Title: Hidden-state-only speciation and extinction models provide accurate tip estimates of diversification rates

Authors: Thaís Vasconcelos[1], Brian C. O'Meara[2], Jeremy M. Beaulieu[1]

[1] Department of Biological Sciences, University of Arkansas, Fayetteville, AR 72701, USA.

[2] Department of Ecology and Evolutionary Biology, University of Tennessee, Knoxville, TN 37996, USA.

Running headline: Hidden-rate-only models of diversification

Abstract

1. State-dependent speciation and extinction (SSE) models provide a framework for testing potential correlations between the evolution of an observed trait and speciation and extinction rates. Recent expansions of these models allow for the inclusion of "hidden states" that, among other things, allow for rate heterogeneity often observed among lineages sharing a particular character state. However, in reality, multiple circumstances and interacting traits related to a focal character play a role in changing diversification dynamics of a lineage over time, restricting the use of available SSE models that require trait information to be assigned at the tips.

2. Here we introduce MiSSE, an SSE approach that infers diversification rate differences from hidden states only. It can be used similarly to other trait-free methods to estimate varying speciation, extinction, but also different functions of these parameters such as net-diversification, turnover rates, and extinction fraction. Given the size of the model space, we also describe an algorithm designed for efficiently searching through a reasonably large set of models without having to be exhaustive.

3. We compare the accuracy of rates inferred at the tips of the tree by MiSSE against popular character-free methods and demonstrate that the error associated with tip estimates is generally low. Due to certain characteristics of the SSE models, this method avoids some of the recent concerns with parameter identifiability in diversification analyses and can be used alongside regular phylogenetic comparative methods in trait-related diversification hypotheses.

4. Finally, we apply MiSSE, with a renewed focus on classic comparative methods, to understand processes happening near the present, rather than deep in the past, to examine how variation in plant height has impacted turnover rates in eucalypts, a species-rich lineage of flowering plants.

Key words: diversification, *Eucalyptus*, hidden-states, Myrtaceae, plant height, turnover, state-dependent speciation and extinction.

1. INTRODUCTION

Molecular phylogenies, when scaled in relation to time, are powerful sources of data to understand the diversification dynamics of organisms and have become crucial in multiple areas of ecology and evolution (Wiens & Donoghue, 2004). As the trees themselves become bigger, more robust, and increasingly complete in proportion to species sampled (e.g. Beaulieu & O'Meara, 2018; Smith & Brown, 2018), the statistical tools used to infer macroevolutionary patterns from them also become more biologically realistic (e.g. Beaulieu & O'Meara, 2016). Modelling the dynamics of lineage origination and extinction through time have allowed us to understand, for example, how shifts in pollination and dispersal strategies are connected to changes in diversification rates of angiosperms (e.g. Lagomarsino et al., 2016; Vasconcelos et al., 2019; Reginato et al., 2020) or the role of environmental instability in the diversification of several vertebrate groups (e.g. Harvey et al., 2020; Morales-Barbero et al., 2021).

Recently, however, a renewed wave of criticisms regarding these methods calls into question whether diversification rates from time-calibrated trees of extant-only organisms should even be estimated at all (Louca & Pennell, 2020). While it is true that phylogenies are often used to address problems beyond their capabilities (Losos, 2011; Cooper et al., 2016; Uyeda et al., 2018), there is still a considerable amount of information that extant-only phylogenies can provide about the diversification process. For instance, O'Meara and Beaulieu (in prep.) demonstrate that state-dependent speciation and extinction (SSE) models can identify different likelihoods for the generating parameters of trees with different topologies but identical lineage through time plots. However, the uncertainty around parameter estimates increases as one moves from tip to root in the tree, due to the decreasing amount of information in ancestral rate reconstructions. A solution, then, could be to continue using SSE models to model diversification dynamics, but focus only on rates estimated near the present or at the tips of the tree, rather than deep in time. Such an approach can have additional advantages for its flexibility in testing for correlations between multistate discrete and continuous traits (Harvey & Rabosky, 2018; Title & Rabosky, 2019).

Here we formally describe MiSSE, a new extension of the SSE framework that uses hidden-states only and provides accurate estimates of various metrics of diversification rates at the tips of a tree. We compare its accuracy against other popular approaches for estimating tip diversification rates, and further demonstrate the capabilities of MiSSE with an empirical

example that examines the correlation between turnover rates and plant height in eucalypts (Eucalypteae, Myrtaceae), a diverse lineage of flowering plants. Finally, we argue why focusing on tip estimates can be advantageous in testing complex hypotheses of diversification and discuss some caveats of MiSSE and possible ways forward for modeling diversification.

2. MATERIALS AND METHODS

2.1. The MiSSE model

State-dependent speciation and extinction models expand the birth-death process to account for speciation (λ), extinction (μ), and trait evolution (q, the transition rates between character states), estimating parameters that maximize the likelihood of observing both the character states at the tips of the tree and the tree itself (Maddison et al., 2007). SSE models were initially developed to overcome three main perceived shortcomings in the field: (1) the need for greater flexibility in tests of key-innovation hypotheses, which, at the time, typically involved sister-clade comparisons that could only measure differences in net-diversification and assumed constant rates within the clades under comparison (Barraclough et al., 1998); (2) the need to incorporate differences in speciation and/or extinction rates associated with a particular character state in analyses of trait and geographical range evolution (Goldberg et al., 2010; Goldberg & Igic, 2012; Ree & Sanmartin, 2018); and (3) the potentially confounding effects of different transition rates and diversification rates when looking at diversification or character evolution, respectively (Maddison, 2006).

All discrete SSE models exist within the following generalized ordinary differential equations (corresponding to eq. 1a-b in FitzJohn, 2012):

$$\frac{dE_i}{dt} = \mu_i - (\lambda_i + \mu_i + \sum_{i \neq j} q_{ij})\, E_i(t) + \lambda_i E_i(t)^2 + \sum_{j \neq i} q_{ij} E_j(t) \tag{1a}$$

$$\frac{dD_{N,i}}{dt} = -(\lambda_i + \mu_i + \sum_{i \neq j} q_{ij})\, D_{N,i}(t) + 2\,\lambda_i\, D_{N,i}(t)\, E_i(t) + \sum_{j \neq i} q_{ij}\, D_{N,j}(t) \tag{1b}$$

The probability $E_i(t)$ is the probability that a lineage starting at time t in state i leaves no descendants at the present day ($t = 0$), and $D_{N,i}(t)$ is the probability of a lineage in state i at time t before the present ($t > 0$) evolved the exact branching structure as observed.

With these ordinary differential equations, any number of states can be included in an SSE model. For example, in character-based models, such as BiSSE (Maddison et al., 2007), *i* and *j* represent two observed states of a focal character (e.g. observed states 0 and 1). A potential issue in this case is that when comparing a simple model where there is no variation in rates among states against a model of trait-dependent diversification there is almost always strong support for a trait-dependent diversification process in empirical settings (Maddison & Fitzjohn, 2014; Rabosky & Goldberg, 2015). The HiSSE model (Hidden-State Speciation and Extinction model; Beaulieu & O'Meara, 2016) partially corrects this issue by harnessing the properties of hidden-Markov models to allow rate heterogeneity to depend not only on the focal trait, but also to be correlated with other factors that were not explicitly scored as character observations at the tips (see also Caetano et al., 2018; Nakov et al., 2019; Boyko & Beaulieu, 2021). Thus, with HiSSE, *i* and *j* represent the different observed and hidden states combinations specified in the model (e.g. observed states 0 and 1, hidden states *A* and *B*).

Although hidden state SSE models made hypothesis testing in diversification studies more realistic, existing SSE models are still generally used to understand the correlation of a particular observed trait or geographical range distribution and the diversification dynamics of a group. However, in reality, it is not a single factor but, rather, a combination of circumstances that are responsible for the heterogeneous diversification rates among clades in a phylogenetic tree (Donoghue & Sanderson, 2015; Nürk et al., 2020). To understand heterogeneous processes that arise from the effect of multiple traits on the dynamics of speciation, extinction, and trait evolution is one of the main utilities of hidden Markov models in phylogenetic comparative methods (Caetano et al., 2018).

It is natural, then, to completely drop the observed trait from the analysis and focus only on the impact of the "unobserved" traits, or the hidden states, in the diversification dynamics of a clade. This is what our MiSSE model, an extension of the HiSSE framework, is intended to do. MiSSE is a direct extension of HiSSE with the main difference between the two models being that with HiSSE, we have an observed character with states 0 or 1, as in BiSSE (Fig.1a), *and* hidden states *A* and *B* (Fig. 1b). Essentially, MiSSE operates in the same way, but simply ignores the observed states altogether and performs the calculations of the hidden rate classes directly (Fig.1c).

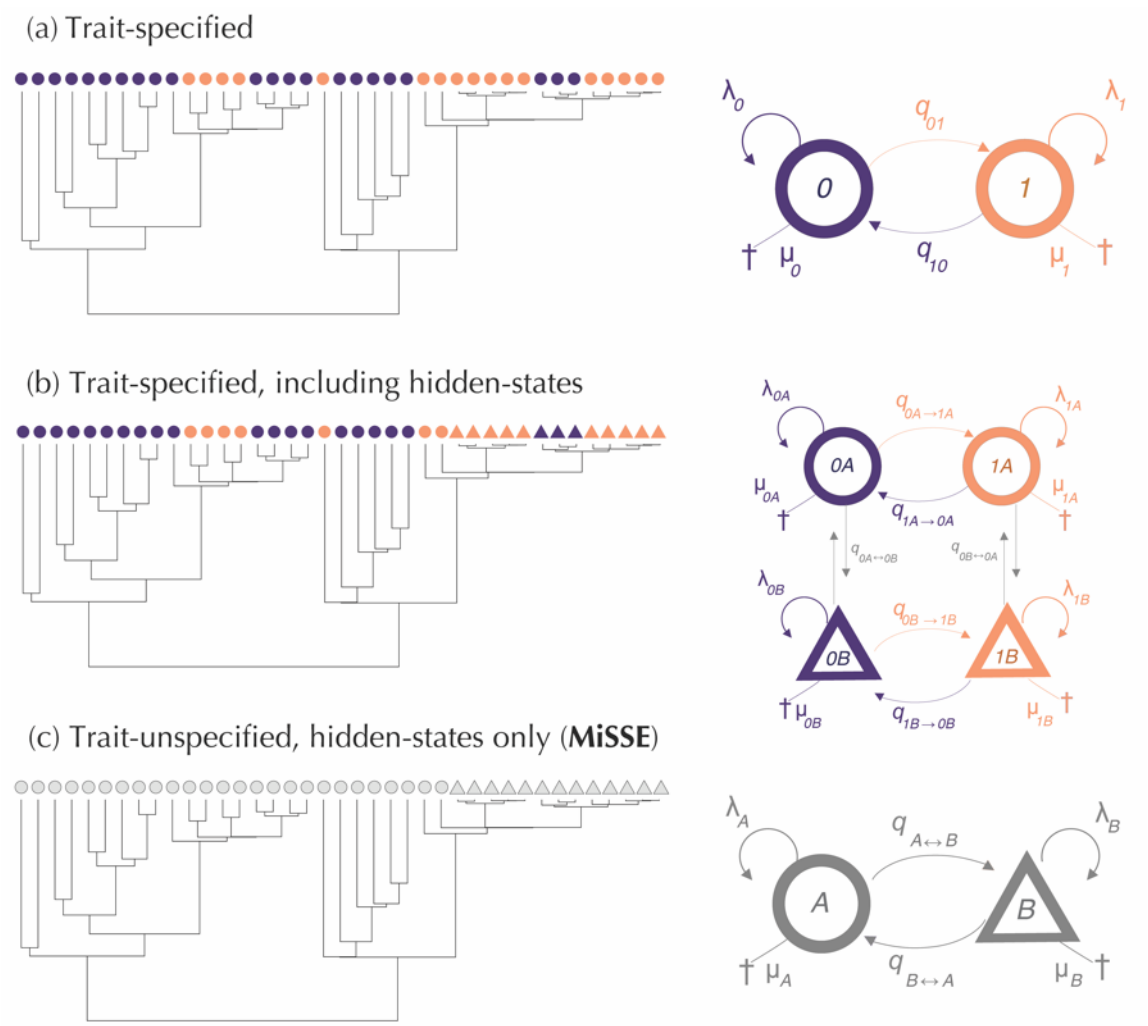

Figure 1: Diagrammatic representations of state-dependent speciation and extinction models that: (a) use only observed trait data in parameter estimation (e.g. BiSSE, MuSSE and GeoSSE), (b) use both observed trait data and hidden-states in parameter estimation (e.g. HiSSE, MuHiSSE, GeoHiSSE), and (c) use only hidden states in parameter estimation (e.g. MiSSE).

The assignment of particular hidden state to each tip is based on probabilities of the data, which, in the case of MiSSE, is just the structure of the tree. This process is like the maximum likelihood calculations in other hidden state methods. These calculations are described thoroughly in Caetano et al. (2018), so here we provide only a brief overview. The first step is to search for rate estimates that maximize the probability of observing the tree at the root by marginalizing over all possible histories at internal nodes and along branches. This search consists of a bounded stochastic simulated annealing, run for 5000 iterations, followed

by a bounded subplex routine that runs until the maximum likelihood is found. The use of simulated annealing helps ease MiSSE away from finding local optima.

Suppose that, in an iteration, rate class *A* has $\lambda = 0.1$ and $\mu = 0.05$, and rate class *B* has $\lambda = 0.2$ and $\mu = 0.1$. The node subtending the branch subtending a tip could have started in *A* giving rise to an *A* or *B*, or it could have started in *B* giving rise to an *A* or a *B*. Each of these scenarios has a probability associated with them. The sister edge has the same set of scenarios and its own set of probabilities. These probabilities are combined at the nodes (which includes the speciation rate to account for the speciation event at a node) and carried down the tree in the same way as in other SSE models. Once at the root, marginal probabilities for whether the root is in *A* or *B* based on a set of rates can be calculated, as well as the overall log-likelihood that these rates produced the observed tree.

At this point, all other branches, nodes, and tips are assumed to be in all possible states. The second step, then, is to take the MLE of the rates and determine which of the states at a given node and given tip is more likely than any of the other possible states. For that, MiSSE simply chooses a node (or tip), fixes it to be in rate class *A*, traverse the tree, and calculates the overall likelihood. Then it does the same for that same node, but this time it fixes it to be in rate class *B*. At the end, MiSSE calculates the marginal probabilities by dividing the probability that a node was in each state divided by the sum of the probabilities across all states. In that way, every tip and node have a probability of being in both states *A* and *B*, but often one state will be more probable than the other. For example, in a clade where speciation occurred more rapidly, the higher λ associated with *B* will be a better fit to the shape of that part of the tree, and so the marginal probabilities will reflect higher support for *B*. In another clade, where speciation is slower, lower λ will be the better fit, so the marginal probabilities will reflect the higher support for hidden state *A* in that part of the tree. And in other clades they may be uncertain, given frequent transitions in and out each hidden state.

Note that *A* and *B* are arbitrary labels that can and will shift positions at the tips in different runs. What really matters is the parameter combination that underlies each label in each run, which should not change their MLE between runs. Note too that MiSSE is still a model of tree and character states like other SSE models, and so the topology matters to assign tips to the correct (hidden) states. Because it uses data from the topology as well as branch lengths, MiSSE likelihood can distinguish between trees that have the same lineage through time plot, but different topologies (O'Meara & Beaulieu, in prep.).

2.2. Implementation

MiSSE is available within the R package *hisse* (Beaulieu & O'Meara, 2016). Some details of MiSSE's implementation differ from other SSE models also implemented in *hisse*. While these differences are summarized here, readers are encouraged to follow the example code available at https://github.com/bomeara/missecomparison when using MiSSE for their empirical analyses. The main difference between MiSSE and other SSE models in *hisse* is how MiSSE models are fit. The user can set and fit MiSSE models individually with a standalone function (hisse::MiSSE) (Table 1). However, it is difficult to know how many different hidden states the models will fit for a given empirical tree. As of now, MiSSE can fit models with up to 26 hidden states each representing a given "rate class", which defines a combination of turnover rates ($\tau = \lambda + \mu$) and extinction fraction ($\varepsilon = \mu / \lambda$). In total there are 676 (26^2) possible model combinations, each with a different combination of rate class number that range from a minimum of two diversification parameters (one hidden state) and a maximum of 52 (26 hidden states).

Performing an exhaustive search by fitting all possible models is not only computationally expensive, but also time-consuming. Furthermore, in general, the complexity of the models with the lowest AICc scale with the size of the tree, meaning that more complex models will tend to be more important in larger trees. In the same way, complex models with many rate classes are likely to be a poor fit for most small trees. To alleviate having to decide which set of models to fit, we developed an heuristic algorithm (hisse::MiSSEGreedy) to search through a reasonably large set of models that terminates after the set of models evaluated no longer provide meaningful support in terms of AICc. The default search starts with a random selection of models. This first set is biased towards simpler models, as initial experimentation showed that these models usually have the lowest AICc for most trees. The subsequent model "chunks" are then models that are neighbors to the ones recovered with the best AICc values in the previous rounds. The path of the MiSSEGreedy search can be visualized with the function hisse::PlotMisseSpace. MiSSEGreedy will continue to fit new "chunk" of models until the lowest ΔAICc of the new "chunk" is higher than that specified by the user (Fig. 2). The user is therefore advised to set the arguments "stop.deltaAICc" and "chunk.size" in a sensible way according to the size of the tree. Based on preliminary tests, we recommend a ΔAIC of 10 for trees with up to 1000 tips, which is the current default setting.

Once MiSSEGreedy terminates, one can select the best fitted model or average across a set of models in relation to their Akaike weight to estimate the diversification parameters at the tips using a marginal reconstruction algorithm with hisse::MarginReconMiSSE. In addition to marginal probabilities of states at nodes, MiSSE will also return marginal probabilities for the

tips. The states can be used as indexes for the estimates of speciation, extinction, and the orthogonal transformations of these rates, namely, net-diversification ($r = \lambda - \mu$), turnover ($\tau = \lambda + \mu$) and extinction fraction ($\varepsilon = \mu / \lambda$) returned by the model fit. A tip rate estimate for a given model, then, becomes the weighted average of the diversification rates with the marginal probabilities used as weights.

Table 1: New functions implemented in R package *hisse* to fit, reconstruct and visualize implementation of MiSSE models on phylogenetic trees.

Function name	**Brief description**
PlotMisseSpace	Imports from package igraph to plot a network showing which models have lower AICc weight in relation to their neighbors.
MiSSEGreedy	Uses greedy algorithm to fit a series of MiSSE models with different complexities in chunks. It will stop when delta AIC does not have significant improvement in the new chunk of models.
MiSSE	Standalone MiSSE function.
MarginReconMiSSE	Uses marginal reconstruction algorithm to reconstruct rates along the tree.
plot.misse.states	"Paints" the averaged rate reconstruction along the branches of the tree. However, inferences of rates through time should not be interpreted literally (see Discussion).
PruneRedundantModels	Prune models that are effectively identical in terms of likelihood and parameter estimates before model averaging.
TipCorrelation	Performs linear regression between phylogenetic independent contrasts of tip rates and continuous traits, with the option to prune out cherries (see Discussion).

We have also added functionality for eliminating redundant models for all models implemented within *hisse*. When calculating model weights, it is best not to include multiple copies of the same model (Burnham & Anderson, 2002). However, such cases can arise inadvertently: for example, a birth death model and a Yule model are essentially the same (other than in number of free parameters) if the best estimate for extinction rate is zero. In the case of MiSSE, where there can be many models compared, including multiple different models

that are in fact identical (in terms of likelihood and parameter estimates) increases the Akaike weights for these models artificially. We thus prune these redundant models with the function hisse::PruneRedundantModels.

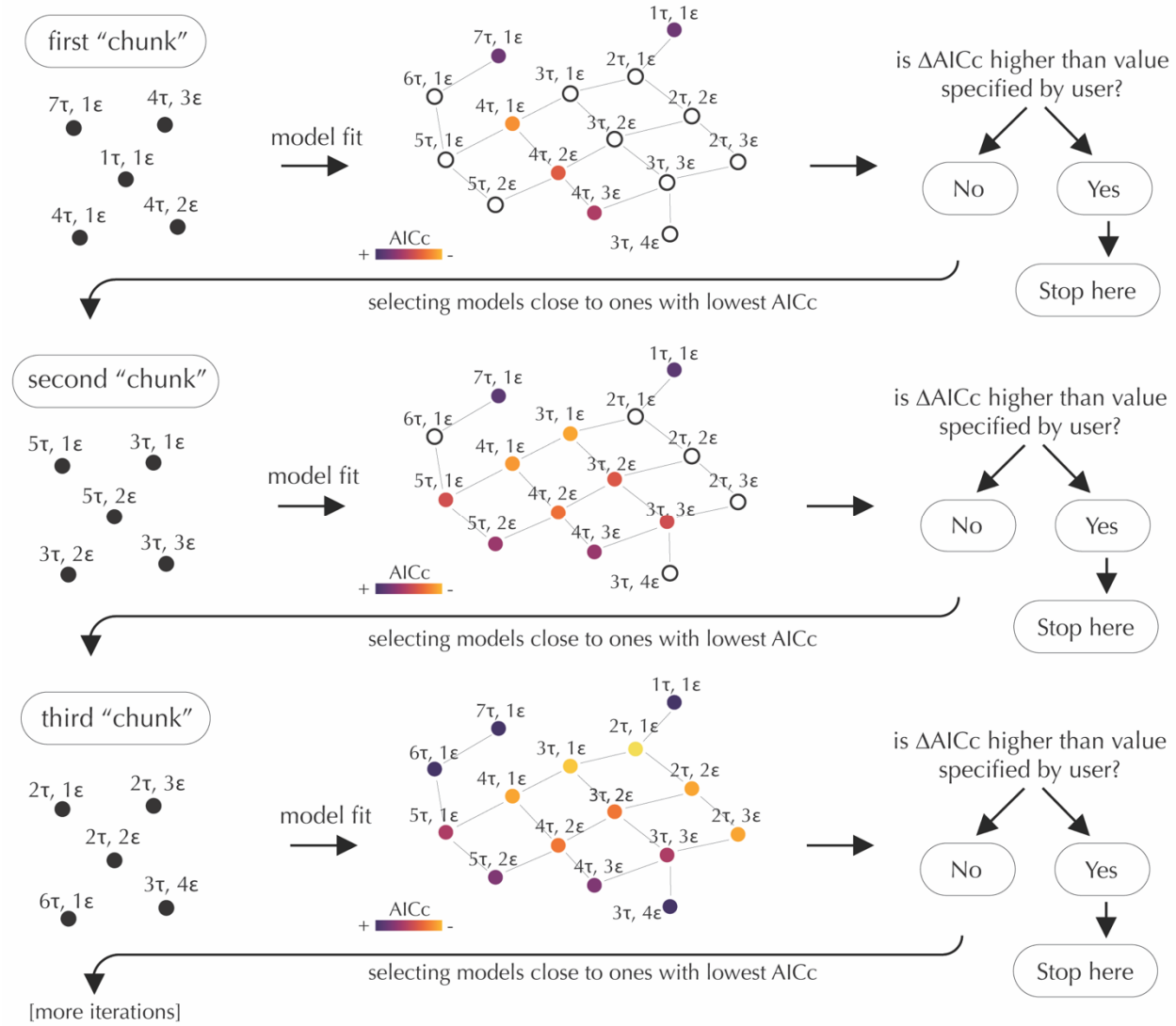

Figure 2: Rationale behind MiSSEGreedy. The search start with a random set of simple models (first "chunk"). After AICc are calculated for this set, the search continues by fitting the models closest to the ones with lowest AICc in previous "chunks". The search continues until the ΔAICc of the new chunk is higher than the threshold specified by the user. Here the search is represented by a "chunk" size of five, where five models are fitted at each time. The mock example in the figure is inspired by the analysis of *Lupinus* (Fabaceae) available as sample code at https://github.com/bomeara/missecomparison.

2.3. Comparing MiSSE's accuracy against other popular trait-independent methods

To test MiSSE's accuracy, we compared tip-estimates of diversification rates in a set of simulated trees extracted from Title and Rabosky (2019). Briefly, these authors compared the accuracy of four popular tip-rate metrics: DR statistics (Redding & Mooers, 2006), BAMM (Rabosky, 2014), node density (ND; Freckleton et al., 2008) and the inverse of terminal branch length (TB; Steel & Mooers, 2010). They used a set of 5200 trees simulated using nine different models to test the accuracy of the metrics under different scenarios. For speed, we selected a random sample of 10% of their original set of simulations, i.e. 521 trees. We then excluded trees simulated under the "multi-regime, constant-rate birth–death" of Meyer and Wiens (2018) because those were represented by only two trees in our set. The final 519 trees analyzed here have different sizes, though most of them are < 500 tips (Table 2).

Besides the range of simulation conditions used for the trees, an advantage of using Title and Rabosky (2019) is that priors and other settings for BAMM were chosen by the authors of BAMM, and thus are likely good approximations to what a skilled user would do. When comparing Bayesian models, either with each other or with non-Bayesian models, priors and MCMC runs may matter greatly in the accuracy of results: one could set priors that exclude the truth and find terrible performance, or set priors at the true values and find excellent performance even in the absence of data. Title and Rabosky (2019) used priors set by BAMMTools (Rabosky et al., 2014) and thus this is a fair test with no special foreknowledge of the true parameters given to BAMM.

We extended the original Title and Rabosky (2019) tip rate comparisons to include MiSSE, where we used the following: (1) tip-rates estimated from the model with overall lowest AICc (i.e. the "best" model); and (2) the tip-rates estimated by averaging all models according to their Akaike weight. We followed the same statistical metric as Title and Rabosky (2019), namely, we compared mean absolute error, given by the formula $\sum_{i=1}^{N_i} |\mathrm{rate}_i - \mathrm{rate}_{TRUE_i}| \,/\, N$, and the RMSE is given by $\sqrt{\sum_{i=1}^{N_i} (\mathrm{rate}_i - \mathrm{rate}_{TRUE_i})^{\,2} / \, N}$, where *i* is a single tip and N is the total number of tips in all trees of a given type of simulation. In all cases, the lower the error values the more accurate is the metric. Note that TB, ND and DR are non-parametric (i.e. not based on an underlying model) and only estimate λ. We then used the function posthoc.kruskal.conover.test from the R package PMCMR (Pohlert, 2014) to perform a pairwise test for multiple comparisons of mean rank sums and calculate whether errors are significantly different between metrics, assuming a significance value of $p < 0.05$. All code required to reproduce this comparison can be found at: https://github.com/bomeara/missecomparison.

Table 2: Simulated trees used in comparisons between tip rate metrics. * The original dataset from Title and Rabosky (2019) contains 1200 diversity-dependent trees, but 117 instead of 120 (10%) of those trees were included here due to our method of radomly sampling trees.

Simulation model	Number of trees	Tree-size	Source
Single-regime, constant-rate birth–death	10	100	Mitchell and Rabosky (2017)
Single- and multi-regime, constant-rate birth–death	10	100	Moore et al. (2016)
Single- and multi-regime, constant-rate birth–death	40	10 - 3157	Rabosky et al. (2017)
Single- and multi-regime, constant-rate birth–death	19	9 - 3458	Mitchell et al. (2019)
Single-regime, constant-rate birth–death, lambda uniform	100	100	Title and Rabosky (2019)
Single-regime, constant-rate birth–death, net diversification uniform	100	100	Title and Rabosky (2019)
Pure birth root regime, 1–4 discrete shifts to diversity-dependent regimes	117*	54 - 794	Rabosky (2014); Mitchell and Rabosky (2017)
Speciation rate evolves via diffusion process	120	25 - 662	Rabosky (2010); Beaulieu and O'Meara (2015); Rabosky (2016); Title and Rabosky (2019)

2.4. Empirical example: plant height and turnover in Eucalypts

We demonstrate the capabilities of MiSSE with an empirical example. Body size is considered an important trait in studies of animal evolution (Cooper & Purvis, 2010) and it has been interpreted as a potential correlate of diversification rates in vertebrates (Cope's rule; FitzJohn, 2010). A similar argument can be made for plants (Boucher et al., 2017). Life span and generation time are often positively correlated with plant height (Westoby, 1998), which would reduce the number speciation and extinction events on a per time basis (Petit & Hampe,

2006); i.e. slower turnover rates. Although these correlations are expected in theory, surprisingly few studies have compared them in practice using model-based approaches (e.g. Boucher et al., 2020). The MiSSE framework allows us to easily test this correlation by using regular comparative phylogenetic methods and tip-rates as a response variable.

We examined turnover rates in relation to tree height in eucalypts, one of the most distinctive components of the Australian landscape (Wilson et al., 2005). Eucalypts are members of the tribe Eucalypteae in the flowering plant family Myrtaceae, a group that includes some of the tallest trees in all angiosperms (e.g. *Eucalyptus regnans* can reach up to 120 m height) (Wilson et al., 2005; Nevill et al., 2010). Eucalypts are also unusual for being considered a relatively speciose (c. 800 species) lineage of trees. It is thought that the tree habit tends to decrease speciation rates, which is offered as reason that clades comprised of predominantly tree growth forms are frequently found to be less diverse than their herbaceous or shrubby relatives (Petit & Hampe, 2006). However, many eucalypts species are also large shrubs to treelets between 1.5 and 5m height (EUCLID, 2015), which led to the hypothesis that the radiation is driven primarily by the smaller representatives of the clade (Petit & Hampe, 2006). Here we test this idea by contrasting tip turnover rates with plant height in this diverse clade of angiosperms.

We used the “ML1” time calibrated phylogenetic tree from Thornhill et al. (2019), which covers 716 out of the c. 800 species of eucalypts. We then collected data on plant height for 673 species available at EUCLID Eucalypts of Australia Edition 4 (2015) (EUCLID, 2015). All measurements represent plant maximum height and are given in meters. Again, we use the default implementation of MiSSE, that is, stop.deltaAICc=10 and chunk.size=10, to estimate tip turnover rates in Eucalypteae. We also pruned redundant models with the function hisse::PruneRedundantModels. Next, we estimated tip turnover rates averaging all models according to their AICc weights with the function hisse::GetModelAveRates. To explore the differences between tip reconstructions in MiSSE and BAMM, we also ran BAMM for this tree using the same configuration settings as Thornhill et al. (2019).

To test correlations between plant height and tip turnover rates, we used the function hisse::TipCorrelation, which performs a regression-through-the-origin of positivized phylogenetic independent contrasts (PICs, Felsenstein, 1985) between tip rates and a continuous trait. This function also gives the option of pruning out PICs from cherries, i.e. two tips that are sister to each other and share the same branch length to the direct ancestral node, before regressions. We reasoned that because cherries theoretically inherit the exact same rate class probabilities in MiSSE they may: (1) present identical tip rates, forcing the slope of the regression to be close

to 0 since all PICs for cherries will be 0; and therefore (2) constitute pseudoreplicates in the analyses. Note that the function will prune PICs and not individual tips that constitute a cherry. Pruning single tips might not be adequate because it would affect all other PIC calculations in the tree; it also can make the results sensitive to which tip is pruned and gives less information for computing contrasts deeper in the tree. Rates and maximum plant height were log scaled before analyses so that they conformed with Brownian motion evolution (Felsenstein, 1985; Garland et al., 1992). All the code necessary to replicate these analyses are available at: https://github.com/bomeara/missecomparison.

3. RESULTS

3.1. Simulation studies

Comparisons between tip rate metrics using simulated data show that, in general, MiSSE has a similar accuracy as BAMM (Fig. 3). We present the results for mean absolute error, though RMSE results are practically the same (SM1; Supplementary Material). Both MiSSE and BAMM are more accurate than the non-parametric methods TB, ND and DR for speciation rates, the only parameter estimated by the latter three metrics. There was no significant difference in error measurements between the best MiSSE model and the AICc-weighted average of MiSSE models in any simulated scenario, and therefore these two results will be henceforward treated as a single MiSSE category.

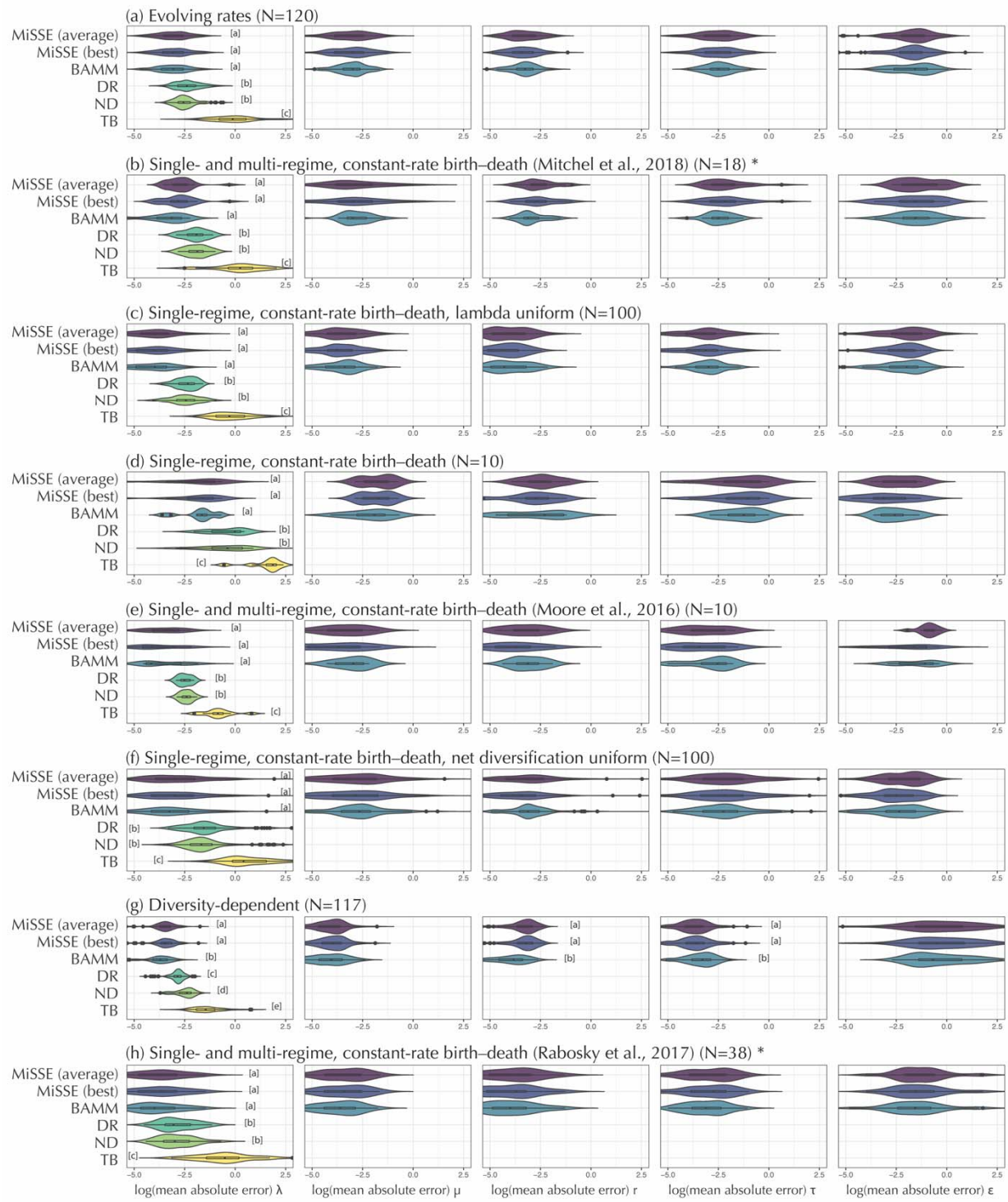

Figure 3: Violin plots showing the distribution of mean absolute error between true and estimated tip rates using six different metrics. Each data point corresponds to the mean absolute error between all tips in one tree. λ = speciation, μ = extinction, = turnover, and r = net-diversification, = extinction fraction. Different letters between brackets indicate when metrics are significantly different from each; same letter indicates mean absolute errors that are

not significantly different. [NB: at the time of submission, rates reconstructions had not finished in three out of the 519 trees. We expect that the inclusion/exclusion of these trees will not affect the general results of the comparisons].

TB is significantly less accurate than other methods in all simulation scenarios, and in seven out of the eight simulation scenarios there was no significant difference between error measurements in the pairs ND and DR (lower accuracy) and BAMM and MiSSE (higher accuracy). The only type of tree where all metrics perform significantly differently from one another is in trees simulated under a diversity dependent scenario (Fig.3g). In that case, DR performs significantly better than ND ($p < 0.01$) for speciation rates, though still significantly worse than BAMM and MiSSE ($p < 0.01$) for the same parameter. BAMM performs significantly better than MiSSE for speciation and net-diversification rates ($p < 0.01$), whereas MiSSE performs significantly better than BAMM for turnover rates ($p < 0.05$). However, the magnitude of differences between BAMM and MiSSE were slight, typically differing by less than 10%. All pairwise comparisons can be found at the Supplementary Material (SM2).

3.2. Tip-correlation between plant height and turnover

Despite the idea that smaller plants have driven the radiation in eucalypts, maximum plant height appears uncorrelated with turnover across the tips of the tree. This correlation is generally weak no matter the metric used to estimate tip rates (MiSSE $R^2 < 0.001$, Fig.4d; BAMM $R^2 < 0.001$, SM3). The general lack of correlation is mainly because: (1) higher turnover rates are not always restricted to clades of small shrubs and treelets, as it can be seeing by comparing the distribution of tip-rates and plant height (Fig.4a-c); and (2) plant height (Fig.4c) appears to be a much more labile trait than turnover rates, regardless of the methods used to infer tip-rates (Fig.4a,b). We note, however, that MiSSE (Fig.4b) recovers more heterogeneity in tip rates than BAMM (Fig.4a), including some accelerations in turnover that are not reconstructed by the latter (clades marked with asterisk, Fig.4a). We note that pruning cherries from the tree minimally changes the slope of the regression line in this example and it does not change our main conclusion that turnover and plant height are likely uncorrelated in eucalypts, given that contrasts of maximum plant height explain less than 0.1% of the contrasts of turnover. In any event, the inclusion of cherries can force the slope of the PIC regression towards 0 in other tip rate comparisons and therefore they should be pruned before analyses.

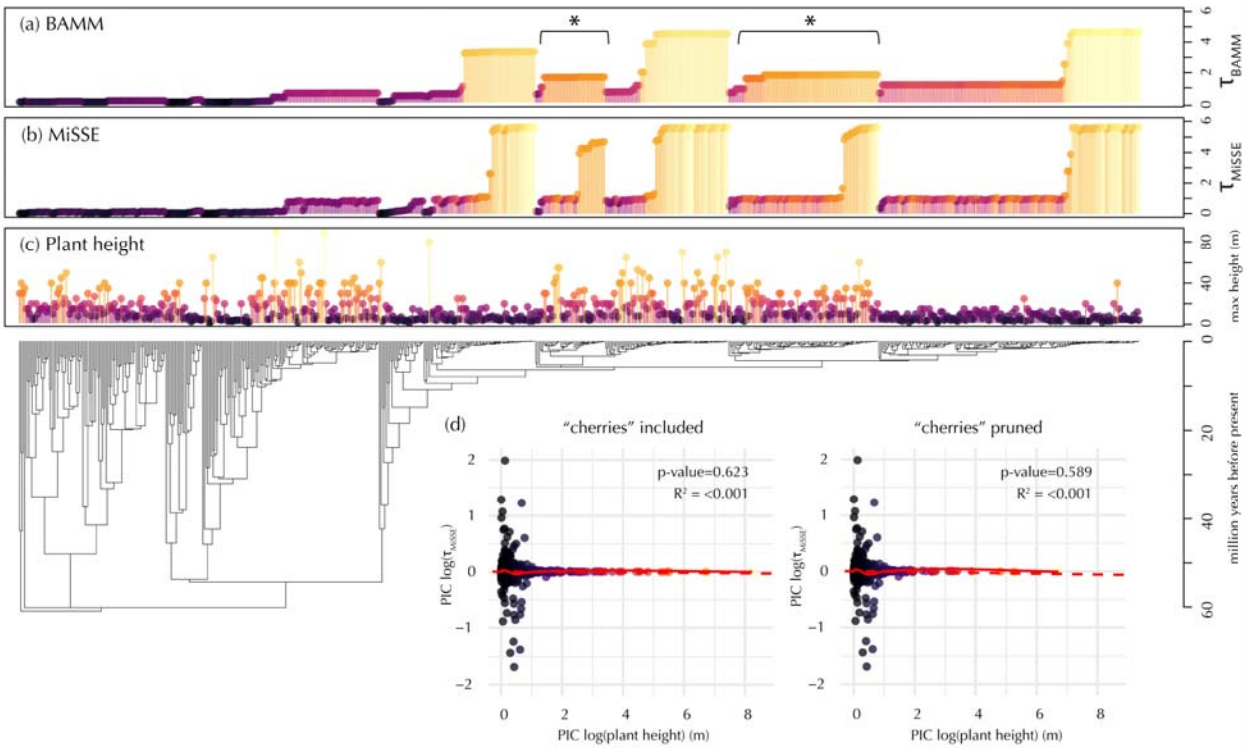

Figure 5: Comparison between maximum plant height and tip turnover rates estimated by MiSSE and BAMM in Eucalypteae. (a) BAMM turnover rates at the tips; (b) MiSSE turnover rates at the tips; (c) Maximum plant height in meters; (d) Regression through the origin between PICs of MiSSE turnover rates and plant height, with cherries included and pruned. Solid line: smoothed line using the loess method; dashed line: regression line.

4. DISCUSSION

4.1. MiSSE provides accurate estimates of diversification rates

Here, we describe our "trait-free" MiSSE framework and show how it can be used to estimate accurate diversification rates at the tips of the tree. Our results show that the difference in accuracy in parameters estimated by averaging all MiSSE models or using only the best MiSSE model seem to be minimal and may depend on the tree (Fig. 3). However, as discussed in previous publications (Beaulieu & O'Meara, 2016; Caetano et al., 2018), the user is strongly encouraged to focus mainly on parameter estimation rather than which model fits "best" when using MiSSE, since the former is more informative about the biology of a clade. We note also that, although extinction fraction and turnover are sometimes used as synonyms in the literature (e.g. Alfaro et al. 2009; Title & Rabosky, 2019), as used here turnover rate is a parameter that is explicitly orthogonal to extinction fraction. We emphasize this difference here because many

diversification hypotheses may be better described by turnover rates rather than net-diversification or extinction fraction (e.g. Vrba, 1993; Vasconcelos et al., 2021; our empirical example). Also, given that extinction and speciation should tend to correlate through time (Marshall, 2017), net-diversification should tend to approach 0. In those scenarios, turnover can naturally become a better metric for understanding the nuances of ups and downs in diversification dynamics than its isolated components.

Key differences between MiSSE and analogous trait-free methods are also highlighted by our analyses of simulated data and our empirical example. Although estimates from MiSSE are as accurate as BAMM, the two methods make very different assumptions. BAMM is based on inferences of rate shifts and has no model underlying rate heritability (Laudanno et al., 2021), so that all tips after a shift inherit a similar rate (e.g. Fig.4). MiSSE may also have some shift-like properties in areas where the probabilities of being in different rate classes change abruptly, but because at every moment in the tree, including at the tips, there is a combined probability of being in all possible rate classes, the evolution of diversification rates becomes smoother. Closely related tips do have similar rates because it is likely that they will be in the same rate class, but they do not inherit the absolute same rate after a shift. The exception is, perhaps the cherries, which will likely inherit the same exact rate. Pseudoreplication from cherries will afflict any model that uses just branch lengths to estimate tip rates. We suspect that, for that reason, it may make sense to prune out cherries from PIC analyses when performing tip-rate correlations using MiSSE, BAMM, DR, or similar tree-only diversification approaches, as we have done in our analysis of eucalypts.

Our main objective with the analyses of simulated data, and comparisons among metrics, is not to demonstrate that MiSSE is necessarily a better alternative to other tip rate metrics, nor to compare MiSSE exhaustively against all other available methods for diversification rate estimation at the tips. Our aim is mainly to take one step forward in understanding the limitations of diversification modeling using SSE and hidden-Markov-models. As we continue learning about the behavior of these models, future comparisons may also include other tip-rate metrics that have become available recently in software such as RevBayes (Höhna et al., 2019) and ClaDS (Maliet et al., 2019). Still, we expect that MiSSE will continue to be a good alternative for prior-free, likelihood-only analysis, since all others currently available methods are based on Bayesian frameworks.

4.2. Plant height and turnover rates are uncorrelated in eucalypts

Our empirical analyses demonstrated that the correlation between plant height and turnover rates is weak within Eucalypteae. The use of tree height as a proxy for longevity is common in plants, but it may be that plant height is not the best proxy for longevity in eucalypts. Tall eucalypt trees are often native to productive environments with fertile soils (Pryor, 1976; Thornhill et al., 2019) and it may be that they present fast life cycles for their size, as observed in other parts of the world where soils are also fertile (Russo et al., 2008). In fact, many eucalypt trees are known to be fast-growing plants (e.g. Barnard et al., 2003; Almeida et al., 2004) and fast growth is linked to higher mortality through the growth-survival trade-off (Reich, 2014), which could potentially lead to higher turnover rates even in lineages of tall trees (Baker et al., 2014). If that is the case, perhaps plant height is not the trait that best captures the variation that we sought to explore in our hypothesis. Other proxies, such as leaf mass area and seed size (Westoby, 1998; Wright et al., 2004) may be a better proxy of longevity in these scenarios and should also be tested.

Alternatively, if plant height is, in fact, a good proxy for longevity in eucalypts, it may be that longevity is generally uncorrelated with turnover in the clade. A similar trend of uncorrelated body size and speciation rates has been observed in other groups (FitzJohn, 2010; Boucher et al., 2020), but it is difficult to ascertain if the lack of correlation is a particularity of the groups that have been analyzed so far or a general rule. Given the heterogeneity of evolution, identifying what is a rule and what is an exception in macroevolution requires analyses of multiple natural replicates. Similar tests in several clades would be ideal to rule out this possibility. A different approach would be to use QuASSE (FitzJohn, 2010) to construct a model of turnover and height; the disadvantage of this is the need to model the relationship and the (current) impossibility of including additional observed or unobserved factors.

One benefit of MiSSE (and similar methods) is that, even though our hypothesis was not supported, we still have new information about the tips: turnover rate (and the other diversification parameters if one wishes). In the same way that looking at two traits, such as height and rainfall, or biome and pollination syndrome, may generate new hypotheses to test, having estimates of turnover rate at tips as in Figure 5 may also lead to new ideas about the factors leading to higher turnover rates in some eucalypts. These can be tested in various ways, including by testing them in other groups.

4.3. Tip rates and best practices in using MiSSE

MiSSE is intended to be used to estimate rates at the tips of the tree, which are also referred to as species-specific diversification rates (Title & Rabosky, 2019; Maliet et al., 2019).

Tip rates are, technically, an emergent property of how diversification dynamics are modeled along the branches of the tree (Freckleton et al., 2008; Title & Rabosky, 2019). However, rate estimates are not unique per tip species, so species should not be analyzed in isolation. In other words, a tip with extinction-fraction of 0.9, for instance, does not necessarily indicate that that tip is about to go extinct, but rather that a lot of the species in that same clade are relatively short-lived in comparison to the rest of the tree. Tip rates can then be seen as a "snapshot" of the diversification potential of the species in a clade. In the MiSSE framework, this diversification potential may be interpreted as the combined effect of the many (hidden) states evolving in the tree, which are represented by the different rate classes.

From a mathematical standpoint, there is an advantage in using tip rates because it is at the tips where the certainty around the probability of being in a particular rate class is the highest (O'Meara & Beaulieu, in prep). The certainty around these probabilities decreases as one moves towards the root of the tree, and so the ability to point out changes in rate classes also decreases. Of course, the model *does* use information from the whole phylogenetic tree in the calculations of tip-rates, but as one moves towards the present in the tree, the more information is available to reconstruct those rates and thus we are more likely to be certain about the probabilities of being in different rate categories (see O'Meara & Beaulieu, in prep.).

The intuitive question that follows is how far back in the past one can go to interpret how the observable diversification potential at the tips came to be and how far into the future can we use it to make predictions about how clades will diversify. The answer is, again, not straightforward and it may depend on the size and age of the phylogenetic tree under analysis. Tip rates will not be able to tell us about mass extinction events deep in the past as these are arguably only measurable from the fossil record (Barnosky et al., 2011), but they may represent a good reflection of diversification dynamics above the species level in times slices close to the present. We therefore recommend great caution with literal interpretations of rate classes deep in time. That is why although MiSSE can "paint" the averaged rate reconstruction along the branches of the tree (i.e. with the function hisse::plot.misse.states), inferences of rates through time should not be interpreted literally. The "painting" should be used just to visualize the rates inferred with MiSSE and should not be used to support narratives of past diversification dynamics for a group. Interpreting past rates is risky given limited information; it could be even riskier with methods like BAMM where there are inheritability problems (Moore et al. 2016). Similarly, modern drivers of diversification dynamics (climate change, invasive species, habitat destruction, and more), which affect especially extinction rates at the present, will not be reflected in the rate estimates returned by MiSSE or similar methods.

Another issue is the non-independence of the tip rates. The identical rates for both taxa in a cherry have been discussed above, but even taxa that do not form a cherry may have somewhat correlated rates. It is important to realize that though a phylogenetic model was used in the estimation of tip rates, they themselves are not corrected for phylogeny (which is a reason we used independent contrasts above to compare turnover and plant height). There is also limited information: a resolved tree of 100 taxa has 198 edges: though MiSSE, BAMM, DR, and other methods give 100 tip estimates, there is not enough information to take each one as known with great certainty, nor to draw a strong conclusion based on a rate shared by a few taxa.

4.4. Caveats and persistent issues in modeling diversification dynamics

MiSSE is a powerful tool for estimating tip-diversification rates given its flexibility, accuracy and possibly avoidance of some of the identifiability issues recently highlighted for similar methods (Kubo & Iwasa, 1995; Louca & Pennell, 2020; see O'Meara & Beaulieu, in prep.). However, there are several caveats that users should consider when using MiSSE: (1) Contrary to other similar methods, MiSSE deliberately does not include an implementation for clade-specific sampling fraction. Clade-specific sampling fraction was found to lead to an incorrect likelihood behavior in *any* diversification method despite its appeal (Beaulieu, 2020), so in MiSSE all sampling fractions are global and only one sampling fraction is given for the whole tree. (2) How species are defined is the most relevant when looking at tip rates. For diversification rate models, the data come from the distribution of branching events across the phylogeny. Since most of these events are nearer the present, lumping and splitting taxonomic entities at the species level will have a greater impact on tip-rate estimation. We suspect that MiSSE, like all other trait-free diversification methods, are particularly sensitive to taxonomic subjectivity. (3) Contrary to previous SSE models, transition rates (q) among rate classes are always set equal. They are informed in the output, but they have no direct interpretable biological meaning like the transition rates in BiSSE, which represent the frequency of state changes among observed character states, or in GeoSSE, which can be interpreted as the frequency of dispersal out of a biogeographic area. Our current implementation will continue to treat these rates as fixed until it is clearer whether transition rates are practically identifiable in MiSSE. An implementation where transition rates are allowed to vary would be ideal given that shifts between rate classes are likely to occur at different speeds throughout the evolution of a group. (4) There may be still issues related to ascertainment bias and underestimation of extinction rates (Beaulieu & O'Meara, 2018). In that sense, the larger and broader taxonomic

sample one uses to test diversification hypotheses, the more biologically realistic extinction rates will tend to be. Ideally, one would be able to correct the extinction estimates in smaller trees, but that is still not possible with the current implementation of MiSSE nor in any other model we are aware of. So, even though extinction rates are estimated accurately relative to the simulated data, they may still be biologically unrealistic in small and young trees. (5) Finally, given that MiSSE does not account for mass-extinction events, it is unclear what impact such events will have on tip-rate estimation in MiSSE, or any “trait-free” approach currently available.

5. CONCLUSIONS

Time-calibrated phylogenetic trees have been important sources of data to understand fundamental aspects of the evolution of organisms. MiSSE provides a novel tool to explore this, and its independence from priors and its ability to explicitly model rate heritability with a finite set of rates makes MiSSE a powerful alternative for diversification analyses with correct likelihoods. There remain many caveats and cautions about its use, but it is an additional tool to understand diversification processes, including a focus on parameter of perhaps great biological relevance such as turnover rate. We suggest that tip-rates estimated by MiSSE will be useful to several questions that were previously addressed by other SSE models. Its versatility is appealing to explore integrative questions linking traits, geographical range distribution, or both in time slices close to the present.

Acknowledgements

The authors thank members of the Beaulieu lab for discussion. This work was funded by the National Science Foundation (grants DEB–1916558 and DEB–1916539).

Author contributions

T.V., B.C.O. and J.M.B. designed the study. J.M.B. led software development, B.C.O. led simulation analyses and T.V. led the empirical example. T.V. wrote the first draft and all authors contributed to the final writing of the manuscript.

References

Alfaro, M. E., Santini, F., Brock, C., Alamillo, H., Dornburg, A., Rabosky, D. L., Giorgio Carnevale, and & Harmon, L. J. (2009). Nine exceptional radiations plus high turnover explain species diversity in jawed vertebrates. *Proceedings of the National Academy of Sciences*, 106(32), 13410–13414. https://doi.org/10.1073/pnas.0811087106

Almeida, A. C., Landsberg, J. J., & Sands, P. J. (2004). Parameterisation of 3-PG model for fast-growing *Eucalyptus grandis* plantations. *Forest Ecology and Management*, *193*(1-2), 179–195. https://doi.org/10.1016/j.foreco.2004.01.029

Baker, T. R., Pennington, R. T., Magallon, S., Gloor, E., Laurance, W. F., Alexiades, M., et al. (2014). Fast demographic traits promote high diversification rates of Amazonian trees. *Ecology letters*, *17*(5), 527–536. https://doi.org/10.1111/ele.12252

Barnard, H. R., & Ryan, M. G. (2003). A test of the hydraulic limitation hypothesis in fast-growing *Eucalyptus saligna*. *Plant, Cell & Environment*, *26*(8), 1235–1245. https://doi.org/10.1046/j.1365-3040.2003.01046.x

Beaulieu, J. M., & O'Meara, B. C. (2015). Extinction can be estimated from moderately sized molecular phylogenies. *Evolution*, *69*(4), 1036–1043. https://doi.org/10.1111/evo.12614

Beaulieu, J. M., & O'Meara, B. C. (2016). Detecting hidden diversification shifts in models of trait-dependent speciation and extinction. *Systematic biology*, *65*(4), 583–601. https://doi.org/10.1093/sysbio/syw022

Beaulieu, J. M., & O'Meara, B. C. (2018). Can we build it? Yes we can, but should we use it? Assessing the quality and value of a very large phylogeny of campanulid angiosperms. *American Journal of Botany*, 105(3), 417–432. https://doi.org/10.1002/ajb2.1020

Beaulieu, J. M. (2020) The Problem with Clade-specific Sampling Fractions. Available at: https://rdrr.io/cran/hisse/f/inst/doc/Clade-specific-sampling.pdf.

Barraclough, T. G., Vogler, A. P., & Harvey, P. H. (1998). Revealing the factors that promote speciation. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *353*(1366), 241–249. https://doi.org/10.1098/rstb.1998.0206

Boucher, F. C., Verboom, G. A., Musker, S., & Ellis, A. G. (2017). Plant size: a key determinant of diversification? *New Phytologist*, *216*(1), 24–31. https://doi.org/10.1111/nph.14697

Boucher, F. C., Quatela, A. S., Ellis, A. G., & Verboom, G. A. (2020). Diversification rate vs. diversification density: Decoupled consequences of plant height for diversification of Alooideae in time and space. *PloS one*, *15*(5), e0233597. https://doi.org/10.1371/journal.pone.0233597

Boyko, J. D., & Beaulieu, J. M. (2021). Generalized hidden Markov models for phylogenetic comparative datasets. *Methods in Ecology and Evolution*, *12*(3), 468–478. https://doi.org/10.1111/2041-210X.13534

Burnham, K.P., & Anderson, D.R. (2002). Model selection and multimodel inference: a practical information-theoretic approach. New York: Springer.

Caetano, D. S., O'Meara, B. C., & Beaulieu, J. M. (2018). Hidden state models improve state-dependent diversification approaches, including biogeographical models. *Evolution*, *72*(11), 2308–2324. https://doi.org/10.1111/evo.13602

Cooper, N., & Purvis, A. (2010). Body size evolution in mammals: complexity in tempo and mode. *The American Naturalist*, *175*(6), 727–738. https://doi.org/10.1086/652466

Cooper, N., Thomas, G. H., & FitzJohn, R. G. (2016). Shedding light on the 'dark side'of phylogenetic comparative methods. *Methods in Ecology and Evolution*, *7*(6), 693–699. https://doi.org/10.1111/2041-210X.12533

Donoghue, M. J., & Sanderson, M. J. (2015). Confluence, synnovation, and depauperons in plant diversification. *New Phytologist*, *207*(2), 260–274. https://doi.org/10.1111/nph.13367

EUCLID Eucalypts of Australia Edition 4 (2015). Internet based, hosted by the Identic Pty Ltd, Brisbane at: https://apps.lucidcentral.org/euclid

FitzJohn, R. G. (2010). Quantitative traits and diversification. *Systematic biology*, *59*(6), 619–633. https://doi.org/10.1093/sysbio/syq053

FitzJohn, R. G. (2012). Diversitree: comparative phylogenetic analyses of diversification in R. *Methods in Ecology and Evolution*, *3*(6), 1084–1092. https://doi.org/10.1111/j.2041-210X.2012.00234.x

Felsenstein, J. (1985). Phylogenies and the comparative method. *The American Naturalist*, *125*(1), 1–15. https://doi.org/10.1086/284325

Freckleton, R. P., Phillimore, A. B., & Pagel, M. (2008). Relating traits to diversification: a simple test. *The American Naturalist*, *172*(1), 102–115. https://doi.org/10.1086/588076

Goldberg, E. E., Kohn, J. R., Lande, R., Robertson, K. A., Smith, S. A., & Igić, B. (2010). Species selection maintains self-incompatibility. *Science*, *330*(6003), 493–495. https://doi.org/10.1126/science.1194513

Goldberg, E. E., & Igić, B. (2012). Tempo and mode in plant breeding system evolution. *Evolution: International Journal of Organic Evolution*, *66*(12), 3701–3709. https://doi.org/10.1111/j.1558-5646.2012.01730.x

Harvey, M. G., & Rabosky, D. L. (2018). Continuous traits and speciation rates: Alternatives to state-dependent diversification models. *Methods in Ecology and Evolution*, *9*(4), 984–993. https://doi.org/10.1111/2041-210X.12949

Harvey, M. G., Bravo, G. A., Claramunt, S., Cuervo, A. M., Derryberry, G. E., Battilana, J., ... & Derryberry, E. P. (2020). The evolution of a tropical biodiversity hotspot. *Science*, 370(6522), 1343–1348. https://doi.org/10.1126/science.aaz6970

Höhna, S., Freyman, W. A., Nolen, Z., Huelsenbeck, J. P., May, M. R., & Moore, B. R. (2019). A Bayesian approach for estimating branch-specific speciation and extinction rates. bioRxiv, 555805. https://doi.org/10.1101/555805

Lagomarsino, L. P., Condamine, F. L., Antonelli, A., Mulch, A., & Davis, C. C. (2016). The abiotic and biotic drivers of rapid diversification in Andean bellflowers (Campanulaceae). *New Phytologist*, *210*(4), 1430–1442. https://doi.org/10.1111/nph.13920

Laudanno, G., Haegeman, B., Rabosky, D. L., & Etienne, R. S. (2021). Detecting lineage-specific shifts in diversification: A proper likelihood approach. *Systematic biology*, *70*(2), 389–407. https://doi.org/10.1093/sysbio/syaa048

Losos, J. B. (2011). Seeing the Forest for the Trees: The Limitations of Phylogenies in Comparative Biology: (American Society of Naturalists Address). *The American Naturalist*, *177*(6), 709–727. https://doi.org/10.1086/660020

Louca, S., & Pennell, M. W. (2020). Extant timetrees are consistent with a myriad of diversification histories. *Nature*, *580*(7804), 502–505. https://doi.org/10.1038/s41586-020-2176-1

Maddison, W. P. (2006). Confounding asymmetries in evolutionary diversification and character change. *Evolution*, 60(8), 1743–1746. https://doi.org/10.1111/j.0014-3820.2006.tb00517.x

Maddison, W. P., Midford, P. E., & Otto, S. P. (2007). Estimating a binary character's effect on speciation and extinction. *Systematic biology*, *56*(5), 701–710. https://doi.org/10.1080/10635150701607033

Maddison, W. P., & FitzJohn, R. G. (2015). The unsolved challenge to phylogenetic correlation tests for categorical characters. *Systematic biology*, *64*(1), 127–136. https://doi.org/10.1093/sysbio/syu070

Maliet, O., Hartig, F., & Morlon, H. (2019). A model with many small shifts for estimating species-specific diversification rates. *Nature ecology & evolution*, 3(7), 1086–1092. https://doi.org/10.1038/s41559-019-0908-0

Marshall, C. R. (2017). Five palaeobiological laws needed to understand the evolution of the living biota. *Nature Ecology & Evolution*, *1*(6), 1–6. https://doi.org/10.1038/s41559-017-0165

Meyer, A. L., & Wiens, J. J. (2018). Estimating diversification rates for higher taxa: BAMM can give problematic estimates of rates and rate shifts. *Evolution*, *72*(1), 39–53. https://doi.org/10.1111/evo.13378

Mitchell, J. S., & Rabosky, D. L. (2017). Bayesian model selection with BAMM: effects of the model prior on the inferred number of diversification shifts. *Methods in Ecology and Evolution*, *8*(1), 37–46. https://doi.org/10.1111/2041-210X.12626

Mitchell, J. S., Etienne, R. S., & Rabosky, D. L. (2019). Inferring diversification rate variation from phylogenies with fossils. Systematic Biology, 68(1), 1–18. https://doi.org/10.1093/sysbio/syy035

Moore, B. R., Höhna, S., May, M. R., Rannala, B., & Huelsenbeck, J. P. (2016). Critically evaluating the theory and performance of Bayesian analysis of macroevolutionary mixtures. *Proceedings of the National Academy of Sciences*, *113*(34), 9569–9574. https://doi.org/10.1073/pnas.1518659113

Morales-Barbero, J., Gouveia, S. F., & Martinez, P. A. (2021). Historical climatic instability predicts the inverse latitudinal pattern in speciation rate of modern mammalian biota. *Journal of Evolutionary Biology*, *34*(2), 339–351. https://doi.org/10.1111/jeb.13737

Nakov, T., Beaulieu, J. M., & Alverson, A. J. (2019). Diatoms diversify and turn over faster in freshwater than marine environments. *Evolution*, *73*(12), 2497–2511. https://doi.org/10.1111/evo.13832

Nevill, P. G., Bossinger, G., & Ades, P. K. (2010). Phylogeography of the world's tallest angiosperm, *Eucalyptus regnans*: evidence for multiple isolated Quaternary refugia. *Journal of Biogeography*, *37*(1), 179–192. https://doi.org/10.1111/j.1365-2699.2009.02193.x

Nürk, N. M., Linder, H. P., Onstein, R. E., Larcombe, M. J., Hughes, C. E., Piñeiro Fernández, et al. (2020). Diversification in evolutionary arenas—Assessment and synthesis. *Ecology and Evolution*, *10*(12), 6163–6182. https://doi.org/10.1002/ece3.6313

O'Meara, B. C., Ané, C., Sanderson, M. J., & Wainwright, P. C. (2006). Testing for different rates of continuous trait evolution using likelihood. *Evolution*, *60*(5), 922–933. https://doi.org/10.1111/j.0014-3820.2006.tb01171.x

Petit, R. J., & Hampe, A. (2006). Some evolutionary consequences of being a tree. *Annual Review of Ecology, Evolution, and Systematics*, *37*, 187–214. https://doi.org/10.1146/annurev.ecolsys.37.091305.110215

Pohlert, T. (2014). The Pairwise Multiple Comparison of Mean Ranks Package (PMCMR). R package, available at: https://CRAN.R-project.org/package=PMCMR.

Pryor, L. D. (1976). *The biology of Eucalypts.* Edward Arnold, London, UK.

Rabosky, D. L. (2010). Extinction rates should not be estimated from molecular phylogenies. *Evolution*, *64*(6), 1816–1824. https://doi.org/10.1111/j.1558-5646.2009.00926.x

Rabosky, D. L. (2016). Challenges in the estimation of extinction from molecular phylogenies: a response to Beaulieu and O'Meara. *Evolution*, *70*(1), 218–228. https://doi.org/10.1111/evo.12820

Rabosky, D. L., Grundler, M., Anderson, C., Title, P., Shi, J. J., Brown, J. W., Huang, H., & Larson, J. G. (2014). BAMM tools: an R package for the analysis of evolutionary dynamics on phylogenetic trees. *Methods in Ecology and Evolution*, *5*(7), 701–707. https://doi.org/10.1111/2041-210X.12199

Rabosky, D. L. (2014). Automatic detection of key innovations, rate shifts, and diversity-dependence on phylogenetic trees. *PloS one*, *9*(2), e89543. https://doi.org/10.1371/journal.pone.0089543

Rabosky, D. L., & Goldberg, E. E. (2015). Model inadequacy and mistaken inferences of trait-dependent speciation. *Systematic biology*, *64*(2), 340–355. https://doi.org/10.1093/sysbio/syu131

Rabosky, D. L., Mitchell, J. S., & Chang, J. (2017). Is BAMM flawed? Theoretical and practical concerns in the analysis of multi-rate diversification models. *Systematic biology*, *66*(4), 477–498. https://doi.org/10.1093/sysbio/syx037

Redding, D. W., & Mooers, A. Ø. (2006). Incorporating evolutionary measures into conservation prioritization. *Conservation Biology*, *20*(6), 1670–1678. https://doi.org/10.1111/j.1523-1739.2006.00555.x

Ree, R. H., & Sanmartín, I. (2018). Conceptual and statistical problems with the DEC+ J model of founder-event speciation and its comparison with DEC via model selection. *Journal of Biogeography*, *45*(4), 741–749. https://doi.org/10.1111/jbi.13173

Reich, P. B. (2014). The world-wide 'fast–slow'plant economics spectrum: a traits manifesto. *Journal of Ecology*, 102(2), 275–301. https://doi.org/10.1111/1365-2745.1221

Reginato, M., Vasconcelos, T. N., Kriebel, R., & Simoes, A. O. (2020). Is dispersal mode a driver of diversification and geographical distribution in the tropical plant family Melastomataceae? *Molecular phylogenetics and evolution*, *148*, 106815. https://doi.org/10.1016/j.ympev.2020.106815

Russo, S. E., Brown, P., Tan, S., & Davies, S. J. (2008). Interspecific demographic trade-offs and soil-related habitat associations of tree species along resource gradients. Journal of Ecology, 96(1), 192–203. https://doi.org/10.1111/j.1365-2745.2007.01330.x

Smith, S. A., & Brown, J. W. (2018). Constructing a broadly inclusive seed plant phylogeny. *American journal of botany*, *105*(3), 302–314. https://doi.org/10.1002/ajb2.1019

Steel, M., & Mooers, A. (2010). The expected length of pendant and interior edges of a Yule tree. *Applied Mathematics Letters*, *23*(11), 1315–1319. https://doi.org/10.1016/j.aml.2010.06.021

Title, P. O., & Rabosky, D. L. (2019). Tip rates, phylogenies and diversification: what are we estimating, and how good are the estimates? *Methods in Ecology and Evolution*, *10*(6), 821–834. https://doi.org/10.1111/2041-210X.13153

Thornhill, A. H., Crisp, M. D., Külheim, C., Lam, K. E., Nelson, L. A., Yeates, D. K., & Miller, J. T. (2019). A dated molecular perspective of eucalypt taxonomy, evolution and diversification. *Australian Systematic Botany*, *32*(1), 29–48. https://doi.org/10.1071/SB18015

Uyeda, J. C., Zenil-Ferguson, R., & Pennell, M. W. (2018). Rethinking phylogenetic comparative methods. *Systematic Biology*, *67*(6), 1091–1109. https://doi.org/10.1093/sysbio/syy031

Vasconcelos, T. N., Chartier, M., Prenner, G., Martins, A. C., Schönenberger, J., Wingler, A., & Lucas, E. (2019). Floral uniformity through evolutionary time in a species-rich tree lineage. *New Phytologist*, *221*(3), 1597–1608. https://doi.org/10.1111/nph.15453

Vasconcelos, T., O'Meara, B.C., & Beaulieu, J.M. (2021). Retiring "cradles" and "museums" of biodiversity. *The American Naturalist*. https://doi.org/10.1086/717412

Vrba, E. S. (1993). Turnover-pulses, the Red Queen, and related topics. *American Journal of Science*, *293*(A), 418.

Westoby, M. (1998). A leaf-height-seed (LHS) plant ecology strategy scheme. *Plant and soil*, *199*(2), 213–227. https://doi.org/10.1023/A:1004327224729

Wiens, J. J., & Donoghue, M. J. (2004). Historical biogeography, ecology and species richness. *Trends in ecology & evolution*, *19*(12), 639–644. https://doi.org/10.1016/j.tree.2004.09.011

Wilson, P. G., O'brien, M. M., Heslewood, M. M., & Quinn, C. J. (2005). Relationships within Myrtaceae sensu lato based on a mat K phylogeny. *Plant Systematics and Evolution,* 251(1), 3–19 https://doi.org/10.1007/s00606-004-0162-y

Wright, I. J., Reich, P. B., Westoby, M., Ackerly, D. D., Baruch, Z., Bongers, F., et al. (2004). The worldwide leaf economics spectrum. *Nature*, *428*(6985), 821–827. https://doi.org/10.1038/nature02403