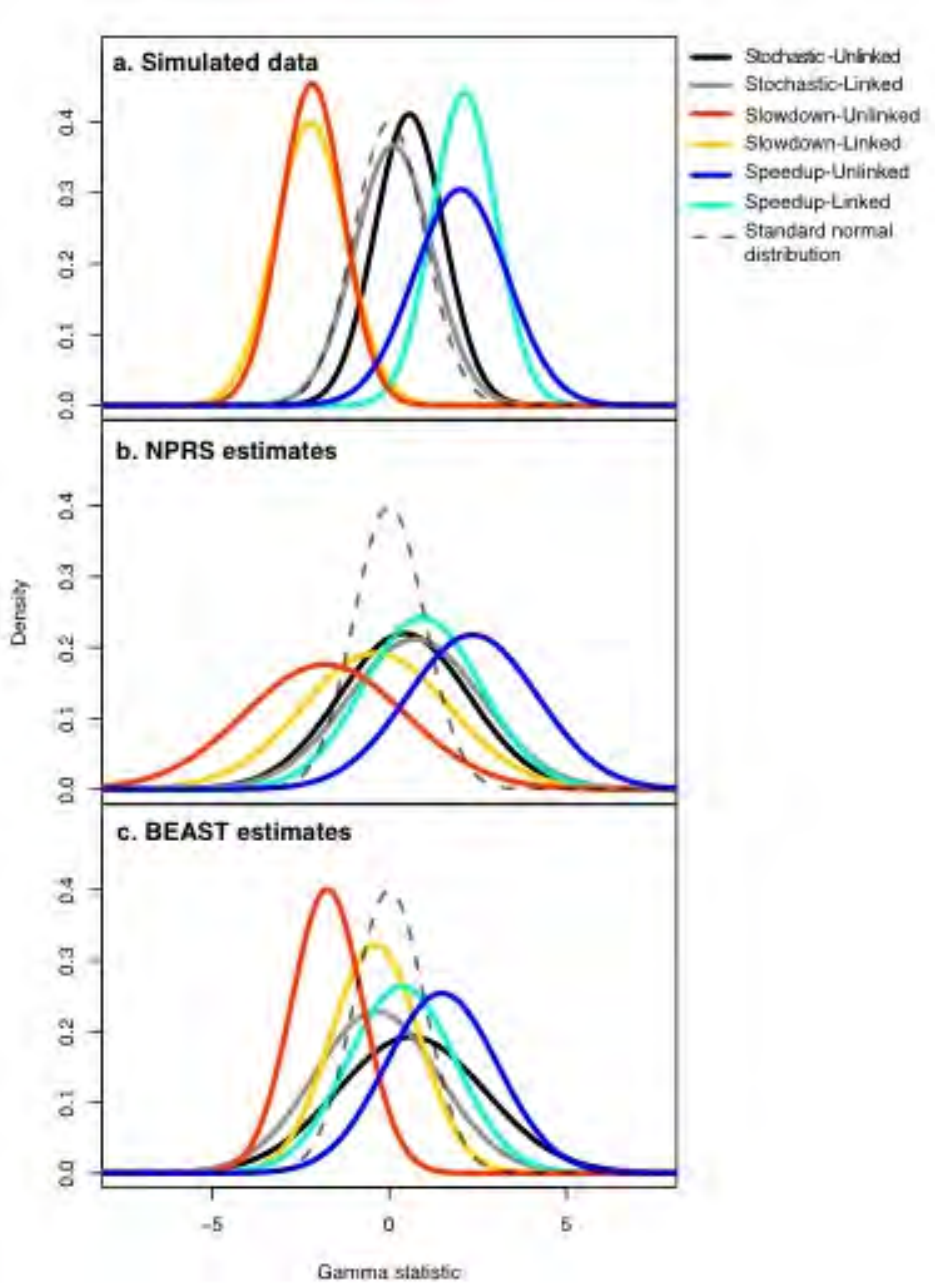

Figure B. Proportion of simulated and estimated phylogenies identified by LASER with each diversification rate model tested by the method: birth-death, slowdown, and speedup. Data are shown for each simulation scheme.

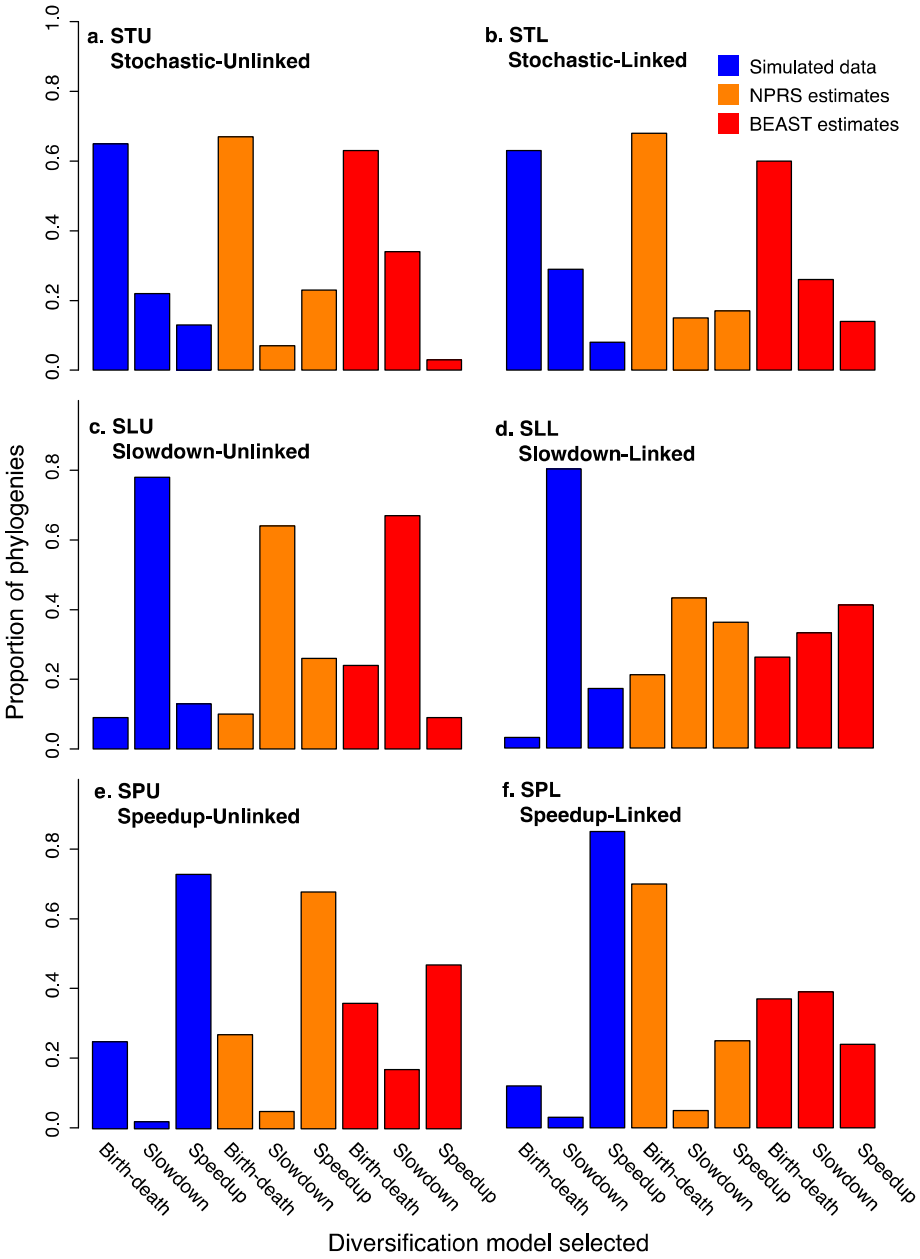

Figure C. Proportion of simulated and estimated phylogenies identified by TESS to be generated by each diversification rate model (birth-death; slowdown; and speedup). Data are shown for each simulation scheme. The diversification rate model is identified as being the most different between simulation and estimation in the schemes with a slowdown or speedup and a link between rates (schemes d and f).

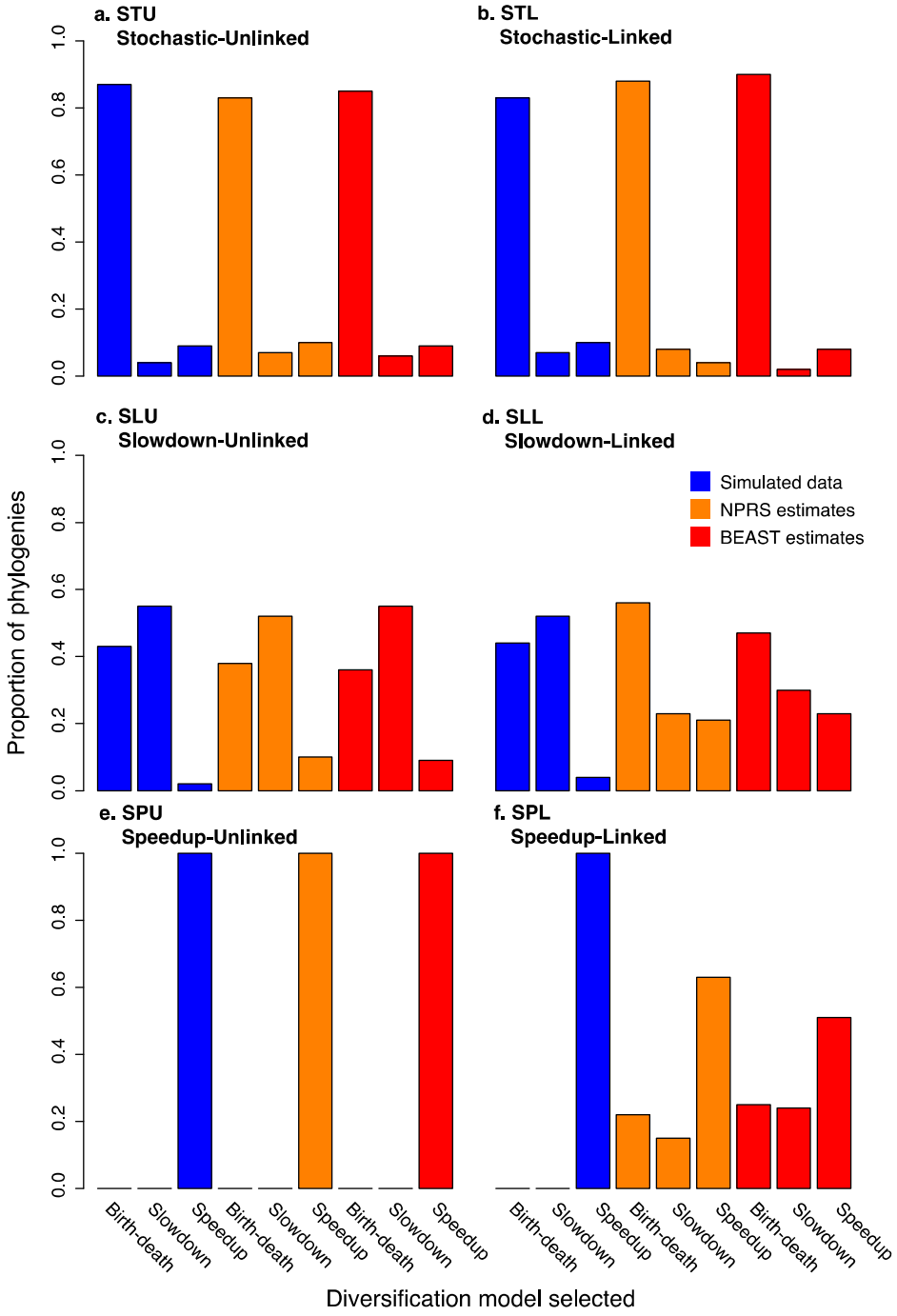

Figure D. Lineages-through time plot for simulated and reconstructed phylogenies for each simulation model. Ten example replicates are shown for each simulation scheme. The greatest departure between the simulations and reconstructions can be observed in the in the schemes with a slowdown or speedup and a link between rates (models SLL and SPL: see Table 1).