

The implications of interrelated assumptions on estimates of divergence times and rates of diversification

Tom Carruthers^{1*} and Robert W. Scotland²

Affiliations:

¹ Royal Botanic Gardens Kew, Richmond, London, TW9 3AE, United Kingdom

² Department of Plant Sciences, University of Oxford, South Parks Road, Oxford, OX1 3RB, United Kingdom

* Correspondence to be sent to: t.carruthers@kew.org

ABSTRACT

Phylogenies are increasingly being used as a basis to provide insight into macroevolutionary history. Here, we use simulation experiments and empirical analyses to evaluate methods that use phylogenies as a basis to make estimates of divergence times and rates of diversification. This is the first study to present a comprehensive assessment of the key variables that underpin analyses in this field – including substitution rates, speciation rates, and extinction, plus character sampling and taxon sampling. We show that in unrealistically simplistic cases (where substitution rates and speciation rates are constant, and where there is no extinction), increased character and taxon sampling lead to more accurate and precise parameter estimates. By contrast, in more complex but realistic cases (where substitution rates, speciation rates, and extinction rates vary), gains in accuracy and precision from increased character and taxon sampling are far more limited. The lack of accuracy and precision even occurs when using methods that are designed to account for more complex cases, such as relaxed clocks, fossil calibrations, and models that allow speciation rates and extinction rates to vary. The problem also persists when analysing genomic scale datasets. These results suggest two interrelated problems that occur when the processes that generated the data are more complex. First, methodological assumptions are more likely to be violated. Second, limitations in the information content of the data become more important.

Keywords: phylogeny, macroevolution, divergence time estimation, diversification rates

Over the last twenty years, there has been a marked increase in the availability of molecular sequence data for constructing phylogenies (Soltis et al. 2000; Panero and Funk 2008; Särkinen et al. 2013; Atchison et al. 2016; Nevado et al. 2016; Contreras-Ortiz et al. 2018; Muñoz-Rodríguez et al. 2018, 2019; Nürk et al. 2018) and considerable developments to methods for making inferences about macroevolution from phylogenies (Sanderson 1997; Drummond et al. 2006; Maddison et al. 2007; FitzJohn 2012; Condamine et al. 2013; Rabosky 2014; Höhna et al. 2016). Given the importance of a phylogenetic context to many macroevolutionary studies, these developments have enabled important insights into many aspects of macroevolution. This includes understanding the rate that different clades have diversified (Baldwin and Sanderson 1998; Hughes and Eastwood 2006; Givnish et al. 2009; Drummond et al. 2012; Nevado et al. 2016); the importance of long-distance-dispersal and continental drift for explaining distribution patterns of extant species (Donoghue et al. 2001; Lavin et al. 2004; Donoghue 2008; Crisp 2009); and the manner by which the earth's major biomes were assembled (Pennington et al. 2004, 2009; Simon et al. 2009; Särkinen et al. 2012; Simon and Pennington 2012; Maurin et al. 2014; Pennington and Hughes 2014, Cardillo et al. 2017; Folk et al. 2019).

However, molecular phylogenies only reflect some aspects of macroevolutionary history (Platnick 1979; Patterson 1981). Methods for making inferences about macroevolution from phylogenies therefore require important assumptions (Sanderson 1997; Drummond et al. 2006; Maddison et al. 2007; Rabosky and Lovette 2008, Rabosky 2014; FitzJohn 2012; Condamine et al. 2013; Höhna et al. 2016; Moore et al. 2016). In our recent phylogenetic and macroevolutionary studies of *Ipomoea* (Muñoz-Rodríguez et al. 2018, 2019, Carruthers et al. 2020), it became apparent that these assumptions are highly interlinked, that they can have a profound impact on parameter estimates, and that they are often based on very limited evidence. For example, different assumptions about the

extremely fragmentary fossil record of close relatives of *Ipomoea* had large effects on age estimates for *Ipomoea*, but this was dependent on assumptions about substitution rates. These assumptions then had important effects on estimates of rates of evolutionary diversification within *Ipomoea* (Muñoz-Rodríguez et al. 2019, Carruthers et al. 2020). Notably, these issues persist even when analysing a large molecular dataset of 100s loci (Muñoz-Rodríguez et al. 2019, Carruthers et al. 2020). Investigating the implications of different assumptions for making inferences about macroevolution is therefore critically relevant even in the genomics era.

Here, we investigate the implications of different assumptions for *divergence time estimation* – where the branches in a molecular phylogeny are placed on an absolute time scale, and *diversification parameter estimation* – where speciation rates (λ), extinction rates (μ), and net diversification rates ($\lambda - \mu$) are estimated. We focus on divergence time estimation and diversification parameter estimation because they are fundamental for interpreting molecular phylogenies in a macroevolutionary context.

To estimate divergence times and infer a *time-calibrated phylogeny*, molecular branch lengths – which reflect the number of substitutions (n) for each branch – are transformed such that they are equal to time (t). This requires assumptions about the substitution rate (r) and t for each branch because the molecular sequence data from which a phylogeny is constructed only provides direct information about n , which is a product of r and t (Britton 2005). Assumptions about t can be made by implementing fossil calibrations, where fossil ages are used as a basis for constraining clade age estimates, and also by implementing branching processes, where the temporal duration of branches is assumed to be controlled by a branching process with constant λ (when branches split) and μ (when branches terminate before the present) (Drummond et al. 2006; Höhna et al. 2016). Meanwhile, assumptions about r can be made by implementing a strict clock where r is assumed to be the same for

every branch, or a relaxed clock where r is assumed to vary throughout the phylogeny according to a specified model (Sanderson 1997, 2002; Thorne et al. 1998; Kishino et al. 2001; Drummond et al. 2006; Drummond and Suchard et al. 2011; Höhna et al. 2016).

In order to estimate diversification parameters, an important initial assumption is that the divergence time estimates in a time-calibrated phylogeny are accurate (for example Rabosky 2014; Höhna et al. 2016). This is because the divergence time estimates are often the basis from which diversification parameters are estimated. However, even assuming they are accurate, a time-calibrated phylogeny only provides direct information about divergence times between sampled extant taxa. As such, there are inevitable limitations in the information provided by a time-calibrated phylogeny for estimating diversification parameters when there are un-sampled extant or extinct taxa (Rabosky 2010, 2014; Stadler 2011a,b; Höhna et al. 2016; Moore et al. 2016; Louca and Pennel 2020).

Numerous studies have sought to characterise the precision and accuracy of methods of divergence time estimation and diversification parameter estimation. For divergence time estimation, it has been shown that among-branch-variation in r leads to inaccurate divergence time estimates, and that attempts to rectify this problem by using relaxed clocks with multiple fossil calibrations can be of limited success, and may in some cases exacerbate the issue (Near and Sanderson 2004; Britton 2005; Near et al. 2005; Warnock et al. 2011, 2015, 2017; Tamura et al. 2012; Magallón et al. 2013; Heath et al. 2014; Ho 2014; Zhu et al. 2015; Carruthers et al. 2020; Carruthers and Scotland 2020). For diversification parameter estimation, it has been shown that inaccurate divergence time estimates have a relatively limited effect on diversification parameter estimates, and that diversification parameter estimates are more strongly affected by limited clade size or taxon sampling (Wertheim and Sanderson 2011). With respect to taxon sampling more specifically, several studies have been undertaken to develop ways of accounting for incomplete taxon sampling (Stadler 2009;

Höhna 2011, 2014; Cusimano et al. 2012; Stadler and Bokma 2013), and it has been shown that non-random sampling can lead to the incorrect inference of rate slowdowns towards the present (Cusimano and Renner 2010). Recently, a particular focus has been to determine the effect of extinct branches on diversification parameter estimates, especially when diversification parameters vary (Rabosky 2010; Moore et al. 2016). Considerable controversy centres on how rate shifts on extinct branches should be accounted for (Rabosky 2014, Moore et al. 2016; Rabosky et al. 2017), whilst in a recent contribution, Louca and Pennel (2020) set out how problems that underpin diversification parameter estimation stem from fundamental issues with model identifiability.

Although previous studies have provided important insights into different methods, they typically consider the effect of the different variables relevant to parameter estimation (such as models for r , branching processes, fossil calibrations, character and taxon sampling) in relative isolation (Britton 2005; Cusimano and Renner 2010; Warnock et al. 2011, 2015, 2017; Rabosky 2014; Zhu et al. 2015; Moore et al. 2016; Carruthers et al. 2020). This is despite the fact that such variables are highly interlinked (Carruthers and Scotland 2020). Because of this, these studies provide a limited perspective of the sensitivity of analyses to different assumptions, and the way that datasets interact with interlinked assumptions to produce parameter estimates. The present study is an attempt to understand and explore these issues in a more comprehensive way.

We approach this problem by using simulation experiments and empirical analyses to evaluate the combined effect of these variables on divergence time estimates and diversification parameter estimates. For divergence times, we consider assumptions relating to r (clock models), and t (branching processes and fossil calibrations), in datasets with different amounts of among-branch-variation in r and diversification parameter variation, and with different levels of character and taxon sampling (Table 1). For diversification

parameters, we consider assumptions relating to diversification parameter variation and extinction, in datasets with different levels of diversification parameter variation and with different levels of taxon sampling (Table 2). We perform analyses in a Bayesian framework and explicitly consider the relationship between prior and posterior probabilities in order to explore the interaction between different datasets and the (prior) assumptions of an analysis.

There are inevitably numerous other factors that could affect parameter estimates that are not directly addressed in the analyses presented here (Appendix 1). For example, phylogenetic inference can affect estimated divergence times and diversification parameters, whilst there are numerous other biological processes that could undermine assumptions about r , t , diversification parameter variation, or extinction. However, this study is not designed to provide an exhaustive overview of all the processes that could potentially affect parameter estimates, but instead it aims to provide a comprehensive overview of the implications and interactions between key methodological assumptions that are central to divergence time estimation and diversification parameter estimation, and the relationship between these assumptions and the nature of the data that is analysed. This in turn can provide pertinent insights into the information about macroevolutionary history that can be extracted from molecular sequence data and molecular phylogenies, and the effect that violation of methodological assumptions has on parameter estimates.

METHODS

Simulation experiments

The purpose of the simulation experiments for divergence time estimation was to 1) investigate the effects of character sampling and taxon sampling on parameter estimates when r and diversification parameters were constant (Table 1); 2) investigate the effects of among-branch-variation in r with different levels of character sampling, and investigate the effectiveness of relaxed clocks and fossil calibrations for accounting for among-branch-variation in r (Table 1); 3) investigate the effects of variation in diversification parameters with different levels of taxon sampling, and investigate the effectiveness of fossil calibrations for accounting for variation in diversification parameters (Table 1). Note that for analyses of the effect of among-branch-variation in r on divergence time estimates, we focus on character sampling because it directly affects the relative number of substitutions for each branch that are actually sampled. By contrast, for the effect of variation in diversification parameters on divergence time estimates, we focus on the effect of taxon sampling because it directly affects the proportion of branching events that are actually sampled. For diversification parameter estimation, the purpose of the experiments was to 1) investigate the effects of taxon sampling on parameter estimation when diversification parameters were constant (Table 2); 2) investigate the effects of taxon sampling on parameter estimation when diversification parameters varied and determine the effectiveness of models that account for variable diversification parameters (Table 2). In all cases, we explicitly considered the effect of extinction on parameter estimation.

Simulations were configured to capture the fundamental links between macroevolutionary history, data, sampling, assumptions, divergence time estimates, and diversification parameter estimates. The simulations were centred on a branching process that is interpreted as the macroevolutionary process that generates a set of extant taxa in a

molecular phylogeny (Fig. 1a). Macroevolutionary processes were simulated with constant diversification parameters, or diversification parameters that changed between two time intervals, and with $\mu=0$ or $\mu>0$ (Table 1).

Molecular sequence data was simulated to evolve along the branches of these macroevolutionary processes (Fig. 1b). Sequences evolved at either the same r for all branches, or different values for each branch that were drawn from a lognormal distribution (Table 1). In some instances, fossils were deposited along the branches (Fig. 1b). This part of the simulation replicates how data is generated from a “real” macroevolutionary process.

Different levels of character sampling and taxon sampling were then used to estimate divergence times from the simulated molecular sequence data (Fig. 1c). There were three levels of taxon sampling and three levels of character sampling (referred to as *low*, *medium*, or *high*) (Table 1). To estimate divergence times, assumptions are also required about r and t (Fig. 1d). For r , it was either assumed to be the same for every branch (strict clock) or to vary among branches, with the value for each branch being drawn from lognormal distribution (relaxed clock) (Table 1). For t , it was assumed that the expected t for each branch was controlled by a branching process with constant diversification parameters. The branching process assumed that either $\mu=0$ or $\mu>0$ (Table 1). In some cases, fossil calibrations were also implemented (as lognormal calibration densities) (Table 1). Assumptions were implemented with the chosen sampling strategy in order to infer a time-calibrated phylogeny (Fig. 1e). These issues of sampling and assumptions reflect the methodological decisions made by biologists when analysing empirical datasets.

Divergence time estimation is statistically unique from phylogenetic inference because the parameter of interest (t) in divergence time estimation is not inferred directly from the sequence data (Britton 2005). We performed further analyses where tree topologies

and branch lengths were estimated in units of n to clarify these differences (Appendix 1 and 2).

In an empirical analysis, the estimated time-calibrated phylogeny would be used, in combination with a set of assumptions (Fig. 1f), to estimate diversification parameters (Fig. 1g). However, the focus of this study was to determine the effects of sampling and assumptions on diversification parameter estimates. If the estimated time-calibrated phylogeny were used, diversification parameter estimates would be confounded by divergence time estimation error. Therefore, in the simulation experiments for diversification parameter estimation, the “correct” time-calibrated phylogeny that corresponds to the simulated macroevolutionary process was used (Fig. 1h). Low, medium, or high taxon sampling was used, in combination with different assumptions about whether diversification parameters were constant or variable, and whether $\mu > 0$ (Table 2).

Given that parameters were estimated in a Bayesian framework, prior parameter distributions were also evaluated. For divergence time estimation, this involved performing analyses with no sequence data (Table 1), and for diversification parameter estimation, this involved performing analyses with no input tree (Table 2).

Empirical analyses

Empirical analyses were performed to clarify the relevance of the variables investigated in the simulation experiments. These were performed with a dataset for the tropical angiosperm genus *Ipomoea*. For divergence time estimation, the age of the crown node (subsequently referred to as the radiation node) was estimated for a clade within which a significant increase in net diversification rates was inferred in Muñoz-Rodríguez et al. (2019) and Carruthers et al. (2020). Different levels of character sampling and taxon sampling were used, and either a strict clock or a relaxed clock (Table 1). For diversification parameter estimation, diversification parameters were estimated across *Ipomoea* with

different levels of taxon sampling, and with different assumptions about whether diversification parameters were constant or variable, and whether $\mu > 0$.

Data availability

Custom R, Revbayes, and Python scripts developed for this study are available in Supplementary Material on Dryad: [http://dx.doi.org/10.5061/dryad.\[NNNN\]](http://dx.doi.org/10.5061/dryad.[NNNN]), and on GitHub: https://github.com/TomCarr/Interrelated_Assumptions. Simulated matrices and other output files are available on request. Further methodological details of all experiments are provided in Appendix 1.

RESULTS

Divergence time estimation

Simulation experiments.- When simulated diversification parameters and r were constant, increased character sampling and increased taxon sampling both caused a reduction in the % error for age estimates (defined as the % difference between the mean posterior estimate and the correct value), and the 95% Highest Posterior Density (HPD) became narrower (Fig. 2 and 3a-f; Table S1; Table S2). In all cases the correct value was included in the 95% HPD for a high % of nodes (Fig. 2 and 3a-f; Table S1; Table S2). When $\mu > 0$, the % error was higher and 95% HPDs wider than when $\mu = 0$ (Fig. 3g-h, Table S2).

When among-branch-variation in r was simulated, and a strict clock was used, increased character sampling caused the % error to fall, although less so than with no among-branch-variation in r (Fig. 4a-c; Table S3). Increased character sampling with a strict clock also caused 95% HPDs to become narrower and a reduction in the % of nodes that included the correct value in the 95% HPD (Fig. 4a-c; Table S3). With a relaxed clock, the % error was higher, 95% HPDs wider, and the correct value was more likely to be included in the 95% HPD (Fig. 4d; Table S3). With node calibrations and a relaxed clock, the % error, 95% HPD widths, and % of nodes that included the correct value in the 95% HPD, were all lower compared to with just a relaxed clock (Fig. 4f, Table S3).

When variation in diversification parameters was simulated, increased taxon sampling caused the % error to fall, although it also caused the ages of older nodes to be underestimated (Fig. 5a-f; Table S4). The 95% HPD also became narrower and there was a reduction in the probability that the correct value was included in the 95% HPD (Fig. 5a-f; Table S4). Notably, with high taxon sampling the correct value was often not included in the 95% HPD for older nodes (Fig. 5c,f). With node calibrations, the % error, 95% HPD widths,

and the % of nodes that included the correct value in the 95% HPD, all fell (Fig. 5g-h, Table S4).

Empirical study.- Increased taxon sampling led to older age estimates for the radiation node (Fig. 6, Table S5). This trend was most noticeable with lower character sampling or a relaxed clock, and was least noticeable with higher character sampling and a strict clock. Increased character sampling did not cause an equivalent trend in age estimates for the radiation node. However, increased character sampling and increased taxon sampling both led to narrower 95% HPDs (Fig. 6, Table S5).

Diversification parameter estimation

Simulation experiments.-When simulated diversification parameters were constant, increased taxon sampling caused a marked fall in the % error and the 95% HPD width for λ , and the % of experiments that included the correct value within the 95% HPD was always high (Table S6). However, with no input tree, higher taxon sampling caused the % error to increase, the 95% HPD to become narrower, and the % of experiments that included the correct value within the 95% HPD to fall (Table S6). When $\mu > 0$, the % error for λ was higher than when $\mu = 0$, and the 95% HPD was wider (Table S6).

When diversification parameters were simulated with an increase λ towards the present but constant diversification parameters were assumed when estimating diversification parameters, increased taxon-sampling caused the % error and 95% HPD width for λ to fall and the correct value was included within the 95% HPD with high taxon sampling (Table S7). In the equivalent experiment with no input tree, increased taxon sampling had little effect on parameter estimates (Table S7). When an episodic model that incorporated shifts in λ was used when estimating diversification parameters, the % error was low for λ at the base and tips of the tree and for the change time, and the correct values were included in the 95%

HPD (Table S7). With no input tree but the same episodic model, the % error was higher, the 95% HPD wider, and the correct values were not included in the 95% HPD (Table S7).

When diversification parameters were simulated with a decrease in μ towards the present, and an episodic model that incorporated a shift in μ was used when estimating diversification parameters, the % error for λ across the entire tree was low and the correct value was included in the 95% HPD (Table S7). By contrast, for μ at the base and tips of the tree and the change time, the % error was far higher, and the correct value was only included in the 95% HPD for the change time (Table S7). With no input tree, the % error for λ was far higher and the 95% HPD far wider (Table S7). For μ and the change time, estimates with no input tree were very similar to where there was an input tree (Table S7).

Empirical study.—Increased taxon-sampling led to more precise diversification parameter estimates (Table S8), although the assumption that $\mu > 0$ led to less precise estimates (Table S8). For branch specific diversification parameter estimates, when branch specific values for λ and μ , or just λ were estimated, shifts in diversification parameters were primarily estimated for a Neotropical clade of *Ipomoea* (Fig. S1a-b). These shifts were estimated to result predominantly from an increase in λ . Where only shifts in μ were permitted, very minor variations in diversification parameters were inferred (Fig. S1c). With no input tree, shifts in diversification parameters were not inferred with any model (Fig. S1d-f).

DISCUSSION

Increased character and taxon sampling lead to more robust divergence time estimates in unrealistically simplistic simulations

In unrealistically simplistic simulations, where the simulated r and diversification parameters were constant and $\mu = 0$, increased character and taxon sampling led to reduced % error for node ages and narrower 95% HPDs (Fig. 2 and 3a-f, Table S1 and S2). Further,

although the 95% HPDs became markedly narrower, there was no reduction in the probability that they contained the correct value.

For increased character sampling, this pattern is likely to result from the fact that increased character sampling means that n for each branch can be estimated with more accuracy and precision (Britton 2005). Given that in this experiment r was correctly assumed to be fixed and known for every branch, increased accuracy and precision in estimates of n led to more accurate and precise estimates of t .

For increased taxon sampling, there are two potential explanations for the pattern. First, increased taxon sampling may cause n to be estimated with more accuracy and precision through its effect on estimates of substitutions along individual branches. Second, increased taxon sampling means that more branching events are sampled. Given that in this experiment divergence times were inferred with the correct assumption that diversification parameters were constant, sampling more branching events may result in a smaller range of probable ts for each branch. This in turn may cause t for each branch to be estimated with more accuracy and precision. It is difficult to distinguish between these explanations. However, when no sequence data was analysed the % error and 95% HPD width also fell with increased taxon sampling (Fig. 3d-f; Table S2). This fall cannot result from increased accuracy and precision in the inferred n because no sequence data is analysed. This indicates that the explanation centred on the effect of sampling more branching events is at least partly responsible for the pattern observed.

In spite of the positive effects of increased character and taxon sampling, even when the simulated r and diversification parameters were constant, and when the correct branching process was used when estimating divergence times, when $\mu > 0$ (leading to un-sampled extinct species) the % error was higher and 95% HPDs were wider compared to when $\mu = 0$ (Fig. 3g-h, Table S2). This finding further highlights the effects of failing to sample species

from a macroevolutionary process, regardless of whether extant species are un-sampled or extinct species are un-sampled.

As with failing to sample extant taxa, the explanation for this result could be that extinction causes n for each branch to be estimated with less accuracy and precision, whilst the alternative explanation is that extinction increases the range of probable ts for each branching event. Given that with no molecular data the % error is also higher and 95% HPDs markedly wider when $\mu > 0$ compared to when $\mu = 0$ (Fig. 3f and h, Table S2), the increase in range of probable ts for each branching event is likely to play at least some role in the increased % error and wider 95% HPDs. The greater range of probable ts in turn stems from the fact that when $\mu > 0$ there is a large quantity of unsampled diversity, the temporal and phylogenetic distribution of which is unknown. This finding and its explanation is consistent with points discussed below in relation to relaxed molecular clocks, which describe how a lack of precision in parameter estimates stem from limitations in the information content of the data from which inferences are made.

Increased character and taxon sampling may not lead to more robust divergence time estimates in more complex simulations

Among-branch-variation in r .— When among-branch-variation in r was simulated, increased character sampling had a far more limited effect in reducing the % error (Fig. 4, Table S3). The % error was still over 25% with high character sampling, regardless of whether divergence times were estimated with a strict clock, a relaxed clock, or a relaxed clock with node calibrations.

Despite this, with a strict clock, increased character sampling caused a dramatic decrease in the 95% HPD width. Consequently, increased character sampling led to a marked fall in the % of nodes that included the correct value in the 95% HPD (Fig. 4a-c, Table S3). This misleading precision results from the fact that the assumptions of the strict clock are

violated by the macroevolutionary process from which the data was generated. By contrast, with a relaxed clock 95% HPDs were considerably broader such that the 95% HPD was more likely to contain the correct value (Fig. 4d, Table S3). In this case, the assumptions of the relaxed clock model are not violated by the underlying macroevolutionary process.

Despite the utility of the relaxed clock for avoiding misleadingly precise parameter estimates, these results highlight that with among-branch-variation in r , such that branch specific values for r are unknown, increased character sampling does not provide a basis for making more accurate estimates of r . Instead, the sequence data only provides a basis for making explicit estimates of n (Britton 2005). Although the problem is partially overcome by using relaxed clock models that lead to less precise estimates, it is inevitable that these estimates are critically sensitive to the assumptions about r expressed by the relaxed clock. If these assumptions are violated, parameter estimates are likely to be misleadingly precise (Carruthers et al. 2020).

When a relaxed clock was used in combination with fossil calibrations, there was a reduction in % error compared to when a relaxed clock was used without fossil calibrations (Fig. 4f, Table S3). Implementing fossil calibrations can therefore constrain age estimates such that they are closer to the correct value. In particular, fossil calibrations appeared to reduce the extent to which node ages were significantly overestimated (Fig. 4f-g).

However, implementing fossil calibrations caused 95% HPD widths and the % of nodes that included the correct value in the 95% HPD to decrease markedly (Fig. 4f-g; Table S3). This indicates that the manner by which fossil calibrations were implemented in this analysis made overly precise assumptions about the ages of the fossils in relation to the ages of the nodes that they calibrate, and therefore caused 95% HPDs to be overly narrow.

Two characteristics of the simulations presented here may mean that problems relating to fossil calibrations are even more acute in empirical datasets. First, the fossils were

simulated at a relatively high and constant rate. They are therefore likely to be markedly closer to the age of the node they calibrate than fossil calibrations used in empirical studies and provide a more reliable temporal framework for calibrating the phylogeny. Second, the manner by which fossil calibrations were implemented in this study made relatively relaxed assumptions about fossil ages in relation to clade ages. Fossil calibrations used in empirical studies often make far more precise assumptions meaning they may be more likely to lead to overly precise estimates (for example Särkinen et al. 2013; Grimm et al. 2015; Magallón et al. 2015, Eserbach et al. 2017; Folk et al. 2019).

Variation in diversification parameters.— When simulated diversification parameters varied over time, increased taxon sampling caused some divergence time estimates to become misleadingly precise. This conclusion is based on the observation – especially pronounced at older nodes – that with increased taxon sampling the 95% HPD became narrower and the % of nodes that included the correct value in the 95% HPD decreased (Fig. 5a-f; Table S4).

This observation can be explained by the fact that when taxon sampling is higher, violation of the assumption of the branching process that diversification parameters are constant (which is inherent to most Bayesian methods of divergence time estimation) has a stronger effect on parameter estimates compared to when taxon sampling is lower. The central role of the branching process for generating this pattern is highlighted by the fact that this pattern is even more pronounced when no sequence data is analysed (Fig. 5d-f, Table S4). This result is consistent with the case where simulated diversification parameters were constant and taxon sampling was increased. In that case, the assumptions of the branching process were consistent with the underlying macroevolutionary process such that increased taxon sampling led to more accurate and precise divergence time estimates (Fig. 3a-f; Table S2). By contrast, where simulated diversification parameters varied such that the assumptions of the branching process were violated by the underlying macroevolutionary process,

increased taxon sampling led to inaccurate and misleadingly precise divergence time estimates (Fig. 5a-f; Table S4).

These explanations have a similar basis to explanations for other findings in this study, and in previous studies (Carruthers et al. 2019), which describe how analysing larger quantities of molecular sequence data with assumptions that are violated by the process that generated the data (for example, using a strict clock when there is among-branch-variation in r) can lead to inaccurate and misleadingly precise parameter estimates. This problem is especially acute in divergence time estimation because the parameter of interest (t) cannot be directly estimated from the data. Estimates will therefore always remain especially sensitive to the assumptions of a given method.

The implementation of fossil calibrations when estimating divergence times, which weaken the assumption that the ages of different clades are entirely controlled by constant diversification parameters, led to reduced % error when simulated diversification parameters varied over time (Fig. 5g-h, Table S4). However, 95% HPDs and the % of nodes that included the correct value in the 95% HPD both showed a marked decrease (Fig. 5g-h, Table S4). This highlights that unless the assumptions underlying fossil calibrations are carefully evaluated, they may not effectively address the problems associated with variation in diversification parameters outlined above.

Careful interpretation of some aspects of the analysis with variable diversification parameters is required. For example, when variation in diversification parameters was simulated, molecular sequence data was simulated such that there was no among-branch-variation in r , and when divergence times were estimated, a strict clock was used with r set to the correct value. In this case, the molecular sequence data may increase the accuracy of divergence time estimates in a manner that is unlikely to be replicated in empirical datasets

where r is unknown and cannot be directly estimated from the molecular sequence data (Britton 2005).

The manner by which fossils were simulated, which resulted in the implementation of an unusually large number (148) of fossil calibrations, is also noteworthy and not necessarily realistic. Nevertheless, the central issue of assumptions about t that are expressed through fossil calibrations, and the way that these assumptions interact with those expressed by branching processes, remains important in datasets with considerably fewer fossil calibrations. For example, in Convolvulaceae and Solanaceae, Carruthers et al. (2020) illustrated how different assumptions about just three fossil calibrations that were implemented in the context of a branching process had a profound impact on divergence time estimates – both in terms of absolute age estimates, and with respect to the relative ages of different clades within the families.

Diversification parameter estimates are also affected by taxon sampling and the nature of the underlying macroevolutionary process

In a simple case where simulated diversification parameters were constant and $\mu=0$, higher taxon sampling caused both the % error and the 95% HPD width for λ to fall markedly, whilst in all cases there was a high probability that the correct value was included within the 95% HPD (Table S6). By contrast, when there were extinct branches in the underlying macroevolutionary process, both the % error and the 95% HPD widths for λ were higher. This is a further illustration of how un-sampled extinct species lead to more uncertain parameter estimates (Table S6).

Where variation in diversification parameters was simulated, but a Yule model that assumes constant diversification parameters was used when estimating diversification parameters, increased taxon sampling also caused both the % error and the 95% HPD width for λ to fall markedly, and with high taxon sampling the correct value was included in the

95% HPD (Table S7). This indicates that at least for λ , model misspecification does not necessarily mean that it is impossible to estimate an accurate and precise value for λ .

However, the extent to which this estimate for λ is informative is debatable. The “correct” value that is being compared to is the average value for λ across all branches in the macroevolutionary process, and the estimate for λ is simply a single estimate across the entire tree. This estimated value is not actually correct for any one part of the tree, and the analysis is incapable of providing any information about shifts in diversification parameters.

When episodic models that incorporate shifts in diversification parameters were used to estimate diversification parameters, and when there was an increase in λ and $\mu=0$ in the simulated macroevolutionary process, diversification parameters were estimated with considerably higher accuracy and precision compared to when shifts resulted from a fall in μ (Table S7). Notably, when there was a fall in μ , estimates for μ and the timing of shifts in μ had a particularly high % error and wide 95% HPDs. Further, parameter estimates were very similar to those estimated with no input tree, indicating little difference between prior and posterior parameter distributions (Table S7). This highlights how the time-calibrated phylogeny provides little information about the dynamics of diversification for extinct (and therefore un-sampled) branches, and that parameter estimates remain uncertain even when the correct model is used. This result, and the implications of it, is consistent with the discussion of the role of relaxed clock models in divergence time estimation outlined above, where even with the correct model the data provides limited information for parameter estimation.

Analyses of divergence times and diversification parameters in *Ipomoea* corroborate findings from simulation experiments

As in our simulation experiments, higher character sampling resulted in more precise divergence time estimates, and the most precise divergence time estimates were obtained

when character sampling was high and a strict clock was used (Fig. 6, Table S5). This observation can be explained according to the same framework as was set out for our simulation experiments. Higher character sampling enables n for each branch to be estimated with more accuracy and precision, with n being used as a basis to estimate t according to a molecular clock model. Although in this empirical experiment r is not assumed known, with a strict clock r is assumed to be the same for every branch such that a relatively narrow parameter space is explored and estimates of t can become very precise with high character sampling. By contrast, with a relaxed clock, a broader parameter space is explored meaning that estimates for t are considerably less precise even with high character sampling.

Taxon sampling also had a notable effect on divergence time estimates through its interaction with the birth-death branching process. Specifically, higher taxon sampling resulted in older age estimates for the radiation node, with the effect being most pronounced with no sequence data, low character sampling, or when a relaxed clock was used (Fig. 6, Table S5). Given this pattern is most pronounced with no sequence data, the characteristics of the birth-death branching process are the most likely explanation. Specifically, the clade descended from the radiation node is significantly more diverse than its sister clade. Given the birth-death branching process assumes constant diversification parameters, branching events in the clade descended from the radiation node will be spread over a longer time interval in order to smooth differences in diversification parameters between the two clades. This will cause age estimates for the radiation node to be older. The effect is most pronounced where taxon sampling is high because the difference in sampled diversity between the two clades is higher. It is also more noticeable where: character sampling is lower because the signal from the sequence data does not override that of the birth-death branching process; and where a relaxed clock is used because branch specific values for r are sensitive to assumptions about t expressed through the birth-death branching process.

Linder et al. (2005) also recovered a pattern where lower taxon sampling led to younger age estimates, in this case for the diverse *Restio* clade of Restionaceae. This pattern was recovered with a method that does not implement a branching process, indicating that other processes have the potential to generate similar patterns to those presented here. However, a generalizable explanation for their finding was not proposed, and it is likely that the explanation presented here plays at least some role – especially as it appears to be most pronounced when character sampling is lower.

With respect to inferring diversification parameters, increased taxon sampling also led to more precise estimates when diversification parameters were assumed to be constant (Table S8). However, when extinction was incorporated into analyses – which is likely to be a realistic component of most macroevolutionary processes – parameter estimates became considerably less precise. In particular, variable extinction rates were very difficult to infer with any precision (Fig. S1). These results are consistent with a recent paper by Louca and Pennel (2020) who discuss how a single macroevolutionary process can be explained with an infinite number of different combinations of diversification parameters.

Overall, the findings from the empirical analyses are consistent with the fundamental finding from the simulation experiments that parameter estimates are highly sensitive to methodological assumptions. More specifically, they also highlight that when methodological assumptions are simplistic (for example by assuming a strict clock, constant diversification parameters, or $\mu = 0$) increased character and taxon sampling can lead to more precise parameter estimates. By contrast, with more complex methodological assumptions (for example by assuming a relaxed clock, variable diversification parameters, or $\mu > 0$) inferences remained highly imprecise. This highlights the limited information content of the data being analysed.

Concluding comments

A robust phylogeny only provides information about certain aspects of macroevolutionary history. The estimation of divergence times and diversification parameters from phylogenies therefore requires important assumptions. Here, we show that in simplistic cases, and where methodological assumptions are not violated by the macroevolutionary processes that generated the data, increased character and taxon sampling can lead to more accurate and precise parameter estimates.

However, because of the overwhelming complexity of macroevolutionary history, the assumptions that underlie methods are often violated by the processes that generated the data. Where assumptions are violated, we show that increased character and taxon sampling can lead to inaccurate and misleadingly precise parameter estimates. Further, even when methods are implemented that are designed to better account for the complexity of macroevolutionary history, and are not therefore violated by the macroevolutionary process, parameter estimates can remain very imprecise, and in some cases misleading. In the analyses presented here, this was most notably the case when using relaxed clock models to estimate divergence times when there was among-branch-variation in r , and when using episodic models to estimate diversification parameters when there were shifts in μ . This finding results from the fact that the sequence data from which a phylogeny is constructed, and the molecular phylogeny itself, does not contain the necessary information from which to make the inferences that are desired. As such, divergence time estimation and diversification parameter estimation are distinct from phylogenetic inference, where molecular sequence data typically provides information that is directly relevant to the inference problem (Britton 2005; dos Reis and Yang 2013; Appendix 2). Because of this difference, divergence time estimates and diversification parameter estimates are likely to remain especially sensitive to the assumptions implemented in a given analysis, regardless of the quantity of data that is analysed.

Nonetheless, studies that incorporate divergence time estimates and diversification parameter estimates are central to our understanding of macroevolution (Baldwin and Sanderson 1998; Spriggs et al. 2015; Atchison et al. 2016; Lagomarsino et al. 2016; Nevado et al. 2016; Cardillo et al. 2017; Contreras-Ortiz et al. 2018; Folk et al. 2019; Muñoz-Rodríguez et al. 2019; Nürk et al. 2019). In order to ensure the validity of biological insights from these studies and future studies, the importance of careful consideration of methodological limitations cannot be overstated.

Given the complexity of macroevolutionary history and the assumptions that underlie analyses, effectively considering the limitations of different methods is not a straightforward task. For divergence times, the task can be somewhat rationalised by grouping assumptions into those that concern t (fossil calibrations, branching processes) and those that concern r (substitution rate models). For diversification parameters meanwhile, the focus will primarily rest on the implications of extinction, and nature and extent of parameter variation across the phylogeny. The impact of these assumptions can then be analysed, for example, by implementing calibrations with differing assumptions about the relationship between fossil ages and clade ages, comparing different branching process models, and testing different models of among-branch-variation in r . In a Bayesian framework, analyses can also be performed without data to further clarify the interactions between assumptions and parameter estimates. When performing such analyses, it is valuable to explicitly consider the likely validity of the assumptions that underlie a method, as well as the information provided by the data with respect to the parameter of interest. Steps such as this are inevitably painstaking and involve the careful repetition and reparameterization of analyses (Sanderson and Doyle 2001; Carruthers et al. 2020).

Further, where there is uncertainty in the topology of the underlying phylogeny, phylogenetic uncertainty must also be accounted for. Although the focus of this study is the

estimation of macroevolutionary parameters from a phylogeny that is assumed to be robust, phylogenetic uncertainty can have important implications for parameter estimation.

Nevertheless, with respect to estimating divergence times and diversification parameters, phylogenetic uncertainty often appears to be less important than subsequent methodological steps (Sanderson and Doyle 2001; Muñoz-Rodríguez et al. 2018; Morris et al. 2018)

In order to provide a degree of focus to the approach outlined above, and to avoid having to evaluate a prohibitively large array of different approaches, we also stress the importance of performing analyses that test specific biological hypotheses rather than attempting to re-construct general narratives about macroevolutionary history. By defining a specific hypothesis, it is easier to determine how methodological limitations are likely to affect the biological conclusions that are drawn from an analysis (Muñoz-Rodríguez et al. 2018; Carruthers and Scotland 2020). This can enable robust biological conclusions to be made, in spite of the fundamental limitations that underpin many analyses in this field.

SUPPLEMENTARY MATERIAL

Supplementary Material available from the Dryad Digital Repository:

[http://dx.doi.org/10.5061/dryad.\[NNNN\]](http://dx.doi.org/10.5061/dryad.[NNNN])

FUNDING

This work was supported by a NERC scholarship granted through the Environmental Research DTP programme to TC.

LITERATURE CITED

Atchison GW, Nevado B, Eastwood RJ, Ortiz-Contreras N, Reynal C, Madriñán S, Filatov DA, Huches CE. 2016. Lost crops of the Incas: Origins of domestication of the Andean pulse crop tarwi, *Lupinus mutabilis*. *Am J Bot.* 103:1592-1606.

Baldwin BG and Sanderson MJ. 1998. Age and rate of diversification of the Hawaiian

silversword alliance (Compositae). *Proc Natl Acad Sci USA*. 95:9402-9406.

Britton T. 2005. Estimating divergence times in phylogenetic trees without a molecular clock. *Syst Biol*. 54:500–507.

Brown JW and Smith SA. 2018. The past sure is tense: on interpreting phylogenetic divergence time estimates. *Syst Biol*. 67:340-353.

Cardillo M, Weston PH, Reynolds ZKM, Olde PM, Mast AR, Lemmon EM, Lemmon AR, Bromham L. 2017. The phylogeny and biogeography of *Hakea* (Proteaceae) reveals the role of biome shifts in a continental plant radiation. *Evolution*. 71:1928-1943.

Carruthers T, Muñoz-Rodríguez P, Wood JRI, Scotland RW. 2020. The temporal dynamics of evolutionary diversification in *Ipomoea*. *Mol Phylogenetics Evol*. 146:106768.

Carruthers T, Sanderson MJ, Scotland RW. 2019. The implications of lineage-specific rates for divergence time estimation. *Syst Biol*.

Carruthers T and Scotland RW. 2020. Estimating divergence times with multiple fossil calibrations and a relaxed clock. *Mol Biol Evol*.

Condamine FL, Rolland J, Morlon H. 2013. Macroevolutionary perspectives to environmental change. *Ecol Lett*. 16:72-85.

Contreras-Ortiz N, Atchison GW, Hughes CE, Madriñán S. 2018. Convergent evolution of high elevation plant growth forms and geographically structured variation in Andean *Lupinus* (Fabaceae). *Bot J Linn Soc.* 187:118-136.

Crisp MD, Arroyo MTK, Cook LG, Gandolfo MA, Jordan GJ, McGlone MS, Weston PH, Westoby M, Wilf P, Linder HP. 2009. Phylogenetic biome conservatism on a global scale. *Nature.* 458:754-756.

Cusimano N and Renner SS. 2010. Slowdowns in diversification rates from real phylogenies may not be real. *Syst Biol.* 59:458-464.

Cusimano N, Stadler T, Renner SS. 2012. A new method for handling missing species in diversification analysis applicable to randomly or nonrandomly sampled phylogenies. *Syst Biol.* 61:785-792.

Donoghue MJ. 2008. A phylogenetic perspective on the distribution of plant diversity. *Proc Natl Acad Sci USA.* 105:11549–11555.

Donoghue MJ, Bell CD, Li J. 2001. Phylogenetic patterns in Northern Hemisphere plant geography. *Int J Pl Sci.* 162(Suppl.):S41–S52.

Drummond AJ, Ho SYW, Phillips MJ, Rambaut A. 2006. Relaxed Phylogenetics and Dating with Confidence. *PLOS Biol.* 4:e88.

Drummond AJ and Suchard MA. 2011. Bayesian random local clocks, or one rate to rule

them all. *BMC Biol.* 8:114.

Drummond CS, Eastwood RJ, Miotto STS, Hughes CE. 2012. Multiple Continental Radiations and Correlates of Diversification in *Lupinus* (Leguminosae): Testing for Key Innovation with Incomplete Taxon Sampling. *Syst Biol.* 61:443-460.

Folk RA, Stubbs RL, Mort ME, Cellinese N, Allen J, Soltis PS, Soltis DE, Guralnick RP. 2019. Rates of niche and phenotype evolution lag behind diversification in a temperate radiation. *Proc Natl Acad Sci USA.* 116: 10874-10882.

Fitzjohn RG. 2012. Diversitree: comparative phylogenetic analyses of diversification in R. *Methods Ecol Evol.* 3:1084-1092.

Givnish TJ, Millam KC, Mast AR, Paterson TB, Theim TJ, Hipp AL, Henss JM, Smith JF, Wood KR, Sytsma KJ. 2009. Origin, adaptive radiation and diversification of the Hawaiian lobeliads (Asterales: Campanulaceae). *Proc R Soc B.* 276:407-416.

Grimm GW, Kapli P, Bomfleur B, McLoughlin S, Renner SS. 2015. Using more than the oldest fossils: dating Osmundaceae with three Bayesian clock approaches. *Syst Biol.* 64:396-405.

Heath TA, Huelsenbeck JP, Stadler T. 2014. The fossilized birth-death process for coherent calibration of divergence-time estimates. *Proc Natl Acad Sci USA.* 111:E2957-E2966.

Ho SYW. 2014. The changing face of the molecular evolutionary clock. *Trends Ecol Evol.*

Höhna S, Landis MJ, Heath TA, Boussau B, Lartillot N, Moore BR, Huelsenbeck JP, Ronquist F. 2016. RevBayes: Bayesian Phylogenetic Inference Using Graphical Models and an Interactive Model-Specification Language. *Syst Biol.* 65:726–736.

Höhna S, Stadler T, Ronquist F, Britton T. 2011. Inferring speciation and extinction rates under different species sampling schemes. *Mol Biol Evol.* 28:2577–2589.

Höhna S. 2014. Likelihood inference of non-constant diversification rates with incomplete taxon sampling. *PLOS ONE.* 9:e84184.

Hughes CE. and Eastwood R. 2006. Island radiation on a continental scale: Exceptional rates of plant diversification after uplift of the Andes. *Proc Natl Acad Sci USA.* 103:10334-10339.

Kishino H, Thorne JL, Bruno WJ. 2001. Performance of a Divergence Time Estimation Method under a Probabilistic Model of Rate Evolution. *Mol Biol Evol.* 18:352–361.

Lagomarsino LP, Condamine FL, Antonelli A, Mulch A, Davis CC. 2016. The abiotic and biotic drivers of rapid diversification in Andean bellflowers (Campanulaceae). *New Phytol.* 210:1430-1442.

Lavin M, Schrire BP, Lewis G, Pennington RT, Delgado-Salinas A, Thulin M, Hughes CE, Matos AB, Wojciechowski MF. 2004. Metacommunity process rather than continental tectonic history better explains geographically structured phylogenies in legumes. *Philos*

Trans Royal Soc. 359:1509-1522.

Louca S and Pennell MW. 2020. Extant timetrees are consistent with a myriad of diversification histories. *Nature*. 580:502-505.

Maddison WP, Midford PE, Otto SP. 2007. Estimating a Binary Character's Effect on Speciation and Extinction. *Syst Biol.* 56:701-710.

Magallón S, Hilu KW, Quandt D. 2013. Land plant evolutionary timeline: Gene effects are secondary to fossil constraints in relaxed clock estimation of age and substitution rates. *Am J Bot.* 100:556-573.

Magallón S, Gomez-Acevedo S, Sanchez-Reyes LL, Hernandez-Hernandez T. 2015. A metacalibrated time-tree documents the early rise of flowering plant phylogenetic diversity. *New Phytol.* 207:437–453.

Maurin O, Davies TJ, Burrows JE, Daru BH, Yessoufou K, Muasya AM, van der Bank M, Bond WJ. 2014. Savanna fire and the origins of the ‘underground forests’ of Africa. *New Phytol.* 204:201–214.

Moore BR, Höhna S, May MR, Rannala B, Huelsenbeck JP. 2016. Critically evaluating the theory and performance of Bayesian analysis of macroevolutionary mixtures. *Proc Natl Acad Sci USA.* 113:9569-9574.

Muñoz-Rodríguez P, Carruthers T, Wood JRI, Williams BRM, Weitemier K, Kronmiller B, Ellis D, Anglin NL, Longway L, Harris SA, Rausher MD, Kelly S, Liston A, Scotland RW. 2018. Reconciling Conflicting Phylogenies in the Origin of Sweet Potato and Dispersal to Polynesia. *Curr Biol.* 28:1246-1256.

Muñoz-Rodríguez P., Carruthers T., Wood J. R. I., Williams B. R. M., Weitemier K., Kronmiller B., Goodwin Z., Sumadijaya A., Anglin N. L., Filer D., Harris D., Rausher M. D., Kelly S., Liston A., Scotland R. W. 2019. A taxonomic monograph of *Ipomoea* integrated across phylogenetic scales. *Nat. Plants.* 5:1136-1144.

Near TJ, Meylan PA, Shaffer HB. 2005. Assessing Concordance of Fossil Calibration Points in Molecular Clock Studies: An Example Using Turtles. *Am Nat.* 165:137-146.

Near TJ and Sanderson MJ. 2004. Assessing the quality of molecular divergence time estimates by fossil calibrations and fossil-based model selection. *Phil Trans R Soc Lond B.* 359:1477-1483.

Nevado B, Atchison GW, Hughes CE, Filatov DA. 2016. Widespread adaptive evolution during repeated evolutionary radiations in New World lupins. *Nat Commun.* 7:1-9.

Nürk NM, Atchison GW, Hughes CE. 2019. Island woodiness underpins accelerated disparification in plant radiations. *New Phytol.* 224:518-531.

Panero JL and Funk VA. 2008. The value of sampling anomalous taxa in phylogenetic studies: Major clades of the Asteraceae revealed. *Mol Phylogenetics Evol.* 47:757-782.

Patterson C. 1981. The significance of fossils in determining evolutionary relationships. *Ann Rev Ecol Syst.* 12:195-223.

Pennington RT, Lavin M, Oliveira-Filho A. 2009. Woody Plant Diversity, Evolution, and Ecology in the Tropics: Perspectives from Seasonally Dry Tropical Forests. 40:437-457. *Annu Rev Ecol Evol. Syst.* 40:437-457.

Pennington RT, Lavin M, Prado DE, Pendry CA, Pell SK, Butterworth CA. 2004. Historical climate change and speciation: neotropical seasonally dry forest plants show patterns of both Tertiary and Quaternary diversification. *Proc R Soc B.* 359:515-537.

Pennington RT and Hughes CE. 2014. The remarkable congruence of New and Old World savanna origins. *New Phytol.* 204:4-6.

Platnick NL. 1979. Philosophy and the transformation of cladistics. *Syst Zool.* 28:537-546.

Rabosky DL. 2010. Extinction rates should not be estimated from molecular phylogenies. *Evolution.* 64:1816-1824.

Rabosky DL. 2014. Automatic detection of key innovations, rate shifts, and diversity-dependence on phylogenetic trees. *PLoS ONE* 9:e89543.

Rabosky DL and Lovette IJ. 2008. Density-dependent diversification in North American wood warblers. *Proc R Soc B.* 275:2363-2371.

Rabosky DL, Mitchell JS, Chang J. 2017. Is BAMM flawed? Theoretical and practical concerns in the analysis of multi-rate diversification models. *Syst Biol.* 66:477-498.

dos Reis M and Yang Z. 2013. The unbearable uncertainty of Bayesian divergence time estimation. *J Syst Evol.* 51:30-43.

Sanderson MJ. 1997. A Nonparametric Approach to Estimating Divergence Times in the Absence of Rate Constancy. *Mol Biol Evol.* 14:1218–1231.

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

Sanderson MJ and Magallón SA. 2005. Angiosperm divergence times: the effect of genes, codon positions and time constraints. *Evolution.* 59:1653-1670.

Sanderson MJ and Doyle JA. 2001. Sources of error and confidence intervals in estimating the age of angiosperms from *rbcL* and 18S rDNA data. *Am J Bot.* 88:1499-1516.

Särkinen T, Bohs L, Olmstead RG, Knapp S. 2013. A phylogenetic framework for evolutionary study of the nightshades (Solanaceae): a dated 1000 tip tree. *BMC Evol Biol.* 13:214.

Särkinen T, Pennington RT, Lavin M, Simon MF, Hughes CE. 2012. Evolutionary islands in the Andes: persistence and isolation explain high endemism in Andean dry tropical forests. *J Biogeogr.* 39:884-900.

Schliep K. P. 2011. phangorn: Phylogenetic analysis in R. *Bioinformatics.* 27:592–593.

Schliep K. P. 2018. phangorn: Phylogenetic Reconstruction and Analysis. R package Version 2.4.0. Available at: <https://CRAN.R-project.org/package=phangorn>.

Simon MF, Grether R, de Queiroz LP, Skema C, Pennington RT, Hughes CE. 2009. Recent assembly of the Cerrado, a neotropical plant diversity hotspot, by in situ evolution of adaptations to fire. *Proc. Natl Acad Sci USA.* 106:20359-20364.

Simon MF and Pennington T. 2012. Evidence for adaptation to fire regimes in the tropical savannas of the Brazilian cerrado. *Int J Plant Sci.* 173:711-723.

Soltis DE, Soltis PS, Chase MW, Mort ME, Albach DC, Zanis M, Savolainen V, Hahn WH, Hoot SB, Fay MF, Axtell M, Swensen SM, Prince LM, Kress WJ, Nixon KC, Farris JS. 2000. Angiosperm phylogeny inferred from *18S* rDNA, *rbcL* and *atpB* sequences. *Bot J Linn Soc.* 133:381-461.

Spriggs EL, Clement WL, Sweeney PW, Madriñán S, Edwards EJ, Donoghue MJ. 2015. Temperate radiations and dying embers of a tropical past: the diversification of *Viburnum*. *New Phytol.* 207:340-354.

Stadler T. 2009. On incomplete sampling under birth-death models and connections to the sampling-based coalescent. *J Theor Biol.* 261:58-66.

Stadler T. 2011a. Mammalian phylogeny reveals recent diversification rate shifts. *Proc Natl Acad Sci USA.* 108:6187–6192.

Stadler T. 2011b. Inferring speciation and extinction processes from extant species data. *Proc Natl Acad Sci USA.* 27:16145-16146.

Stadler T. 2019. TreeSim: Simulating Phylogenetic Trees. R package version 2.4. Available at: <https://CRAN.R-project.org/package=TreeSim>.

Stadler T and Bokma F. 2013. Estimating speciation and extinction rates for phylogenies of higher taxa. *Syst Biol.* 62:220-230.

Tamura K, Battistuzzi FU, Billings-Ross P, Murillo O, Filipowski A, Kumar S. 2012. Estimating divergence times in large molecular phylogenies. *Proc Natl Acad Sci USA.* 109:19333-19338.

Thorne JL, Kishino H, Painter IS. 1998. Estimating the rate of evolution of the rate of molecular evolution. *Mol Biol Evol.* 15:1647–1657.

Warnock RCM, Parham JF, Joyce WG, Tyler R, Donoghue PCJ. 2015. Calibration uncertainty in molecular dating analyses: there is no substitute for the prior evaluation of time priors. *Proc R Soc B.* 282:20141013.

Warnock RCM, Yang ZH, Donoghue PCJ. 2011. Exploring uncertainty in the calibration of the molecular clock. *Biol Lett.* 8:156 – 159.

Warnock RCM, Yang Z, Donoghue PCJ. 2017. Testing the molecular clock using mechanistic models of fossil preservation and molecular evolution. *Proc R Soc B.* 284:20170227.

Wertheim JO and Sanderson MJ. 2010. Estimating diversification rates: how useful are divergence times? *Evolution.* 65:309-320.

Zhu T, dos Reis M, Yang Z. 2015. Characterization of the Uncertainty of Divergence Time Estimation under Relaxed Molecular Clock Models Using Multiple Loci. *Syst Biol.* 64:267–280.

CAPTIONS

Figure 1. Summary of the configuration of the simulation experiments performed in this study. These were designed to capture the relationship between macroevolutionary history, data, sampling, assumptions, divergence time estimates, and diversification parameter estimates. a) shows macroevolutionary processes that generate the extant species sampled in a molecular phylogeny. These macroevolutionary processes may or may not have extinct branches, and may or may not have diversification parameter variation. b) shows the types of data that evolve according to the macroevolutionary processes and which can be used to infer a time-calibrated phylogeny. c) shows the different sampling strategies that can be used when inferring divergence times and d) shows the different assumptions that can be implemented. The implementation of a sampling strategy and assumptions enables a time-calibrated phylogeny (e) to be inferred. In an empirical study, the inferred time-calibrated phylogeny would be combined with different assumptions (f) in order to estimate diversification parameters (g). However, in the simulation experiments presented here, the “correct” time-calibrated phylogeny that corresponds to the simulated macroevolutionary process was used, in combination with different sampling strategies (h) and assumptions (f).

Figure 2. The implications of increased character sampling (bp) for estimating divergence times when there is no among-branch-variation in r , diversification parameters are constant, and $\mu = 0$. Black points indicate % error, grey bars indicate 95% highest posterior density (HPD) intervals. a) low character sampling, b) medium character sampling, c) high character sampling, d) no sequence data.

Figure 3. The implications of increased taxon sampling for estimating divergence times when there is no among-branch-variation in r and diversification parameters are constant. Black points and grey bars are the same as for Figure 2. a) low taxon sampling and $\mu = 0$, b) medium taxon sampling and $\mu = 0$, c) high taxon sampling and $\mu = 0$, d) low taxon sampling with no sequence data and $\mu = 0$, e) medium taxon sampling with no sequence data and $\mu = 0$, f) high taxon sampling with no sequence data and $\mu = 0$, g) high taxon sampling and $\mu > 0$, h) high taxon sampling with no sequence data and $\mu > 0$.

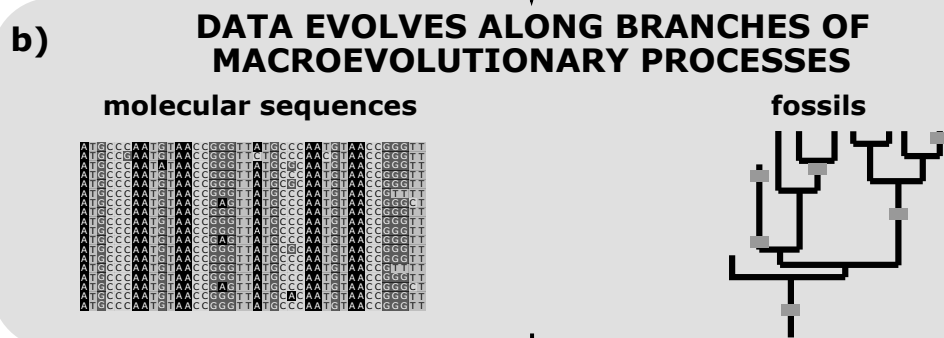
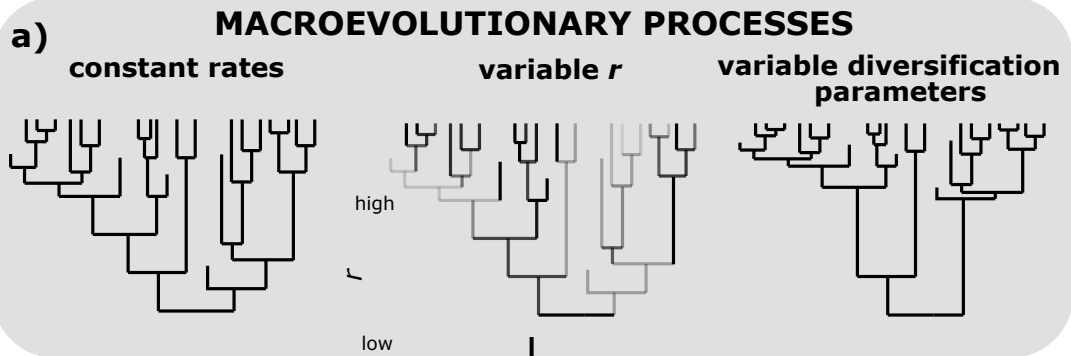
Figure 4. The implications of increased character sampling (bp) for estimating divergence times when there is among-branch-variation in r , but diversification parameters are constant, and $\mu = 0$. Black points and grey bars are the same as for Figure 2. a) low character sampling, b) medium character sampling, c) high character sampling, d) high character sampling with a relaxed clock, e) no sequence data and no node calibrations, f) high character sampling with a relaxed clock and node calibrations, g) no sequence data with node calibrations.

Figure 5. The implications of increased taxon sampling for estimating divergence times when there is variation in diversification parameters, $\mu = 0$, and no among-branch-variation in r . Black points and grey bars are the same as for Figure 2. a) low taxon sampling, b) medium taxon sampling, c) high taxon sampling, d) low taxon sampling with no sequence data, e) medium taxon sampling with no sequence data, f) high taxon sampling with no sequence data, g) high taxon sampling with node calibrations, h) high taxon sampling with no sequence data but with node calibrations.

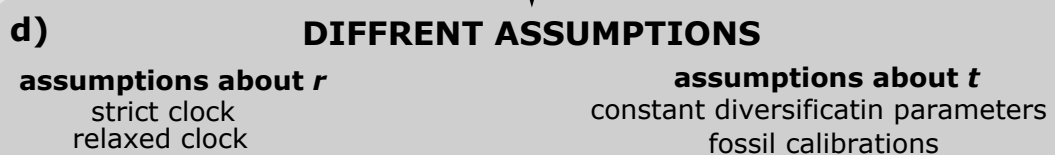
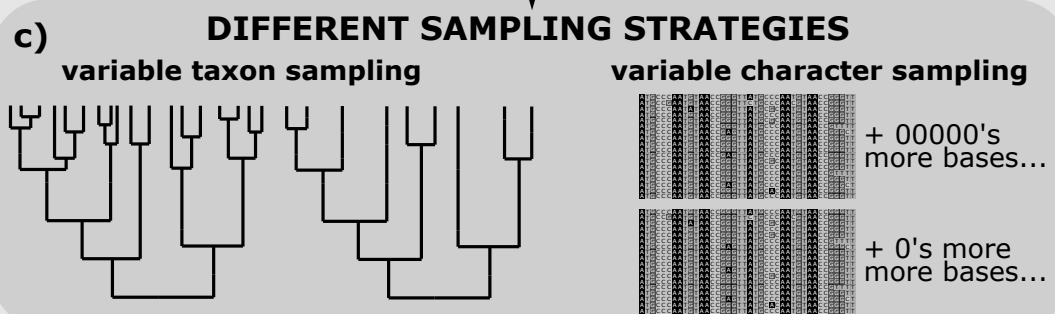
Figure 6. The implications of different levels of taxon sampling, character sampling, and molecular clock models for estimating the age of the radiation node in *Ipomoea*. Points refer to the mean posterior estimate (MPE) for the age of the radiation node, grey bars refer to the

95% HPD interval for the age of the radiation node, *lts* refers to low taxon sampling, *mts* refers to medium taxon sampling, and *hts* refers to high taxon sampling. For low character sampling and high character sampling, black points refer to the MPE with a strict clock, and black rings with a white centre refer to the MPE with a relaxed clock.

h) in simulations here, correct time-calibrated phylogeny is used (equivalent to macroevolutionary process without extinct branches) with different levels of taxon sampling

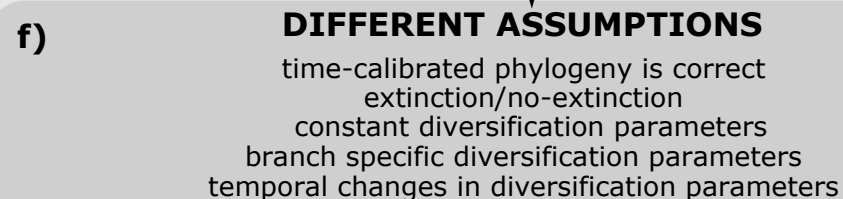


DIVERGENCE TIME ESTIMATION

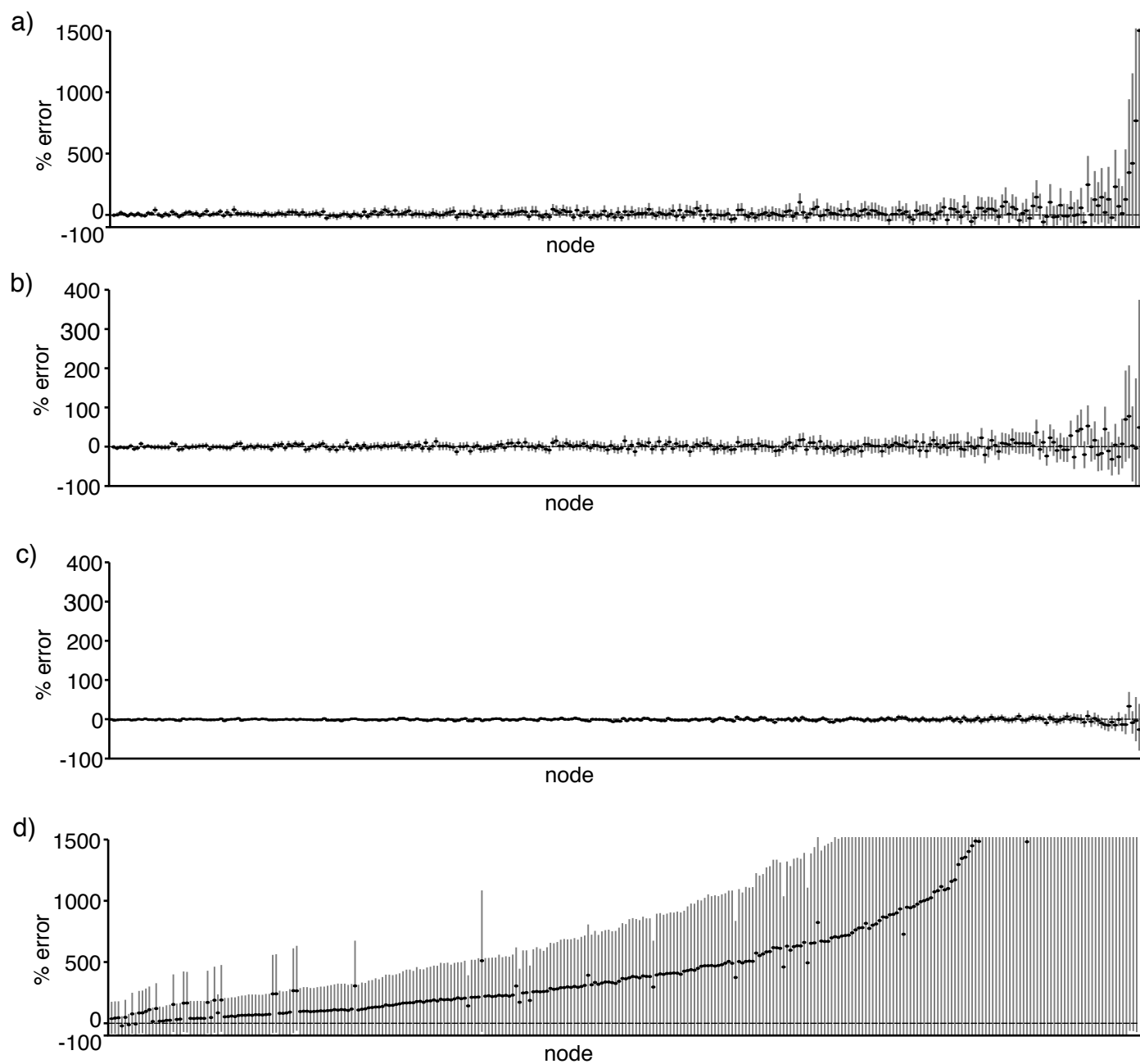


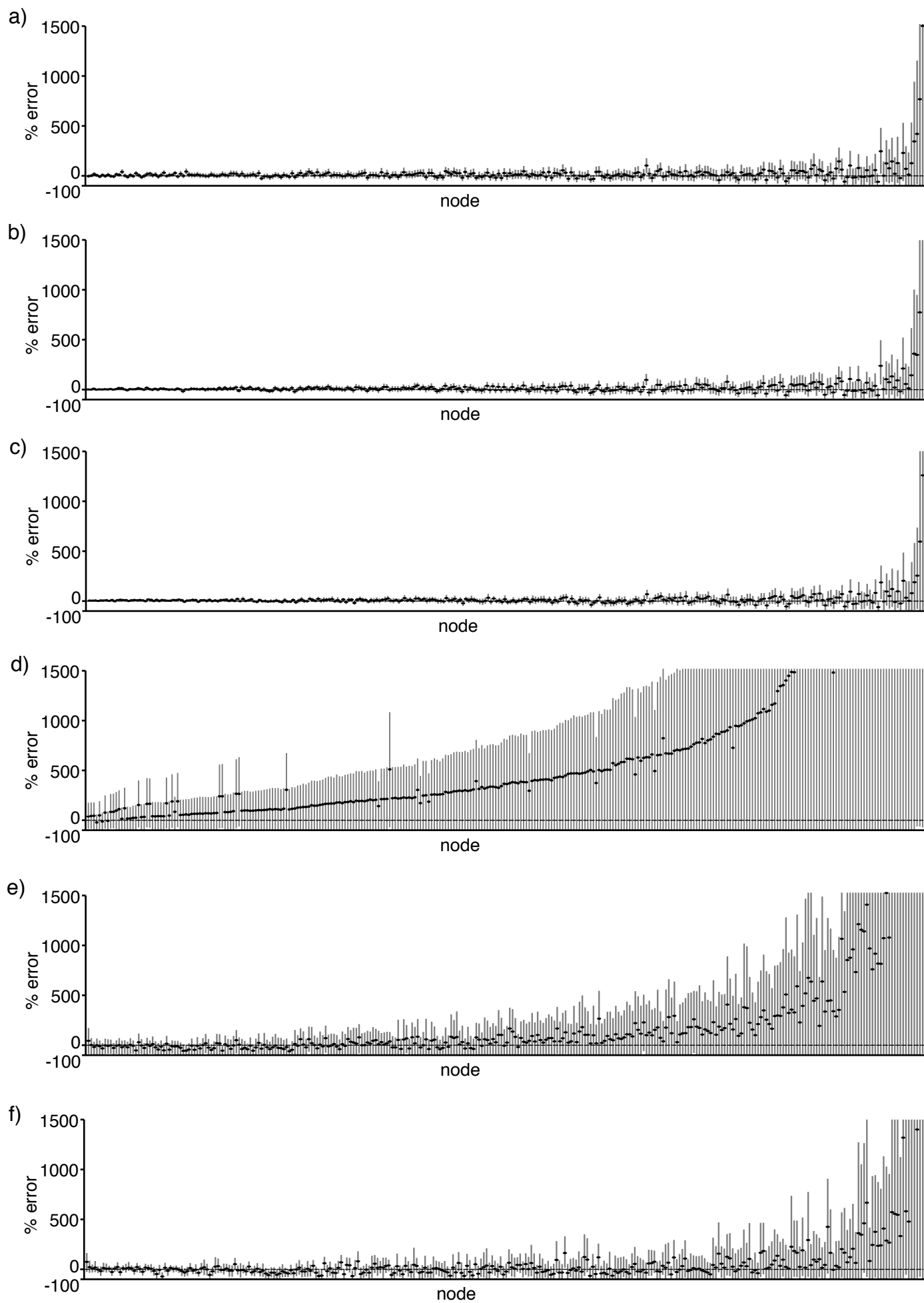
e) TIME-CALIBRATED PHYLOGENY

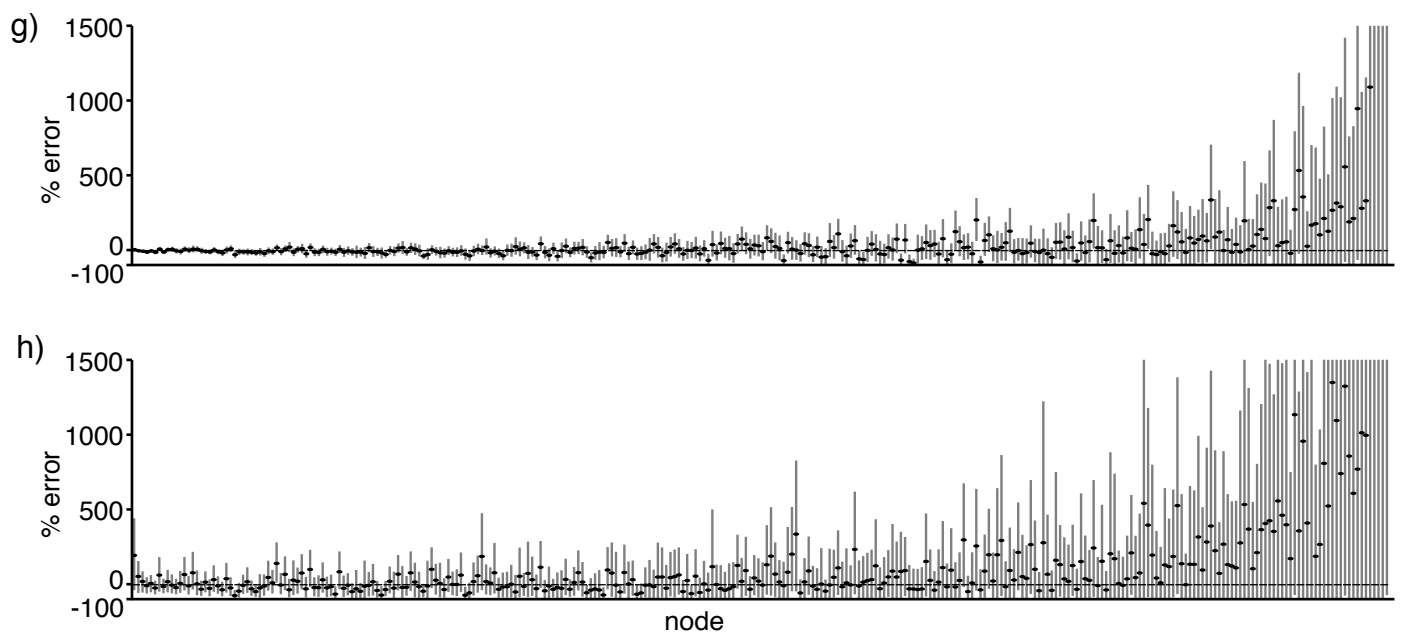
DIVERSIFICATION PARAMETER ESTIMATION

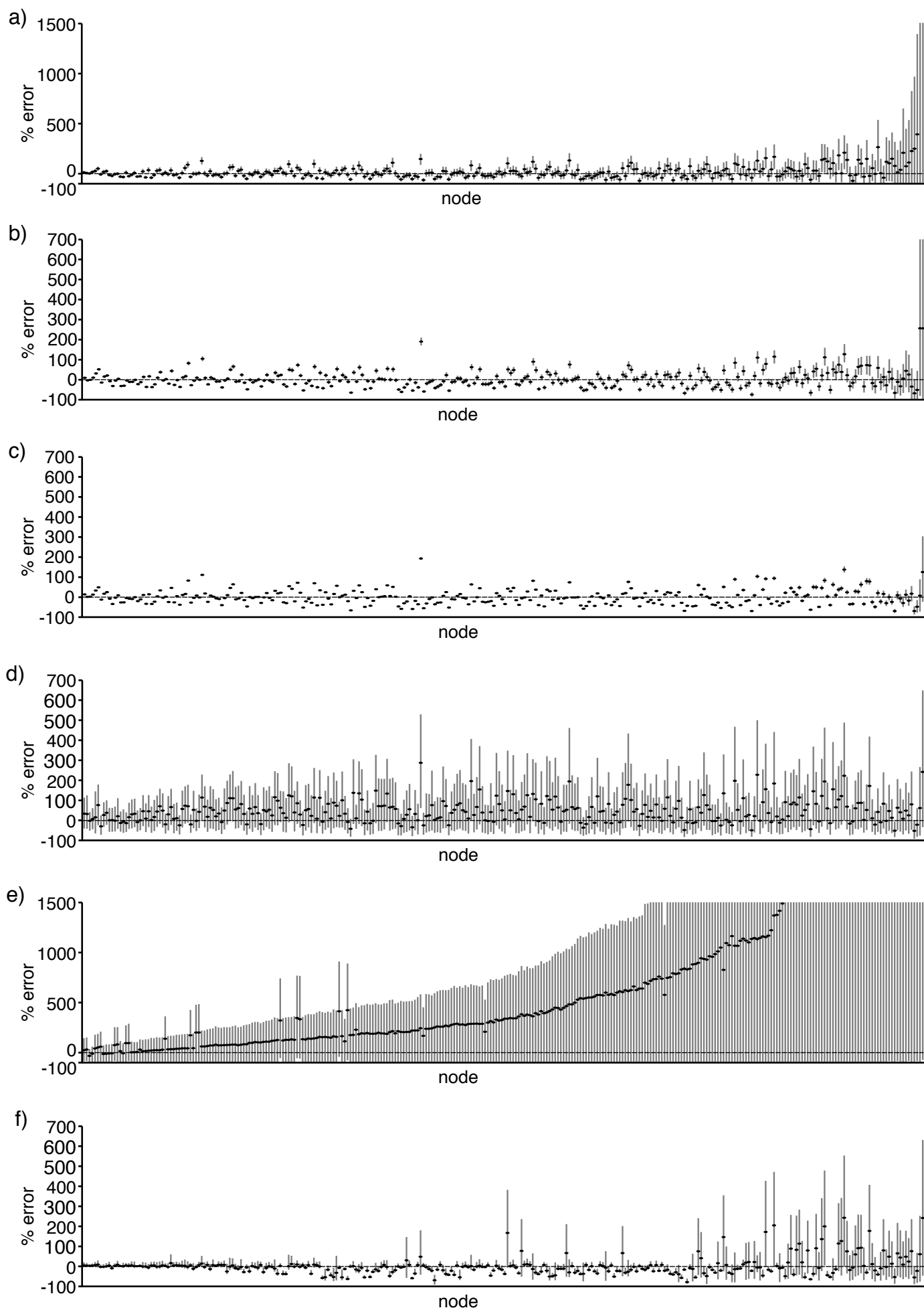


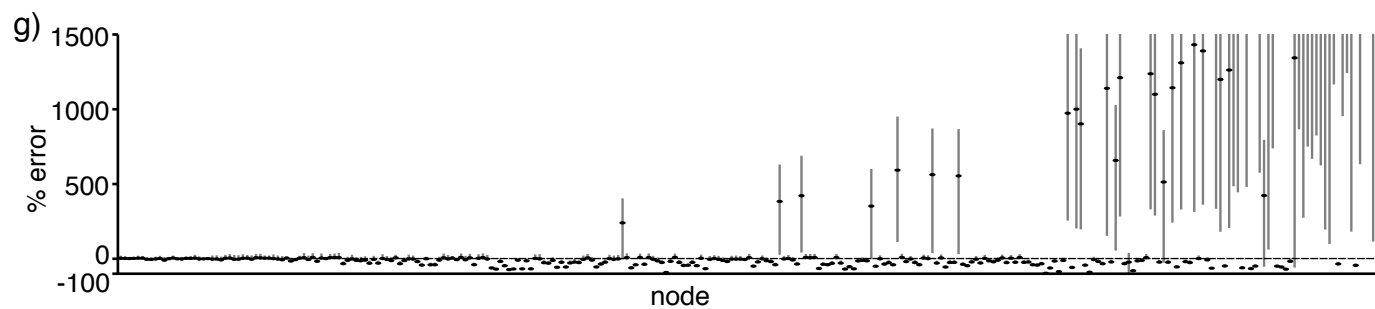
g) DIVERSIFICATION PARAMETERS ESTIMATED

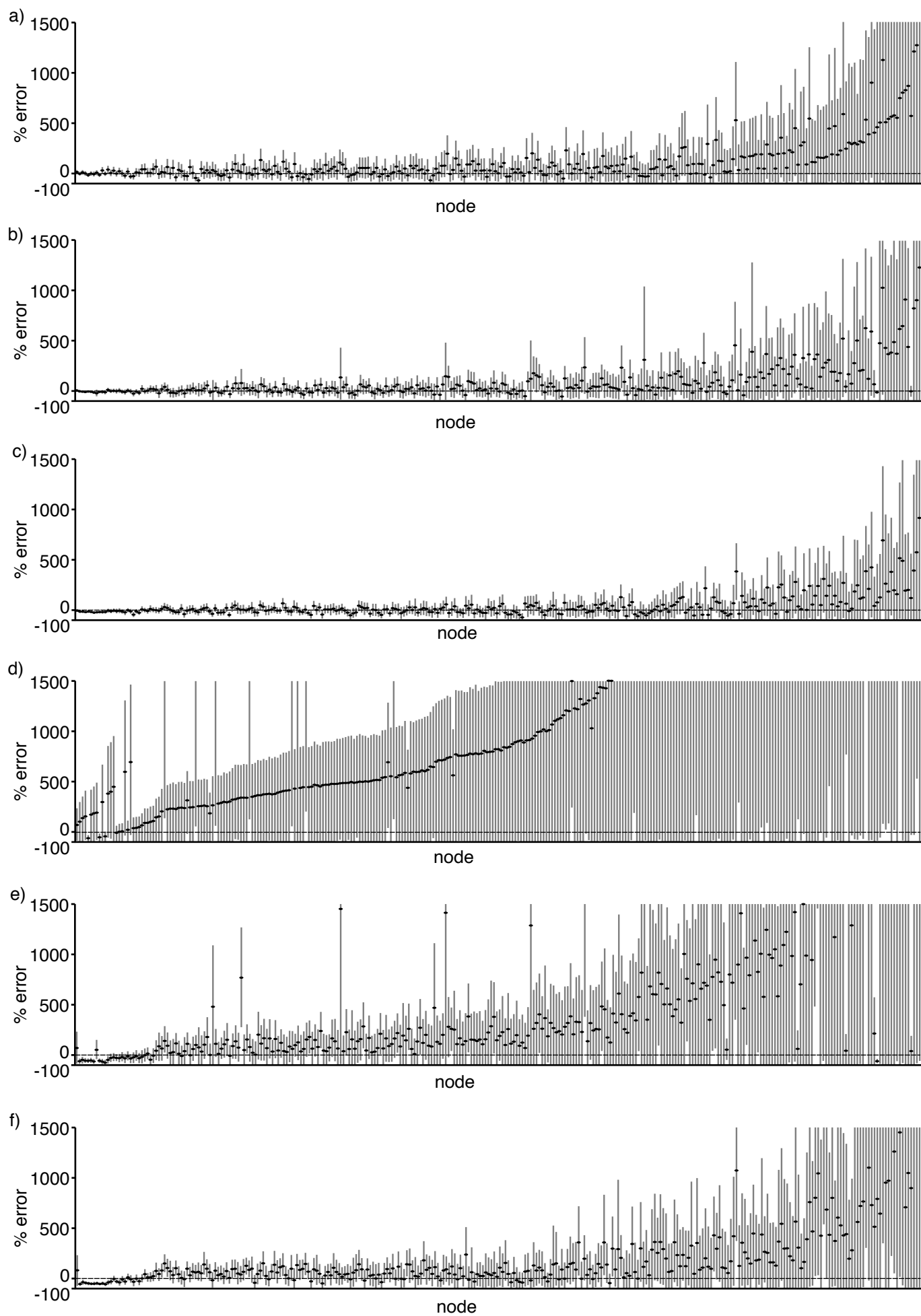


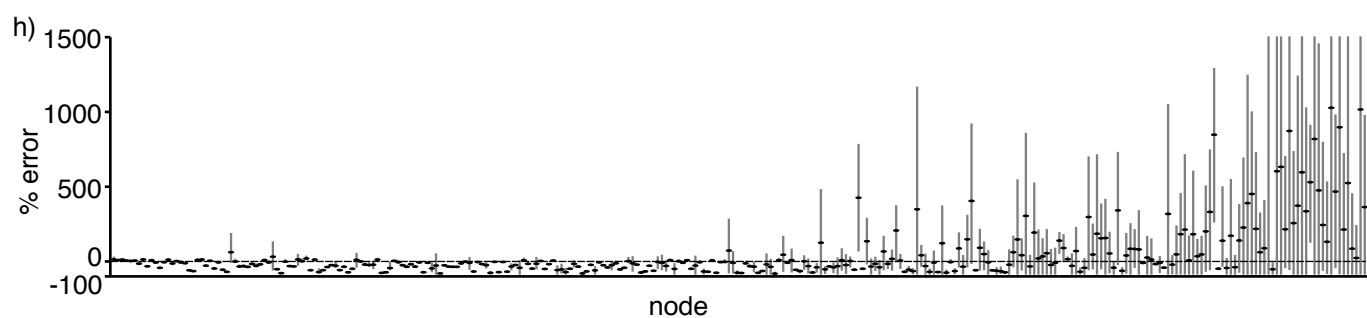
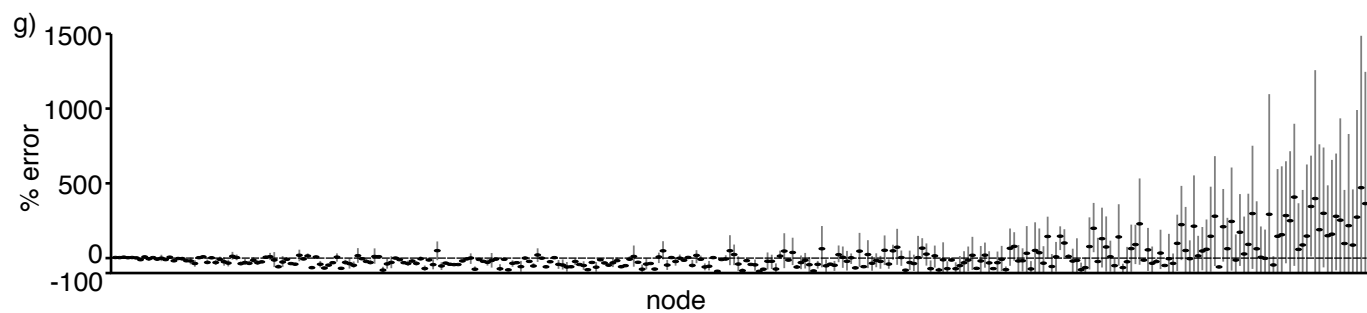












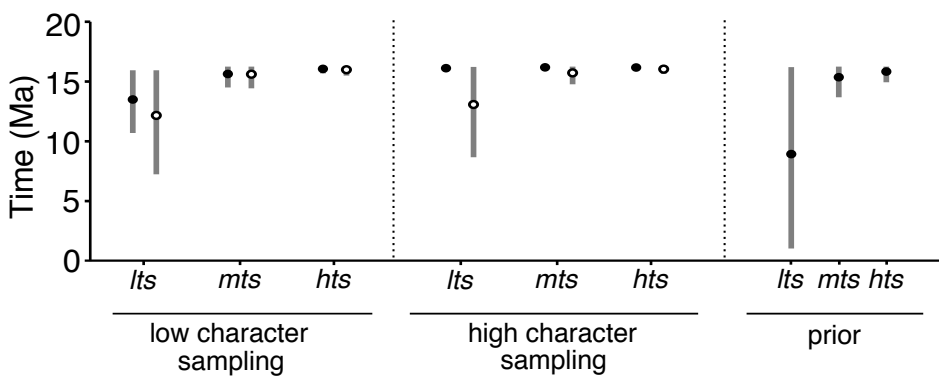


Table 1. A summary of experiments performed to investigate the implications of different assumptions and different levels of character and taxon sampling for divergence time estimation.

Simulated diversification parameters	Simulated r	Taxon Sampling	Character Sampling	Assumptions about r	Assumptions about t
con λ , $\mu = 0$	con	low	low	con	con λ , $\mu = 0$
con λ , $\mu = 0$	con	low	medium	con	con λ , $\mu = 0$
con λ , $\mu = 0$	con	low	high	con	con λ , $\mu = 0$
con λ , $\mu = 0$	con	low	low	con	con λ , $\mu = 0$
con λ , $\mu = 0$	con	medium	low	con	con λ , $\mu = 0$
con λ , $\mu = 0$	con	high	low	con	con λ , $\mu = 0$
con λ , con μ	con	high	low	con	con λ , con μ
con λ , $\mu = 0$	abv	low	low	con	con λ , $\mu = 0$
con λ , $\mu = 0$	abv	low	medium	con	con λ , $\mu = 0$
con λ , $\mu = 0$	abv	low	high	con	con λ , $\mu = 0$
con λ , $\mu = 0$	abv	low	high	abv	con λ , $\mu = 0$
con λ , $\mu = 0$	abv	low	high	abv	con λ , $\mu = 0$, foss
ts λ , $\mu = 0$	con	low	low	con	con λ , $\mu = 0$
ts λ , $\mu = 0$	con	medium	low	con	con λ , $\mu = 0$
ts λ , $\mu = 0$	con	high	low	con	con λ , $\mu = 0$
ts λ , $\mu = 0$	con	high	low	con	con λ , $\mu = 0$, foss
empirical analyses		low	low	con	con λ , con μ
		low	low	abv	con λ , con μ
		low	high	con	con λ , con μ
		low	high	abv	con λ , con μ
		medium	low	con	con λ , con μ
		medium	low	abv	con λ , con μ
		medium	high	con	con λ , con μ
		medium	high	abv	con λ , con μ
		high	low	con	con λ , con μ
		high	low	abv	con λ , con μ
		high	high	con	con λ , con μ
		high	high	abv	con λ , con μ

*con refers to constant, abv refers to among branch variation, ts refers to time specific.

Table 2. A summary of experiments performed to investigate the implications of different assumptions and different levels of taxon sampling for diversification parameter estimation.

Simulated diversification parameters	Taxon Sampling	Assumptions about diversification parameters
con λ , $\mu = 0$	low	con λ , $\mu = 0$
con λ , $\mu = 0$	medium	con λ , $\mu = 0$
con λ , $\mu = 0$	high	con λ , $\mu = 0$
con λ , con μ	high	con λ , con μ
ts λ , $\mu = 0$	low	con λ , $\mu = 0$
ts λ , $\mu = 0$	medium	con λ , $\mu = 0$
ts λ , $\mu = 0$	high	con λ , $\mu = 0$
ts λ , $\mu = 0$	high	ts λ , $\mu = 0$
con λ , ts μ	high	con λ , ts μ
empirical analyses	low	con λ , $\mu = 0$
	medium	con λ , $\mu = 0$
	high	con λ , $\mu = 0$
	low	con λ , con μ
	medium	con λ , con μ
	high	con λ , con μ
	high	bs λ
	high	bs λ , con μ
	high	con λ , bs μ

*con refers to constant, ts refers to time specific, bs refers to branch specific.