Running head: Hidden state speciation and extinction

Detecting hidden diversification shifts in models of trait-dependent speciation and extinction

Jeremy M. Beaulieu[1,2] and Brian C. O'Meara[2]

[1]*National Institute for Biological and Mathematical Synthesis, University of Tennessee, Knoxville, TN 37996, USA*

[2]*Department of Ecology and Evolutionary Biology, University of Tennessee, Knoxville, TN, 37996-1610, USA*

Email for correspondence: jbeaulieu@nimbios.org

Abstract – The distribution of diversity can vary considerably from clade to clade. Attempts to understand these patterns often employ state speciation and extinction models to determine whether the evolution of a particular novel trait has increased speciation rates and/or decreased their extinction rates. It is still unclear, however, whether these models are uncovering important drivers of diversification, or whether they are simply pointing to more complex patterns involving many unmeasured and co-distributed factors. Here we describe an extension to the popular state speciation and extinction models that specifically accounts for the presence of unmeasured factors that could impact diversification rates estimated for the states of any observed trait. Specifically, our model, which we refer to as HiSSE (Hidden State Speciation and Extinction), assumes that related to each observed state in the model are "hidden" states that exhibit potentially distinct diversification dynamics and transition rates than the observed states in isolation. Under rigorous simulation tests and when applied to empirical data, we find that HiSSE performs reasonably well, and can at least detect net diversification rate differences between observed and hidden states. We also discuss the remaining issues with state speciation and extinction models in general, and the important ways in which HiSSE provides a more nuanced understanding of trait-dependent diversification.

Key words – hidden states, diversification, comparative methods, speciation, extinction

A key question in biology is why some groups are much more diverse than others. Discussions of such questions are often focused on whether there is something unique about exceptionally diverse lineages, such as the presence of some novel trait, which has increased their speciation rate and/or decreased their extinction rates. The BiSSE model (Binary State Speciation and Extinction; Maddison et al. 2007) was derived specifically as a means of examining the effect that the presence or absence of a single character could have on diversification rates, while also accounting for possible transitions between states. In theory, this model could be used not only for identifying differences in diversification, but also detecting differences in transitions between character states, or even some interplay of the two. In practice, however, it has mainly been used to infer diversification rate differences.

It is somewhat surprising, perhaps, that studies that employ BiSSE often find that the prediction of a trait leading to higher diversification rates is supported by the data. In fact, all sorts of traits have been implicated as potential drivers of diversity patterns, ranging from the evolution of herbivory in mammals (Price et al. 2012), to the evolution of extra-floral nectaries in flowering plants (Weber and Agrawal 2014), to even the evolution of particular body plans in fungi (i.e., gasteroid vs. nongasteroid forms; Wilson et al. 2011). Some important caveats are needed, however. First, Rabosky and Goldberg (2015) recently showed that for realistically complex datasets, BiSSE methods almost always incorrectly show a trait is correlated with higher diversification. Maddison and FitzJohn (2014) raise a more philosophical concern regarding the inability of these methods to properly account for independence. Consider, for instance, that the carpel, which encloses the seeds of angiosperms, has evolved only once. They argue that the

inheritance of a single event becomes problematic – even if BiSSE uncovers a significant correlation between the carpel and diversification, it is unclear whether the carpel is an important driver of the immense diversity of flowering plants, or whether this diversity is simply coincidental. Finally, it was pointed out by Beaulieu and Donoghue (2013) that even in the case of multiple origins of a trait, it could be that only one clade actually has a higher diversification rate associated with the focal trait, which is strong enough to return higher diversification rates for that trait as a whole. In their case, it appeared that plants with an achene (a fruit resembling a bare seed, as in "sunflower seeds") had a higher diversification rate, but upon subdividing the tree it appeared that this was from the inclusion of one clade in particular, the highly diverse sunflower family, the Asteraceae. They argue that it is far more likely that some combination of the achene and another unmeasured, co-distributed trait led to a higher diversification rate (on this point, also see Maddison et al. 2007).

All these caveats more or less relate to the broader issue of whether the proximate drivers of diversification are really ever just the focal trait themselves. At greater phylogenetic scales this issue seems the most relevant, where the context of a shared trait is unlikely to be consistent across many taxonomically distinct clades. In other words, a characters' effect on diversification will often be contingent on other factors, such as the assembly of particular combinations of characters and/or movements into new geographic regions (e.g., de Querioz 2002; Vamosi and Vamosi 2004; Moore and Donoghue 2007). Recent generalizations of the BiSSE model (i.e., MuSSE: Multistate Speciation and Extinction) do allow for additional binary characters to be accounted for when examining the correlation between a binary trait and net diversification. However, it may not always

be clear what the exact characters might be, and in the absence of such information, it should be difficult to ever confidently view any character state as the *true* underlying cause of increased diversification rates.

Here we describe a potential way forward by extending the BiSSE framework to account for the presence of unmeasured factors that could impact diversification rates estimated for the states of any observed trait. Our model, which we refer to as HiSSE (Hidden State Speciation and Extinction), assumes that related to each observed state in the model are "hidden" states that exhibit potentially distinct diversification dynamics and transition rates than the observed states in isolation. In this way, HiSSE is simply a state speciation and extinction form of a Hidden Markov model (HMM). As we will show, HiSSE can adequately distinguish higher net diversification rates within clades exhibiting a particular character state when they exist, and can provide a much more a refined understanding of how particular observed character states may influence the diversification process.

THE HIDDEN STATE SPECIATION AND EXTINCTION MODEL

In our view, despite the important methodological advancement and power afforded by the BiSSE model, it can still provide a rather coarse-grained view of trait-based patterns of diversification. Specifically, what may appear like a causal connection when examining particular characters in isolation may actually be due to other unmeasured factors, or because the analysis included a nested clade that exhibits both the focal character plus "something" else (Maddison et al. 2007; Beaulieu and Donoghue 2013; Beaulieu and O'Meara 2014). This particular point is illustrated in Figure 1. Here the true underlying model is one in which state 0 and state 1 have identical diversification

rates, but related to state 1 is some unknown third variable that is associated with a doubling of the speciation rate. Thus state 1 actually takes on two states: 1*A* when the hidden character is absent, and 1*B* when the hidden state is present. As indicated by the model, transitions to this unmeasured variable produce nested shifts towards higher rates of diversification *within* clades comprised of species in state 1. Necessarily, BiSSE can only infer parameters for characters that we can observe, and since all origins of state 1 are lumped together, the model incorrectly infers state 1 as being associated with significantly higher diversification rates.

We attempt to solve this problem by deriving an extension to the BiSSE model that allows for the inference of these hidden states. For the example in Figure 1, we can assume that all observations of state 1 are actually ambiguous for being in each of possible hidden states, 1*A* (i.e., hidden state absent) or state 1*B* (i.e., hidden state present). We then include transition rates and parameters for the diversification process associated with this hidden state. Our model, which we refer to hereafter as the HiSSE model (Hidden State Speciation and Extinction model), is actually a modified form of the MuSSE model (Multi-State Speciation and Extinction) that extends BiSSE type analyses to allow for multiple binary characters or characters with more than two states (FitzJohn 2012). The likelihood $D_{N,i}(t)$ is proportional to the probability that a lineage in state i (e.g., i = 0, 1*A*, or 1*B*) at time t before the present (t=0) evolved the exact branching structure and character states as observed. Changes in $D_{N,i}$ over time along a branch are described by the following ordinary differential equation:

$$\frac{\mathrm{d}D_{N,i}(t)}{\mathrm{d}t} = -\left(\lambda_i + \mu_i + \sum_{j \neq i} q_{ij}\right) D_{N,i}(t) + 2\lambda_i E_i(t) D_{N,i} + \sum_{j \neq i} q_{ij} D_{N,j}(t) \qquad \text{Eq. (1a)}$$

where $E_i(t)$ is the probability that a lineage in state i in time, t, goes completely extinct by the present, and is described by:

$$\frac{\mathrm{d}E_i(t)}{\mathrm{d}t} = \mu_i - \left(\lambda_i + \mu_i + \sum_{j \neq i} q_{ij}\right) E_i(t) + \lambda E_i(t)^2 + \sum_{j \neq i} q_{ij} E_j(t) \qquad \text{Eq. (1b)}$$

These series of equations are solved numerically along each edge starting tipward and moving rootward. The initial conditions for $D_{N,i}(0)$ are set to 1 when the trait is consistent with the observed data, and 0 otherwise. In the case of a hidden state, we set the probability to 1 for both state 1A and 1B for all species exhibiting state 1: that is, the probability of observing a tip demonstrating state 1 is 100% if the true underlying state is 1*A* or 1*B*. The initial conditions for $E_i(0)$ are all set to zero (i.e., we observe the tip in the present). Incomplete sampling can be allowed by incorporating a state-specific sampling frequency, f, and setting the initial conditions for $D_{N,i}(0)$ as f_i, and 0 otherwise, and for 1-f_i for $E_i(0)$ (FitzJohn et al., 2009). We assume the sampling frequency of the hidden state to be identical to the state with which it is associated (i.e., $f_{1A} = f_{1B}$).

At each ancestral node, A, that joins descendant lineages N and M, we multiply the probabilities of both daughter lineages together with the rate of speciation:

$$D_{A,i}(t) = D_{N,i}(t)\, D_{M,i}(t)\, \lambda_i \qquad \text{Eq. (2)}$$

and the resulting values become the initial conditions for the subtending branch. The overall likelihood is the sum of $D_{R,i}(t)$ calculated at the root. However, we condition this likelihood by $(1-E_i(t))^2$, which is the probability that the two descendant lineages of the root, evolving under the same estimates of speciation and extinction, survived to the present and were sampled (Nee et al. 1994). Finally, we follow the procedure described by FitzJohn et al. (2009) and also weight the overall likelihood by the likelihood that each possible state gave rise to the observed data. Note that in the absence of a hidden state, our likelihood calculation reduces exactly to a two-state BiSSE model.

It is important to point out that for present purposes we only focus on models that contain a single hidden state that is associated with a single observed state (i.e., state 1; Fig. 1). This need not always be the case. In fact, our HiSSE model can be generalized to include any number of hidden states associated with the observed states of a binary character. For example, it is straightforward to generate the SSE equivalent of the "precursor" model described by Marazzi et al. (2012) or the "hidden rates" model (HRM) described by Beaulieu et al. (2013). In the latter case, consider a model with "two rate classes", *A* and *B*. We can define speciation and extinction parameters for states 0*A*, 1*A*, 0*B*, and 1*B*, and then define a set of transitions to account for transitioning between all character state combinations:

$$\mathbf{Q} = \begin{bmatrix} - & q_{0A\to 1A} & q_{0A\to 0B} & 0 \\ q_{1A\to 0A} & - & 0 & q_{1A\to 1B} \\ q_{0B\to 0A} & 0 & - & q_{0B\to 1B} \\ 0 & q_{1B\to 1A} & q_{1B\to 0B} & - \end{bmatrix}$$

Thus, HiSSE can be used to account for differences in the diversification process while simultaneously identifying different classes of transition rates within a binary character. Our implementation allows for this and more complex models, but, again, this paper focuses on a more restricted case.

IMPLEMENTATION

We implemented the above model in the R package "*hisse*" available through CRAN. As input all that *hisse* requires is a phylogeny with branch lengths and a data file that contains the observed states of a binary character. Note that this is an entirely new implementation, not a fork of the existing *diversitree* package, as we employ modified optimization procedures and model configurations. For example, rather than optimizing λ_i and μ_i separately, *hisse* optimizes orthogonal transformations of these variables: we let $\tau_i = \lambda_i+\mu_i$ define "net turnover", and we let $\varepsilon_i = \mu_i / \lambda_i$ define the extinction fraction. This reparameterization alleviates problems associated with overfitting when λ_i and μ_i are highly correlated, but both matter in explaining the diversity pattern (e.g., Goldberg et al. 2010, Beaulieu and Donoghue 2013). With empirical data we often see good estimates for diversification rate but correlations for birth and death rate estimates. For example, Beaulieu and Donoghue (2013) showed that the confidence region for birth and death rates tightly follows a diagonal line, with different characters having lines of the same slope but different intercepts. This leads to a behavior where looking at the confidence or credibility intervals for birth or death show overlap between the characters but intervals for diversification rates do not overlap between characters. However, the only way to fit this in the normal parameterization is with two birth rates and two death rates.

Reparameterizing, as we do, allows us to have the same turnover rate for both states but estimate different diversification rates, resulting in a less complex, but better fitting, model. Thus, users specify configuration of models that variously fix τ_i and ε_i. Note that estimates of τ_i and ε_i can be easily backtransformed to reflect estimates of λ_i and μ_i by

$$\lambda_i = \tau_i / (1 + \varepsilon_i), \qquad \text{Eq. (3a)}$$

$$\mu_i = (\tau_i \, \varepsilon_i) / (1 + \varepsilon_i), \qquad \text{Eq. (3b)}$$

which can be used to derive other measures of diversification rate, such as net diversification (λ_i-μ_i).

The *hisse* package implements a marginal reconstruction method to assign the probability of each ancestral state to an internal node. These probabilities can be used to "paint" areas of a phylogeny where the diversification rate has increased or decreased due to some unmeasured character. In our implementation the marginal probability of state *i* for a focal node is simply the overall likelihood of the tree and data when the state of the focal node is fixed in state *i*. Note that the likeliest tip states can also be estimated: we observe state 1, but the underlying state could either be state 1*A* or 1*B*. Thus, for any given node or tip we traverse the entire tree as many times as there are states in the model. As the size of the tree grows, however, these repeated tree traversals can slow the calculation down considerably. For this reason, we also include a function that allows the marginal calculation to be conducted in parallel across any number of independent computer processors.

SIMULATIONS

We evaluated the performance of our HiSSE model by simulating trees and characters states under various scenarios and then estimating the fit and bias of the inferred rates from these trees. Our simulations included scenarios with and without hidden states – the known parameters for each scenario are described in detail in Table 1. With the BiSSE scenarios, our goal was to test whether we could correctly conclude that there was no support for a HiSSE model in the absence of hidden states in the generating model. For each scenario, trees were simulated to contain 50, 100, 200, or 400 species, with 100 replicates for each taxon set. When the generating model included a hidden state, we simulated trees that could transition between three possible states, 0, 1, or 2. After each simulation replicate was completed, we created the hidden state by simply switching the state of all tips observed in state 2 to be in state 1.

Each simulated data set was evaluated under the generating model, as well as 13 additional models, that variously added, removed, or constrained certain parameters. The entire model set is described in Table 1. For all models under a given scenario, model fit was assessed by calculating the average AIC weight (w_i), which represents the relative likelihood that model i is the best model given a set of models (Burnham and Anderson 2002). We also calculated a null expectation of the AIC weight across our model set, as the average AIC weight, assuming an equal likelihood across all models. Thus, our null expectation is based solely on the penalty term in the AIC calculation. Comparing the model-average of the parameters against the known parameters provided an assessment of the bias and precision of the inferred parameters. However, rather than averaging

parameters across the entire model set, we only averaged across models that included similar parameters. For example, when estimating the bias in the HiSSE scenarios we only model-averaged parameters for models that included the hidden state. This required reformulating the AIC weights to reflect the truncated model set. Finally, we assessed the reliability of the ancestral state reconstructions by comparing the true node states from each simulated tree to the marginal probabilities calculated from the model-averaged parameter estimates.

From a model comparison perspective, our simulations indicated that data sets that lack hidden states could generally be distinguished from those that do, especially with larger datasets (Table 1). When the generating model is a BiSSE model with only two observable states, there were low levels of support for all seven HiSSE models. In fact, as sample size increased, the average AIC weight of the HiSSE models converged towards the null AIC weight based only on the penalty term (Table 1). When evaluating data sets that included a hidden state, the ability to correctly favor a HiSSE model over any of the BiSSE models depended not just on the size of the dataset, but also the underlying generating model. For example, when the generating model assumed a hidden state with higher speciation rates (HiSSE Scenarios 1-2), data sets that contained 200 or more taxa were required to provide strong support for a HiSSE model that varied the turnover rate (Table 1). However, when the main effect of the generating model assumed either variable extinction fractions (HiSSE Scenario 3) or variable transition rates (HiSSE Scenario 4), there remained strong support for a BiSSE model that assumed both equal turnover rates and extinction fractions. In the case of HiSSE Scenario 4, there was fairly strong support for a BiSSE model even after we extended our simulations to include trees

containing 800 species. Interestingly, when we simulated under a HiSSE Scenario 6, which combines the processes of higher speciation rates and asymmetrical transition rates, the issue of incorrectly favoring a BiSSE model completely disappears (Table 1).

It is important to point out that in all simulation scenarios we never recovered strong support for a model consistent with the model that actually generated the data. That is, whether we vary the speciation rate or the extinction rate, which affects both net turnover and extinction fraction, we always found support for a simplified model that either allowed τ to vary (e.g., HiSSE Scenario 1-2) or ε to vary (e.g., HiSSE Scenario 3), but not both. This is likely a consequence of the uncertainty and upward biases in estimates of extinction fraction (Fig. 2), making it difficult for the model to infer multiple extinction fractions, at least with smaller data sets.

It is well known that there can be difficulties, generally, in obtaining precise estimates of the extinction rates (i.e., μ_i) under the BiSSE framework (Maddison et al. 2007). Backtransforming estimates of net turnover and extinction fraction shows that HiSSE suffers the same precision issues in regards to the rate of extinction (Fig. 3, Fig. S2). Recently, it was reported that biases in the tip state ratios could also generally impact all parameter estimates (see Davis et al. 2013). Our simulations did not specifically test issues related to tip state biases. However, we did find that with HiSSE, the precision in extinction rate estimates remains fairly low regardless of the ratio of the states at the tips, especially when estimating extinction for the hidden state (Fig. 3). Interestingly, the lack of precision for extinction rates seems to have a relatively minor impact on estimates of net turnover or net diversification (Fig. 2; Fig. S1), though there is a general bias for net diversification to be underestimated as a consequence of inflated extinction rates.

Nevertheless, it appears HiSSE correctly and qualitatively distinguishes differences in diversification among the various states in the model. And, as the number of species increases, the trajectory of the downward bias in net diversification suggests it will eventually disappear, and rate differences can be distinguished even if the differences are trivial (e.g., Fig. 2B).

One issue of concern from our simulations is that estimates of transition rates are almost always overestimated. This behavior appears unique to the HiSSE model (Fig. 2), given that when evaluating data sets under BiSSE scenarios, transition rates are estimated reasonably well (Fig. S3). At the moment we are unsure as to what might be specifically causing this bias in HiSSE. There are relatively large confidence intervals surrounding each of the transition rate estimates that naturally favor models that assume equal transition rates, which should be reflected in the model-averaged rates. Indeed, as in the case of Scenario 5, which assumed pronounced differences in the transitions rates, even 800 taxa was still not enough to unequivocally reject models that assumed equal transition rates (Table 1). We also examined the impact each parameter had on a fixed set of equilibrium frequencies, by randomly sampling sets of values and retaining those that estimated the same frequency within a small measure of error. The proportional range of values for the transition rates were more than double those found for the speciation rate, indicating that state frequencies are fairly resilient to changes in the transition rates. Taken together, estimating unique transition rates (and to some extent the rate of extinction) appears to be a naturally difficult problem, which is not made any easier by HiSSE. That is to say, conceptually, HiSSE requires more from the data by including additional parameters without providing any more observable information. It is likely that

in many cases far larger data sets with many more state origins than the ones we have generated here may be required to adequately estimate these particular parameters.

Finally, in regards to inference of ancestral states, the simulations indicate that HiSSE correctly identifies and locates regions of the tree where supposed diversification rate differences have taken place (Fig. 4). The degree of reliability does, of course, depend on the size of the data set. For HiSSE Scenario 1, for example, data sets comprised of 50 taxa, 84.1% of the nodes, on average, have the state correctly inferred, and data sets comprised of 400 taxa 92.4% are correct. However, we note that for smaller data sets there is also general tendency for HiSSE to infer high marginal probabilities for the incorrect state (e.g., Fig. 4A), which could provide misleadingly confident state reconstructions.

THE EVOLUTION OF ACHENE FRUITS

The development of this model was inspired by results from recent empirical work that applied BiSSE to understand the macroevolutionary consequences of evolving particular fruit types within a large flowering plant clade (i.e., campanulids; Beaulieu and Donoghue 2013). This study investigated whether diversification rates differences could explain why more than 80% of campanulid species exhibit fruits that are indehiscent (i.e., do not open mechanically), dry, and contain only a single-seed. From a terminological standpoint, these fruits were broadly referred to as “achene” or “achene-like” to unify the various terms used to identify the same basic fruit character configuration (e.g., “cypselas” of the Asteraceae – sunflowers and their relatives – or the single-seeded “mericarps” of the Apiaceae – carrots and their relatives). According to the BiSSE model, the preponderance of achene fruits within campanulids can be explained by strong

differences in diversification rates, with achene lineages having a rate that was roughly three times higher than non-achene lineages.

While these results are seemingly straightforward, they are complicated by the fact that the correlation between net diversification rates and the achene character state differed among the major campanulid lineages and was driven entirely by the inclusion of the Asterales clade (Beaulieu and Donoghue 2013). Within the Apiales and the Dipsacales, the two remaining major achene-bearing clades, the diversification rate differences were not technically significant. However, in both these clades there were qualitative differences in the predicted direction, likely as a consequence of one or more shifts in diversification nested within one of the major achene-bearing clades. Together, these point to a more complex scenario for the interaction between achene fruits and diversification patterns that is not being adequately explained by BiSSE.

We illustrate an empirical application of HiSSE by examining the Dipsidae (Paracryphiales+Dipsacales) portion of the achene data set of Beaulieu and Donoghue (2013). Specifically, we used HiSSE to locate and "paint" potential areas within Dipsidae that may be inflating the estimates of net diversification rates for achene lineages as a whole. We modified the original data set of Beaulieu and Donoghue (2013) in two important ways. First, we re-scaled the original branch lengths for Dipsidae in units of time using *treePL* (Smith and O'Meara 2012), an implementation of the penalized likelihood dating method of Sanderson (2002) specifically designed for large trees. We applied the same temporal constraints for Dipsidae as in the original Beaulieu and Donoghue (2013) study, and used cross-validation to determine the smoothing value, λ, that best predicted the rates of terminal branches pruned from the tree. Second, we

conservatively removed various taxa of dubious taxonomic distinction, taxa considered varietals or subspecies of a species already contained within the tree, and tips that *treePL* assigned very short branch lengths (i.e., <0.50 Myr) – all of which have the tendency to negatively impact diversification rates. The exclusion of these taxa resulted in a data set comprised of 418 species from the original 457.

We fit 15 different models to the achene data set for Dipsidae. Seven of these models corresponded to various BiSSE models that either removed or constrained particular parameters, and eight corresponding to various HiSSE models that assumed a hidden state associated with the observed achene state (i.e., achene+). In all cases, we incorporated a unique sampling frequency scheme to the model. Rather than assuming random sampling for the entire tree (see above), we included the sampling frequency for each major clade included in the tree (see Supplemental Materials). Finally, in order to generate a measure of confidence for all parameters estimated under a given model, we implemented an "adaptive search" procedure that provides an estimate of the parameter space that is some pre-defined likelihood distance (e.g., 2lnL units) from the maximum likelihood estimate (MLE). All analyses were carried out in *hisse*.

The best model, based on Akaike weights, was a full HiSSE model, which allowed all associated diversification and transition rates to be estimated freely (Table 2). This model had a pronounced improvement over the set of BiSSE models, where none had relative probability that was greater than $2x10^{-22}$. However, before describing the parameter estimates of the best model, we note that, despite a modified tree and character set, the parameter estimates under the BiSSE models were similar to those reported by Beaulieu and Donoghue (2013). In our case, however, the higher diversification rates

estimated for achene lineages (r_{achene}=0.104, support range: [0.080,0.127]) compared to non-achene lineages ($r_{non\text{-}achene}$=0.056, support range: [0.037,0.078]) were indeed significant based on a sampling of points falling within 2lnL units away from the MLE (Fig. 5). The parameters estimated under the HiSSE model, on the other hand, suggest a more nuanced interpretation of this result. The higher diversification rates of clades bearing achenes as a whole is likely due to a hidden state nested within some of these lineages that is associated with exceptionally high diversification rates ($r_{achene+}$=0.306, support region: [0.280,0.380]). In fact, achene lineages not associated with the hidden state have a diversification rate that is very similar to the non-achene diversification rate ($r_{non\text{-}achene}$=0.057 vs. $r_{achene\text{-}}$=0.058; Fig. 5).

Character reconstructions identified two transitions to the "hidden" variable, and thus a higher diversification rate: one shift occurred along the stem leading to crown Dipsacaceae and the other occurred along the stem leading to "core Valerianaceae" (the most inclusive clade that excludes *Patrinia*, *Nardostachys*, and *Valeriana celtica*). It is important to note, however, that in both cases there is a fair amount of uncertainty in regards to the states at the tips, and in the case of Dipsacaceae, there is even uncertainty as to whether or not the shift to the hidden state includes the genus *Bassecoia* (Fig. 5). The reconstruction also suggests that the hidden state may have been lost in several contemporary tip taxa (e.g., *Succisella*, *Succisa*, *Lomelosia*).

Nevertheless, the general location of these inferred shifts is intriguing, as they appear to coincide with specific clades that exhibit specialized structures related to achene dispersal. In the Dipsacaceae, for example, there is tremendous diversity in the shape of the "epicalyx", a tubular structure that subtends the calyx and encloses the ovary

(Donoghue et al. 2003; Carlson et al. 2009), which is often associated with elaborated structures (e.g., “wings”, “pappus-like” bristles) that accompany their achene fruits. Interestingly, many of these same general forms are observed throughout the core Valerianaceae, although they arise from modifications to the calyx (Donoghue et al. 2003; Jacobs et al. 2010). While the significance of these structures is thought to improve protection, germination, and dispersal of the seed, we emphasize that, at this stage, it is difficult to confidently rule out other important factors, such as traits that are completely unrelated to the achene fruits, important biogeographic movements (Moore and Donoghue 2007), or even purely genetic changes, such as gene and genome duplications (Hildago et al. 2010; Carlson et al. 2011). But, we can at least confidently conclude that the achene by itself is likely not a strong correlate of diversity patterns within the Dipsidae clade.

DISCUSSION

Progress in biology comes from confronting reality with our hypotheses and either confirming that our view of the world is correct, or, more excitingly, finding out that we still have a lot left to discover. With studies of diversification, we have mostly been limited to doing the former – we have an idea about a trait that may affect diversification rates based on intuition about its potential effect (e.g., an achene might allow greater dispersal, and thus easier colonization of new habitats to form new species) as well as some knowledge of its distribution (e.g., some very large plant clades have achenes) and then run an analysis. Typical outcomes are yes, there is a difference in diversification in the way we expected, or there is no difference but maybe we just lack the necessary power. In any case, chances are we are at least vaguely correct that the trait

we think credibly has a mechanism for increasing diversification rate actually has an effect (at least once there is enough power).

Surprise is a necessary part of discovery that, to put it bluntly, has been relatively lacking in trait-dependent diversification studies until now. With HiSSE we can still test our intuitions about a particular character, but we can also discover that rates are driven by some unknown and unmeasured character state, allowing the data to help us generate new hypotheses. Is the diversification rate correlated with the achene, or is the achene simply a necessary precursor to some other trait that is more likely to be driving diversification? This lets us go from a scenario where we simply reject trivial nulls, such as whether diversification rates of clades with and without some focal trait are precisely equal, to being potentially surprised by the results – No, it is not the achene *per se*, but it is something else nested within these particular clades in addition to the achene fruit type.

For all intents and purposes, HiSSE is an enhancement of BiSSE-type models, but also provides a solution to some of the recent important criticisms levied against state speciation and extinction models (Maddison and FitzJohn 2014, Rabosky and Goldberg 2015). Specifically, our method no longer requires the assumption that a focal character state is the one associated with diversification rate. Instead, it allows this assumption to be explored as part of a more flexible overall model, as opposed to relying on separate tests for uncovering character-dependent and character-independent rates of diversification (e.g., Beaulieu and Donoghue 2013; Rabosky and Goldberg 2015).

We highlight two important ways, however, in which HiSSE requires significant cautions in its use. The first is its indifference to number of changes (Maddison and FitzJohn 2014). As with BiSSE, we should find it more credible that a particular

character state enhances diversification rates if we see a rate increase in each of the 10 times it evolves than if it evolved just once but had the same magnitude of rate increase. A good solution to this problem has not yet been proposed, and we urge a healthy skepticism of any result based on a trait that has evolved only a few times. An important issue was also raised by Rabosky and Goldberg (2015), who demonstrated that on empirical trees even traits evolving under a neutral, diversification-independent model will still tend to be best fit by a BiSSE model. The HiSSE model makes this issue somewhat less problematic – namely, we expect under these circumstances that models allowing a hidden state will provide a better fit to the data than a standard BiSSE model. This should reduce ones belief that their focal trait is the one driving diversification, which, of course, is still different from returning the true model where the diversification rate does not depend on the trait.

On this point, it is important to bear in mind that BiSSE, HiSSE, and any other model of state speciation and extinction are *not* models of trait evolution, but rather are joint models for the underlying tree and the traits. A trait evolution model like those in Pagel (1994) or Hansen (1997) maximizes the probability of the observed states at the tips, given the tree and model – the tree certainly affects the likelihood of the tip data, but that is the only way it enters the calculation. A trait-based diversification model, on the other hand, maximizes the probability of the observed states at the tips and the observed tree, given the model. If a tree violates a single regime birth-death model due to any number of causes (e.g., mass extinction events, trait-dependent speciation or extinction, maximum carrying capacity, climate change affecting speciation rates, etc.), then even if the tip data are perfectly consistent with a simple model, the tip data *plus* the tree are not.

In such a case, it should not be surprising that a more complex model will tend to be chosen over a nested simpler model, particularly if the underlying tree is large enough.

Furthermore, as is well known in statistics, rejecting the null model does not imply that the alternative model is true. It simply means that it fits better and always will, especially when looking at models in any complex system where the true model may not be one of the included models. For example, Rabosky and Goldberg (2015) showed that datasets simulated under a model with characters having no effect on diversification rate on an empirical tree of cetaceans (i.e., whales, dolphins, and relatives) almost always rejected the null model. Though presented as a Type I error (i.e., incorrectly rejecting a true null), it is not. While the chosen character model is wrong, the cetacean tree is almost certainly not evolving with a single speciation and extinction rate for the entire clade. BiSSE is correct in saying that the simple model is not correct, but it is very wrong in assigning rate differences to the simulated traits. Users of BiSSE, and now HiSSE, must be careful not to put too much weight into an alternate model, even one with strong support, unless other credible models have been thoroughly investigated. Finding evidence for an alternate model must not simply be accepted as the truth, but rather be the start of further investigation of additional lines of evidence.

These are far from the only potential pitfalls in empirical applications of the HiSSE model. While it could be used over a set of trees (i.e., bootstrap or Bayesian postburnin tree samples), the model assumes that the branch lengths, topology, and states are known without error. Certain kinds of phylogenetic errors, such as terminal branch lengths that are too long (as may occur with sequencing errors in the data used to make the tree) can result in particular biases in estimates of speciation and/or extinction (also

see Beaulieu and O'Meara 2015). Similarly, if one clade were reconstructed to be younger than it actually is, perhaps due to a substitution rate slowdown caused by some other trait (e.g., life-history; Smith and Donoghue 2008), it could be interpreted as having a faster diversification rate, perhaps even inferred to have its own hidden state. The HiSSE model assumes discrete characters, whether they be hidden or observed, but it could be that a continuous parameter is the cause of a diversification rate difference (e.g., perhaps extinction risk varies inversely with mean of the dispersal kernel). However, this may only enter the model as an unseen discrete character, perhaps corresponding to low and high values of the continuous character.

There is no question that state speciation and extinction models are an important advancement for understanding characters' impact on diversification patterns. They have greatly improved statistical power over older, simpler sister-clade comparisons, and the explicit inference of differences in speciation and extinction has the potential for a much more fine-grained analysis of diversification. But, in a way, these models have also allowed us to retreat to the old comforts of reducing complex organisms into units of single, independently evolving characters, and offering adaptive interpretations to each (c.f., Gould and Lewontin 1979). To be fair, of course, it is unlikely that any trait of great interest to biologists has exactly zero effect on speciation and/or extinction rates, but it is certainly unlikely that this trait acts in isolation. Thus, we hope HiSSE is viewed as a step away from this line of thinking, as we no longer have to necessarily focus analyses, or even interpret the results, by reference to the focal trait by itself. It is in this way that analyses focused on "hidden" factors promoting diversification will afford us a more refined understanding of why certain clades become extraordinarily diverse.

Acknowledgements

We thank Nathan Jackson, Nick Matzke, Kathryn Massana, Chuck Bell, and Andrew Leslie for helpful discussions. We also thank Jeffrey Oliver and Elizabeth Spriggs for feedback on the method implementation. JMB would like to thank Michael Donoghue for wonderful discussions over the years regarding some of the concepts described here. Support for JMB has been provided by the National Institute for Mathematical and Biological Synthesis, an Institute sponsored by the National Science Foundation, the U.S. Department of Homeland Security, and the U.S. Department of Agriculture through NSF Award #EF-0832858, with additional support from The University of Tennessee, Knoxville.

References

Beaulieu J.M., Donoghue M.J. 2013. Fruit evolution and diversification in campanulid angiosperms. Evolution. 67: 3132-3144.

Beaulieu J.M., O'Meara B.C., Donoghue M.J. 2013. Identifying hidden rate changes in the evolution of a binary morphological character: the evolution of plant habit in campanulid angiosperms. Systematic Biology, 62:725-737.

Beaulieu J.M., O'Meara B.C. 2014. Hidden Markov models for studying the evolution of binary morphological characters. In *Modern Phylogenetic Comparative Methods and their Application in Evolutionary Biology*, Garamszegi L.Z. editor. London:Springer.

Beaulieu J.M., O'Meara B.C. 2015. Extinction can be estimated from moderately sized molecular phylogenies. Evolution. In press.

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

Carlson S.E., Mayer V., Donoghue M.J. 2009. Phylogenetic relationships, taxonomy, and morphological evolution in Dipsacaceae (Dipsacales) inferred by DNA sequence data. Taxon, 58:1075-1091.

Carlson S.E., Howarth D.G., Donoghue M.J. 2011. Diversification of CYCLOIDEA-like genes in Dipscaceae (Dipsacales): implications for the evolution of capitulum inflorescences. BMC Evolutionary Biology, 11:325.

Davis M.P., Midford P.E., Maddison W. 2013. Exploring power and parameter estimation of the BiSSE method for analyzing species diversification. BMC Evolutionary Biology 13:38.

de Querioz A. 2002. Contingent predictability in evolution: Key traits and diversification. Systematic Biology 51:917-929.

Donoghue M.J., Bell C.D., Winkworth R.C. 2003. The evolution of reproductive characters in Dipsacales. International Journal of Plant Sciences. 164:S453-S464.

FitzJohn R.G. 2012. Diversitree: comparative phylogenetic analyses of diversification in R. Methods in Ecology and Evolution. 3:1084-1092.

FitzJohn R.G., Maddison W.P., Otto S.P. 2009. Estimating trait-dependent speciation and extinction rates from incompletely resolved phylogenies. Systematic Biology. 58:595-611.

Goldberg E.E., Kohn J.R., Lande R., Robertson K.A., Smith S.A., Igíc B. 2010. Species selection maintains self-incompatibility. Science 330:493-495.

Gould S.J., Lewontin R.C. 1979. The spandrels of San Marco and the Panglossian paradigm: a critique of the adaptationist programme. Proceedings of the Royal Society, B. 205:581-598.

Hansen T.F. 1997. Stabilizing selection and the comparative analysis of adaptation. Evolution 51:1341-1351.

Hildago O., Mathez J., Garcia S., Garnatje T., Pellicer J., Valles J. 2010. Genome size study in the Valerianaceae: First results and new hypotheses. Journal of Botany. doi:10.1155/2010/797246.

Jacobs B., Bell C., Smets E. 2010. Fruits and seeds of the Valeriana clade (Dipsacales): Diversity and evolution. International Journal of Plant Sciences. 171:421-434.

Maddison W.P., Midford P.E., Otto S.P. 2007. Estimating a binary character's effect on speciation and extinction. Systematic Biology. 56:701-710.

Maddison W.P., FitzJohn R.G. 2014. The unsolved challenge to phylogenetic correlation tests for categorical characters. Systematic Biology. 64:127-136.

Marazzi B., Ane C., Simon M.F., Delgado-Salinas A., Luckow M., Sanderson M.J. 2012. Locating evolutionary precursors on a phylogenetic tree. Evolution 66:3918-3930.

Moore B.R., Donoghue M.J. 2007. Correlates of diversification in the plant clade Dipsacales: geographic movement and evolutionary innovations. American Naturalist. 170:S28-S55.

Nee S., Holmes E.C., May R.M., Harvey P.H. 1994. Extinction rates can be estimated from molecular phylogenies. Philosophical Transactions of the Royal Society, B. 344:77-82.

Pagel M. 1994. Detecting correlated evolution on phylogenies: a general method for the comparative analysis of discrete characters. Proceedings of the Royal Society, B. 255:37-45

Price S.A., Hopkins S.S.B, Smith K.K., Roth V.L. 2012. Tempo of trophic evolution and its impact on mammalian diversification. Proceedings of the National Academy of Sciences of the U.S.A. 109: 7008-7012.

Rabosky D.L., Goldberg E.E. 2015. Model inadequacy and mistaken inferences of trait-dependent speciation. Systematic Biology. In press.

Sanderson M.J. 2002. Estimating absolute rates of molecular evolution and divergence times: a penalized likelihood approach. Molecular Biology and Evolution. 19:101-109.

Smith S. A., Donoghue M.J. 2008. Rates of molecular evolution are linked to life history in flowering plants. Science 322: 86-89.

Smith S.A. O'Meara B.C. 2012. treePL: Divergence time estimation using penalized likelihood for large phylogenies. Bioinformatics. 28:2689-2690.

Vamosi J.C., Vamosi S.M. 2004. The role of diversification in causing the correlates of dioecy. Evolution. 58:723-731.

Weber M.G., Agrawal A.A. 2014. Defense mutualisms enhance plant diversification. Proceedings of the National Academy of Sciences of the U.S.A. 111: 16442-16447.

Wilson A.W., Binder M., Hibbett D.S. 2011. Effects of gasteroid fruiting body morphology on diversification rates in three independent clades of fungi estimated using binary state speciation and extinction analysis. Evolution 65:1305-1322.

TABLE 1. Summary of the model support for each simulated scenario, where calculating the average AIC weight (w_i) for all models assessed the fit. We also calculated a null expectation of the AIC weight as the average AIC weight if we assumed an equal likelihood across all models. Thus, the null expectation is based solely on the penalty term in the AIC calculation.

		BiSSE							HiSSE						
				ε0=ε1		τ0=τ1		τ0=τ1			ε0=ε1A=ε1B		τ0=τ1A=τ1B		τ0=τ1A=τ1B
Generating Model	Ntaxa	Free	q's equal	q's equal	τ0=τ1	q's equal	ε0=ε1	ε0=ε1	Free	q's equal	q's equal	τ0=τ1A=τ1B	q's equal	ε0=ε1A=ε1B	ε0=ε1A=ε1B
Null expectation		0.0503	0.0830	0.1368	0.0830	0.1368	0.0830	0.1368	0.0068	0.0305	0.0830	0.0185	0.0830	0.0185	0.0503
BiSSE Scenario 1	50	0.0154	0.0611	***0.1996***	0.0587	***0.2257***	0.0685	***0.2514***	0.0001	0.0042	0.0571	0.0009	0.0484	0.0019	0.0070
λ0=λ1=0.1	100	0.0219	0.0697	0.1713	0.0514	0.1676	0.0738	***0.2339***	0.0008	0.0152	0.1043	0.0030	0.0656	0.0102	0.0112
μ0=μ1=0.03	200	0.0233	0.0482	0.0988	0.0783	0.1059	0.0803	***0.2825***	0.0018	0.0246	0.1053	0.0111	0.0843	0.0260	0.0296
q01=0.015, q10=0.005	400	0.0370	0.0674	0.0497	0.0981	0.0687	0.0949	***0.3305***	0.0010	0.0139	0.0813	0.0102	0.0796	0.0169	0.0507
BiSSE Scenario 2	50	0.0109	0.0553	***0.2885***	0.0558	0.2299	0.0691	0.1681	0.0101	0.0035	0.0577	0.0016	0.0415	0.0034	0.0047
λ0=0.1, λ1=0.2	100	0.0163	0.0668	***0.2940***	0.0499	0.2091	0.0884	0.1180	0.0002	0.0081	0.0903	0.0019	0.0470	0.0067	0.0034
μ0=μ1=0.03	200	0.0194	0.0711	***0.3329***	0.0557	0.2043	0.1077	0.0445	0.0034	0.0059	0.0786	0.0023	0.0464	0.0265	0.0014
q01=q10=0.01	400	0.0347	0.0930	***0.2903***	0.0769	0.2029	0.1270	0.0106	0.0028	0.0108	0.0618	0.0060	0.0497	0.0292	0.0043
BiSSE Scenario 3	50	0.0210	0.0671	***0.1969***	0.0601	***0.1845***	0.0747	***0.2594***	0.0012	0.0071	0.0636	0.0013	0.0475	0.0085	0.0073
λ0=λ1=0.1	100	0.0244	0.0821	***0.1686***	0.0636	***0.2174***	0.0630	***0.2271***	0.0004	0.0047	0.0683	0.0067	0.0593	0.0062	0.0082
μ0=0.06, μ1=0.03	200	0.0397	0.0957	0.1448	0.0659	***0.2434***	0.0574	0.1667	0.0005	0.0146	0.0704	0.0015	0.0683	0.0057	0.0253
q01=q10=0.01	400	0.0486	0.1181	0.1093	0.0829	***0.2270***	0.0649	0.1437	0.0031	0.0120	0.0630	0.0035	0.1029	0.0050	0.0160
HiSSE Scenario 1	50	0.0206	0.0540	***0.2157***	0.0749	***0.1943***	0.0914	***0.2212***	0.0003	0.0089	0.0635	0.0012	0.0356	0.0028	0.0156
λ0=λ1A=0.01, λ1B=0.2	100	0.0328	0.0693	***0.1611***	0.0585	***0.1430***	0.0788	0.1435	0.0034	0.0181	***0.1798***	0.0055	0.0873	0.0073	0.0117
μ0=μ1A=μ1B=0.03	200	0.0567	0.0472	0.0813	0.0685	0.0869	0.0680	0.0967	0.0022	0.0247	***0.2891***	0.0094	0.1276	0.0269	0.0148
q01A=q1A0=q1A1B=q1B1A=0.01	400	0.0271	0.0287	0.0428	0.0216	0.0298	0.0274	0.0268	0.0091	0.0629	***0.4114***	0.0200	0.2043	0.0858	0.0023
HiSSE Scenario 2	50	0.0277	0.0590	***0.2211***	0.0667	0.1516	0.1060	0.1296	0.0024	0.0268	***0.1442***	0.0101	0.0274	0.0240	0.0035
λ0=λ1A=0.01, λ1B=0.3	100	0.0326	0.0476	0.1236	0.0498	0.0960	0.0835	0.0874	0.0022	0.0345	***0.2902***	0.0059	0.1105	0.0251	0.0109
μ0=μ1A=μ1B=0.03	200	0.0352	0.0258	0.0386	0.0241	0.0490	0.0552	0.0150	0.0070	0.0625	***0.5343***	0.0147	0.0787	0.0595	0.0004
q01A=q1A0=q1A1B=q1B1A=0.01	400	0.0140	0.0133	0.0070	0.0097	0.0088	0.0539	0.0100	0.0076	0.1136	***0.5980***	0.0079	0.0722	0.0836	0.0004
HiSSE Scenario 4	50	0.0224	0.0781	***0.1921***	0.0686	***0.1896***	0.0845	***0.2480***	0.0001	0.0107	0.0548	0.0008	0.0429	0.0018	0.0055
λ0=λ1A=λ1B=0.1	100	0.0270	0.0568	0.1539	0.0732	0.1527	0.0786	***0.2654***	0.0015	0.0168	0.0528	0.0057	0.0603	0.0156	0.0397
μ0=μ1A=0.06, μ1B=0.03	200	0.0317	0.0470	0.1169	0.0863	***0.1516***	0.0790	***0.1918***	0.0018	0.0282	0.0913	0.0071	0.1039	0.0274	0.0360
q01A=q1A0=q1A1B=q1B1A=0.01	400	0.0338	0.0362	0.0445	0.0893	0.0729	0.1035	***0.1644***	0.0181	0.0436	0.0876	0.0338	***0.2010***	0.0200	0.0513
HiSSE Scenario 5	50	0.0268	0.0512	0.1844	0.0714	0.1582	0.0776	***0.2971***	0.0022	0.0055	0.0681	0.0007	0.0379	0.0041	0.0147
λ0=λ1A=λ1B=0.1	100	0.0369	0.0441	0.1182	0.0936	0.0937	0.1039	***0.3213***	0.0015	0.0072	0.0698	0.0027	0.0646	0.0192	0.0233
μ0=μ1A=μ1B=0.03	200	0.0337	0.0231	0.0648	0.1003	0.0817	0.0928	***0.3247***	0.0023	0.0166	0.0699	0.0108	0.0708	0.0324	0.0760
q01A=0.03, q1A0=0.005,	400	0.0560	0.0120	0.0180	0.1402	0.0225	0.1321	***0.4033***	0.0026	0.0196	0.0622	0.0068	0.0623	0.0097	0.0528
q1A1B=0.01, q1B1A=0.00	800	0.0441	0.0045	0.0163	0.1110	0.0180	0.1022	***0.3015***	0.0150	0.0123	0.0295	0.0590	0.0283	0.0557	***0.2027***
HiSSE Scenario 6	50	0.0169	0.0489	***0.1928***	0.0640	***0.1705***	0.0823	***0.2030***	0.0023	0.0328	0.1148	0.0073	0.0468	0.0125	0.0049
λ0=λ1A=0.01, λ1B=0.3	100	0.0194	0.0476	0.0805	0.0448	0.0764	0.0746	0.1140	0.0055	0.0339	***0.3299***	0.0065	0.0701	0.0823	0.0143
μ0=μ1A=μ1B=0.03	200	0.0353	0.0156	0.0149	0.0361	0.0189	0.0505	0.0783	0.0074	0.0841	***0.4054***	0.0073	0.1225	0.1183	0.0055
q01A=0.015, q1A0=q1A1B=0.005,	400	0.0069	0.0014	0.0044	0.0096	0.0091	0.0296	0.0164	0.0517	0.1347	***0.3887***	0.0171	0.0576	***0.2724***	0.0004
q1B1A=0.01	800	0.0015	0.0000	0.0000	0.0012	0.0000	0.0009	0.0062	0.0592	0.1242	***0.3752***	0.0093	0.0710	***0.3513***	0.0001

Bold italics indicate significance based on 95% confidence intervals

TABLE 2. The fit of alternative models of achene fruit evolution in the Dipsidae. The best model (denoted in bold), based on ΔAIC and Akaike weights (w_i) was a full HiSSE model, which allowed all associated diversification and transition rates to be estimated freely.

Model	lnLik	AIC	ΔAIC	w_i
BiSSE: All free	-1391.2	2794.3	102.5	<0.001
BiSSE: q's equal	-1392.3	2794.5	102.7	<0.001
BiSSE: ε0=ε1, q's equal	-1394.9	2797.9	106.1	<0.001
BiSSE: τ0=τ1	-1391.5	2793.0	101.2	<0.001
BiSSE: τ0=τ1, q's equal	-1392.9	2793.8	102.0	<0.001
BiSSE: ε0=ε1	-1393.0	2795.9	104.1	<0.001
BiSSE: τ0=τ1, ε0=ε1	-1394.5	2797.0	105.2	<0.001
HiSSE: All free	***-1335.9***	***2691.8***	***0.0***	***0.714***
HiSSE: q1B1A=0	-1338.8	2695.6	3.8	0.107
HiSSE: q's equal	-1340.1	2694.3	2.5	0.206
HiSSE: ε0=ε1A=ε1B, q's equal	-1357.5	2725.0	33.3	<0.001
HiSSE: τ0=τ1A=τ1B	-1352.1	2720.1	28.3	<0.001
HiSSE: τ0=τ1A=τ1B, q's equal	-1356.4	2722.8	31.0	<0.001
HiSSE: ε0=ε1A=ε1B	-1347.7	2711.4	19.6	<0.001
HiSSE: τ0=τ1A=τ1B, ε0=ε1A=ε1B	-1393.4	2798.8	107.0	<0.001

FIGURE LEGENDS

FIGURE 1. A conceptual problem with the presence of hidden states in trait-dependent models of speciation and extinction. Here the true underlying model is one in which state 0 and state 1 have identical diversification rates (dashed black line), but related to state 1 is an unmeasured third variable that is associated with a doubling of the speciation rate (dashed red line). In fact, state 1 actually takes on two states: 1*A* when the hidden character is absent, and 1*B* when the hidden state is present. As shown in the example tree, transitions to this unmeasured variable naturally produces nested shifts towards higher rates of diversification within clades comprised of species observed in state 1. When we run 100 simulations of this particular model and fit the resulting data sets in BiSSE, the model incorrectly infers state 1 as being associated with a significantly higher diversification rates.

FIGURE 2. The uncertainty surrounding estimates of net diversification rate (λ-μ), extinction fraction (μ/λ), and transition rates as a function of tree size. Each row represents a different simulation scenario under the HiSSE model, all of which are described in detail in Table 1 from the main text. For both net diversification and extinction fractions the solid green line represents and green region represent the mean and 95% confidence interval for state 0, the solid blue line and blue region represents mean and 95% confidence interval for state 1*A*, and the solid red line and red region represents the mean and 95% confidence interval for state 1*B*. In the panels depicting the log-transformed transition rates, the solid purple line and purple region represent mean

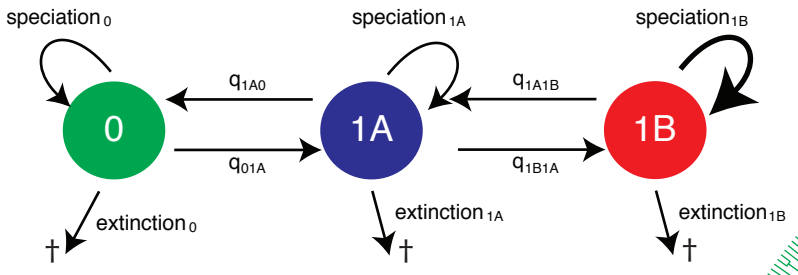
and 95% confidence interval for transition to and from the "hidden" state (i.e., 1*B*). In all panels, the color of dashed line corresponds to the true value under the generating model.

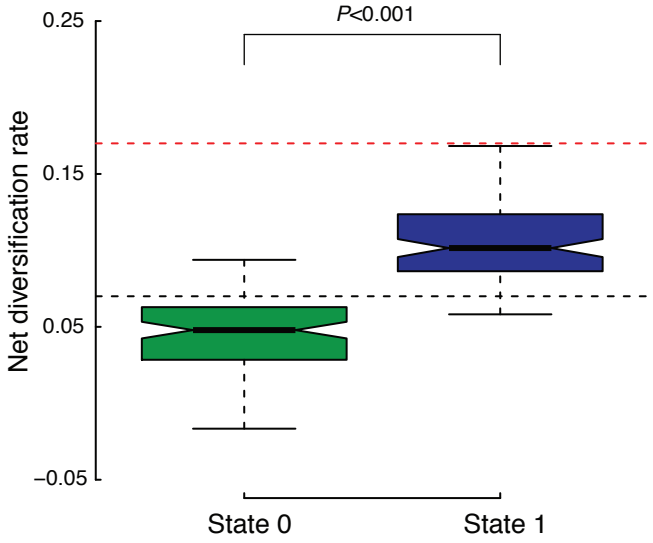
FIGURE 3. The parameter estimates, under a generating model that assumes transitions to state 1*B* is associated with a doubling of the speciation rate (HiSSE Scenario 1), transformed to reflect speciation (λ) and extinction (μ) rates as a function of the tip state frequencies. The dashed lines in each panel correspond to the true value under the generating model.

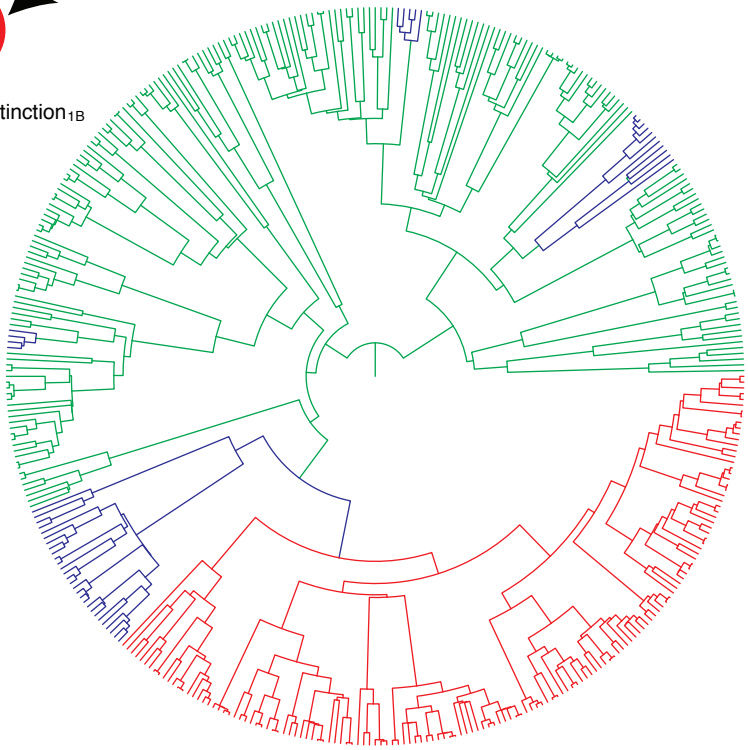
FIGURE 4. The reliability of ancestral state reconstructions as a function of tree size based on the model-averaged parameter estimates from HiSSE Scenario 1. The trend lines represent the relationship between the rolling median of the maximum marginal probability and the rolling mean of proportion of nodes inferred to be in the correct state across all simulated trees. The rolling mean and median are based on a sliding window of 500 points, and the dotted line denotes the 1:1 relationship between the two variables. While the trend lines are meant to reflect the relationship across a range of values, the gray histograms on each of the axes show the underlying distribution of the actual data used to the produce the lines.

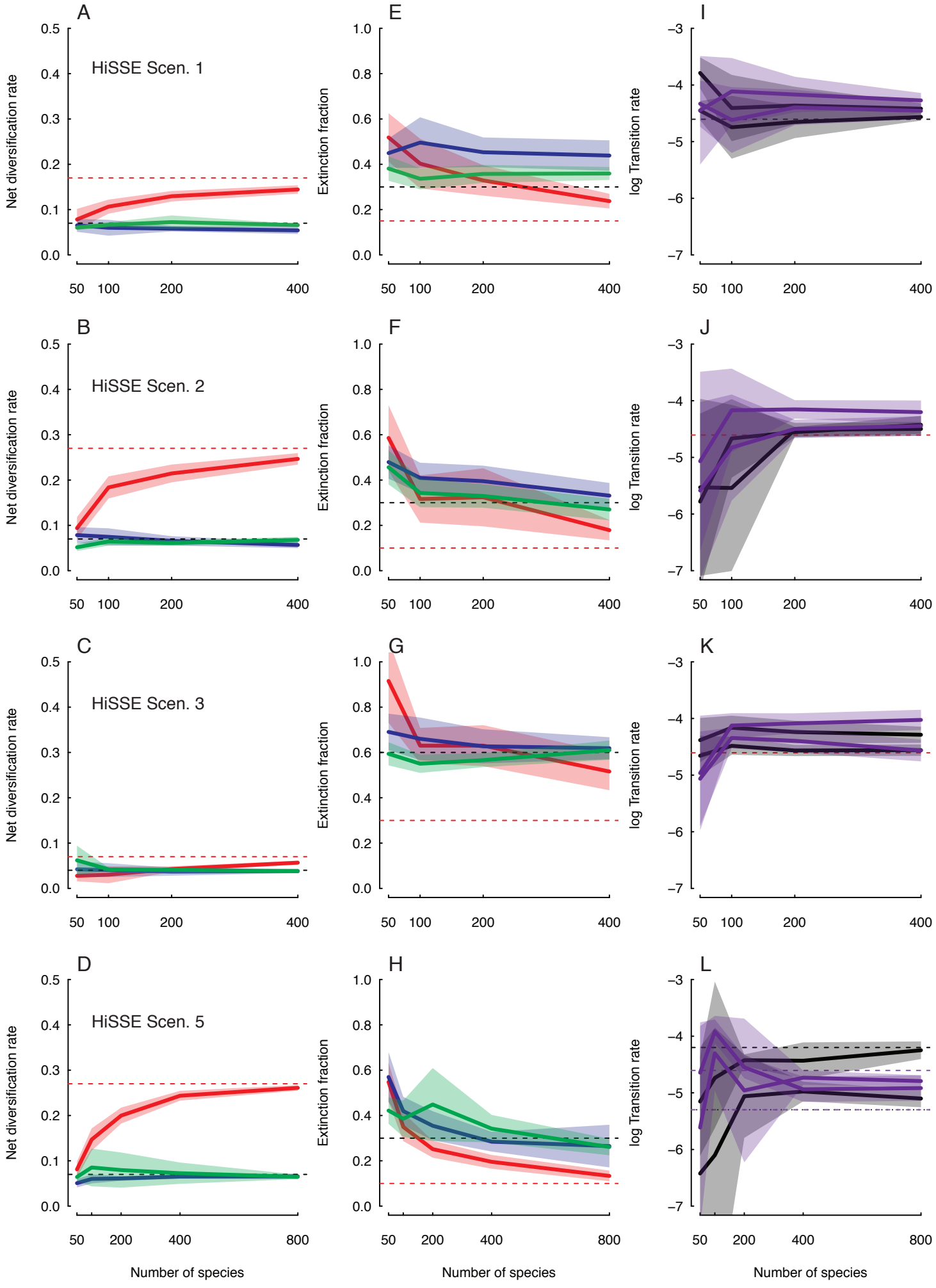
FIGURE 5. Comparisons of character reconstructions and estimates of net diversification rate when both the BiSSE model and HiSSE model are applied to a large empirical data set of achene fruits (i.e., dry, indehiscent, and single-seeded fruits) for the Dipsidae. The major clades are labeled and estimates of the most likely state are based on the marginal

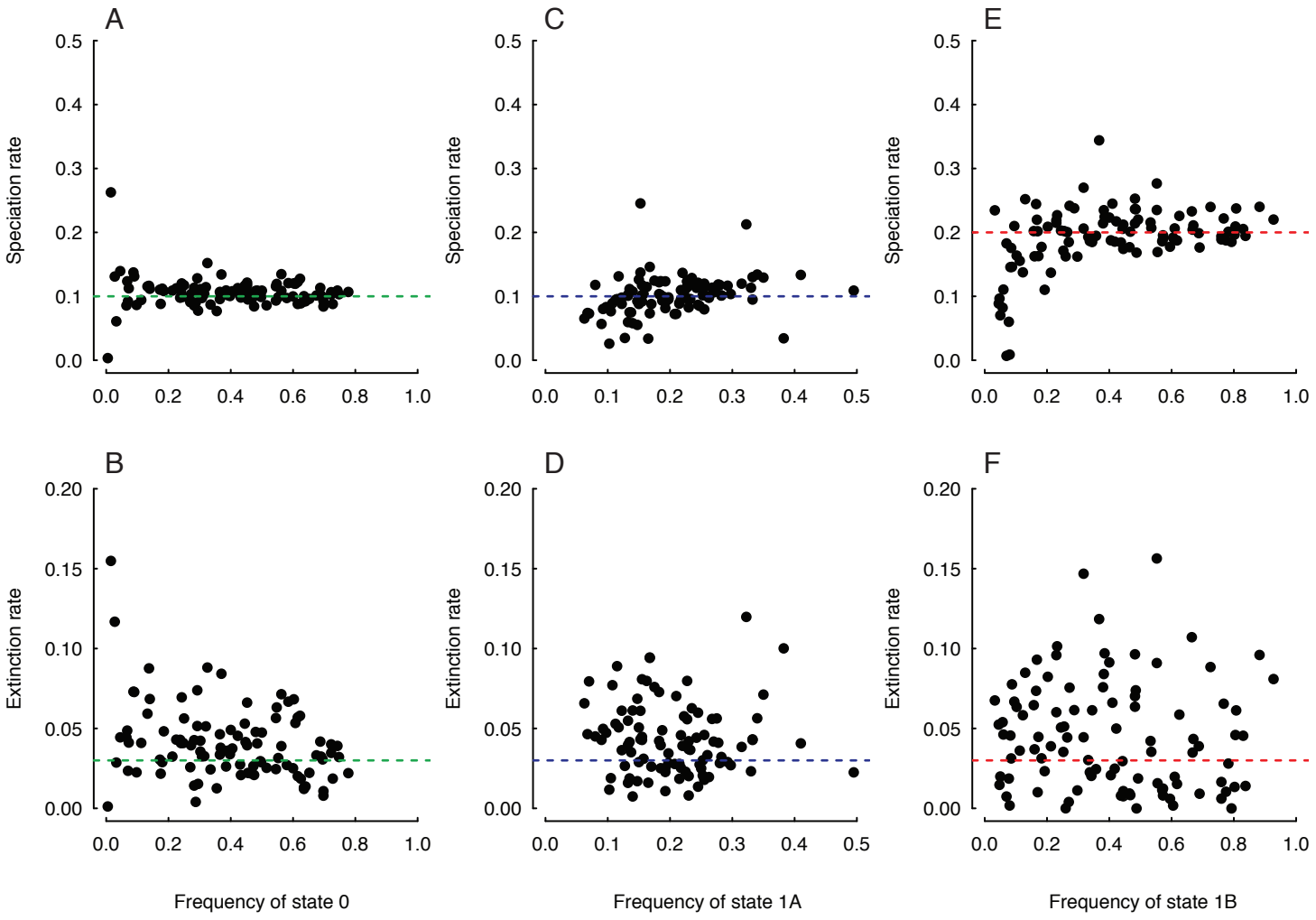
reconstruction using model-averaged parameters inferred under the BiSSE model and the HiSSE model. In the case of the HiSSE character reconstruction, branches are colored based on the marginal probability; the branches showing high probability in the achene state minus the hidden state (achene-) will appear blue, whereas achene branches showing high probability in the achene state plus the hidden state (achene+) will appear red. Thus, purple branches indicate uncertainty in the reconstruction. Para=Paracryphiales, Adox=Adoxaceae, Dier=Diervilleae, Capr+*Hepta*=Caprifolieae+*Heptacodium*, Linn=Linnaeeae, *Zab*=*Zabelia*, Morin=Morinaceae, *Triplo*+Dipsac=*Triplostegia*+Dipsacaceae,Valerian=Valerianaceae.

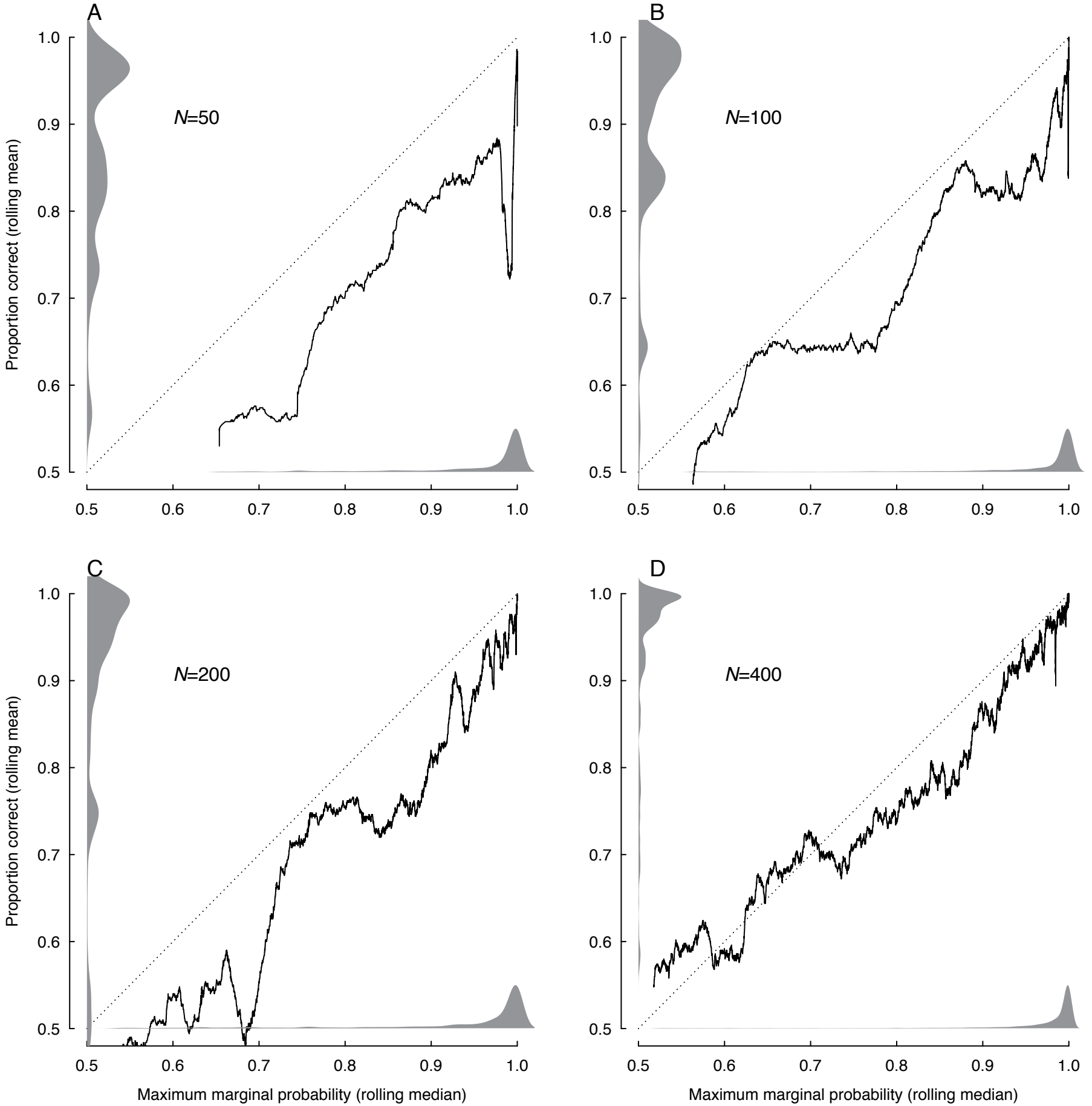

speciation 0
speciation 1A
speciation 1B
q1A0
q1A1B
0
1A
1B
q01A
q1B1A
extinction 0
extinction 1A
extinction 1B
†
†
†

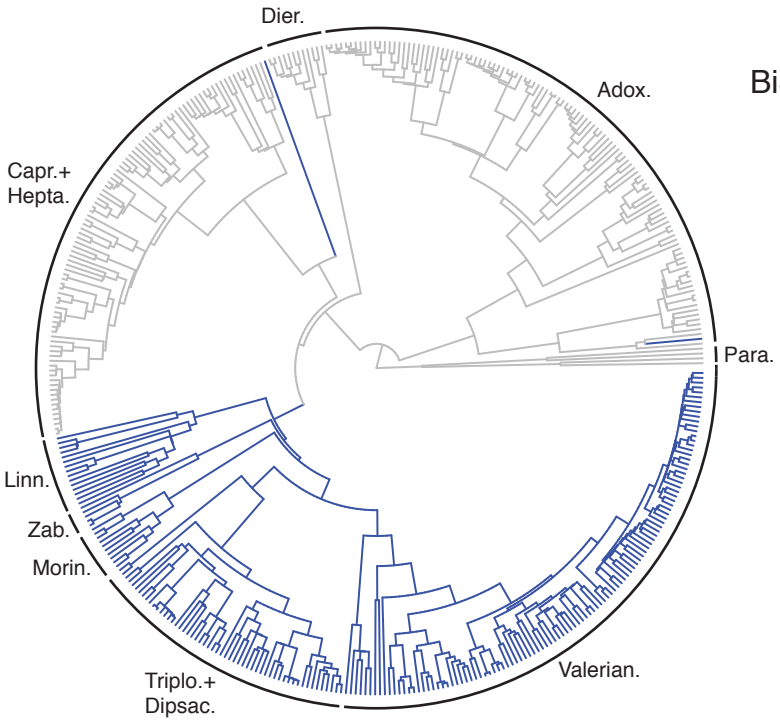

P<0.001
0.25
0.15
0.05
−0.05
Net diversification rate
State 0
State 1



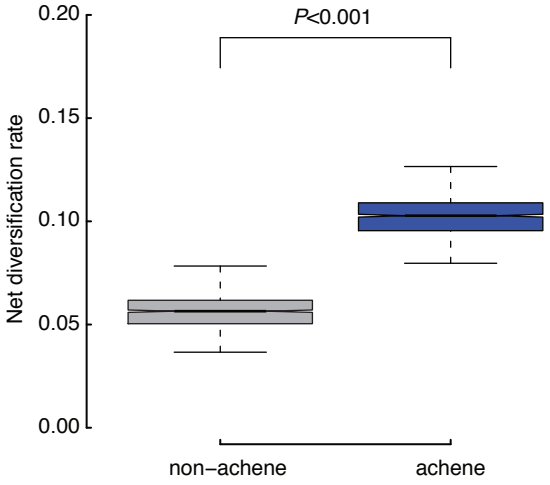

A
HiSSE Scen. 1
Net diversification rate
E
Extinction fraction
I
log Transition rate
B
HiSSE Scen. 2
F
J
C
HiSSE Scen. 3
G
K
D
HiSSE Scen. 5
H
L
Number of species


A
B
C
D
E
F
Speciation rate
Extinction rate
Frequency of state 0
Frequency of state 1A
Frequency of state 1B

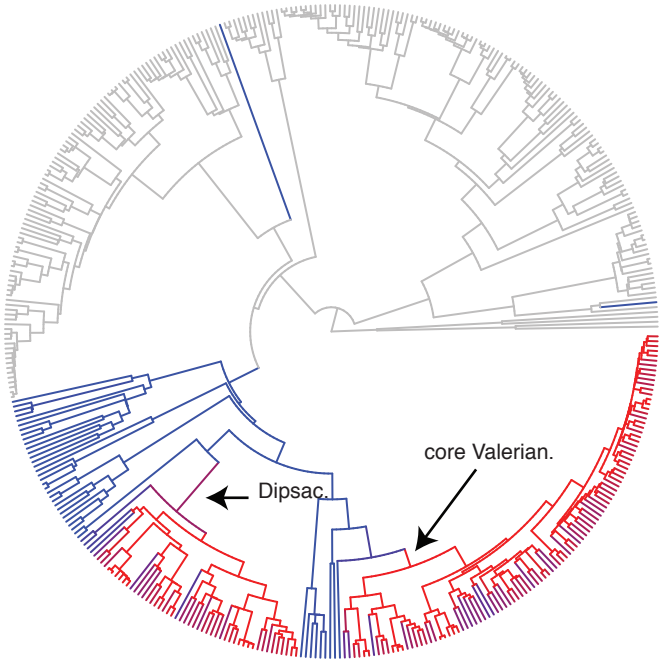

A
B
C
D
N=50
N=100
N=200
N=400
Proportion correct (rolling mean)
Maximum marginal probability (rolling median)
1.0
0.9
0.8
0.7
0.6
0.5

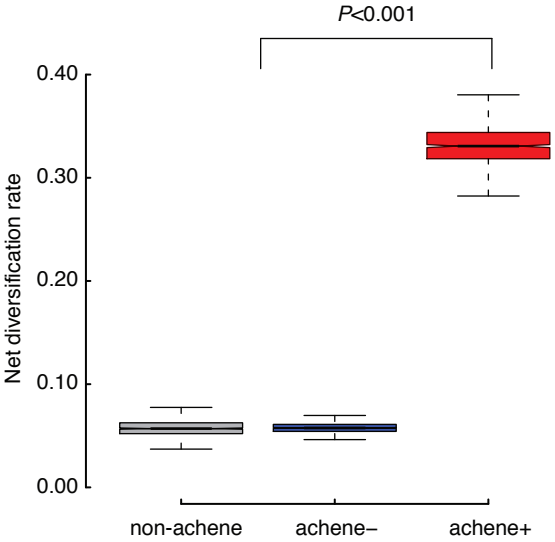

Dier.
Adox.
Capr.+
Hepta.
Para.
Linn.
Zab.
Morin.
Triplo.+
Dipsac.
Valerian.

BiSSE


P<0.001
Net diversification rate
0.20
0.15
0.10
0.05
0.00
non−achene
achene

HiSSE


core Valerian.
Dipsac.


P<0.001
Net diversification rate
0.40
0.30
0.20
0.10
0.00
non-achene
achene−
achene+