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-dependent 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-dependent 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, solving at least one major criticism of BiSSE methods. 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. We also demonstrate how our model can be used as null models that allow for a complex diversification process that is independent of the evolution of a character. 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 and detect when diversification rate differences do not correlate with observed states. We also discuss the remaining issues with state-dependent 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 focus on diversification rate differences (e.g., Goldberg et al. 2010; Wilson et al. 2011; Price et al. 2012; Beaulieu and Donoghue 2013; Weber and Agrawal, 2014).

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 crucial caveats have recently been identified, however. First, Maddison and FitzJohn (2014) raise a statistical 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 and pseudoreplication 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. It was also pointed out by Beaulieu and Donoghue (2013) that even with many 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, Asterales, which contain the highly diverse Asteraceae, the sunflower family. They argue that it is far more likely that some combination of the achene and another unmeasured, co-distributed trait within Asterales led to a higher diversification rate for achenes as a whole (on this point, also see Maddison et al. 2007). Finally, Rabosky and Goldberg (2015) recently showed that for realistically complex data sets, BiSSE methods almost always incorrectly find that a neutral trait is correlated with higher diversification. This is, however, largely a consequence of using fairly "trivial" null models that assume equal rates of diversification between character states, which cannot discern the possibility that a complex diversification process is entirely independent from the focal trait.

All these caveats relate to the broader issue of whether the proximate drivers of diversification are really ever just the focal traits 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 (e.g., a "synnovation" as defined by Donoghue and Sanderson 2015) and/or movements into new geographic regions (e.g., de Querioz 2002; Moore and Donoghue 2007). Recent generalizations of the BiSSE model (i.e., MuSSE: Multistate Speciation and Extinction; FitzJohn 2012) 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 one character state as the *true* underlying cause of increased diversification rates.

Here we describe a potential way forward for trait-dependent models of diversification 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 a state-dependent speciation and extinction form of a Hidden Markov model (HMM). As we will show, HiSSE can adequately distinguish higher net diversification rates nested within clades exhibiting a particular character state, can provide more meaningful tests of character-dependent diversification through its use as a character-independent null model, and thus can provide a much more refined understanding of how particular observed character states may influence the diversification process.

The Hidden State Speciation and Extinction Model

Despite the important methodological advancement and power afforded by the BiSSE model, it can provide a rather coarse-grained view of trait-based patterns of diversification. Specifically, what may appear like a 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 1 have identical diversification rates. However, there is some other trait with states *A* and *B*, and state *B* has twice the diversification rate of *A*. We call this trait a “hidden” trait, because it is not observed in the tip data (though it could be observable if we knew what it was). If state 1 happens to be a prerequisite for evolving state *B* (or even by chance), all the state 0 branches will have state *A*, but some branches in state 1 will have state *A* and some will have state *B*. Thus state 1 actually takes on two states, 1*A* when the hidden state with higher diversification rate is absent, and 1*B* when the hidden state with higher diversification rate 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 the 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, is actually a modified form of the MuSSE model (Multi-State Speciation and Extinction; FitzJohn 2012) that extends BiSSE type analyses to allow for multiple binary characters or characters with more than two states. Thus, the HiSSE model can, in theory, have a number of observed states and a number of hidden states (i.e., observed states 0, 1, 2, and hidden states *A*, *B*, *C*, resulting in nine possible state combinations).

Formally, the state space in our model is defined as *o* being the index of the observed state, $o \in 0,1,...,\alpha$, and *h* as the index of the hidden state, $h \in A,B,...,\beta$. A lineage at time *t* evolves under diversification rates $\lambda_i = \lambda_{oh}$ and $\mu_i = \mu_{oh}$. Thus, the model has, in general, αβ different diversification rate categories. The likelihood $D_{N,i}(t)$ is proportional to the probability that a lineage in state *i* 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 thus described by the following ordinary differential equation:

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

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

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

These series of equations are solved numerically along each edge starting at the tips 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. For example, we would set the probability to 1 for both state 1*A* and 1*B* for all species exhibiting state 1 – in other words, the probability of observing a tip demonstrating state 1 is 1 if the true underlying state is 1*A* or 1*B* (note this could be done with states 1*C*, 1*D*, and so forth in this model, although the current implementation in software currently only allows a binary observed character with four or fewer total hidden state combinations). The initial conditions for $E_i(0)$ are all set to zero (i.e., we observe the tip at the present). Incomplete sampling can be allowed by incorporating a state-specific or even clade-specific sampling frequency, *f*, and setting the initial conditions for $D_{N,i}(0)$ as f_i if the corresponding tip, *n*, is in state *i*, and 0 otherwise, and for $1\text{-}f_i$ 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 (e.g., $f_{1A} = f_{1B}$, $f_{0A} = f_{0B}$).

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 product of $D_{R,i}(t)$ calculated at the root. We condition this

likelihood by $(1\text{-}E_i(t))^2$, which is the probability that the two descendant lineages of the root, evolving under the same estimates of speciation and extinction rates, survived to the present and were sampled (Nee et al. 1994). Finally, we follow the procedure described by FitzJohn et al. (2009) and weight the overall likelihood by the probability that each possible state gave rise to the observed data, although other weights, such as assuming equal probabilities or even "known" fixed probabilities, can also be used. Note that in the absence of a hidden state, our likelihood calculation reduces exactly to the BiSSE model.

We also note that it is rather straightforward to use this framework to implement a SSE version 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 HiSSE 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:

$$\mathbf{Q} = \begin{array}{c} 0A \\ 1A \\ 0B \\ 1B \end{array} \begin{bmatrix} - & q_{0A\to1A} & q_{0A\to0B} & 0 \\ q_{1A\to0A} & - & 0 & q_{1A\to1B} \\ q_{0B\to0A} & 0 & - & q_{0B\to1B} \\ 0 & q_{1B\to1A} & q_{1B\to0B} & - \end{bmatrix} \qquad \text{Eq. (3)}$$

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, including those that allow for dual transitions between both the observed trait and the hidden trait (e.g., $q_{0A} \leftrightarrow q_{1B}$).

TYPE I ERRORS AND THE ISSUE OF APPROPRIATE NULL MODELS

An important issue was recently 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 [a similar issue of MuSSE falsely assigning diversification rate effects to a character with no such effect was discussed by FitzJohn (2012)]. While this behavior is seemingly troubling, it is important to bear in mind that BiSSE, HiSSE, and any other model of state-dependent 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 the alternative model fits better. This will often be the case 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, among several other empirical examples, that data sets 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 (see Slater et al. 2010; Morlon et al. 2011). 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.

The broader point here is that in comparative biology we often compare trivial nulls. For instance, we often ask questions like, "is there a single rate of evolution over the whole tree?" The answer is often an unsurprising "no", because for most complex real data a comparison between a simple model and one with even slightly more complexity will very likely yield strong support for the latter. In our view, a fairer comparison would involve some sort of null model that contains the same degree of complexity in terms of numbers of parameters to allow for comparisons among any complex models of interest. Such tests should avoid falling into the trap of choosing a model purely on the basis of it having more parameters, and, in theory, should solve the issue raised by Rabosky and Goldberg (2015).

Here we propose two null models that are devised to equal the complexity with respect to the number of parameters for the diversification process [i.e., same number of free speciation and extinction rates under the Maddison et al. (2007) parameterization] as a general BiSSE or HiSSE model, but without actually linking them to the observed traits. Thus, these null models explicitly assume that the evolution of a binary character is

independent of the diversification process without forcing the diversification process to be constant across the entire tree, which is the normal null used in these types of analyses. The first model, which we refer to as "null-two", contains four diversification process parameters that account for trait-dependent diversification solely on the two states of an unobserved, hidden trait (e.g., for speciation rates, $\lambda_{0A}= \lambda_{1A}$, $\lambda_{0B}= \lambda_{1B}$). In this way, null-two contains the same amount of complexity in terms of diversification as a BiSSE model. However, the possible transition rates under this model take the form described in Eq. (3) to account for the possible transitions among all state combinations. These transition rates can be general by allowing them all to be freely estimated, or simplified in various ways such as assuming they are all equal. The second model, which we refer to as "null-four" contains the same number of diversification parameters as in the general HiSSE model that are linked across four hidden states (e.g., for speciation rates, $\lambda_{0A}= \lambda_{1A}$, $\lambda_{0B}= \lambda_{1B}$, $\lambda_{0C}=\lambda_{1C}$, $\lambda_{0D}= \lambda_{1D}$). The transition rates under this model are set up to account for transitions between the four different rate categories, as well as between the states of the binary character and, in matrix form, they are set up as follows:

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

Eq. (4)

To simplify the number of transitions in the model there are two natural assumptions: assume either all transition rates are equal, or assume there are three distinct transition rates: one rate describing transitions among the different hidden states (i.e., the rates in columns and rows 1-4, and columns and rows 5-8), and two rates for transitions between the observed character states (i.e., one rate for columns 5-8, rows 1-4, and one rate for columns 1-4, and rows 5-8).

IMPLEMENTATION

We implemented the above models 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 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 over-fitting 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) (their Figure S2) 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 common 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 configurations 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

$$\lambda_i = \tau_i / (1 + \varepsilon_i), \qquad \text{Eq. (5a)}$$

$$\mu_i = (\tau_i \, \varepsilon_i) / (1 + \varepsilon_i), \qquad \text{Eq. (5b)}$$

which can be used to derive other measures of diversification process parameters, such as net diversification (λ_i-μ_i). One of the *diversitree* authors (Sally Otto, personal communication) noted that Hardy and Otto (2014) tested a hidden model that only affected transition rates within the MuSSE framework. Note that we have not directly compared the two implementations, though we have confirmed that MuSSE and HiSSE return the same likelihood when all states are observed.

The *hisse* package assigns the probability of each ancestral state to an internal node using marginal ancestral state reconstruction (Yang et al.1995; Schluter et al., 1997). 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 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 the HiSSE model by simulating trees and character states under various scenarios and then estimating the fit and bias of the inferred rates from these trees. Our initial simulations first tested scenarios of the single hidden state situation described in Figure 1; the known parameters for each scenario are described in detail in Table 1. We included BiSSE scenarios to test whether we could correctly conclude that there was no support for a HiSSE model in the absence of a hidden state in the generating model. For each of these scenarios, trees and trait data were simulated using *diversitree* (FitzJohn 2012) 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 Akaike 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 Akaike weight across our model set,

as the average Akaike weight assuming an equal likelihood across all models. Thus, our null expectation is based solely on the penalty term in the AIC calculation – in the absence of information from the model, we would expect to see these weights returned, rather than equal weights for all models. Moreover, since BiSSE is nested within HiSSE, they could return the same likelihood, so even with infinite amounts of data the weight of the HiSSE model should drop to this null expectation, but, counterintuitively, not to zero.

We also conducted a second set of simulations that specifically evaluated the performance of the general HiSSE model and our two null models of character-independent diversification. Our goal was to test our ability to accept or properly rule out the trait-independent models under scenarios of trait-dependent diversification. We were concerned at the outset that our null models would always fit at least as well, if not better, than a trait-dependent BiSSE or HiSSE model. That is to say, these models were constructed such that they are not constrained by character states and can assign rates wherever they want in order to maximize the likelihood. Table 2 describes the known parameters for the three scenarios evaluated. Again we relied on *diversitree* to simulate trees and trait data that contained 400 species, repeated 100 times, with transitions allowed between four possible states, 0, 1, 2, or 3. 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 0, and all tips in state 3 to be in state 1.

We also included a scenario that was designed to test whether the general HiSSE model is immune to empirical issues of spurious assignment of importance to state combinations that have no actual effect on diversification. In other words, is HiSSE still favored in situations where the trees and traits evolved under a very different model than

the one used for inference? Here we generated trees containing 400 taxa using code from Rabosky (2010). A symmetric Markov model for trait evolution alone (no influence by diversification) was used to simulate binary traits on this tree (using the R package *geiger*; Harmon et al. 2008; Pennell et al. 2014). The Rabosky (2010) model was used, as it is very different from the model assumed by BiSSE/HiSSE; speciation rates evolve gradually on branches, rather than moving discretely between distinct levels based on a trait (hidden or not). Though the rate change is gradual under the Rabosky (2010) model, the speciation rate does not evolve under a Brownian, Ornstein-Uhlenbeck, or similar processes, but in a heterogeneous way that depends on the timing of speciation events (Beaulieu and O'Meara 2015). Thus, as with many empirical data sets, this diversification model is quite different from HiSSE's model, providing a difficult challenge. The Rabosky (2010) model has also been very influential in affecting biologists' attitudes towards estimating extinction rates, and so we include it as a semi-realistic "worst-case" scenario. In this scenario, and the three others described above, each simulated data set was evaluated under the generating model, as well as ten additional models, that variously added, removed, or constrained certain parameters, with model fit again being assessed by calculating the average Akaike weight (w_i). The entire model set is described in Table 2. For simplicity, all models in this set assumed equal transition rates.

In regards to the parameter estimates in all our simulations, comparisons between 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 Akaike weights to reflect the truncated model set. Finally, we also 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.

ASSESSING MODEL FIT

From a model comparison perspective, our first set of simulations indicated that data sets that lack hidden states could generally be distinguished from those that do, especially with larger data sets (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 Akaike weight of the HiSSE models converged towards the null Akaike 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 data set, but also the underlying generating model. For example, when the generating model assumed a hidden state with higher speciation rates, data sets that contained 200 or more taxa were required to provide strong evidence for a HiSSE model that varied the turnover rate (Table 1). However, when the main effect of the generating model assumed lower extinction fractions for the hidden state there remained strong support for a BiSSE model that assumed both equal turnover rates and extinction fractions. Interestingly, when we simulated under a HiSSE scenario that combines the processes of higher speciation rates

and asymmetrical transition rates, the issue of incorrectly favoring a BiSSE model disappears (Table 1).

We also found that HiSSE was able to distinguish between generating models that assumed diversification rates differences are trait-dependent (e.g., when speciation is two-fold greater for state 1*B*), or when the diversification rate difference is trait-independent and due to simply to the presence of a hidden trait only (e.g., when speciation is two-fold greater for hidden state *B*) (Table 2; Table S1). And, contrary to our concerns, when the generating model assumed some form of trait-dependent diversification neither the null-two nor the null-four model had much support.

For the "worst-case" scenario, which simulated a neutral binary character along trees generated from a complex heterogeneous rate branching process, it clearly falls into the zone where BiSSE has issues (94% of data sets favored a BiSSE model over a single-rate diversification model, and the average Akaike weight for state-dependent models across data sets was 92.0%; see Table S1). Thus, these simulation conditions create difficult data sets of the sort used by Rabosky and Goldberg (2015). The addition of merely the null-two model dramatically improved performance, where the average weight for the BiSSE models went from 92.0% to 1.3%, and all but one data set chose the null-two model as best (see Table S1). When the full HiSSE model set was examined, we found that the null-four trait independent model had the highest support on average across the entire set of models, and was the best overall in 49% of the simulations (Table 2; Table S1). However, if we combine support both this null model and the null-two model, they cumulatively account for 70% of the total model weight. Both of these models are independent of any particular character state, but assume shifts among distinct

levels, and therefore more closely resemble the conditions by which these data sets were generated.

We note, however, that there was some support for a trait-dependent HiSSE model with the same number of diversification parameters (Table 2; Table S1). And, in spite of the Akaike weight across all data sets showing greater support for the null-four model on average, 29% of the individual data sets would favor some form of the general HiSSE model (Table S1). However, there was only a 1.02% difference in the mean diversification rate between states 0 and 1 (though, focusing on just the set of trees for which a HiSSE model was best, the percent difference is 16% – still relatively small for diversification studies, but clearly not as good). Taken together, these results indicate that in these "worst-case" trait-independent scenarios, likely encountered in many empirical data sets, the inclusion of the null models can attenuate the issue of spuriously accepting trait-dependent diversification. The addition of models that allow some dependence with the observed traits and some with hidden traits, the rates of falsely assessing some state-dependence somewhat increases, but with the benefit of being able to better detect hidden effects when they are true (see earlier simulation results).

ESTIMATING MODEL PARAMETERS

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 or ε to vary, but not both. This is likely a consequence of the uncertainty and upward biases in estimates of extinction fraction (Fig. 2-4), 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 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. 2-4; 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. 2c).

For parameter estimates from the simulations testing both the null models and the general HiSSE model we found very similar results as those described above, in that the model-averaged net turnover and net diversification parameters largely resembled the generating model (Fig. 4, S4). However, similar to the first set of simulations, there is a

large amount of uncertainty and upward biases in estimates of extinction fraction. Thus, HiSSE generally has very low power in distinguishing multiple extinction fractions regardless of the number of states included in the model.

One issue of concern from our first set of simulations is that transition rates are almost always overestimated. This behavior appears unique to the HiSSE model in our simulations (Fig. 2), given that when evaluating data sets under BiSSE scenarios, transition rates are estimated reasonably well (Fig. S3). One suggestion from Rich FitzJohn (personal communication) is that this can occur when some states are present in low frequency, and since HiSSE has more states than BiSSE, it is likely that many state combinations are in very low frequencies. There are also 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 the HiSSE scenario that assumed pronounced differences in the transition 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 (see Supplemental Materials). 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 difficult problem, which is not made any easier by HiSSE's increase in state space. That is to say, 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 from the model-averaged parameter estimates, the simulations indicate that HiSSE correctly identifies and locates regions of the tree where supposed diversification rate differences have taken place (Fig. 5). The degree of reliability does, of course, depend on the size of the data set. For the HiSSE scenario that assumed a doubled speciation rate for state 1*B*, 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 there is also a general tendency for HiSSE to infer high marginal probabilities for the incorrect state (e.g., Fig. 5), 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 Asteraceae – sunflowers and their relatives – or the single-seeded

“mericarps” of 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 Asterales clade (Beaulieu and Donoghue 2013). Within Apiales and Dipsacales, the two remaining major achene-bearing clades, the diversification rate differences were not 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; Tank and Donoghue 2010) 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 three important ways. First, we re-estimated the original molecular branch lengths using PAUP* (Swofford 2000), as opposed to relying on the branch lengths from the original RAxML (Stamatakis 2006) analysis, because PAUP* provides better optimization precision. Second, the molecular branch lengths were re-scaled 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 that best predicted the rates of terminal branches pruned from the tree. Third, 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., <1.0 Myr) – all of which have the tendency to negatively impact accuracy of estimating diversification rates (though in this case, rerunning the analyses including such tips did not have a qualitative effect on the results). The exclusion of these taxa resulted in a data set comprised of 417 species from the original 457.

We fit 24 different models to the achene data set for Dipsidae. Four of these models corresponded to BiSSE models that either removed or constrained particular parameters, 16 corresponded to various HiSSE models that assumed a hidden state associated with both the observed states (i.e., non-achene+, achene+), and four corresponded to various forms of our trait-independent null models (i.e., null-two, null-four). 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 Table S2; note, that the results are qualitatively similar regardless of sampling scheme used). In order to generate a measure of confidence for 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), which follows from Edwards (1992). We also took into account both state uncertainty and uncertainty in the model when "painting" diversification rates across the tree. Our procedure first calculated a weighted average of the likeliest state and rate combination for every node and tip for each model in the set, using the marginal probability as the weights, which were then averaged across all models using the Akaike weights. All analyses were carried out in *hisse*.

The best model, based on Akaike weights, was a relatively simple HiSSE model with regards to the number of free parameters it contained (Table 3). The model suggests character-dependent diversification with fruit type, where only rates for non-achene+ and achene+ are allowed to be free, and where transitions between state 0*B* and 1*B* were disallowed. This model had a pronounced improvement over the set of BiSSE models, where none had an Akaike weight that was greater than 0.001. However, before describing the parameter estimates of the best model, we note that, with a modified tree and character set, the parameter estimates under the BiSSE models were different from those reported by Beaulieu and Donoghue (2013). Specifically, the higher diversification rates estimated for achene lineages (r_{achene}=0.148, support range: [0.121,0.161]) compared to non-achene lineages ($r_{\text{non-achene}}$=0.065, support range: [0.0498,0.083]) were indeed significant based on a sampling of points falling within 2lnL units away from the MLE. 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 ($r_{\text{achene+}}$=0.199,

support region: [0.179,0.221]). In fact, the model suggests that achene lineages not associated with the high diversification hidden state have a diversification rate that is indistinguishable from the non-achene diversification rate ($r_{\text{non-achene-}}= r_{\text{achene-}}$=0.059, support range: [0.049,0.068]). Non-achene lineages associated with the high diversification hidden state also show elevated diversification rates ($r_{\text{non-achene+}}$=0.158, support region: [0.138,0.185]), relative to the rate of the non-achene state with the other hidden state, suggesting strong rate heterogeneity even in Dipsidae lineages that bear other fruit types.

Character reconstructions identified two transitions to the hidden fast state in achene lineages, 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*)(Fig. 6). It is important to note that in the case of non-achene lineages, the model identified four shifts to the hidden fast state – one at the base of Caprifolieae (*Lonicera*, *Symphoricarpus*, *Leycesteria*, and *Triosteum*), and three within *Viburnum* (Fig. 6). The several shifts detected in *Viburnum* are noteworthy in that they correspond almost exactly with the inferred shifts detected in a more focused study of the genus that applied various character-independent models of diversification (Spriggs et al. 2015).

In regards to achene lineages, the general location of the inferred shifts is intriguing, as they appear to coincide with specific clades that exhibit specialized structures related to achene dispersal. In 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 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 important biogeographic movements due to increased dispersal abilities (i.e., “key opportunities” *sensu* 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 Dipsidae.

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), or, less compellingly, clades we identify for

such a test from eyeballing the data return a significant result despite no underlying reality.

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.

Currently, diversification models are divided between those that look at one or more focal traits only, integrating over any other factors (e.g., BiSSE, Maddison et al. 2007; BiSSE-ness, Magnuson-Ford and Otto, 2012; ClaSSE, Goldberg and Igic, 2012; sister group comparisons, Mitter et al. 1988), and those that fit rates to trees but ignore trait information altogether (e.g., MEDUSA, Alfaro et al. 2009; BAMM, Rabosky 2014). Our HiSSE framework spans this range. If the true model is strictly trait-dependent diversification, it can detect this (Fig. 1-5), as a BiSSE analysis would. If the true model is rates varying due to some other unexamined factor (a physical trait, a property of the environment, etc.) HiSSE recovers this too, as in the case of achene evolution in the Dipsidae clade (Fig. 6). Uniquely, HiSSE can also give an intermediate answer where a focal trait explains some but not all of the diversification difference.

The HiSSE framework also provides a solution to some of the recent important criticisms levied against state speciation and extinction models (Rabosky and Goldberg 2015). Specifically, our method no longer requires the assumption that focal character states are associated with diversification rate differences. 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; Weber and Agrawal 2014; Spriggs et al. 2015). Including our null models of trait-independent diversification in a set of alternative trait-dependent models should also alleviate concerns of spurious assignments of diversification rate differences between observed character states in cases where trees are evolving separately from the focal trait (c.f., Rabosky and Goldberg 2015). These character-independent diversification models are designed specifically to be as complex as competing BiSSE or HiSSE models, which provides a fairer comparison over more trivial "equal-rate" nulls.

We do, however, highlight one important statistical concern in which HiSSE requires significant caution in its use. This involves its indifference to number of changes (c.f., 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.

There are other practical concerns in empirical applications of the HiSSE model. While it could be used over a set of trees (i.e., bootstrap or Bayesian post-burnin 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, 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. For trees that come from Bayesian dating analyses, whether a birth-death or Yule prior was used may affect results unless the data is strong enough to overwhelm the prior (Condamine et al. 2015). Future simulation studies will focus on better understanding the impact that these and other branch length error scenarios can have on the interpretation and estimation of various model parameters. Also, we point out that 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.

Our HiSSE model is part of a long tradition in comparative methods of identifying, and then fixing, perceived shortcomings. For instance, Felsenstein (1985) pointed out the pitfalls of treating species values as independent data points and provided two alternatives (i.e., independent contrasts and dividing a tree into pairs of tips) to address those issues. Maddison (2006) realized the issue of not accounting for

diversification in transition-based methods (i.e., Pagel 1994) and not accounting for differential transitions in diversification models, which led to the development of state-dependent speciation and extinction models (Maddison et al. 2007; FitzJohn et al. 2009). Recently, Rabosky and Goldberg (2015) pointed out a serious problem with interpretations in state speciation and extinction analyses in general, and our work largely addresses these concerns. We caution users that it is important to keep a perspective: all methods have flaws, and all will fail given a strong enough violation of their underlying assumptions. For example, it appears difficult for HiSSE to adequately estimate different transition rates when the model assumes any number of hidden states, and so many estimates will be biased. Even with the increasing efforts to test new methods (in our case, over 17,000 computer-days were devoted to conducting simulations and analyses) there will be flaws that may have gone undetected. We urge skepticism towards all models, but also skepticism towards statements of fatal flaws in some models while leaving newer, competing models relatively untested.

There is no question that state-dependent 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, but can instead estimate how important it is as a component of diversification overall. 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, while still allowing us to examine our hypotheses about effects of observed traits.

Supplemental Material

Data available from the Dryad Digital Repository: http://dx.doi.org/10.5061/dryad.52hp1

Funding

This work was supported by the National Institute for Mathematical and Biological Synthesis [to J.M.B.], 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.

Acknowledgements

We thank Nathan Jackson, Nick Matzke, Kathryn Massana, Chuck Bell, and Andrew Leslie for helpful discussions. Rich FitzJohn, Sally Otto, and two anonymous reviewers provided very useful comments on the manuscript, and we are deeply indebted to Tanja

Stadler for her suggestion to consider null models. 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.

References

Alfaro M.E., Santini F., Brock C., Alamillo H., Dornbug A., Rabosky D.L., Carnevale G., Harmon L.J. 2009. Nine exceptional radiations plus high turnover explain species diversity in jawed vertebrates. Proceedings of the National Academy of Sciences of the U.S.A. 106:13410-13414.

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 69:1036-1043.

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.

Condamine F.L., Nagalingum N.S., Marshall C.R., Morlon H. 2015. Origin and diversification of living cycads: a cautionary tale on the impact of the branching process prior in Bayesian molecular dating. BMC Evolutionary Biology 15:65.

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.

Donoghue M.J., Sanderson M.J. 2015. Confluence, synnovation, and depauperons in plant diversification. New Phytologist doi:10.1111/nph.13367.

Edwards A.W.F. 1992. Likelihood. Cambridge:Cambridge University Press.

Felsenstein J. 1985. Phylogenies and the comparative method. The American Naturalist 125:1-15.

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.

Goldberg E. E., Igic B. 2012. Tempo and mode in plant breeding system evolution. Evolution 66:3701-3709.

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.

Hardy N.B. and Otto S.P. 2014. Specialization and generalization in the diversification of phytophagous insects: tests of the musical chairs and oscillation hypotheses. Proceedings of the Royal Society B. 281: 20132960.

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

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. 2006. Confounding asymmetries in evolutionary diversification and character change. Evolution 60:1743–1746.

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.

Magnuson-Ford K., Otto S.P. Linking the investigations of character evolution and species diversification. American Naturalist 180:225-245.

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.

Mitter C., Farrell B., Wiegmann B. 1988. The phylogenetic study of adaptive zones: Has phytophagy promoted insect diversification? American Naturalist 132:107-128.

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.

Morlon H., Parsons T.L., Plotkin J.B. 2011. Reconciling molecular phylogenies with the fossil record. Proceedings of the National Academy of Sciences of the U.S.A. 108:16327-16332.

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.

Pennell M.W., Eastman J.M., Slater G.J., Brown J.W., Uyeda J.C., FitzJohn R.G., Alfaro M.E., Harmon L.J. 2014. Geiger v2.0: an expanded suite of methods for fitting

macroevolutionary models to phylogenetic trees. Bioinformatics doi: 10.1093/bioinformatics/btu181

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

Rabosky D.L. 2014. Automatic detection of key innovations, rate shifts, and diversity-dependence on phylogenetic trees. PLoS One doi: 10.1371/journal.pone.0089543

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.

Schluter D., Price T., Mooers A.O., Ludwig D. 1997. Likelihood of ancestor states in adaptive radiation. Evolution 51:1699-1711.

Slater G.J., Price S.A., Santini F., Alfaro M.E. 2010. Diversity versus disparity and the radiation of modern cetaceans. Proceedings of the Royal Society, B. 277:3097-3104.

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.

Spriggs E.L., Clement W.L., Sweeney P.W., Madriñán, Edwards E.J., Donoghue M.J. 2015. Temperate radiations and dying embersof a tropical past: the diversification of Viburnum. New Phytologist. In press, doi:10.1111/nph.13305.

StamatakisA. 2006. Maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics 22:2688–2690.

Swofford D.L. 2000. PAUP*: Phylogenetic analysis using parsimony (and other methods) version 4. Sinauer Associates, Sunderland, Massachusetts.

Tank D.C., Donoghue M.J. 2010. Phylogeny and phylogenetic nomenclature of the Campanulidae based on an expanded sample of genes and taxa. Systematic Botany 35:425-441.

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.

Yang Z., Kumar S., Nei M. 1995. A new method of inference of ancestral nucleotide and amino acid sequences. Genetics 141:1641–1650.

TABLE 1. Summary of model support for simulated scenarios with and without a hidden trait, where calculating the average Akaike weight (w_i) for all models assessed the fit. We also calculated a null expectation of the Akaike weight as the average Akaike 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 w_i expectation		0.0260	0.0708	0.1923	0.0708	0.1923	0.0707	0.1923	0.0004	0.0096	0.0708	0.0035	0.0708	0.0035	0.0260
BiSSE, transition rates vary	50	0.0416	0.0821	***0.1504***	0.0864	***0.1639***	0.0925	***0.1737***	0.0022	0.0179	0.0739	0.0076	0.0725	0.0010	0.0257
λ0=λ1=0.1	100	0.0464	0.0837	0.1322	0.0791	0.1319	0.0930	***0.1641***	0.0046	0.0289	0.0944	0.0123	0.0811	0.0185	0.0298
μ0=μ1=0.03	200	0.0499	0.0623	0.0927	0.0969	0.0982	0.0989	***0.1812***	0.0069	0.0374	0.0919	0.0198	0.0874	0.0313	0.0450
q01=0.015, q10=0.005	400	0.0662	0.0635	0.0578	0.1147	0.0668	0.1119	***0.2116***	0.0070	0.0320	0.0771	0.0227	0.0801	0.0284	0.0599
BiSSE, state 1 2x speciation	50	0.0370	0.0848	***0.1896***	0.0803	***0.1681***	0.0930	0.1339	0.0117	0.0160	0.0781	0.0075	0.0685	0.0113	0.0203
λ0=0.1, λ1=0.2	100	0.0450	0.0918	***0.1920***	0.0783	0.1545	0.1066	0.0997	0.0030	0.0218	0.0933	0.0089	0.0706	0.0182	0.0163
μ0=μ1=0.03	200	0.0508	0.1006	***0.2129***	0.0820	0.1544	0.1194	0.0508	0.0069	0.0226	0.0837	0.0113	0.0643	0.0314	0.0086
q01=q10=0.01	400	0.0690	0.1172	***0.2028***	0.0911	0.1503	0.1324	0.0186	0.0091	0.0268	0.0637	0.0161	0.0615	0.0356	0.0057
BiSSE, state 1 .5x extinction	50	0.0480	0.0866	***0.1465***	0.0866	***0.1451***	0.0958	***0.1714***	0.0040	0.0219	0.0769	0.0092	0.0684	0.0132	0.0264
λ0=λ1=0.1	100	0.0502	0.0934	***0.1332***	0.0879	***0.1545***	0.0874	***0.1576***	0.0039	0.0214	0.0784	0.0128	0.0779	0.0147	0.0267
μ0=0.06, μ1=0.03	200	0.0596	0.1015	0.1219	0.0904	***0.1652***	0.0826	0.1324	0.0048	0.0286	0.0717	0.0101	0.0824	0.0144	0.0344
q01=q10=0.01	400	0.0710	0.1145	0.1008	0.0996	***0.1605***	0.0825	0.1113	0.0082	0.0325	0.0618	0.0156	0.0983	0.0163	0.0270
HiSSE, state 1B 2x speciation	50	0.0486	0.0808	***0.1573***	0.0936	***0.1488***	0.1069	***0.1520***	0.0030	0.0200	0.0768	0.0089	0.0621	0.0128	0.0285
λ0A=λ1A=0.01, λ1B=0.2	100	0.0533	0.0786	***0.1264***	0.0777	***0.1147***	0.0931	0.1074	0.0095	0.0386	***0.1438***	0.0146	0.0957	0.0248	0.0216
μ0A=μ1A=μ1B=0.03	200	0.0666	0.0553	0.0719	0.0726	0.0742	0.0791	0.0743	0.0134	0.0611	***0.2158***	0.0225	0.1263	0.0469	0.0202
q0A1A=q1A0A=q1A1B=q1B1A=0.01	400	0.0321	0.0275	0.0359	0.0254	0.0283	0.0370	0.0236	0.0315	0.1094	***0.3119***	0.0420	0.1876	0.1024	0.0053
HiSSE, state 1B 3x speciation	50	0.0524	0.0796	***0.1594***	0.0821	0.1218	0.1124	0.1003	0.0091	0.0364	***0.1295***	0.0156	0.0544	0.0317	0.0153
λ0A=λ1A=0.01, λ1B=0.3	100	0.0504	0.0604	0.0998	0.0629	0.0817	0.0893	0.0686	0.0112	0.0660	***0.2250***	0.0156	0.1048	0.0464	0.0177
μ0A=μ1A=μ1B=0.03	200	0.0414	0.0316	0.0405	0.0304	0.0397	0.0556	0.0146	0.0241	0.1101	***0.4102***	0.0215	0.0906	0.0871	0.0026
q0A1A=q1A0A=q1A1B=q1B1A=0.01	400	0.0220	0.0140	0.0112	0.0140	0.0090	0.0437	0.0078	0.0347	0.1654	***0.4646***	0.0175	0.0703	0.1244	0.0015
HiSSE, state 1B .5x extinction	50	0.0487	0.0911	***0.1453***	0.0918	***0.1472***	0.1019	***0.1711***	0.0023	0.0230	0.0694	0.0074	0.0667	0.0096	0.0246
λ0A=λ1A=λ1B=0.1	100	0.0518	0.0745	0.1251	0.0939	0.1236	0.0968	***0.1743***	0.0068	0.0286	0.0697	0.0147	0.0728	0.0223	0.0450
μ0A=μ1A=0.06, μ1B=0.03	200	0.0538	0.0648	0.1002	0.0951	***0.1128***	0.0918	***0.1402***	0.0083	0.0435	0.0877	0.0207	0.1032	0.0347	0.0433
q0A1A=q1A0A=q1A1B=q1B1A=0.01	400	0.0591	0.0461	0.0464	0.0986	0.0660	0.1018	***0.1200***	0.0235	0.0604	0.0920	0.0441	***0.1523***	0.0352	0.0546
HiSSE, state 1B 2x speciation, transition rates vary	50	0.0432	0.0743	***0.1452***	0.0834	***0.1338***	0.0981	***0.1403***	0.0070	0.0391	0.1095	0.0156	0.0682	0.0212	0.0209
λ0A=λ1A=0.01, λ1B=0.3	100	0.0379	0.0518	0.0737	0.0594	0.0672	0.0778	0.0835	0.0181	0.0701	***0.2457***	0.0222	0.0823	0.0886	0.0216
μ0A=μ1A=μ1B=0.03	200	0.0399	0.0184	0.0190	0.0399	0.0202	0.0500	0.0535	0.0310	0.1267	***0.3259***	0.0229	0.1090	0.1323	0.0111
q0A1A=0.015,q1A0A=q1A1B=0.005,	400	0.0098	0.0027	0.0048	0.0108	0.0067	0.0233	0.0118	0.0883	0.1706	***0.3261***	0.0291	0.0596	***0.2545***	0.0019
q1B1A=0.01	800	0.0020	0.0002	0.0004	0.0022	0.0004	0.0021	0.0038	0.1040	0.1647	***0.3156***	0.0230	0.0638	***0.3172***	0.0005

Bold italics indicate significance based on 95% confidence intervals

TABLE 2. Summary of the model support for simulated scenarios that tested the performance of the general HiSSE model and two null models (null-two, null-four; see text). All data sets tested contained 400 species, and calculating the average Akaike weight (w_i) for all models assessed the fit. As with Table 1, we calculated a null expectation of the Akaike weight as the average Akaike weight if we assumed an equal likelihood across all models.

	Equal rates	BiSSE		Null-two		Null-four		HiSSE			
	τ0=τ1			τA, τB	τA, τB	τA, τB, τC, τD	τA, τB, τC, τD			τ0A=τ1A=τ0B	τ0A=τ1A=τ0B
Generating Model	ε0=ε1	Free	ε0=ε1	εA, εB	εA=εB	εA, εB, εC, εD	εA=εB=εC=εD	Free	ε0A=ε1A=ε0B=ε1B	ε0A=ε1A=ε0B	ε0A=ε1A=ε0B=ε1B
Null w_i expectation	0.1426	0.0865	0.1426	0.0865	0.1426	0.0193	0.0865	0.0117	0.0525	0.0865	0.1426
BiSSE, state 1 2x speciation λ0=0.1, λ1=0.2 μ0=μ1=0.03 All q's = 0.01	0.0085	0.1995	***0.3313***	0.0046	0.0089	0.0001	0.0012	0.0255	0.0917	0.1298	0.1989
Null-two, state B 2x speciation λ0A=λ1A=0.1, λ0B=λ1B=0.2 μ0A=μ1A=μ0B=μ1B=0.03 All q's = 0.01	0.0505	0.0445	0.0568	0.1836	***0.3107***	0.0063	0.0656	0.0322	0.1461	0.0407	0.0631
HiSSE, state 1B 2x speciation λ0A=λ1A=λ0B=0.1, λ1B=0.2 μ0A=μ1A=μ0B=μ1B=0.03 All q's = 0.01	0.0049	0.0957	0.1327	0.0073	0.0135	0.0002	0.0032	0.0417	0.1199	0.2265	***0.3544***
"Worst-case" trait-independence Initial λ=0.1, drawn from lognormal ε =0.3, q01=q10=0.01	0.0005	0.0003	0.0007	0.0897	0.1562	0.0544	***0.3996***	0.0563	0.2213	0.0067	0.0142

Bold italics indicate significance based on 95% confidence intervals

TABLE 3. The fit of alternative models of achene fruit evolution in the flowering plant clade Dipsidae. The best model, based on ΔAIC and Akaike weights (w_i) is denoted in bold.

Model	np	lnLik	AIC	ΔAIC	w*i*
BiSSE: All free	6	-1486.5	2984.9	90.3	<0.001
BiSSE: ε0=ε1	5	-1488.6	2987.2	92.6	<0.001
BiSSE: q's equal	5	-1487.3	2984.5	89.9	<0.001
BiSSE: ε0=ε1, q's equal	4	-1490.2	2988.4	93.8	<0.001
Null-two: q's equal	5	-1449.1	2908.3	13.7	0.001
Null-two: ε's, q's equal	4	-1449.7	2907.4	12.8	0.001
Null-four: q's equal	9	-1445.1	2908.1	13.5	0.001
Null-four: ε's equal, q's equal	6	-1445.3	2902.6	7.98	0.015
HiSSE: q's equal	9	-1446.6	2911.2	16.6	<0.001
HiSSE: ε's equal, q's equal	6	-1447.2	2906.5	11.9	0.002
HiSSE: τ0A=τ1A=τ0B, ε0A=ε1A=ε0B, q's equal	5	-1463.3	2936.6	42.0	<0.001
HiSSE: τ0A=τ1A=τ0B, ε's equal, q's equal	4	-1471.0	2950.0	55.4	<0.001
HiSSE: q0B1B=0, q1B0B=0, all other q's equal6	9	-1451.7	2921.5	26.9	<0.001
HiSSE: ε's equal, q0B1B=0, q1B0B=0, all other q's equal	6	-1451.7	2915.5	20.9	<0.001
HiSSE: τ0A=τ1A=τ0B, ε0A=ε1A=ε0B, q0B1B=0, q1B0B=0, all other q's equal	5	-1462.2	2934.3	39.7	<0.001
HiSSE: τ0A=τ1A=τ0B, ε's equal, q0B1B=0, q1B0B=0, all other q's equal	4	-1469.5	2947.1	52.5	<0.001
HiSSE: τ0A=τ0B, ε0A=ε0B, q's equal	7	-1459.6	2933.2	38.6	<0.001
HiSSE: τ0A=τ0B, ε's equal, q's equal	5	-1466.4	2942.8	48.2	<0.001
HiSSE: τ0A=τ0B, ε0A=ε0B, q0B1B=0, q1B0B=0, all other q's equal	7	-1458.4	2930.8	36.2	<0.001
HiSSE: τ0A=τ0B, ε's equal, q0B1B=0, q1B0B=0, all other q's equal	5	-1464.7	2939.4	44.8	<0.001
HiSSE: τ0A=τ1A, ε0A=ε1A, q's equal	7	-1446.7	2907.3	12.7	<0.001
HiSSE: τ0A=τ1A, ε's equal, q's equal	5	-1469.1	2948.3	53.7	<0.001
HiSSE: τ0A=τ1A, ε0A=ε1A, q0B1B=0, q1B0B=0, all other q's equal	7	-1442.2	2898.3	3.71	0.130
HiSSE: τ0A=τ1A, ε's equal, q0B1B=0, q1B0B=0, all other q's equal	5	***-1442.3***	***2894.6***	***0.00***	***0.848***

FIGURE LEGENDS

FIGURE 1. The conceptual problem with the presence of hidden states in the application of state-dependent models of speciation and extinction. Here related to state 0 and 1 is an unmeasured third variable with states *A* and *B*, and state *B* has twice the diversification rate of *A*. This trait is "hidden" because it is not observed in the tip data. If state 1 happens to be a prerequisite for evolving state *B*, all the state 0 branches will have state *A*, but some branches in state 1 will have state *A* and some will have state *B*. Thus state 1 actually takes on two states, 1*A* when the hidden state with higher diversification rate is absent, and 1*B* when the hidden state with higher diversification rate is present. As shown in the example tree from a simulated tree and trait data, 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, obtained from backtransforming net turnover rates and extinction fraction; see main text), 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 and green region represent the mean and 95% confidence interval for state 0*A*, the solid blue line and blue region represents the 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, the solid light purple line and light purple region, represent mean and 95% confidence interval for transition to and from the "hidden" state (i.e., 1*B*), respectively. 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.

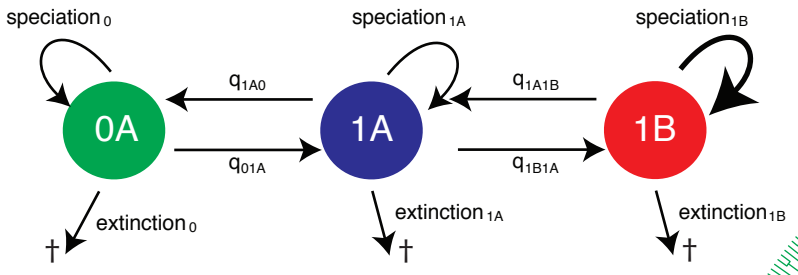
FIGURE 4. The uncertainty surrounding estimates of net diversification rate (λ-μ; obtained from backtransforming net turnover rates and extinction fraction; see main text) and extinction fraction (μ/λ) when the generating model assumes (a,b) diversification is independent of the observed characters (i.e., null-two model), or (c,d) a hidden state underlies both observed states (i.e., a general HiSSE model). The specifics of the different simulation scenarios are described in detail in Table 2. In all panels, the color of dashed line corresponds to the true value under the generating model.

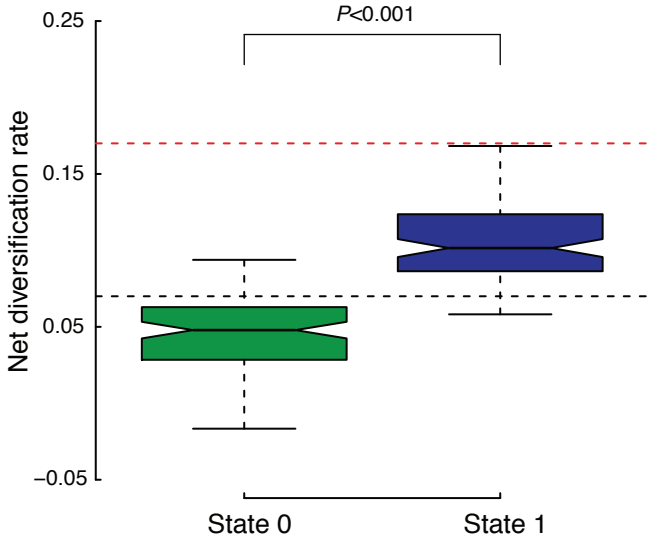
FIGURE 5. The reliability of ancestral state reconstructions as a function of tree size based on the model-averaged parameter estimates from HiSSE Scenario 1. The trend lines

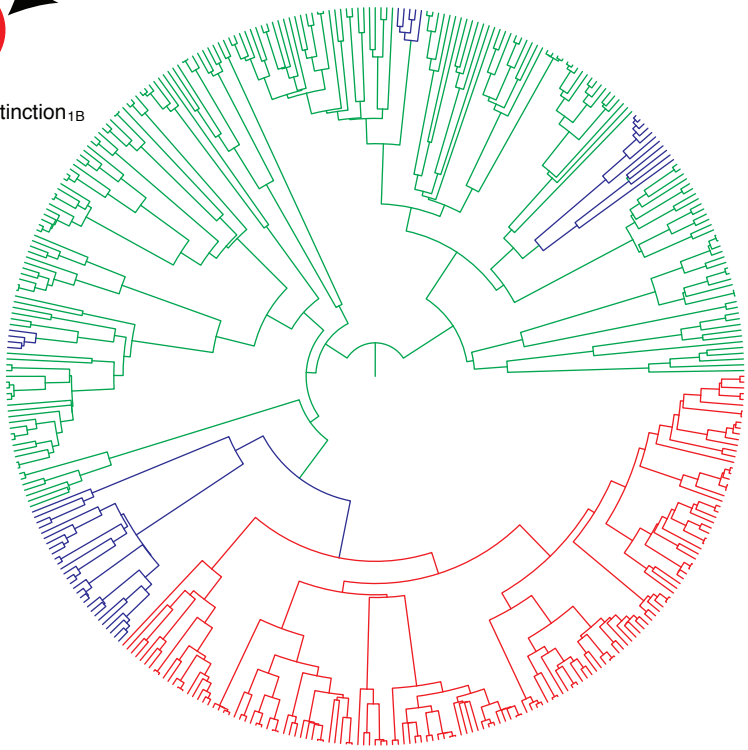
represent the relationship between the rolling median of the maximum marginal probability and the rolling mean of proportion of nodes inferred to be in the correct state across all simulated trees. The rolling mean and median are based on a sliding window of 500 points, and the dotted line denotes the 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: note, for instance, that most nodes fall near 1 in terms of actual accuracy of the reconstruction.

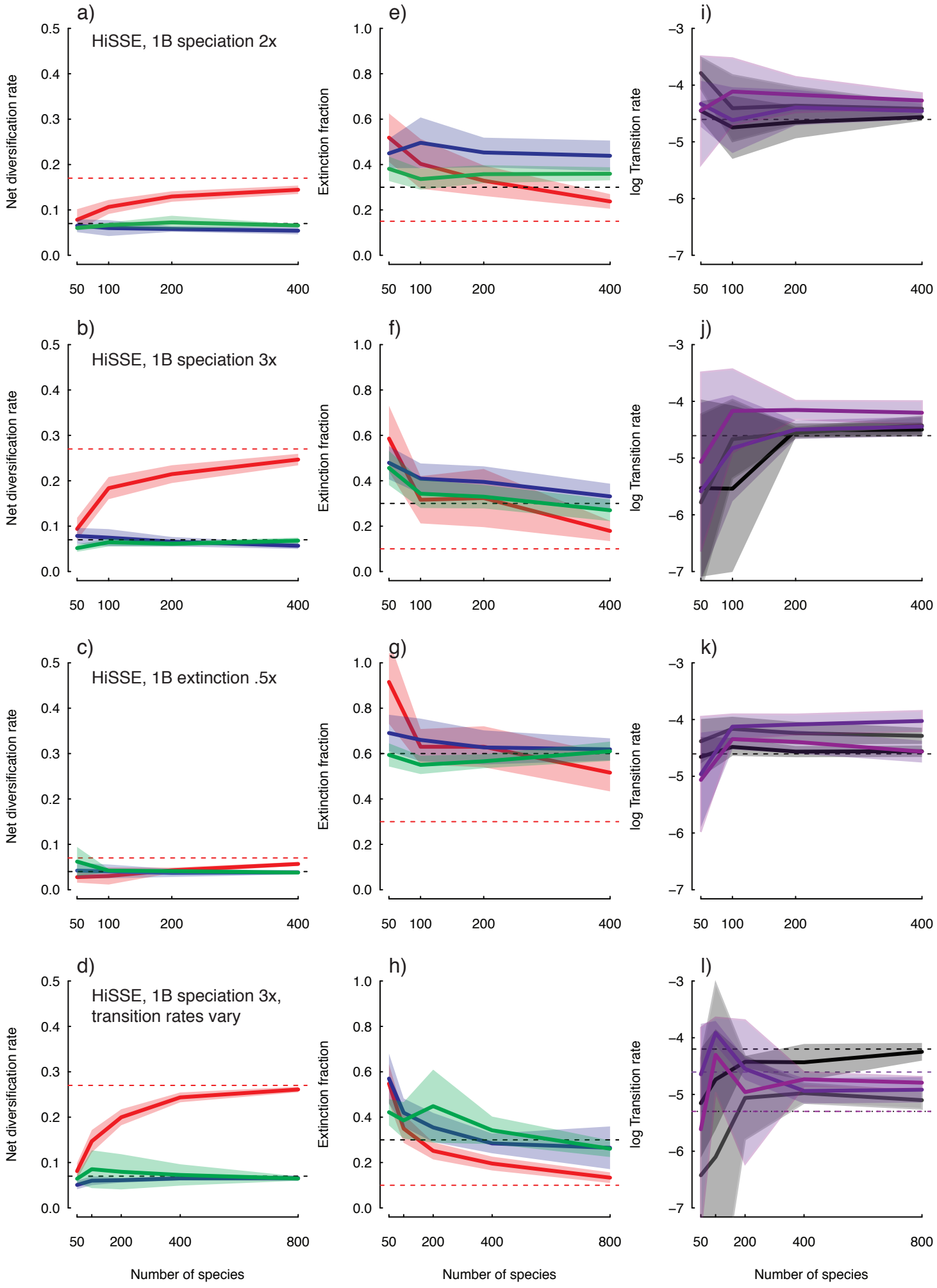
FIGURE 6. Comparisons of character reconstructions of states and net diversification rates under both the BiSSE model and HiSSE models, applied to a large empirical data set of non-achene (black branches) and achene fruits (white branches) for the Dipsidae. The major clades are labeled and estimates of the most likely state and rate are based on the model-averaged marginal reconstructions inferred under a set of models (see main text). Rate colors are comparable between figures. (a) With BiSSE, there are only two distinct rate categories associated with each observed state, which suggests that all achene-bearing lineages have an elevated net diversification rate relative to non-achenes. The histograms to the right of the tree show the location of the rates on a gradient of rates, as well as the frequency of both these rates and states for each contemporary tip taxa. (b) However, the HiSSE character reconstructions reveal strong heterogeneity in the diversification rates underlying each observed state, with contemporary taxa showing a large range of diversification rates, once uncertainty in the model and the reconstruction is taken into accounted. We highlight and discuss several interesting clades that contain

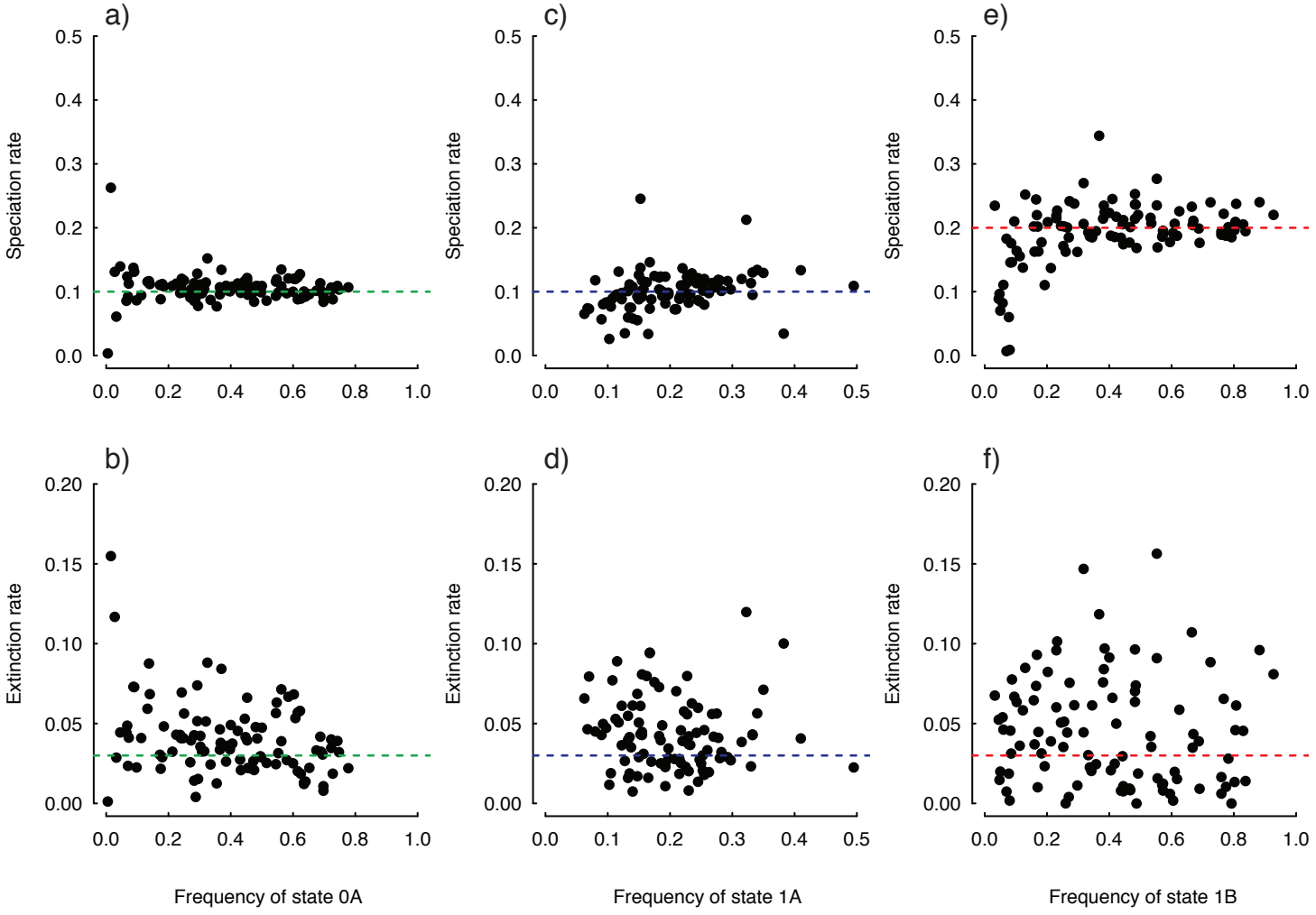
some of the highest diversification rates across the Dipsidae, particularly those within achene lineages (i.e., Dipsacaceae and core Valerianaceae). Para=Paracryphiales, Adox=Adoxaceae, Dier=Diervilleae, Capr+*Hepta*=Caprifolieae+*Heptacodium*, Linn=Linnaeeae, *Zab*=*Zabelia*, Morin=Morinaceae, *Triplo*+Dipsac=*Triplostegia*+Dipsacaceae, Valerian=Valerianaceae.

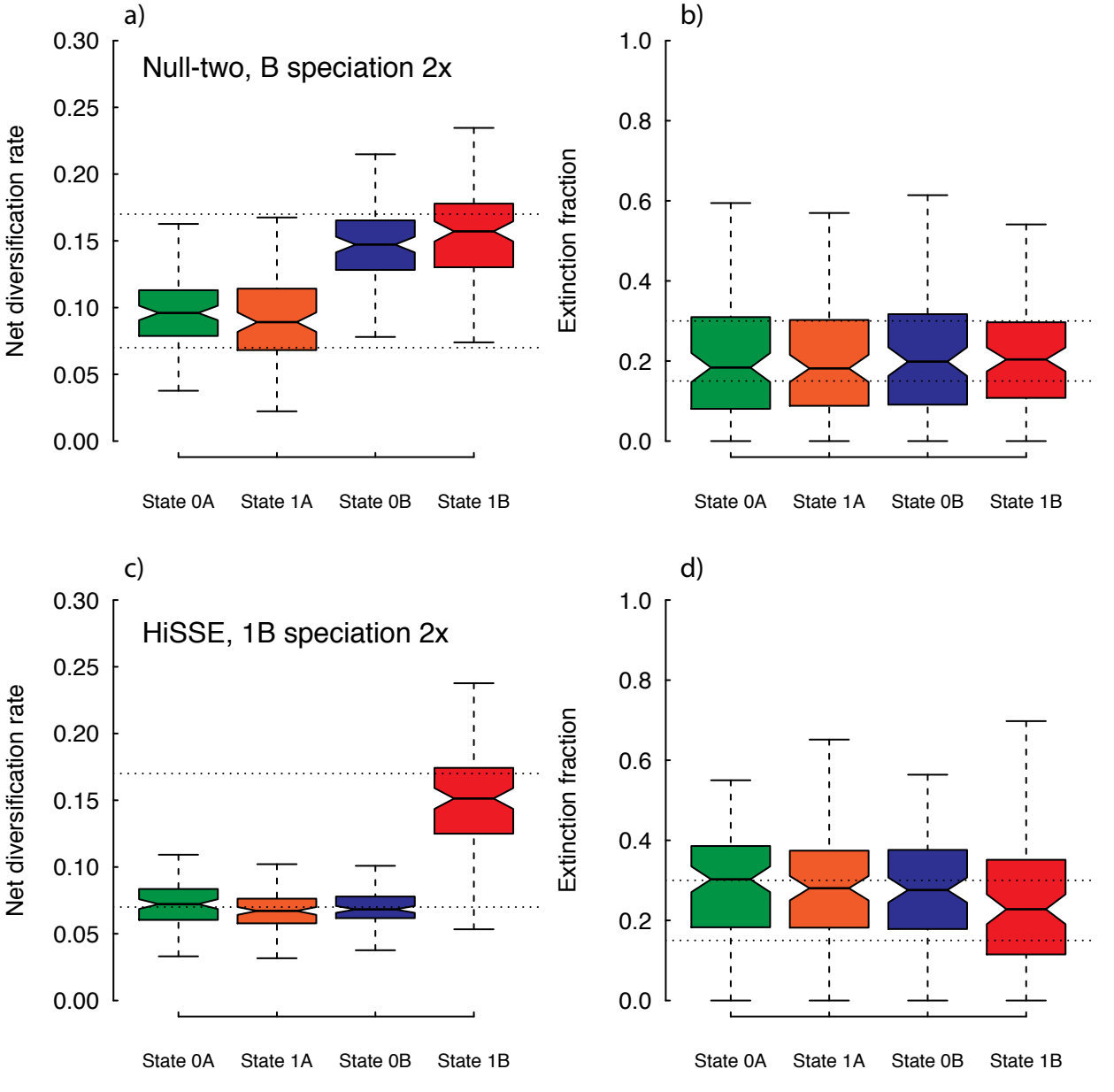

speciation 0
speciation 1A
speciation 1B
q1A0
q1A1B
0A
1A
1B
q01A
q1B1A
extinction 0
extinction 1A
extinction 1B
†
†
†

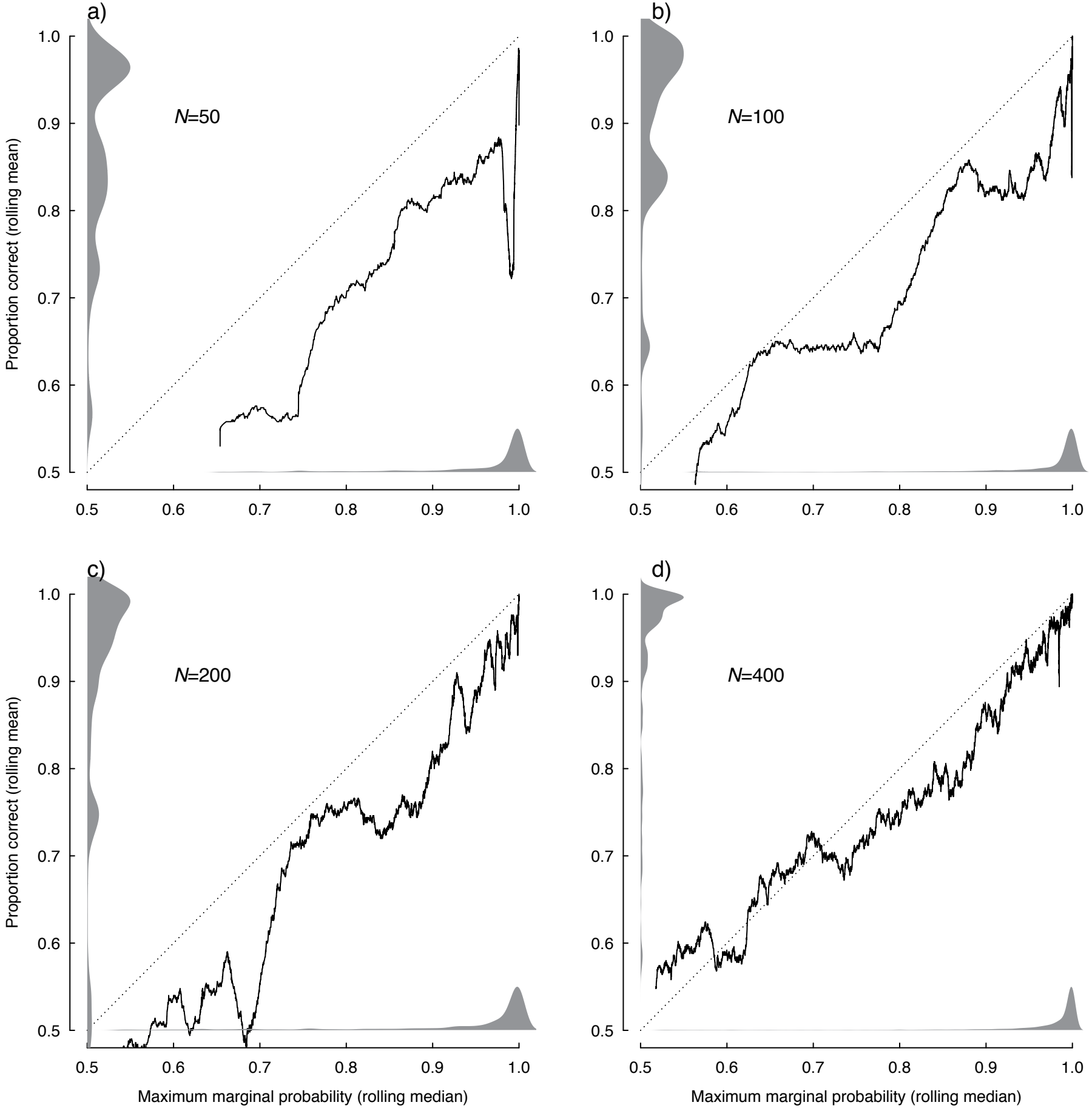

P<0.001
0.25
0.15
0.05
−0.05
Net diversification rate
State 0
State 1

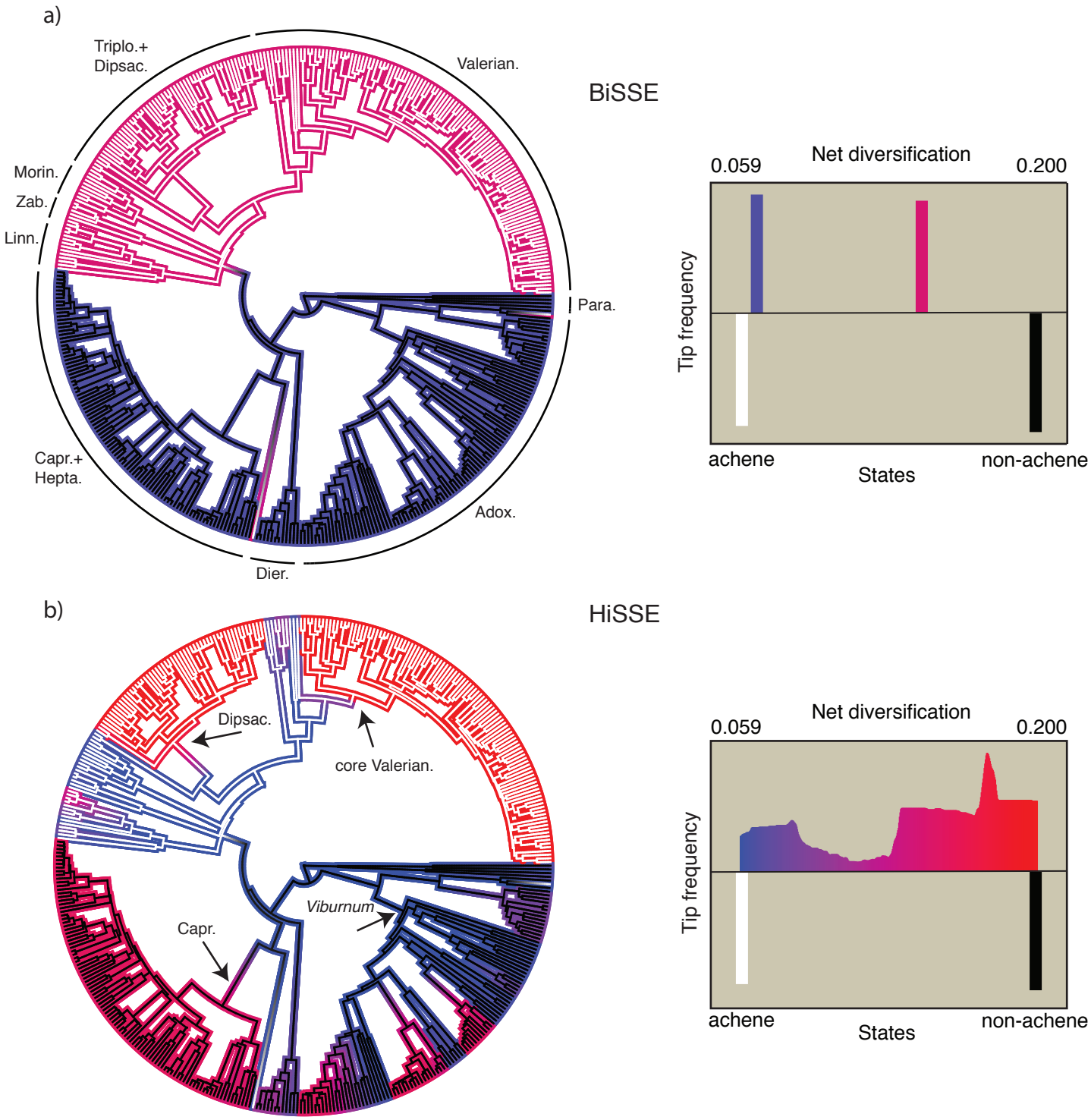



a)
HiSSE, 1B speciation 2x
b)
HiSSE, 1B speciation 3x
c)
HiSSE, 1B extinction .5x
d)
HiSSE, 1B speciation 3x,
transition rates vary
e)
f)
g)
h)
i)
j)
k)
l)
Net diversification rate
Extinction fraction
log Transition rate
Number of species


a)
b)
c)
d)
e)
f)
Speciation rate
Extinction rate
Frequency of state 0A
Frequency of state 1A
Frequency of state 1B


a)
Null-two, B speciation 2x
Net diversification rate
0.30
0.25
0.20
0.15
0.10
0.05
0.00
State 0A
State 1A
State 0B
State 1B
b)
Extinction fraction
1.0
0.8
0.6
0.4
0.2
0.0
State 0A
State 1A
State 0B
State 1B
c)
HiSSE, 1B speciation 2x
Net diversification rate
0.30
0.25
0.20
0.15
0.10
0.05
0.00
State 0A
State 1A
State 0B
State 1B
d)
Extinction fraction
1.0
0.8
0.6
0.4
0.2
0.0
State 0A
State 1A
State 0B
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


a)
Triplo.+
Dipsac.
Valerian.
BiSSE
Morin.
Zab.
Linn.
Para.
Capr.+
Hepta.
Adox.
Dier.
Net diversification
0.059
0.200
Tip frequency
achene
States
non-achene
b)
HiSSE
Dipsac.
core Valerian.
Viburnum
Capr.
Net diversification
0.059
0.200
Tip frequency
achene
States
non-achene