

Estimating diversification rates for higher taxa: BAMM can give problematic estimates of rates and rate shifts

Andreas L. S. Meyer^{1,2} and John J. Wiens^{2,3}

¹Graduate Program in Zoology, Universidade Federal do Paraná, Curitiba, Paraná 81531, Brazil

²Department of Ecology and Evolution, University of Arizona, Tucson, Arizona 85721

³E-mail: wiensj@email.arizona.edu

Received September 15, 2016

Accepted October 4, 2017

Estimates of diversification rates are invaluable for many macroevolutionary studies. Recently, an approach called BAMM (Bayesian Analysis of Macro-evolutionary Mixtures) has become widely used for estimating diversification rates and rate shifts. At the same time, several articles have concluded that estimates of net diversification rates from the method-of-moments (MS) estimators are inaccurate. Yet, no studies have compared the ability of these two methods to accurately estimate clade diversification rates. Here, we use simulations to compare their performance. We found that BAMM yielded relatively weak relationships between true and estimated diversification rates. This occurred because BAMM underestimated the number of rates shifts across each tree, and assigned high rates to small clades with low rates. Errors in both speciation and extinction rates contributed to these errors, showing that using BAMM to estimate only speciation rates is also problematic. In contrast, the MS estimators (particularly using stem group ages), yielded stronger relationships between true and estimated diversification rates, by roughly twofold. Furthermore, the MS approach remained relatively accurate when diversification rates were heterogeneous within clades, despite the widespread assumption that it requires constant rates within clades. Overall, we caution that BAMM may be problematic for estimating diversification rates and rate shifts.

KEY WORDS: BAMM, diversification, extinction, macroevolution, simulations, speciation.

Diversification rates are fundamental components of many macroevolutionary studies (Ricklefs 2007; Morlon 2014). The diversification rate of a clade is the rate at which it accumulates species over time, or the rate of speciation minus the rate of extinction. Diversification rates may be particularly important for understanding evolutionary and ecological origins of species richness patterns among clades and habitats, and for testing the impact of different factors (e.g., ecology, genetics, morphology, physiology, behavior, development) on species diversity (review in Wiens 2017). Estimated diversification rates have been used to address numerous ecological and evolutionary questions, such as why species richness is higher in the tropics (e.g., Condamine et al. 2012; Pyron and Wiens 2013; Rolland et al. 2014) and in terrestrial environments (e.g., Wiens 2015b; Miller and Wiens 2017). These estimates have also been used to test the association between diversification and the evolution of body size and shape

(e.g., Adams et al. 2009; Rabosky et al. 2013), sexual-size dimorphism (De Lisle and Rowe 2015), herbivory (e.g., Price et al. 2012; Wiens et al. 2015), plant-defense mutualisms (e.g., Weber and Agrawal 2014), genome-size evolution (e.g., Puttick et al. 2015), climatic-niche evolution (e.g., Kozak and Wiens 2010; Cooney et al. 2016), habitat (Alfaro et al. 2007; Wiens 2015a; Moen and Wiens 2017; Bars-Closel et al. 2017; Tedesco et al. 2017), reproductive isolation (e.g., Rabosky and Matute 2013), or many morphological and ecological traits simultaneously (e.g., Jezkova and Wiens 2017).

The importance of estimated diversification rates to many macroevolutionary studies is clear, but the best way to estimate these rates is not. Numerous methods have been proposed to estimate diversification rates (review in Morlon 2014). One widely used approach is the method-of-moments estimator (Magallón and Sanderson 2001; MS estimator hereafter). This approach

follows from the idea that a clade's net diversification rate (r) is $\ln(n)/t$, where n is the clade's extant species richness and t is the age (e.g., Stanley 1979). The clade age used can be the stem group age (age of the split between the clade and its sister taxon) or the crown group age (age of the oldest split among species within the clade). Magallón and Sanderson (2001) proposed a correction for the bias caused by the failure to sample extinct clades (leading to overestimation of diversification rates), for both crown and stem-group ages. This correction involves the relative extinction fraction (epsilon, or the extinction rate divided by the speciation rate), which is typically assumed rather than estimated. The MS approach is particularly useful because it does not require a detailed phylogeny within each clade. Using the stem-group estimator, only one species per clade is needed, and only two are needed for the crown-group estimator (if they encompass the clade's crown-group). This approach has been used to address many diverse questions related to diversification (e.g., Alfaro et al. 2007; Adams et al. 2009; Rabosky and Matute 2013; Gómez-Rodríguez et al. 2015; Wiens 2015a,b; Scholl and Wiens 2016; Cooney et al. 2016; Tedesco et al. 2017).

Nevertheless, several articles have suggested that the MS estimators give inaccurate estimates of diversification rates (e.g., Rabosky 2009, 2010; Rabosky and Adams 2012; Rabosky et al. 2012). Surprisingly, these articles did not actually test the accuracy of the MS approach (e.g., whether estimated rates matched known, simulated rates). Instead, they reported that simulations of net diversification rates often yielded significant, positive relationships between clade ages and their species richness, whereas observed relationships between clade age and richness were often nonsignificant or negative. It is unclear how these results were related to the accuracy of the MS estimators (as opposed to unrealistic simulation assumptions, see Kozak and Wiens 2016; Scholl and Wiens 2016). Similarly, these authors argued that the MS estimators should only be used when clade age and species richness are positively related (e.g., Rabosky and Matute 2013; Rabosky et al. 2013). A recent simulation study found strong correlations between true and estimated rates using the crown-group MS estimator, regardless of the relationship between clade age and richness (Kozak and Wiens 2016).

Some of the authors who claimed that the MS estimators were inaccurate argued that a different approach should be used instead, called Bayesian Analysis of Macro-evolutionary Mixtures (BAMM hereafter; Rabosky 2014). BAMM is intended to detect significant shifts in diversification rates across a phylogeny, and provide estimated rates of speciation, extinction, and diversification for each branch, using reversible jump Markov Chain Monte Carlo methods. BAMM has since become widely used in empirical studies (e.g., Espeland et al. 2015; Peña et al. 2015; Schilling et al. 2015; Spriggs et al. 2015; Blair and Sanchez-Ramirez 2016; Davis et al. 2016; Huang 2016; Larson-Johnson 2016).

Unfortunately, the accuracy of these two diversification rate estimators (MS and BAMM) has not been adequately compared. Rabosky (2014) found that BAMM gave more accurate estimates of speciation rates than an approach called MEDUSA (from Alfaro et al. 2009), but did not address the accuracy of these methods for estimating diversification rates. Recent simulations suggest that MEDUSA may be relatively inaccurate at estimating rates and rate shifts (May and Moore 2016). Similarly, simulation studies of the accuracy of the MS method have been limited to analyses of the crown-group estimator with complete sampling for only one backbone tree (Kozak and Wiens 2016). To our knowledge, no studies have compared the performance of BAMM and the MS estimators under matched, simulated conditions.

A recent article highlighted several potential flaws in BAMM (Moore et al. 2016). For example, they used simulations to test the accuracy of estimated diversification rates from BAMM (with simulated trees that were complete but relatively small, with ~50–150 species). They found that BAMM gave accurate estimates of diversification rates when true rates were constant, but relatively inaccurate estimates when rates varied across the tree (i.e., on average, a correlation between true and estimated rates of zero). However, more recently, Rabosky et al. (2017) suggested that the results of Moore et al. (2016) were incorrect and that BAMM yields accurate estimates of speciation rates (but without addressing diversification rates). Given these conflicting views, it is presently unclear whether BAMM should be used in empirical studies or not.

Here, we use simulations to compare the accuracy of BAMM and the MS estimators for estimating diversification rates of clades. We are particularly interested in estimating rates for higher taxa (e.g., genera, families, orders, phyla), with incomplete phylogenetic species sampling within each clade, to test relationships between net diversification rates of clades and their ecological and/or morphological traits across large phylogenies. This was the focus of many of the studies on diversification cited above. However, we acknowledge that this is not the focus of every study using BAMM. We analyze whether each method yields accurate rate estimates given complete taxon sampling (i.e., all extant species in the clade are sampled in the phylogeny) and when only 50% or 25% of the species are sampled in each clade. We especially compare statistical relationships between the true and estimated diversification rates for clades among different methods. These relationships may be especially relevant for those wishing to test for correlations between diversification rates and traits among clades: methods that yield weak relationships between true and estimated diversification rates might yield weak or misleading relationships between diversification rates and traits. We also test for parallels between our simulation results and results from empirical data.

Methods

SIMULATING TREES

We simulated trees with known speciation (birth) and extinction (death) rates for each clade. We then compared the ability of each method (BAMM, MS estimators) to estimate the known diversification rates (birth - death). We analyzed 20 replicate backbone trees, each with 10 higher level clades, for a total of 200 clades overall, each with different rates of speciation and extinction. The trees are summarized in online Figure S1 and are provided in newick format in online Supplementary Files 1–3. The true speciation, extinction, and diversification rates, along with clade ages and species richness are given in Table S1. Although 20 trees and 200 clades may seem limited, many trees and clades were relatively large (see below), thus making analyses very computationally intensive. Furthermore, each tree and clade was analyzed three times, once for each level of taxon sampling (100%, 50%, 25%).

We simulated 20 backbone trees using the package PHYTOOLS version 0.5–00 (Revell 2012) in R version 3.2.2 (R Core Team 2015). Backbone trees were each set to have 10 terminal taxa and a total length of 100 million years, with each terminal taxon at least 15 million years (Myr) old. These terminal branch lengths then became the stem ages of the 10 higher level clades within each backbone tree. We used the minimum of 15 Myr because preliminary analyses showed that younger ages would not allow enough time to generate the minimum number of species within these clades (10 or more per clade), given the speciation and extinction rates we simulated. We chose a minimum of 10 species because smaller numbers would be problematic for BAMM when combined with subsampling only 50% or 25% of the species (i.e., estimating speciation and extinction rates for a clade of multiple species based on a single species).

We then used the age of the terminal branch of each higher level clade to simulate trees under a birth-death process. Trees were simulated using the *sim.bdtree* function in the R package GEIGER version 2.0.6 (Harmon et al. 2008; Pennell et al. 2014). For each clade, we randomly selected speciation rates from 0.01 to 0.30 speciation events/Myr (from a uniform distribution). Extinction rates were chosen from a uniform distribution such that the relative extinction fraction (extinction rate divided by the speciation rate) ranged from 0.01 to 0.90 for each higher level clade. This combination of speciation and extinction rates yields an expected mean net diversification rate of ~ 0.08 , with a potential range from 0.01 to 0.30. These diversification rates generated clades that were of sufficient size (10 or more species) but not computationally intractable. The mean size of the 200 clades was ~ 200 species per clade.

These seem to be realistic net diversification rates, at least based on empirical estimates from MS estimators. For both ani-

mals and plants, the mean diversification rate across families is 0.05 species/Myr (maximum of 0.51 for animals and 0.88 for plants), and rates across the Tree of Life are generally higher in younger clades and lower in older clades (Scholl and Wiens 2016). For example, from our own work, estimated mean diversification rates across clades include ~ 0.01 species/Myr for animal phyla (Wiens 2015b), ~ 0.03 for insect orders (Wiens et al. 2015) and ~ 0.09 for major clades within plethodontid salamanders (Kozak and Wiens 2010). The mean diversification rate simulated might be somewhat higher than in some empirical datasets because we excluded clades with fewer than 10 species (presumably clades having lower diversification rates). Nevertheless, we think the mean and ranges of rates simulated here are otherwise realistic.

When simulating phylogenies within each clade, a polytomy was created between the clade's sister group and the two basal clades within the group. To eliminate this polytomy, we randomly selected and deleted one of the two basal clades. This also ensured that there was a distinct stem-group age and crown-group age for each clade. This deletion should not affect the speciation and extinction rates, given that these rates are constant across the tree within each higher level clade.

Overall, we used this two-part structure to the simulations (i.e., 20 backbone trees of 10 clades each) so that we could generate the desired number of clades of adequate size with realistic diversification rates. Again, generating smaller clades (i.e., < 10 species) would make it more difficult to address the effects of incomplete taxon sampling. Thus, simply generating entirely random trees with randomly varying rates would not allow us to address our questions of interest here.

ESTIMATING DIVERSIFICATION RATES

We then used BAMM version 2.5.0 (downloaded 21 January 2016) and the MS estimators to estimate diversification rates for each clade. For BAMM, we estimated shifts in diversification rates across the tree (i.e., changes in rate regimes) and the diversification rates for each rate regime, following standard practice in empirical studies. The configuration of rates with the highest posterior probability was chosen for each tree. Typically, a single rate was estimated within a clade. We chose tree-specific priors using the R package BAMMTOOLS version 2.1.0 (Rabosky et al. 2014). When we analyzed each tree, we set the expected number of shifts to 10, given that each tree had 10 clades, each with random and independent diversification rates. Thus, we seeded the BAMM analyses with a number close to the actual number of rate regimes, even though this number would be unknown in empirical analyses. We address the consequences of changing the number of expected shifts below, and show that this has little impact on the results. Speciation rates were allowed to vary over time in all analyses (the default setting).

We performed multiple runs on each tree using the Markov Chain Monte Carlo approach, choosing both the number of generations and sampling frequency according to the number of tips (species) in the tree: 10 million generations and sampling every 2000 generations for trees with up to 999 species, 50 million generations and sampling every 10,000 for trees with 1000–1999 species, and 100 million generations and sampling every 20,000 for trees with 2000 or more species. Following standard practice, we discarded the first 10% of the posterior distribution of samples as burn-in. To assess whether discarding more samples would affect our conclusions, we reanalyzed the results (for full sampling) discarding 50% of the posterior as burn-in (results shown in Table S2). The relationship between estimated rates obtained using burnins of 10 and 50% was extremely strong ($r^2 = 0.97$), which indicated good convergence. Therefore, only results obtained using the 10% burn-in were used in the following analyses.

We then calculated mean speciation and extinction rates for each of the 10 higher level clades on each of the 20 trees using the function *getCladeRates* in the R package BMMTOOLS. Diversification rates for each clade were then calculated as the mean speciation rate minus the mean extinction rate for each clade.

Importantly, we note that the use of the mean speciation, extinction, and diversification rates is standard practice both for empirical studies and for simulation studies that address the performance of BMM (e.g., Rabosky 2014; Moore et al. 2016; Rabosky et al. 2017). It is not standard practice to merely test whether the true rate is within the 95% highest posterior density interval. Nevertheless, we do present some results along these lines as well.

To evaluate whether the posterior was influenced by the choice of priors on the number of rate shifts, we re-ran BMM for the first three backbone trees ($n = 30$ clades, full sampling), but at this time selecting three different numbers of rates shifts as priors: 5, 3, and 1 shifts (instead of 10). Other priors remained the same. Priors on rate parameters were set using the *setBMMpriors* function in the BMMTOOLS package. We then assessed the relationship between diversification rates obtained with a prior of 10 rate shifts and those obtained using 5, 3, and 1 shifts (i.e., testing 10 vs 5, 10 vs 3, and 10 vs 1, across the set of 30 clades). Rates were very tightly related (Table S3), regardless of the number of rate shifts ($r^2 > 0.99$ in all cases). Moreover, the number and position of shifts remained the same across priors. We also evaluated whether estimated rates varied through time in response to the number of shifts set as priors. The results show that rates were virtually the same (Fig. S2), suggesting that the posterior was not influenced by the choice of priors.

We assessed the effect of incomplete taxon sampling by randomly selecting and then pruning 50% and then 75% of the species in each higher level clade. The same sampling level was applied across all clades in all trees. Diversification rates were then reesti-

mated from the subsampled trees, as described above. The proportional sampling of each clade was input into the BMM analyses. In other words, the total number of species in each clade was assumed to be known, and BMM was allowed to compensate for incomplete sampling. BMM control files are available on Dryad (<https://doi.org/10.5061/dryad.db561>).

Net diversification rates were then estimated separately for each higher level clade using both stem and crown group MS estimators. Following standard practice in empirical studies (e.g., Adams et al. 2009; Gómez-Rodríguez et al. 2015; Wiens 2015a,b), we used three assumed values of the relative extinction fraction (epsilon): low extinction (0), intermediate extinction (0.5), and relatively high extinction (0.9). MS estimators were implemented in GEIGER 2.0.6. We also implemented these approaches on the subsampled trees with only 50% and 25% of the species included. However, the impact of incomplete sampling on the MS estimators should be limited, since they depend only on clade ages and species numbers. Rates were identical for the stem-group estimator, since the estimated stem age of the clade will remain the same with incomplete taxon sampling within that clade. In contrast, rates for the crown-group estimator could be influenced if incomplete sampling caused incorrect estimation of the crown group age. Specifically, diversification rates will be overestimated if the oldest split within the clade is not sampled, and the crown-group age is therefore estimated to be younger than it actually is (i.e., a younger clade age and the same number of species yields a higher diversification rate).

EVALUATING ACCURACY

To assess accuracy, we first estimated the mean percentage deviance between the true and estimated rates across all 200 clades for each method. Specifically, for each rate estimate, we divided the estimated rate by the true rate, subtracted the resulting value from 1, and multiplied that value by 100. We then calculated the mean across all 200 clades. Rates could be overestimated (estimated rates were higher than the true rates used to simulate the data), yielding positive mean percentage deviance values, or underestimated (estimated rates lower than the true rates), yielding negative values. Analysis of mean deviance values will reveal whether a method has an overall systematic bias, but a method could also be highly inaccurate without being systematically biased. We also present results from the mean of the absolute percentage deviance values.

As another approach for evaluating accuracy, we tested the relationships between the true and estimated diversification rates for each method for each set of conditions, using linear regression in R. The full set of 200 clades was compared for each set of conditions. More accurate methods should yield stronger positive, relationships between the true and estimated rates. This relationship is particularly important, because many studies use estimated

diversification rates to test relationships between diversification rates and traits (see Introduction), and weak relationships between true and estimated rates might yield weak or misleading relationships between rates and traits.

We did not focus on estimating diversification rates for backbone trees (i.e., the tree linking the 10 focal clades in each phylogeny). Diversification rates were clearly lower for the backbone tree than for higher level clades (i.e., 10 species in 100 million years yields a rate of ~ 0.02 species/Myr, much slower than the expected average of ~ 0.08 for higher level clades). However, the slower diversification rate for the backbone tree should simply be estimated as a different rate regime by BAMM. Thus, there is no a priori reason why different rates in the backbone tree and individual clades should be problematic for this method, and similar shifts were simulated by Rabosky (2014). For the MS estimators, we did not estimate rates for the backbone tree, following standard practice in empirical studies.

Finally, we note that some may consider the sample size for BAMM to be smaller than that for the MS estimators because BAMM is applied to entire trees (i.e., 20) and the MS estimators to individual clades (i.e., 200). Regardless, BAMM is estimating rates for each branch in every tree, potentially using data from hundreds of species per clade and thousands of species per tree. Perhaps more importantly, rate estimates from BAMM for each clade are not fully independent across clades within a given backbone tree. Our results illustrate how this nonindependence among clades impacts accuracy, and we also explore how accuracy is changed by allowing BAMM to estimate each clade's rates independently of other clades in the tree (see below).

ADDITIONAL SIMULATION ANALYSES

The analyses described above generated the primary results of this study, focused on comparing the accuracy of diversification rate estimates from BAMM and the MS estimators (and relationships between true and estimated rates for each method). However, we also performed a series of analyses to further explore these main results. Specifically, we tested if the accuracy of BAMM could be improved by estimating rates for clades having a constant rate (rather than estimating rate shifts across a rate-heterogeneous tree), as predicted by Moore et al. (2016). We also addressed whether the accuracy of the MS estimators was decreased when there was heterogeneity in rates within clades. Previous studies have stated that the MS estimators assume diversification rates are constant within clades (e.g., Rabosky et al. 2012), but without actually comparing their accuracy given constant versus heterogeneous rates within clades (and previous analyses of their accuracy simulated only constant rates within clades; Kozak and Wiens 2016). We also addressed how errors in estimated speciation and extinction rates contributed to the errors in diversification rates estimated by BAMM, given that some empirical studies only use speciation

rates from BAMM. In addition, we addressed whether incorrect diversification rates in BAMM were related to differences in the true diversification rates. Finally, we tested how often each method (BAMM vs MS estimators) correctly identified which of a given pair of sister clades had a higher diversification rate. We describe these additional analyses in detail below. However, we think the most relevant measure of accuracy overall is the statistical relationship between the true and estimated rates.

We first tested if the accuracy of BAMM could be improved by estimating diversification rates separately for each higher level clade (similar to how the MS estimators were applied), instead of using BAMM to estimate both rates and rate shifts across the entire tree (how BAMM is normally applied). To estimate diversification rates separately for each clade, we deleted all species in a tree except for that focal clade. When we analyzed each clade separately, we set the expected number of rate shifts to 1 (instead of 10), given that each clade had a single true diversification rate. All other settings were the same as in the main analyses. Because these analyses were very computationally intensive, we did this for all clades in the first three backbone trees (full sampling), totaling 30 clades (note that since diversification rates are generated randomly and independently for each clade, it should not matter which trees or clades are chosen). Clade-specific mean diversification rates we extracted using the function *getCladeRates* in the R package BAMMTOOLS. In some cases, the individual clades were so small that it was not possible to use BAMM to estimate diversification rates when only 50% or 25% of the species were included. These clades were simply left out when calculating average accuracies among clades.

Next, we tested if the accuracy of the diversification rates estimated by the MS estimators was impacted by heterogeneity in diversification rates within clades. Several articles have referred to the MS estimators as “constant rate” estimators, given the supposition that they require that diversification rates be constant within clades to accurately estimate rates (e.g., Rabosky and Adams 2012; Rabosky et al. 2012). However, this assumption was not actually tested. In fact, these estimators depend only on the age and richness of clades. Therefore, given the same age and richness, they will yield identical net diversification rates regardless of whether instantaneous diversification rates are increasing, declining, or are otherwise heterogeneous within clades.

Nevertheless, we tested how heterogeneity in rates within clades influenced the accuracy of the MS estimators by combining pairs of clades with different diversification rates. We focused on sister pairs of clades, since their diversification rates are independent, and therefore just as different as any two clades in a tree. In theory, more clades could be included (e.g., groups of four clades instead of two). However, the underlying rates for these clades would not be any more different (on average) than those for the sister pairs.

We identified 66 independent sister pairs of clades across the 20 trees, together encompassing the majority (132) of the 200 clades simulated. For each sister pair, we applied the MS estimators for the pair of clades as if they were a single clade. Specifically, we recalculated diversification rates for each pair of clades using the combined richness of the two clades and the stem and crown ages for the common ancestor of the two clades. We then treated the average of the known diversification rates for the pair of clades as the “true” rate for that combined clade. We then compared the mean accuracy of the separately estimated rates for the pairs of clades (i.e., rates constant within clades) to the mean accuracy of the estimated rates for the pairs of clades combined (i.e., rates heterogeneous within the pair of clades).

We did not attempt to compare true and estimated diversification rates of clades when the instantaneous diversification rates were changing over time within the clade (e.g., as in the case when diversification slows over time). Even though it is straightforward to estimate the net diversification rate for such a clade (as always, it is based on the clade’s age and richness), it is less clear what the true rate(s) would be, since this would depend on the specific time slice considered (i.e., the instantaneous rates). Estimates of net diversification rates address the diversification rate based on the entire history of the clade, not a single time slice. Therefore, we did not test the ability of net rate estimators to correctly estimate instantaneous rates at different time slices, since this would be a misapplication of the MS estimators. Furthermore, simulations that incorporate limited carrying capacities and declining diversification over time suggest that net diversification rates may still be informative in explaining richness patterns under these conditions (Pontarp and Wiens 2017).

Given that our main results showed that BAMM often gave erroneous estimates of diversification rates, we addressed the contribution of errors in speciation and extinction rates to errors in diversification rates. Using the MS estimators, only diversification rates are estimated, and the contributions of speciation and extinction to diversification rates are not estimated. Using BAMM, the net diversification rate is estimated based on the difference between the estimated speciation and extinction rates. We therefore used regression analyses (in R) to test the relationship between absolute error in diversification rates and absolute error in speciation rates, and then between absolute errors in diversification and extinction rates (“absolute” meaning all deviations from the true rates were treated as positive). If most error in estimated diversification rates was explained by error in estimating extinction rates, this might justify the practice of focusing only on speciation rates estimated by BAMM, and not diversification rates.

We also evaluated how different the true diversification rates of sister clades must be for BAMM to recognize that these clades belong to different rate regimes. We used a binomial GLM (in R) to compare the number of times BAMM assigned different diver-

sification regimes within a pair of sister clades and the magnitude of the difference of the true diversification rate of these clades. The response variable was the number of rate regimes assigned to each sister pair of clades (i.e., BAMM could assign either a single regime for both clades or a separate regime for each clade) and the predictor was the magnitude of the difference between the true diversification rates of each pair. Since binomial GLMs do not have r^2 values we used the R package PSCL version 1.4.9 (Jackman 2015) to estimate a pseudo r -squared (using the $pR2$ function).

Finally, we tested how often each method (BAMM vs MS estimators) correctly detected which of two sister clades had higher diversification rates. We compared the 66 pairs of sister clades across the 20 trees and calculated the frequency with which BAMM and the MS estimators correctly detected which sister clade had the higher true diversification rate. Note that all clades had different true rates, so inferring two sister clades as having the same rate was considered incorrect. To prevent this analysis from being influenced by very small differences in estimated diversification rates, values of estimated rates were rounded to the second decimal place for all methods for this comparison. Overall, however, we consider this test less relevant than the comparison of the relationships between true and estimated rates for each method.

EMPIRICAL ANALYSES

To complement the results of our simulations, we used an empirical phylogeny (from snakes; Zheng and Wiens 2016) to compare BAMM estimates of diversification rates for individual clades when diversification rates are estimated from the entire phylogeny and from each clade separately. These rates should be very similar, if not identical. If they are not similar, then we know that at least one of these estimates must be incorrect, even if we do not know which one. We chose snakes because they offer a set of well-sampled clades that are similar to our simulations, in terms of the number of clades and their size (15 clades per tree, mean = 227 species per clades vs 10 clades per tree and mean = 200 species per clade in the simulations). For these comparisons, we focused on 15 clades that are typically recognized as families, but we also included subfamilies within the megadiverse family Colubridae (clades recognized as families by some authors; Uetz et al. 2017). Since the subfamily Colubrinae was not monophyletic in the tree of Zheng and Wiens (2016), we treated it as two separate clades. We included only families with >5 species in the tree for these comparisons, since BAMM estimation failed on clades with 5 or fewer species (smaller families were included in the overall tree, but not used in these comparisons). We used the Reptile Database (Uetz et al. 2017) to obtain the number of described species in each clade and to seed BAMM with the sampling fraction (proportion of described species in the clade included in the tree) across the phylogeny (0.35) and for each clade (Table S4).

Table 1. Regression results (r^2) between true and estimated diversification rates for BAMM and the MS estimator (using stem and crown group ages and different epsilon values) under different sampling regimes (all species in the clade sampled in the tree, 50% of species sampled, 25% of species sampled).

Method	Full sampling	50% sampling	25% sampling
BAMM	0.37	0.34	0.24
MS Stem			
Epsilon = 0.0	0.75	0.75	0.75
Epsilon = 0.5	0.77	0.77	0.77
Epsilon = 0.9	0.76	0.76	0.76
MS Crown			
Epsilon = 0.0	0.66	0.58	0.37
Epsilon = 0.5	0.70	0.62	0.42
Epsilon = 0.9	0.79	0.76	0.66

All regressions were highly significant ($P < 0.001$).

First, we applied BAMM to the entire tree, setting priors using the *setBAMMpriors* function in BAMMTOOLS. Given no a priori expectations regarding the number of diversification regimes, we arbitrarily set the expected number of shifts to three (this should have little impact on the results; see above). We then calculated diversification rates for each family using the *getCladeRates* function in BAMMTOOLS.

Second, we selected all families with >5 species and applied BAMM to each family separately, with a separate sampling fraction for each clade (Table S4). In these within-clade analyses, the expected number of shifts was set to one, and other priors were set using the *setBAMMpriors* function in BAMMTOOLS.

We then tested the relationship between these two sets of estimates. Again, these estimates should be very similar, or at least tightly related. However, these estimates might not be strictly identical given that the sampling fraction for the entire tree can differ from the sampling fraction for individual clades. The snake trees used are available on Dryad (<https://doi.org/10.5061/dryad.db561>).

Results

MAIN SIMULATION RESULTS

The main results, including all simulated and estimated rates are given in online Table S1. The main results are summarized in Figure 1 and Table 1. Across all 200 clades, we found that BAMM frequently gave biased (overestimated) estimates of diversification rates, even with complete taxon sampling (Fig. 1A). In some cases (31 of 200 for full sampling), the true rates were outside the 95% highest posterior density intervals for the estimated rates (Fig. S3), even though these intervals were very broad. Further,

the mean BAMM estimates were often less accurate than those from some of the MS estimators, in some cases by threefold or more (especially when using the stem-group MS estimator with an intermediate epsilon and the crown-group MS estimator with a high epsilon). The results in Fig. 1 show only the mean deviances, which indicate potential systematic biases in the methods (e.g., overestimation and underestimation of rates). However, estimates can be wrong without being systematically biased. Estimates based on the means of the absolute deviances between true and estimated rates gave concordant results (Fig. S4), with BAMM again performing poorly relative to these same two MS estimators.

Similarly, even though there was a significant relationship between true and estimated diversification for all methods (Table 1), this relationship was much weaker for BAMM than for the MS estimators, typically by roughly twofold ($r^2 = 0.24$ – 0.37 for BAMM, $r^2 = 0.75$ – 0.77 for the MS stem estimator). The relationship between true and estimated rates was consistently strong for the MS estimator using stem-group ages, whereas the relationship for crown-group ages was more sensitive to sampling levels and assumed relative extinction fractions (epsilon values).

The errors with BAMM seemed to occur because it underestimated the number of rate shifts, and assigned smaller clades with low diversification rates to the same rate regimes as larger clades with higher diversification rate. (leading to an overall overestimation of diversification rates when averaged across all clades). Within each tree, we simulated 10 clades, each with a different combination of speciation and extinction rates. However, the optimal configuration of rate shifts estimated by BAMM across each tree was on average 2.35 (range = 0–5) diversification regimes (Table S1). Therefore, it underestimated the number of rate shifts across each tree, typically by more than fourfold. Furthermore, the specific cause of the overestimation of rates was seemingly the assignment of small clades to the same diversification rate regime as larger clades. This was apparent when we compared BAMM rate estimates from large clades and small clades (arbitrarily considering clades with 150 or more species to be large). For the 61 larger clades (Fig. 1B), BAMM rate estimates were quite accurate. For the 139 smaller clades (Fig. 1C), the estimates were far less accurate and were strongly biased toward overestimation. Importantly, the majority of clades were small clades.

The accuracy of the MS estimators depended on the particular estimator used (stem vs crown group), the epsilon value assumed, and the taxon sampling (Fig. 1; Table 1). The stem-group MS estimator was generally more accurate than the crown-group estimator, and was insensitive to limited taxon sampling (because the stem group estimate of clade age remains the same regardless of how few taxa are sampled within the clade). In contrast, when only 50% or 25% of the species in a clade were sampled, there was the potential to sample only within a subclade (rather than spanning

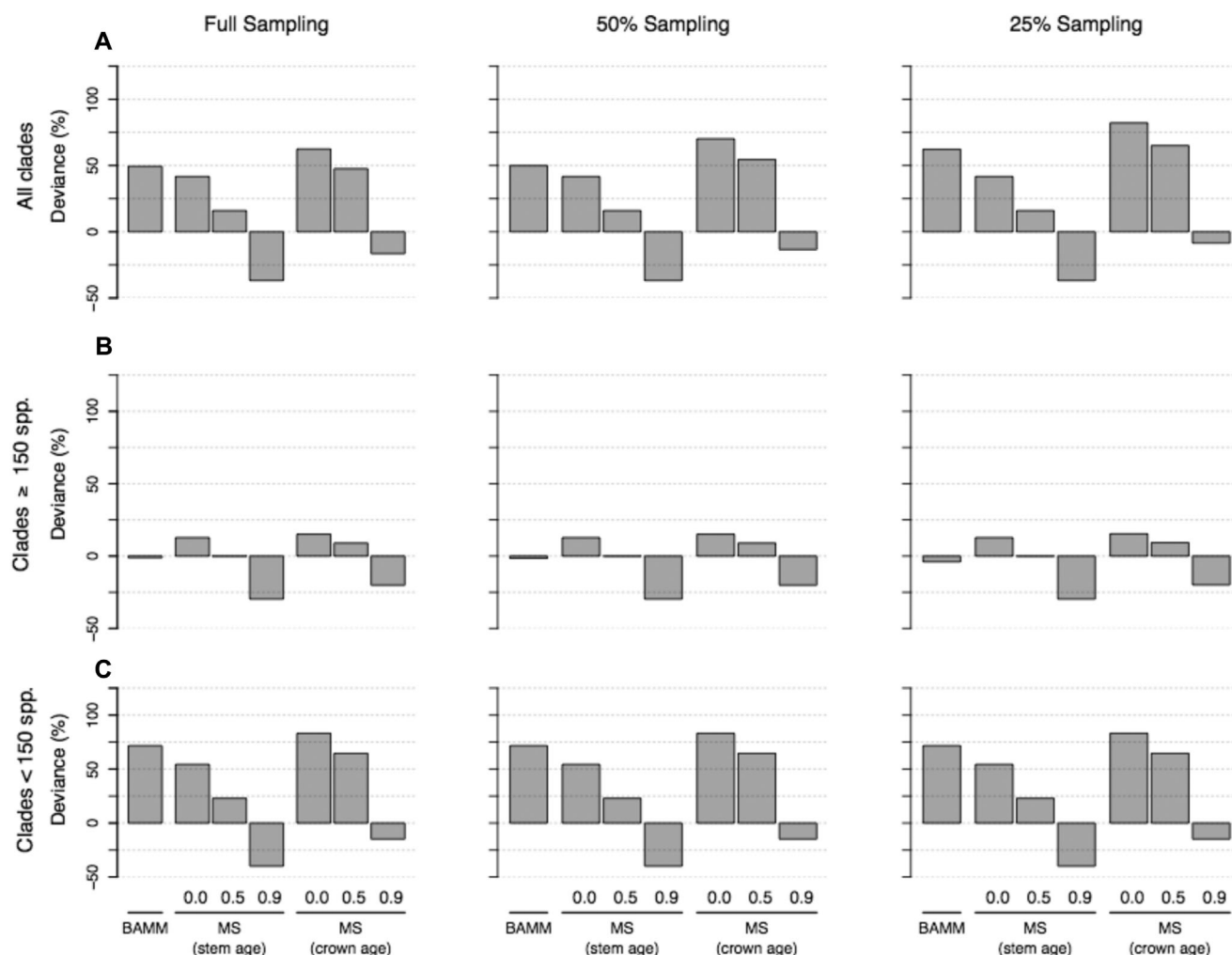


Figure 1. Accuracy of BAMM and MS estimators for estimating diversification rates of clades, including stem and crown group MS estimators with different values of epsilon (0, 0.5, and 0.9). (A) The y-axis shows the mean percentage deviation of the estimated diversification rates from the true diversification rates across all 200 clades. The first column of graphs (Full Sampling) shows the results based on complete sampling of all species in each clade, while the second and third columns show the results when only 50% and 25% of the species in each clade were included in the tree when estimating diversification rates. (B) Separate results for the 61 clades with 150 species or more. (C) Separate results for the 139 clades with less than 150 species. Full results are given in online Table S1.

the basal split and the true crown age) and thereby underestimate the crown group age and overestimate diversification rates. Most importantly, limited taxon sampling weakened the relationships between true and estimated crown-group rates (Table 1).

The assumed epsilon values also influenced the accuracy of the MS estimators. The mean simulated value of epsilon was ~ 0.45 (mean of a uniform distribution from 0 to 0.90). The use of the stem-group estimator with an epsilon of 0.5 gave the most accurate estimates overall, with only some overestimation of rates on average (for smaller clades, Fig. 1C). For the stem MS estimator, an epsilon of 0 led to more strongly overestimating the diversification rate (especially for smaller clades), whereas an epsilon of 0.9 led to more strongly underestimating the rate. However, the relationships between true and estimated diversification rates

remained strong and very similar using different epsilon values (Table 1).

For the crown-group MS estimator, epsilon values of 0 and 0.5 led to strongly overestimating the diversification rates in smaller clades (Fig. 1C). In contrast, a value of 0.9 led to relatively accurate estimates that were only slightly underestimated on average. Note that this high epsilon value did not match the true ratio of speciation to extinction, even though it yielded more accurate estimates of diversification rates. For the crown-group estimator, the relationship between true and estimated rates was often more strongly impacted by taxon sampling than epsilon values (Table 1).

For the MS estimators (as with BAMM), accuracy was generally higher for larger clades (Fig. 1B) than smaller clades

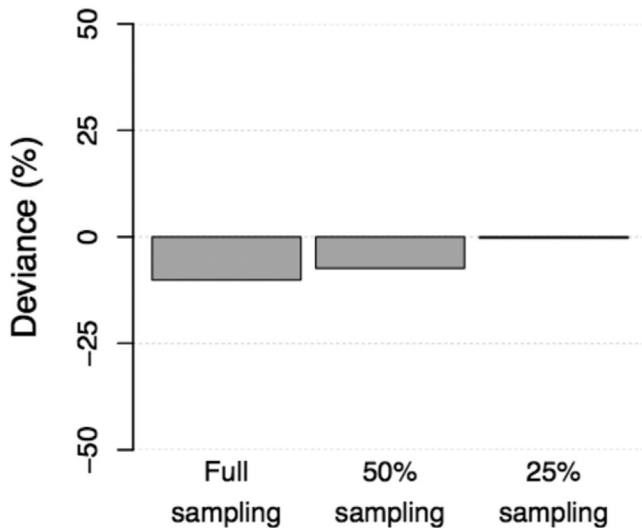


Figure 2. Accuracy of BMM for estimated diversification rates for 30 clades individually (all clades from the first three trees), with rates estimated after excluding all species in the tree not belonging to the focal clade, and setting the number of rate shifts to 1. Full results are given in online Table S2.

(Fig. 1C), with rates typically being overestimated in smaller clades (especially for crown-group estimators). However, the MS estimators continued to underestimate diversification rates given a high epsilon value (0.9) and larger clades (> 150 species).

ADDITIONAL SIMULATION ANALYSES

To further investigate the cause of the errors in BMM, we estimated diversification rates for 30 sampled clades individually (Fig. 2; Table S5). That is, instead of letting BMM estimate rate shifts across each tree, we included only the focal clade, and set the number of rate shifts to 1. We found that this greatly improved the accuracy of the estimated diversification rates (see also Moore et al. 2016), supporting our hypothesis that inaccurate rate estimates from BMM are primarily caused by underestimating the number of rate shifts across each tree.

We also evaluated whether the accuracy of the MS estimators is influenced by the presence of heterogeneity in rates within clades (Fig. 3; Table S6). These analyses were conducted on all 66 pairs of sister clades across the 20 trees, with complete taxon sampling. We first calculated the accuracy of each MS estimator when rates were estimated separately for each clade in the pair (i.e., no heterogeneity in rates within each clade; Fig. 3A). We then calculated their accuracy when applied to each pair of clades, treating each pair as a single clade despite their different diversification rates, and compared the estimated rates to the average of the known rates for the pair (Fig. 3B). For matched methods (same epsilon, stem vs crown), accuracy was generally similar given constant versus heterogeneous rates within clades

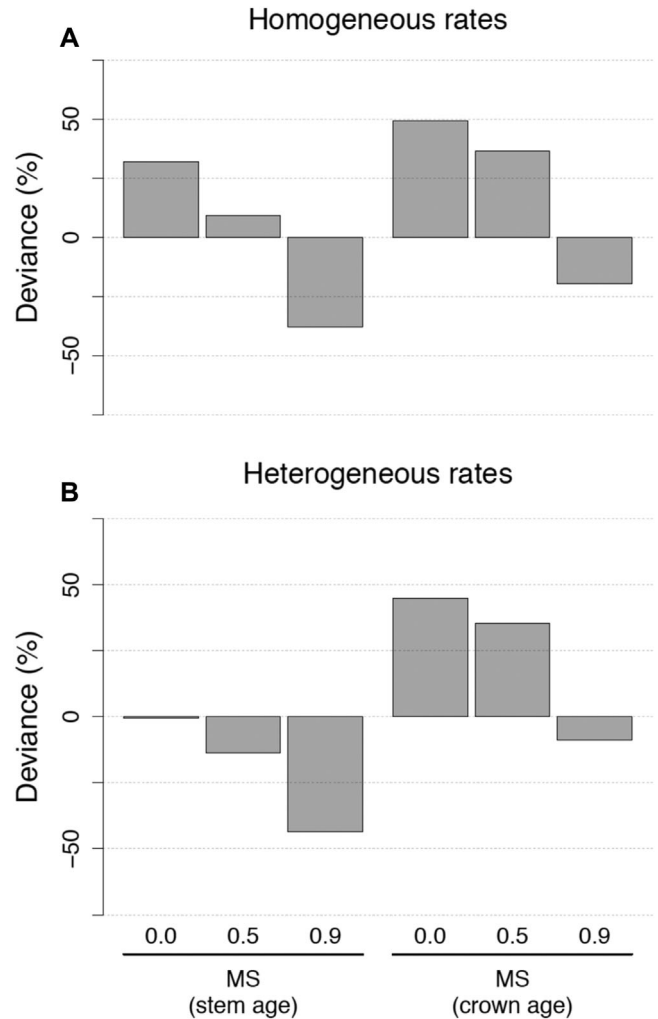


Figure 3. Comparison of the accuracy of MS estimators when the true diversification rates (speciation – extinction) are constant within clades (A) versus heterogeneous (B). Analyses are focused on all pairs of sister clades across all 20 trees ($n = 66$ pairs, including 132 of all 200 clades). (A) Accuracy when rates are constant within clades. The accuracy for the pair of clades was based on averaging their mean accuracy when rates are estimated separately for each clade. The values shown are the means of the averaged accuracy values across the 66 pairs. (B) Accuracy when rates are heterogeneous within clades, because each clade has a different rate, but each pair is treated as one clade (i.e., diversification rates were recalculated for each pair of clades using the combined richness of the two clades and the stem and crown ages for the common ancestor of each sister clade). Comparison of (A) and (B) shows that the accuracy of the MS estimators is similar when rates are constant versus heterogeneous within clades.

(Fig. 3A, B). Overall, the changes in error caused by heterogeneous rates were small relative to those between different MS estimators, and between MS estimators and BMM (Fig. 1).

We found that errors in both speciation and extinction rates contributed to errors in diversification rates with BMM. Errors

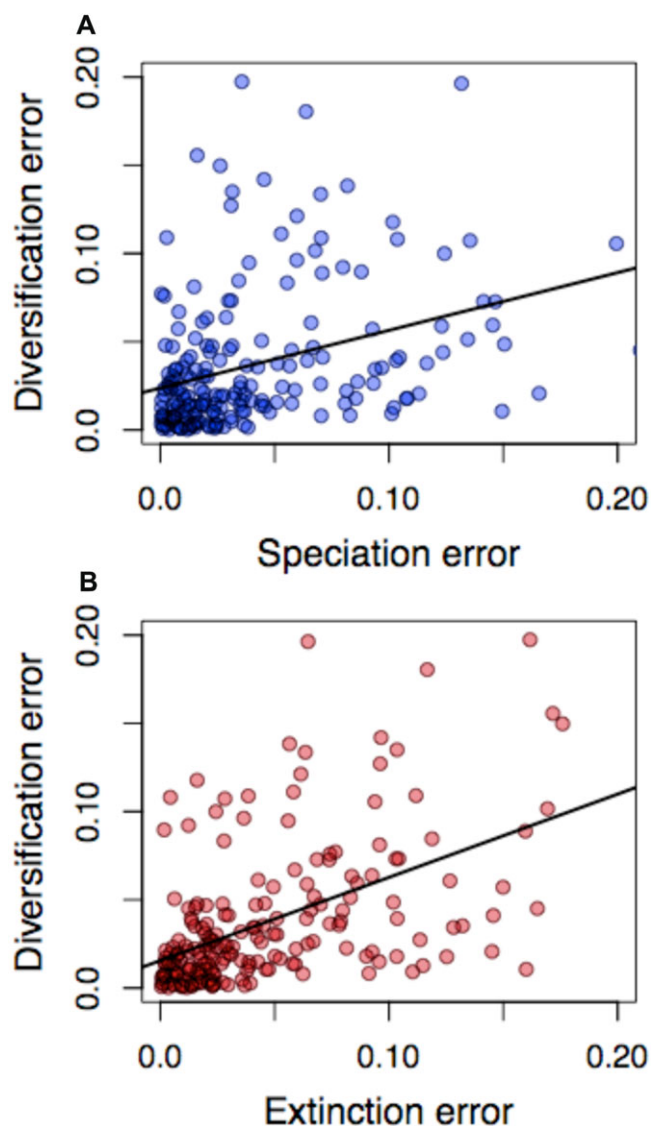


Figure 4. Relationship between absolute error in diversification rates and absolute error in (A) speciation rates, and (B) extinction rates, where all three rates are estimated using BAMM. Errors and rates are given in Table S1.

in diversification rates were significantly related to errors in both speciation and extinction rates, and to similar degrees (Fig. 4A: absolute error in diversification rate vs absolute error in speciation rate: $r^2 = 0.22$, $P < 0.001$; Fig. 4B: absolute diversification error vs absolute extinction error: $r^2 = 0.26$, $P < 0.001$). This result further supports the idea that the primary problem with BAMM is that it underestimates shifts in rate regimes across each tree, and thus fails to recognize variation among clades in both speciation and extinction rates. Most importantly, these results suggest that using BAMM only to estimate speciation rates does not ensure that those estimated speciation rates will be accurate.

We also evaluated how different the true diversification rates of sister clades must be for BAMM to recognize that these clades

belong to different rate regimes (Fig. S5). GLM analysis showed that the relationship between the difference in true rates and the probability of them being recognized as different by BAMM is not significant ($P = 0.74$). Using maximum likelihood, the estimated r^2 was 0.002, indicating very poor fit. Therefore, BAMM often failed to correctly identify different diversification-rate regimes, even when the magnitude of the true differences in rates was large.

Similarly, we also tested how often each method can correctly detect which of two sister clades has a higher diversification rate. The results (Fig. 5) show that BAMM failed to correctly identify the clade with the higher diversification rate much more frequently than any of the MS estimators.

EMPIRICAL ANALYSES

Finally, we compared BAMM-estimated diversification rates for 15 clades of snakes (each with >5 species) to see if estimates for each clade from an analysis of the entire tree were similar to those when the same clades were analyzed individually. In theory, these estimates should be similar, if not identical. Instead, the results (Fig. 6) showed only a weak relationship between these estimates that was not actually significant ($r^2 = 0.21$; $P = 0.08$). Two clades clearly played a large role in weakening this relationship: Pareatidae and Sybonophiinae (Table S4). These two clades are notable in having the smallest number of species among the 15 included clades. Their diversification rates were estimated to be relatively low when they were analyzed in isolation but were estimated to be much higher when the entire tree was analyzed simultaneously. Thus, these empirical results seem to closely parallel those from the simulations: BAMM overestimates diversification rates of smaller clades (Fig. 1).

Discussion

In this article, we show that a widely used method (BAMM) can give biased and inaccurate estimates of diversification rates, even for relatively simple simulated datasets. Our results show that this occurs because BAMM underestimates the number of rate shifts across trees, and assigns high diversification rates to smaller clades with low rates. We find that BAMM can be less accurate than the MS diversification estimators (Magallón and Sanderson 2001), yielding consistently weaker relationships between true and estimated diversification rates (Table 1). Several studies have claimed that the MS estimators are inaccurate, but without directly evaluating their ability to estimate known rates (e.g., Rabosky 2009, 2010; Rabosky and Adams 2012; Rabosky et al. 2012). Overall, given our results, we caution against using BAMM in empirical studies, even to estimate only speciation rates (Fig. 4). Similarly, we caution against using methods based on BAMM (e.g., STRAPP; Rabosky and Huang 2015). Our analyses also show problematic results from BAMM in empirical data, with

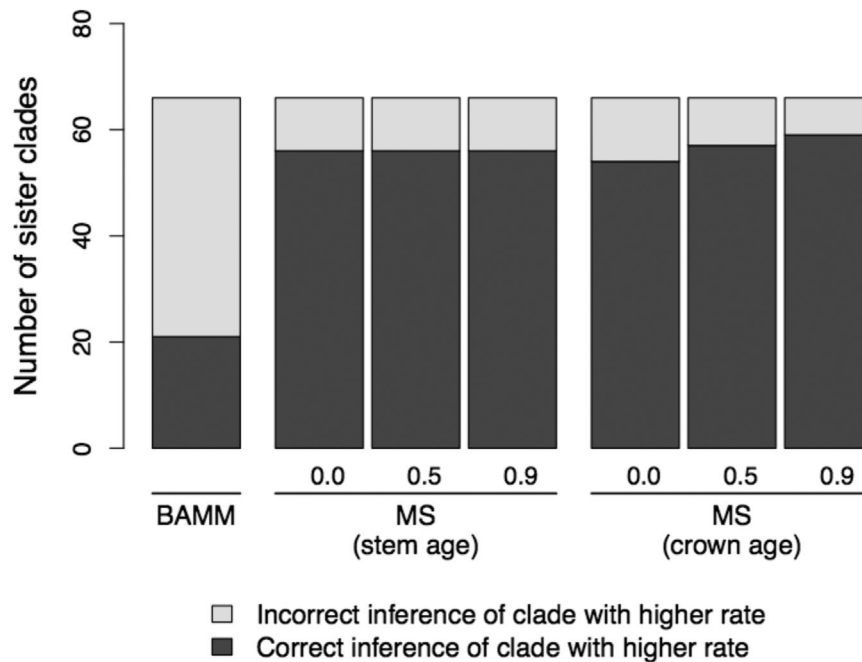


Figure 5. Barplot illustrating how often BMM and the MS estimators correctly identify which of two sister clades has the higher diversification rate. Dark gray bars show the frequency with which each method correctly identified the clade with the higher rate. Light gray bars show the frequency with which each method incorrectly identified the clade with the higher rate (including inferring that rates were equal when they were not).

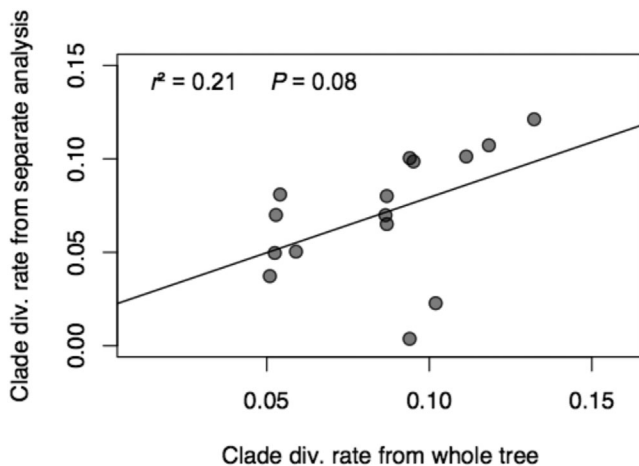


Figure 6. Relationships between diversification rates estimated by BMM in 15 clades of snakes (all with >5) species, in which rates are estimated across the entire tree of snakes and in which rates are estimated separately within each clade. Raw data are given in Table S4.

patterns that parallel those from simulations (Fig. 6). Another empirical study has found that BMM was less informative than the MS estimators in explaining species richness patterns among regions (Hutter et al. 2017). Our results also document important variation in the performance of the MS estimators (Fig. 1; Table 1).

Interestingly, our results are concordant with those of Moore et al. (2016) in showing that BMM may be problematic, but

the specific problems we highlight differ considerably. Here, we find that the number of diversification rate shifts across trees is consistently underestimated, leading to biased (overestimated) diversification rate estimates for smaller clades. Moore et al. (2016) did not focus on the accuracy of diversification rates estimated for higher level clades across large phylogenies. Instead, they estimated the overall relationship between true and estimated speciation and extinction rates for all branches of each tree (finding relationships to be very weak), focusing on relatively small simulated trees (~50–150 species). Rabosky et al. (2017) suggested that the weak relationships found by Moore et al. (2016) were an artifact of using many datasets with relatively invariant rates. Our results suggest that there can be significant relationships between true and estimated diversification rates using BMM (using demonstrably realistic rates), but that these relationships are consistently weaker than those using the MS estimators, sometimes by twofold or more (Table 1).

Our results also suggest a potential source of bias in analyses of the performance of diversification-rate estimators. By focusing on the rates of higher level clades (as do many empirical analyses), our simulations revealed that rates were often overestimated for small clades (Fig. 1). However, this problem could be missed in simulation studies that focus on the accuracy of rate estimates for individual branches (e.g., Rabosky 2014; Moore et al. 2016). The reason is that most species (and branches) may belong to large clades (which are more likely to be estimated correctly;

Fig. 1, Fig. S4). The impact of the substantial errors in estimating diversification rates for small clades may be overlooked if accuracy is assessed on a per-branch basis, instead of on a per-clade basis. Again, our interest here is in correctly estimating diversification rates of clades (e.g., higher taxa), both large and small, as in many empirical studies (e.g., Adams et al. 2009; De Lise and Rowe 2015; Puttick et al. 2015; Wiens 2015a,b; Cooney et al. 2016; Jezkova and Wiens 2017; Tedesco et al. 2017).

Our results show that the MS estimators can potentially outperform BAMM (Fig. 1; Table 1). However, our point here is not that these MS estimators are universally the best approach for estimating diversification rates (although they may be very useful given limited taxon sampling within clades, since many other methods require detailed trees within each clade). Instead, a more general point is that before new methods become widely adopted, there should be an unbiased appraisal of their accuracy, especially in comparison to other available methods. It is unclear whether this happened with BAMM, even as this method became widely used. Unfortunately, our results here suggest that empirical studies based on BAMM should be re-evaluated (see also Moore et al. 2016). Remarkably, despite how computationally intensive BAMM is, we found much stronger relationships between true and estimated rates by simply using the natural log of a clade's richness divided by its stem age (i.e. the stem MS estimator with $\epsilon = 0$; Table 1). We also found that heterogeneity in rates within clades had relatively little impact on the accuracy of the MS estimators (Fig. 3), despite the untested claim that these estimators require constant rates within clades (e.g., Rabosky et al. 2012).

In some ways, the goals of the MS estimators are much less ambitious than those of BAMM. Specifically, the MS estimators do not directly identify the specific location of diversification rate shifts on trees, nor the contribution of speciation and extinction rates to differences in diversification rates among clades. However, based on our results, BAMM performs poorly at these more challenging tasks, and its inability to perform them successfully reduces its ability to perform the simpler task of estimating diversification rates of individual clades. BAMM performs at least as well as the MS estimators for estimating the diversification rate of a single clade in isolation (Figs. 1 and 2). Moore et al. (2016) also found that BAMM was reasonably accurate when speciation and extinction rates were constant across a tree. Nevertheless, BAMM still has a serious disadvantage when applied in this simple case: it cannot be used to estimate diversification rates of single clades (in isolation) that have small numbers of species. Given that numerous higher taxa at all levels of the Tree of Life have 5 or fewer species (e.g., from phyla to genera), BAMM cannot be used to address why higher taxa have different diversification rates, at least not without seriously biasing the analysis by excluding the least diverse clades (or using BAMM in a way that generates

inaccurate rate estimates). Similarly, BAMM cannot be applied unless at least five species are sampled in the tree for each clade, regardless of the actual number of species. Thus, we reiterate our caution against using BAMM in analyses of diversification rates of higher-level taxa, even in the simple case of estimating a rate for a single clade with no rate shifts.

Our results also revealed important variation in the performance of the MS estimators. First, we found that the stem-group estimators were generally more accurate than the crown-group estimators, even with perfect sampling within clades (Fig. 1; Table 1), despite claims that crown-group estimators should be preferred (e.g., Stadler et al. 2014). Further, the error for the crown-group estimators generally increased with decreased taxon sampling. Nevertheless, we also found that the crown-group MS estimator performed surprisingly well using an ϵ of 0.90 (assuming high extinction rates), even when this ϵ value was clearly incorrect. This high ϵ value seemed to compensate for the general tendency of the crown-group estimators to overestimate diversification rates (in general and due to incomplete taxon sampling). Overall, we found that the stem-group estimator with an intermediate ϵ value or the crown-group estimator with high ϵ performed best across many different conditions (Table 1; Fig. 1).

As with any simulation study, there are many unrealistic aspects of our study and many potential sources of bias. We acknowledge that in our primary analyses, the clades among which rates varied were also the clades for which rates were estimated using the MS approach. Although this may favor the MS approach, we found that the MS approach was similarly accurate when pairs of clades (with potentially very different diversification rates) were treated as belonging to the same clade (Fig. 3). Furthermore, the BAMM analyses across each tree of 10 clades encountered considerable heterogeneity in rates, and yet failed to recognize much of this heterogeneity (i.e., underestimating rate shifts). Thus, including more heterogeneity in rates across each tree would not necessarily lead to more accurate rate estimates with BAMM (see also Moore et al. 2016).

In summary, our results show that BAMM can give biased and inaccurate estimates of diversification rates. This seemingly occurs because it underestimates the number of rate shifts across trees, and assigns high diversification rates to small clades with low rates. The relationship between true and estimated rates was substantially weaker than that of a much less computationally intensive approach (i.e., Magallón and Sanderson 2001), which performed relatively well. We suggest that published results based on BAMM should be viewed with considerable caution.

AUTHOR CONTRIBUTIONS

ALSM and JJW designed the study, performed analyses, and wrote the paper.

ACKNOWLEDGMENTS

We thank E.C. Miller, J. Tobias, and two anonymous reviewers for helpful comments on the manuscript. A.L.S.M. thanks CAPES Foundation, Ministry of Education of Brazil, for the fellowship (grant 10193/14-6) that supported his visit to Arizona to work on this project with J.J.W. Wiens was supported by U.S. National Science Foundation grant DEB 1655690.

DATA ARCHIVING

All data are provided in the Supporting Information and Dryad (<https://doi.org/10.5061/dryad.db561>).

LITERATURE CITED

- Adams, D. C., C. M. Berns, K. H. Kozak, and J. J. Wiens. 2009. Are rates of species diversification correlated with rates of morphological evolution? *Proc. R. Soc. Lond. B* 276:2729–2738.
- Alfaro, M. E., F. Santini, and C. D. Brock. 2007. Do reefs drive diversification in marine teleosts? Evidence from the pufferfishes and their allies (Order Tetraodontiformes). *Evolution* 61:2104–2126.
- Alfaro, M. E., F. Santini, C. D. Brock, H. Alamillo, A. Dornburg, G. Carnevale, D. L. Rabosky, and L. J. Harmon. 2009. Nine exceptional radiations plus high turnover explain species diversity in jawed vertebrates. *Proc. Natl. Acad. Sci. USA* 106:13410–13414.
- Bars-Closel, M., T. Kohlsdorf, D. S. Moen, and J. J. Wiens. 2017. Diversification rates are more strongly related to microhabitat than climate in squamate reptiles (lizards and snakes). *Evolution* 71:2243–2261.
- Blair, C., and S. Sanchez-Ramirez. 2016. Diversity-dependent cladogenesis throughout western Mexico: evolutionary biogeography of rattlesnakes (Viperidae: Crotalinae: *Crotalus* and *Sistrurus*). *Mol. Phylogenet. Evol.* 97:145–154.
- Condamine, F. L., F. A. H. Sperling, N. Wahlberg, J. Y. Rasplus, and G. J. Kergoat. 2012. What causes latitudinal gradients in species diversity? Evolutionary processes and ecological constraints on swallowtail biodiversity. *Ecol. Lett.* 15:267–277.
- Cooney, C. R., N. Seddon, and J. A. Tobias. 2016. Widespread correlations between climatic niche evolution and species diversification in birds. *J. Anim. Ecol.* 85:869–878.
- Davis, A. M., P. J. Unmack, R. P. Vari, and R. Betancur-R. 2016. Herbivory promotes dental disparification and macroevolutionary dynamics in grunners (Teleostei: Therapontidae), a freshwater adaptive radiation. *Am. Nat.* 187:320–333.
- De Lisle, S. P., and L. Rowe. 2015. Independent evolution of the sexes promotes amphibian diversification. *Proc. R. Soc. Lond. B* 282:20142213.
- Espeland, M., J. P. W. Hall, P. J. DeVries, D. C. Lees, M. Cornwall, Y.-F. Hsu, L.-W. Wu, D. L. Campbell, G. Talavera, R. Vila, et al. 2015. Ancient Neotropical origin and recent recolonisation: phylogeny, biogeography and diversification of the Riodinidae (Lepidoptera: Papilionoidea). *Mol. Phylogenet. Evol.* 93:296–306.
- Gómez-Rodríguez, C., A. Baselga, and J. J. Wiens. 2015. Is climatic niche width related to diversification rate? *Global Ecol. Biogeogr.* 24:383–395.
- Harmon, L. J., J. Weir, C. D. Brock, R. E. Glor, and W. Challenger. 2008. GEIGER: Investigating evolutionary radiations. *Bioinformatics* 24:961–964.
- Huang, J. P. 2016. The Great American Biotic Interchange and diversification history in beetles (Scarabaeidae; Dynastinae). *Zool. J. Linn. Soc.* 178:88–96.
- Hutter, C. R., S. M. Lambert, and J. J. Wiens. 2017. Rapid diversification and time explain amphibian species richness at different scales in the Tropical Andes, Earth's most biodiverse hotspot. *Am. Nat.* 190:828–843.
- Jackman, S. 2015. *pscl: classes and methods for R* developed in the Political Science Computational Laboratory, Stanford University. Department of Political Science, Stanford University, California. <http://pscl.stanford.edu/>
- Jezkova, T., and J. J. Wiens. 2017. What explains patterns of diversification and richness among animal phyla? *Am. Nat.* 189:201–212.
- Kozak, K. H., and J. J. Wiens. 2010. Accelerated rates of climatic-niche evolution underlie rapid species diversification. *Ecol. Lett.* 13:1378–1389.
- . 2016. Testing the relationships between diversification, species richness, and trait evolution. *Syst. Biol.* 65:975–988.
- Larson-Johnson, K. 2016. Phylogenetic investigation of the complex evolutionary history of dispersal mode and diversification rates across living and fossil Fagales. *New Phytol.* 209:418–435.
- Magallón, S., and M. J. Sanderson. 2001. Absolute diversification rates in angiosperm clades. *Evolution* 55:1762–1780.
- May, M. R., and B. R. Moore. 2016. How well can we detect lineage-specific diversification-rate shifts? A simulation study of sequential AIC methods. *Syst. Biol.* 65:1076–1084.
- Miller, E. C., and J. J. Wiens. 2017. Extinction and time help drive the marine-terrestrial biodiversity gradient: is the ocean a deathtrap? *Ecol. Lett.* 20:911–921.
- Moen, D. S., and J. J. Wiens. 2017. Microhabitat and climatic-niche change explain patterns of diversification among frog families. *Am. Nat.* 190:29–44.
- Moore, B. R., S. Höhna, M. R. May, B. Rannala, and J. P. Huelsenbeck. 2016. Critically evaluating the theory and performance of Bayesian analysis of macroevolutionary mixtures. *Proc. Natl. Acad. Sci. USA* 113:9569–9574.
- Morlon, H. 2014. Phylogenetic approaches for studying diversification. *Ecol. Lett.* 17:508–525.
- Peña, C., H. Witthauer, I. Klečková, Z. Fric, and N. Wahlberg. 2015. Adaptive radiations in butterflies: evolutionary history of the genus *Erebia* (Nymphalidae: Satyrinae). *Biol. J. Linn. Soc.* 116:449–467.
- Pennell, M. W., J. M. Eastman, G. J. Slater, J. W. Brown, J. C. Uyeda, R. G. FitzJohn, M. E. Alfaro, and L. J. Harmon. 2014. *geiger v2.0: an expanded suite of methods for fitting macroevolutionary models to phylogenetic trees*. *Bioinformatics* 15:2216–2218.
- Pontarp, M., and J. J. Wiens. 2017. The origin of species richness patterns along environmental gradients: uniting explanations based on time, diversification rate, and carrying capacity. *J. Biogeogr.* 44:722–735.
- Price, S. A., S. S. B. Hopkins, K. K. Smith, and V. L. Roth. 2012. Tempo of trophic evolution and its impact on mammalian diversification. *Proc. Natl. Acad. Sci. USA* 109:7008–7012.
- Puttick, M. N., J. Clark, and P. C. J. Donoghue. 2015. Size is not everything: rates of genome size evolution, not C-value, correlate with speciation in angiosperms. *Proc. R. Soc. Lond. B* 282:20152289.
- Pyron, R. A., and J. J. Wiens. 2013. Large-scale phylogenetic analyses reveal the causes of high tropical amphibian diversity. *Proc. R. Soc. Lond. B* 280:20131622.
- R Core Team. 2015. *R: a language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
- Rabosky, D. L. 2009. Ecological limits and diversification rate: alternative paradigms to explain the variation in species richness among clades and regions. *Ecol. Lett.* 12:735–743.
- . 2010. Primary controls on species richness in higher taxa. *Syst. Biol.* 59:634–645.
- . 2014. Automatic detection of key innovations, rate shifts, and diversity-dependence on phylogenetic trees. *PloS One* 9:e89543.

- Rabosky, D. L., and D. C. Adams. 2012. Rates of morphological evolution are correlated with species richness in salamanders. *Evolution* 66:1807–1818.
- Rabosky, D. L., and D. R. Matute. 2013. Macroevolutionary speciation rates are decoupled from the evolution of intrinsic reproductive isolation in *Drosophila* and birds. *Proc. Natl. Acad. Sci. USA* 110:15354–15359.
- Rabosky, D. L., and H. Huang. 2015. A robust semi-parametric test for detecting trait-dependent diversification. *Syst. Biol.* 65:181–193.
- Rabosky, D. L., G. J. Slater, and M. E. Alfaro. 2012. Clade age and species richness are decoupled across the eukaryotic tree of life. *PLoS Biol.* 10:e1001381.
- Rabosky, D. L., F. Santini, J. Eastman, S. A. Smith, B. Sidlauskas, J. Chang, and M. E. Alfaro. 2013. Rates of speciation and morphological evolution are correlated across the largest vertebrate radiation. *Nat. Commun.* 4:1958.
- Rabosky, D. L., M. Grundler, C. Anderson, P. Title, J. J. Shi, J. W. Brown, H. Huang, and J. G. Larson. 2014. BAMMtools: an R package for the analysis of evolutionary dynamics on phylogenetic trees. *Methods Ecol. Evol.* 5:701–707.
- Rabosky, D. L., J. S. Mitchell, and J. Chang. 2017. Is BAMM flawed? Theoretical and practical concerns in the analysis of multi-rate diversification models. *Syst. Biol.* 66:477–498.
- Revell, L. J. 2012. Phytools: an R package for phylogenetic comparative biology (and other things). *Methods Ecol. Evol.* 3:217–223.
- Ricklefs, R. E. 2007. Estimating diversification rates from phylogenetic information. *Trends Ecol. Evol.* 22:601–610.
- Rolland, J., F. L. Condamine, F. Jiguet, and H. Morlon. 2014. Faster speciation and reduced extinction in the tropics contribute to the mammalian latitudinal diversity gradient. *PLoS Biol.* 12:e1001775.
- Schilling, E. E., J. L. Panero, B. S. Crozier, R. W. Scott, and P. Dávila. 2015. Bricklebrush (*Brickellia*) phylogeny reveals dimensions of the great Asteraceae radiation in Mexico. *Mol. Phylogenet. Evol.* 85: 161–170.
- Scholl, J. P., and J. J. Wiens. 2016. Diversification rates and species richness across the Tree of Life. *Proc. R. Soc. Lond. B* 283:20161335.
- Spriggs, E. L., W. L. Clement, P. W. Sweeney, S. Madrinan, E. J. Edwards, and M. J. Donoghue. 2015. Temperate radiations and dying embers of a tropical past: the diversification of *Viburnum*. *New Phytol.* 207:340–354.
- Stadler, T., D. L. Rabosky, R. E. Ricklefs, and F. Bokma. 2014. On age and species richness of higher taxa. *Am. Nat.* 184:447–455.
- Stanley, S. M. 1979. *Macroevolution*. W. H. Freeman, San Francisco, CA.
- Tedesco, P. A., E. Paradis, C. Lévêque, and B. Hugueny. 2017. Explaining global-scale diversification patterns in actinopterygian fishes. *J. Biogeogr.* 44:773–783.
- Uetz, P., P. Freed, and J. Hošek. 2017. The Reptile Database, <http://www.reptile-database.org>, accessed in 22 January 2017.
- Weber, M. G., and A. Agrawal. 2014. Defense mutualisms enhance plant diversification. *Proc. Natl. Acad. Sci. USA* 111:16442–16447.
- Wiens, J. J. 2015a. Explaining large-scale patterns of vertebrate diversity. *Biol. Lett.* 11:20150506.
- . 2015b. Faster diversification on land than sea helps explain global biodiversity patterns among habitats and animal phyla. *Ecol. Lett.* 18:1234–1241.
- . 2017. What explains patterns of biodiversity across the Tree of Life? *BioEssays* 39:1600128.
- Wiens, J. J., R. T. Lapoint, and N. K. Whiteman. 2015. Herbivory increases diversification across insect clades. *Nat. Commun.* 6:8370.
- Zheng, Y., and J. J. Wiens. 2016. Combining phylogenomic and supermatrix approaches, and a time-calibrated phylogeny for squamate reptiles (lizards and snakes) based on 52 genes and 4,162 species. *Mol. Phylogenet. Evol.* 94:537–547.

Associate Editor: J. Tobias
Handling Editor: M. Noor

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's website:

Figure S1. Summary of the 20 phylogenies and 200 clades that were simulated and analyzed in this study. Labels near the tips indicate the number of species in each clade. Branch colors show the simulated diversification rates of the clades. Black indicates the backbone tree, which had relatively low, invariant diversification rates. The overall length of each tree is 100 million years.

Figure S2. BAMM estimates of diversification rates through time (black lines) with confidence intervals (gray). Rates were estimated for the first three trees (columns). Each line shows estimates with different number of shifts (i.e. 10, 5, 3, and 1 shift).

Figure S3. Summary of the 95% higher posterior interval (HPI) for diversification rates estimated using BAMM. Panels show estimates for all 200 clades (1–100 in the top panel, 101–200 in the bottom panel). The bars represent the upper and lowest limits of the estimated 95% HPI. Circles show the estimated mean diversification rate. Triangles represent the true diversification rate. Triangles are blue when the true diversification rate is within the estimated 95% HPI, and red when the rate is outside.

Figure S4. Accuracy of BAMM and MS estimators for estimating diversification rates of clades, including stem and crown group MS estimators with different values of epsilon (0, 0.5, and 0.9), but using the mean absolute deviance between true and estimated rates (rather than the mean deviance, as in Fig. 1). (A) The y-axis shows the mean absolute percentage deviance of the estimated diversification rates from the true diversification rates across all 200 clades. The first column of graphs (Full Sampling) shows the results based on complete sampling of all species in each clade, while the second and third columns show the results when only 50% and 25% of the species in each clade were included in the tree when estimating diversification rates. (b) Separate results for the 61 clades with 150 species or more. (c) Separate results for the 139 clades with less than 150 species. Full results are given in online Table S1.

Figure S5. Probability that BAMM will estimate a regime shift as a function of the magnitude of the difference in true diversification rates between sister clades ($n = 66$ pairs of sister clades). Points show pairs of sister clades in which BAMM assigned a single regime (0) or two different regimes (1). The black line shows the function curve.

Table S1. Main results including clade ages, species richness, simulated speciation, extinction, and diversification rates, and estimated speciation, extinction, and diversification rates from BAMM (and their errors) and estimated diversification rates from the MS estimators (and errors in those estimates). Effective sample size and number of shifts with the maximum a posteriori probability are also shown for BAMM estimates.

Table S2. Results from BAMM comparing the impact of discarding 10% or 50% of the posterior distribution of trees as burnin.

Table S3. Results from BAMM comparing the impact of using a prior of 10 rate shifts and those obtained using 5, 3, and 1 rate shifts.

Table S4. Results and data for BAMM analyses of 15 snake clades, including the number of species in each clade that were included in the phylogeny used, the total number of described species in each clade, the sampling fraction (the proportion of described species included in the tree), the diversification rate estimated by BAMM when analyzing the whole tree simultaneously, and the diversification rate from BAMM when analyzing each clade separately.

Table S5. Estimates of speciation, extinction, and net diversification rates (and errors in all three rates), when each clade was analyzed separately using BAMM, along with effective sample sizes. Because these analyses were very computationally intensive, they were performed only for the clades in the first three backbone trees (total of 30 clades). When clade size was equal to or smaller than five species (i.e. when sampling only 50% or 25% of the species), BAMM could not estimate speciation, extinction, and net diversification rates. These are listed as NA.

Table S6. Estimates of net diversification rates using MS estimators for all pairs of sister clades across all 20 trees ($n = 66$ pairs, including 132 of all 200 clades). Results include analyses in which each pair is treated as one clade (i.e. rates are heterogeneous within clades), and analyses in which rates were estimated separately for each clade (i.e. rates are constant within clades) and the error is averaged for the two clades.

Supplementary File 1. The 20 trees with full species-level sampling.

Supplementary File 2. The 20 trees with 50% species-level sampling.

Supplementary File 3. The 20 trees with 25% species-level sampling.