doi:10.1111/evo.12614

Extinction can be estimated from moderately sized molecular phylogenies

Jeremy M. Beaulieu[1,2] and Brian C. O'Meara[3]

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

[2]*E-mail: jbeaulieu@nimbios.org*

[3]*Department of Ecology and Evolutionary Biology, University of Tennessee, Knoxville, Tennessee 37996*

Received July 3, 2014
Accepted January 3, 2015

Hundreds of studies have been dedicated to estimating speciation and extinction from phylogenies of extant species. Although it has long been known that estimates of extinction rates using trees of extant organisms are often uncertain, an influential paper by Rabosky (2010) suggested that when birth rates vary continuously across the tree, estimates of the extinction fraction (i.e., extinction rate/speciation rate) will appear strongly bimodal, with a peak suggesting no extinction and a peak implying speciation and extinction rates are approaching equality. On the basis of these results, and the realistic nature of this form of rate variation, it is now generally assumed by many practitioners that extinction cannot be understood from molecular phylogenies alone. Here, we reevaluated and extended the analyses of Rabosky (2010) and come to the opposite conclusion—namely, that it is possible to estimate extinction from molecular phylogenies, even with model violations due to heritable variation in diversification rate. Note that while it may be tempting to interpret our study as advocating the application of simple birth–death models, our goal here is to show how a particular model violation does not necessitate the abandonment of an entire field: use prudent caution, but do not abandon all hope.

KEY WORDS: **Birth–death, diversification, extinction, molecular phylogenies, speciation.**

The diversity of life—both in its narrow meaning of number of species and its general meaning of life's astounding morphological, behavioral, and biochemical variety—is shaped by the complementary forces of speciation and extinction. Several hundred studies have been dedicated to investigating these forces, using phylogenies of extant species, information from fossils, or both. Although uncertainty in estimates of extinction using trees of extant organisms is well recognized (Nee et al. 1994), an influential paper by Rabosky (2010) made a far blunter point, summarized by its title: "Extinction rates should not be estimated from molecular phylogenies." This paper's impact (>150 citations since its publication) came from its important topic and its seemingly clear results. It also serves as a surprising but good example of a scientist who had developed many of the key methods in an area (e.g., Rabosky 2006, 2009; Rabosky and Lovette 2008), thinking carefully about potential problems, and, based on analyses, changing his mind about the appropriateness of his own and others' methods to the most available data type. However, we find that this conclusion was far too sweeping. In fact, it *is* possible to estimate extinction from molecular phylogenies, even with model violations due to heritable variation in diversification rate and drastic subsampling, as long as the phylogenies are of ample size and appropriate corrections for subsampling are incorporated.

There are a number of likelihood-based models for characterizing diversification through time from molecular data (e.g., Nee et al. 1992; Harmon et al. 2003; Moore et al. 2004; Rabosky 2006, 2009; Bokma 2008; Rabosky and Lovette 2008; Morlon et al. 2010, 2011). The fundamental assumption in all these models is that functions of speciation and extinction leave distinct signatures on the branching structure in any given phylogeny (Nee et al. 1994). For instance, as the ratio of the extinction rate (μ) to the speciation rate (λ)—or the extinction fraction (denoted by ε)—increases, species become increasingly vulnerable to extinction, resulting in increased rates of species turnover through time. The net diversification rate (speciation

© 2015 The Author(s). *Evolution* © 2015 The Society for the Study of Evolution.

EVOLUTION INTERNATIONAL JOURNAL OF ORGANIC EVOLUTION

rates minus extinction rates, denoted by r), on the other hand, controls the pace by which species accumulate.

In theory, diversification rates and the extinction fraction should have a distinguishable effect on age structuring (i.e., tree shape) and branch length distributions (Nee et al. 1994). Thus, it should be possible to estimate each. However, many methods assume these rates have been constant through time and across lineages in a group of interest (but see Morlon et al. 2011; Stadler 2011; Etienne et al. 2012; Rabosky 2014). Yet, the observed diversity distributed across clades is often highly imbalanced, reflecting differences in the diversification rate. Rabosky (2010) demonstrated that continuous variation in net diversification rates could severely affect the ability to recover true extinction fractions. Such heterogeneity within a tree can cause the age structuring of clades to be biased with respect to the true underlying extinction rate. For example, in the absence of extinction, descendants of a lineage with an increased diversification rate will inherit this new rate, and cause a disproportionate number of lineages to arise toward the present (Rabosky 2010). This can often lead to erroneously high estimates of ε. In Rabosky's analysis, the distribution of ε estimates can appear strongly bimodal, with a peak suggesting no extinction (i.e., $\varepsilon = 0$) and a peak implying speciation and extinction rates are approaching equality (i.e., $\varepsilon = 1$).

In the context of evaluating the performance of our own newly developed diversification model, and given the continued development and use of approaches that still include estimation of extinction from molecular trees (Morlon et al. 2011; FitzJohn 2012; Rabosky 2014), we replicated and extended the analyses of Rabosky (2010), and found very different conclusions. Briefly, trees were generated by first starting from a root node that had two living descendants, and then stepping through time allowing speciation rates (λ) to vary along branches while keeping ε constant. Speciation rates for each branch were drawn from a lognormal distribution with a mean equal to the value of the parent branch and a SD equal to $\sigma\sqrt{T}$, where σ describes the magnitude of rate heterogeneity among branches and T is the length of the parent branch. Note our use of $\sigma\sqrt{T}$ as opposed to σT used by Rabosky (2010). When speciation rates are allowed to evolve under a continuous-time process analogous to Brownian motion (see Heard 1996), the variance should increase as $\sigma^2 T$—thus, the SD should increase as $\sigma\sqrt{T}$. The parameterization of Rabosky (2010) biases longer branches to accumulate more variance than if the variance had accumulated along piecewise segments of the same branch. Take the expected variance at three species, A, B, and C, all sharing a common ancestor at 10 million years ago (MYA). If A has had no speciation events leading to it, its expected variance is $10^2\sigma^2$ under the original formulation, $10\sigma^2$ under ours. If on the path to B there was a speciation event one million year (MY) after the root node, its expected variance is $1^2\sigma^2 + 9^2\sigma^2 = 82\sigma^2$ under the original formulation, $1\sigma^2 + 9\sigma^2 = 10\sigma^2$ under ours. If on the path to C there was a speciation event 5 MY after the root node, its expected variance is $5^2\sigma^2 + 5^2\sigma^2 = 50\sigma^2$ under the original formulation, $5\sigma^2 + 5\sigma^2 = 10\sigma^2$ under ours.

Following Rabosky (2010), simulations were conducted under four relative extinction scenarios (i.e., $\varepsilon = 0.0$, 0.25, 0.50, and 0.75) with ε being kept constant by reparameterizing the extinction rate, μ, on each branch once a new speciation rate was drawn. The initial speciation rates for a given value of ε used by Rabosky (2010) were based on the assumption that the average diversity is roughly $N = 50$ after 50 MY. However, with the exception of when $\varepsilon = 0.0$, these speciation rates do not fit this expectation. To correct this, we estimated a new set of initial speciation rates for each relative extinction rate by optimizing equation (5) in Magallón and Sanderson (2001): $\lambda = 0.064$, 0.084, 0.118, and 0.196 speciation events per million years for $\varepsilon = 0.0$, 0.25, 0.50, and 0.75, respectively. Also, because our primary interest was to examine the behavior of the two extremes of heritable from the original Rabosky (2010) study, we only performed simulations with σ set to 0 and 0.06 speciation events per $MY^{3/2}$ (the SD of speciation rates is in units of speciation events per MY and is equal to σ times the square root of the branch length, so the units of σ are events per $MY^{3/2}$). For each combination of σ and ε, 2000 trees were simulated using modified R scripts used in the original study of Rabosky (2010). Estimates of ε were obtained by fitting a constant birth–death model to each tree using the R package *diversitree* (FitzJohn 2012) with parameters optimized using the subplex algorithm without bounds. To assess the possible influence of different optimization routines, we also estimated ε using the Nelder–Mead algorithm with the parameter search on ε bounded on the interval (0,1).

In the absence of heritable rate variation (i.e., $\sigma = 0.00$), estimates of ε were consistent with those of Rabosky (2010)—that is, large confidence intervals surrounded reasonably unbiased estimates of ε. For instance, under a simulation model without extinction ($\varepsilon = 0.0$), the median estimate was 0.055, but with a 95% confidence interval that included both 0 and 1 (95% CI = [0.00,1.12]). When speciation rates were allowed to vary (i.e., $\sigma = 0.06$), there was an upward bias in our estimates of ε, with the median value of 0.20 being above the true value of zero. Interestingly, simulations with the greatest extinction fraction (i.e., $\varepsilon = 0.75$) showed the least amount of bias, with the median estimate of ε centered on 0.799, although this was accompanied by a large confidence interval (95% CI = 0.00–1.12).

It is also worth noting that there were no differences in the actual results when the parameter search of ε was bounded on the interval [0,1] (see Fig. 1), though the binning of estimates of ε that were greater than 1 creates a peak at 1, as in Rabosky (2010) Figures 2 and 3. We note that, biologically, having an extinction rate greater than the speciation rate is unlikely but entirely possible, and has been shown for groups with particular traits (e.g.,

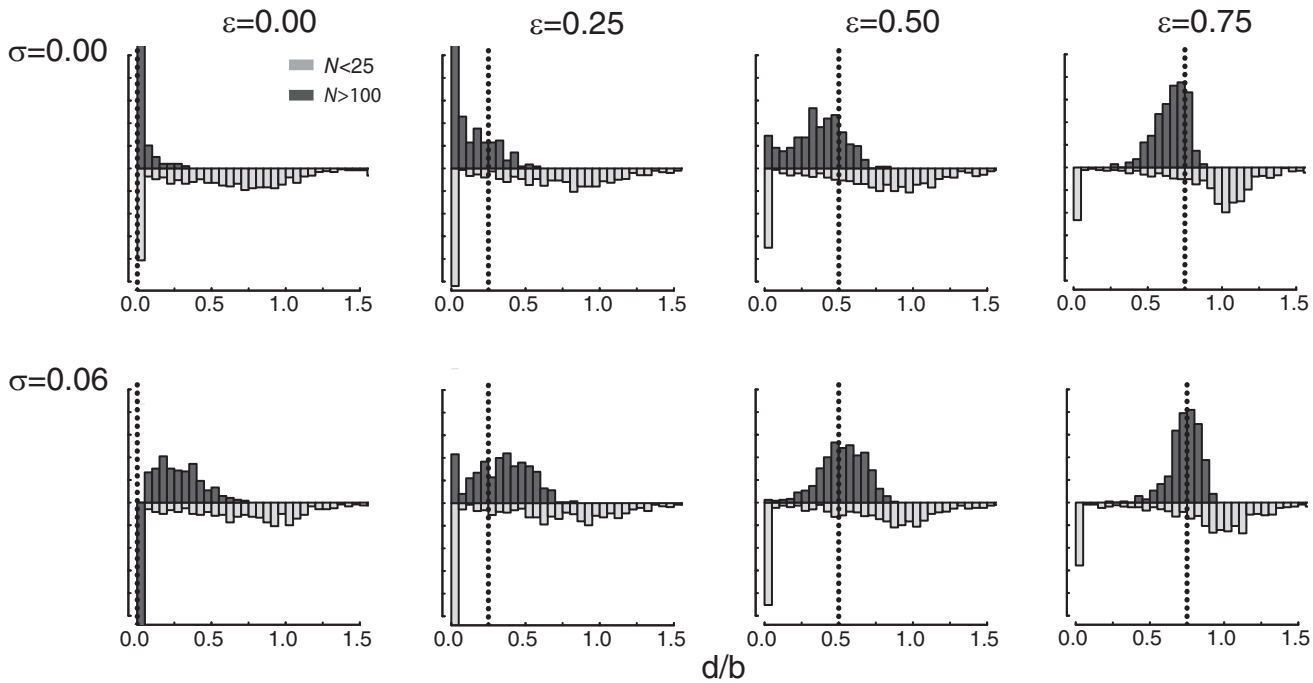

Figure 1. Comparisons of ε estimated with and without speciation rates varying among lineages (i.e., $\sigma = 0.0$ vs. $\sigma = 0.06$ speciation events/MY$^{3/2}$, respectively). Speciation rates are allowed to evolve under a continuous-time process analogous to Brownian motion that assumes σ increases as $\sigma\sqrt{T}$, as opposed to σT used by Rabosky (2010). Histograms depict frequencies of ε according to differences separated whether the tree was $N < 25$ (light gray) or $N > 100$ (dark gray). The dotted line represents the values of ε used in the simulation model (ε = 0.00, 0.25, 0.50, and 0.75). As tree size increases, there is a considerable decrease in the bias even in presence of heritable variation in speciation, the magnitude of which clearly increases as the simulated value of ε increased.

Goldberg et al. 2010). Over the long term, the expectation for any group with a positive extinction rate is complete extinction, but this does not preclude seeing such groups at any moment in time.

We also observed a general influence of tree size in our simulations. When we examined the distribution of ε from simulated trees that contain >100 taxa, there was a considerable decrease in the bias in the presence of heritable rate variation, the magnitude of which increased as the simulated value of ε increased (Fig. 1). There was also a considerable improvement in the width of the associated confidence intervals, with the upper bound consistently falling well below 1 (e.g., $\varepsilon = 0.0$, $\sigma = 0.06$: 95% CI = 0.00–0.58). Even under a simulation model with ε set to 0.75, there was not a single estimate of $\varepsilon > 1$ (max $\varepsilon = 0.95$; Fig. 1). By contrast, when we examined trees that contained fewer than 25 taxa, estimates of ε generally tended toward 0 or >1, regardless of the values of the simulation model (Fig. 1).

There is the possibility, however, that the influence of tree size in this context is somewhat misleading. The distribution of ε from simulated trees that contain diversity either above or below a certain expectation may largely reflect randomly fluctuating speciation and extinction rates. For example, for a simulated tree to exceed the expected diversity over some time interval, a run of speciation events would be necessary, which could drive estimates of ε down. Conversely, a run of extinction events would result in a tree that contained much less diversity than expected, leading to generally higher estimates of ε for a given simulated model. Such biases due to tree size could, therefore, mask the true direction and magnitude of the bias due to model misspecification when examining estimates of ε in the context of varying tree sizes.

To correct for this bias, and to better examine the influence of tree size, we conducted a different set of simulations. We used the same combinations of σ and ε described above, but here we specified a new set of initial speciation rates to simulate trees with an expected diversity of $N = 25$ ($\lambda = 0.051, 0.066, 0.091, 0.145$, for $\varepsilon = 0.0, 0.25, 0.50$, and 0.75, respectively) and $N = 100$ ($\lambda = 0.078, 0.103, 0.145, 0.249$). However, rather than examining the distribution of ε across all simulated trees in each scenario, we only examined the likelihood profile from 20 simulated trees that contained exactly the expected number of species (i.e., $N = 25$ or 100 taxa), discarding all others.

Results are shown in Figure 2. The filled contour lines, corresponding to the average two-unit support region, clearly show that even in the presence of heritable rate variation the estimates of ε are centered on the true value (rather than around 0 or 1), albeit occasionally with large confidence intervals. They also show that as tree size increases there is also a clear contraction in the

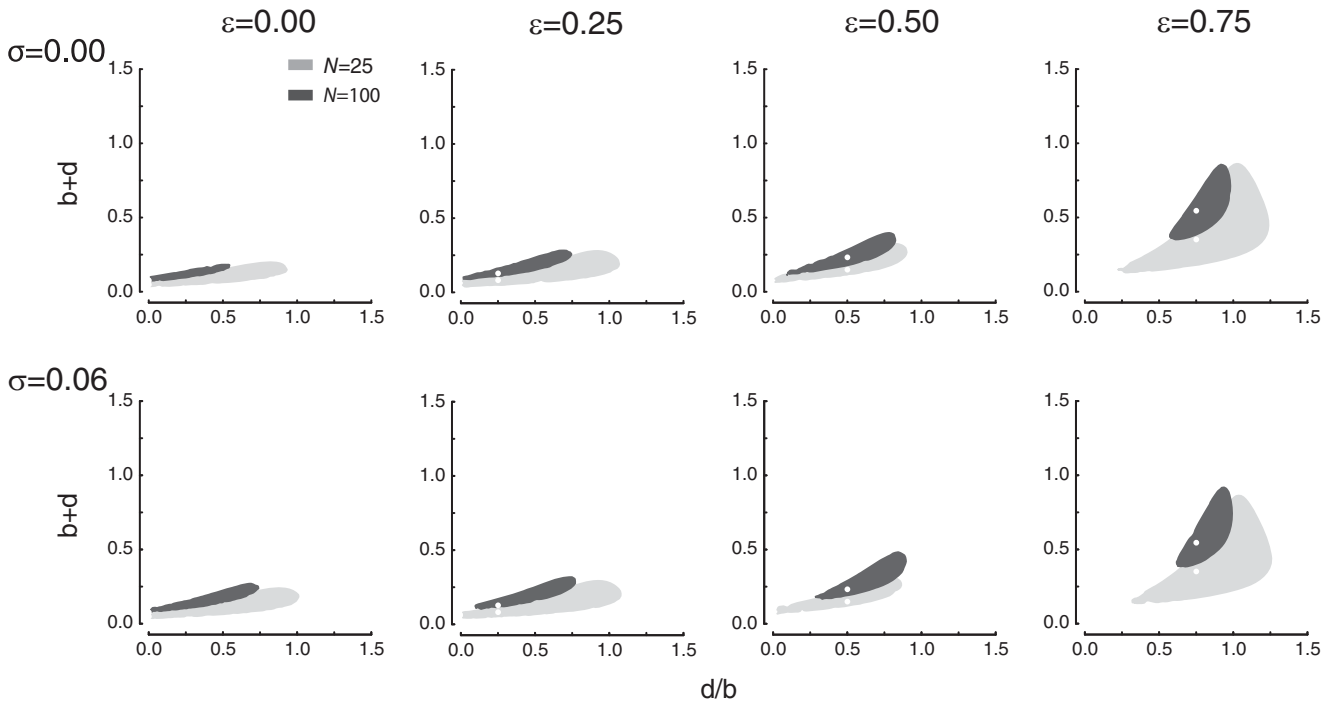

Figure 2. The influence of tree size on the likelihood profile of ε. Here we specified a different set of initial speciation rates in order simulate trees with an expected diversity of $N = 25$ ($\lambda = 0.051$, 0.067, 0.101, and 0.202 speciation events per MY, for $\varepsilon = 0.00$, 0.25, 0.50, and 0.75, respectively) and $N = 100$ ($\lambda = 0.078$, 0.104, 0.157, and 0.313 speciation events per MY). The filled contour lines correspond to the average two-unit support region from simulated trees that contained exactly the expected number of species (i.e., $N = 25$ or 100 taxa). Results indicate that even in the presence of heritable variation in speciation estimates of ε are centered on the true value (denoted by a white dot), with increases in tree size also resulting in a reduction in the confidence limits implied by the support region.

confidence limits implied by the support region. In the case of trees simulated under a high extinction fraction ($\varepsilon = 0.75$), we further explored the effect of heritable rate variation by incrementally doubling the value of σ up to a large value (see below). Overall, as σ increases the support region does appear to be moving away from the true value toward $\varepsilon = 1$ (Fig. 3). However, only when σ is 0.24 speciation events per $MY^{3/2}$, which is a rate so high that the SD of speciation rate is equal to the speciation rate after just one million year, does the support region actually exclude the true value.

Of course, this raises an important question about σ—namely, what is an appropriate or even biological meaningful value in all these cases? Most methods are robust to small deviations between the true model and the assumed model, but extreme enough differences can result in failure for almost any method. In our case, we need to show that values of σ that are likely higher than those found in nature still result in reasonable estimates of ε. To provide some context to the overall spread in speciation rates as a consequence of σ, we estimated the average speciation rate across a large set of flowering plant clades using equation (5) in Magallón and Sanderson (2001). We assumed $\varepsilon = 0.75$ and estimated speciation rates using information compiled by Magallón and Castillo (2009), detailing divergence time estimates of many exceptionally diverse angiosperm crown clades and their known standing diversity. We then converted these speciation rates to reflect the average wait time between speciation events (i.e., $1/\lambda$), which we compared to the overall spread in the average wait times from 100 trees simulated assuming a mean diversity of $N = 50$ described above (Fig. 3B, C).

Before proceeding, it is important to expand a bit more on what the wait times of Magallón and Castillo (2009) are meant to demonstrate. The effect of heritable variation in speciation rate is to have species sorting: those with the fastest wait times have more offspring, which causes the overall mean to decline. More variation due to higher values of σ only accelerates this process. So, here we are using the datasets of Magallón and Castillo (2009) to provide a baseline rate for determining what values of σ would produce the number of species we observe in real world clades of flowering plants. To put it another way, consider an estimate of variation in the interbirth interval of human females. It may vary from 2 to 10 years, with an unknown amount of variation within a population. However, if a simulation generates an interbirth interval of just two months, it clearly indicates too much variation is being generated.

It is quite clear that under these conditions remarkably high speciation rates can accumulate rather quickly under ever increasing values of σ. The model of heritable variation in speciation,

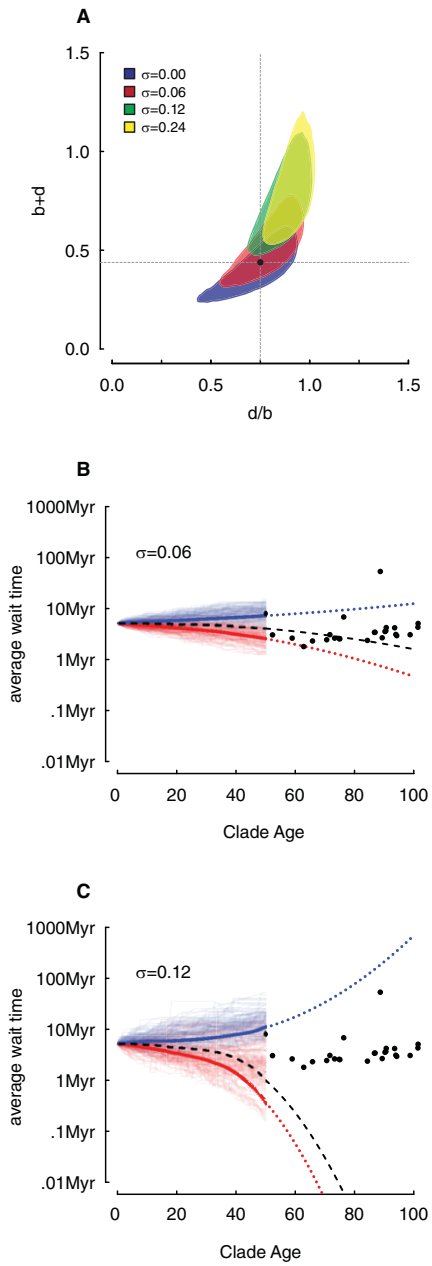

Figure 3. (A). Likelihood profiles of ε for further increases in σ (σ = 0.00, 0.06, 0.12, and 0.24 speciation events per $MY^{3/2}$) when the true ε = 0.75 and *N* = 100. Only when σ = 0.24 events per $MY^{3/2}$ does the support region exclude the true value. (B and C). The trajectories of the lowest (solid blue lines), highest (solid red lines), and mean (dashed black lines) average wait times per lineage (1/λ) obtained from 100 simulated trees that assume ε = 0.75 and a mean diversity of *N* = 50 (refer to Figure 1); these wait times were also interpolated across an additional 50 MY interval (dashed red and blue lines). In the case of σ = 0.12 events/$MY^{3/2}$, it would only take the simulation an additional 20 MY for lineages to evolve rates that implied speciation events occur, on average, once every 10,000 years per lineage. These extreme rates result in wait times far exceeding the average wait times inferred for a large set of exceptionally diverse flowering plant clades of nearly the same age (black dots).

both here and in Rabosky (2010), has no net bias in increasing or decreasing speciation rate within a lineage. However, by keeping extinction fraction constant, it causes selection for diversification rate: for a given ε of 0.5, for example, a species with λ of 3 would have a net diversification rate (i.e., λ − ελ) of 1.5, whereas a species with a λ of 1 would have a net diversification rate one-third as high. As the model has heritable speciation rates, and species with higher speciation rates have more offspring on average, this creates a runaway species selection process, leading to faster diversification rates and faster turnover rates. However, even when we assume a fairly high σ (i.e., σ = 0.06 events/$MY^{3/2}$) that still allows reasonable estimation of ε, the wait times until speciation estimated for all major angiosperm clades—including exceptionally diverse groups, such as Asterales (i.e., sunflowers and relatives; approximately 27,000 spp.), Lamiales (i.e., mints and relatives; approximately 23,000 spp.), and Asparagales (i.e., orchids and relatives; approximately 29,000 spp.)—never exceed the higher end of wait times across the simulated trees (Fig. 3B, C). This indicates that even a value of 0.06 generates rates more extreme than reality. Furthermore, in the case of σ = 0.12 events/$MY^{3/2}$, had we allowed the simulation to exceed 50 MY, we estimate that it would only take roughly an additional 20 MY for lineages to have evolved rates that implied speciation events occur once every 10,000 years per lineage on average. Thus, this particular model, where variance in speciation rates accumulates linearly with time, clearly represents a biologically extreme scenario. Even if this model is operating, however, most systems will be in an area of parameter space where estimation using traditional approaches will work (Fig. 3A).

On the whole, we find that larger tree sizes alone are enough to overcome the inherent biases associated with estimates of ε brought about by heritable variation in diversification rates. For trees comprised of only 100 taxa, even widespread heterogeneity in speciation rates does not prevent likelihood-based methods from estimating extinction fraction with acceptable precision. And even for very small trees, ε will often be distinguishable, even though such estimates will be associated with a large degree of uncertainty.

We caution that the model we, and Rabosky (2010), simulated under is only one of many ways to include heterogeneity in rates. It matches reality in that some clades have higher rates, but it effectively allows the rate of diversification to evolve ad infinitum. Those lineages that happen to have higher diversification rates generally leave more descendants, who then inherit this rate, with modification. Over the long run, however, this means either incredibly high diversification rates evolve or the mutation rate (and thus heterogeneity) must be quite low to slow the rate of evolution. There may be other methods for modeling rate heterogeneity without this behavior. For example, Maddison (2006) showed that if rate heterogeneity due to a discrete character state

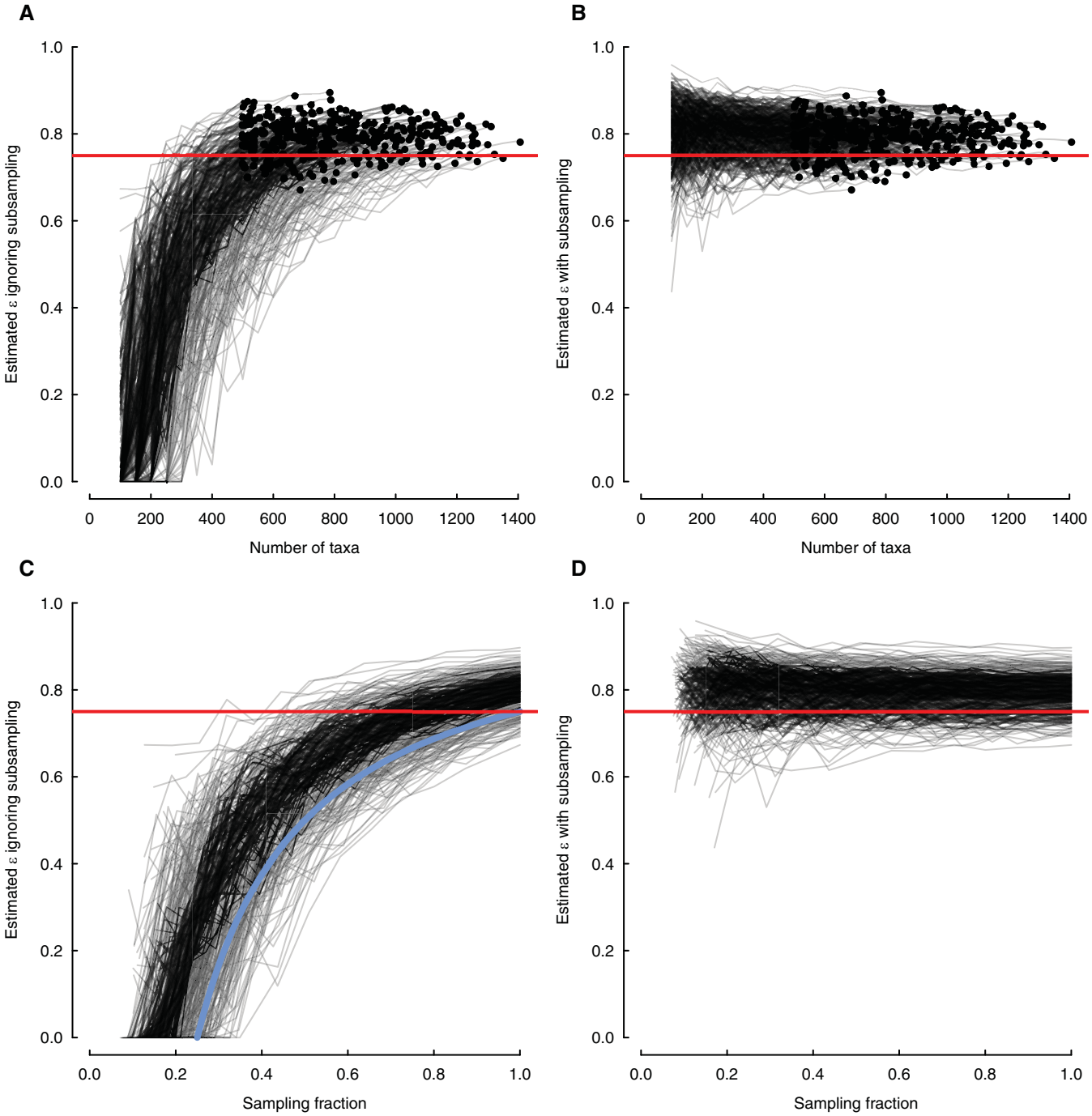

Figure 4. The robustness of subsampling corrections to model violations. Trees were simulated using a heterogeneity model (initial $\lambda = 0.376$ speciation events per million year, $\sigma = 0.06$ speciation events per $MY^{3/2}$) with $\varepsilon = 0.75$ (red line) and only retaining trees that contain 500 or more taxa; the black dots in (A) and (B) represent the final diversity contained in each tree examined. Blue line in (C) represents the predicted estimate of ε given by ignoring sampling using a model with no heterogeneity, calculated by manipulating equations from Stadler (2013). Estimates of ε were obtained from subsampled trees with and without incorporating the fraction of species they represent. Whether viewed in terms of the number of taxa (A, B), or in terms of the sampling fraction (C, D), estimating ε on empirical data coming from trees subsampled from a larger whole, and not including the sampling fraction, the reconstructed rates will clearly be biased toward underestimating ε.

is ignored methods for estimating diversification rate could perform poorly, even for large trees. However, when given the proper model (Maddison et al. 2007), extinction can still be estimated reasonably well from molecular phylogenies under these circumstances (FitzJohn et al. 2009; Davis et al. 2013).

On the basis of information from the fossil record, we know that most of life is extinct (Raup 1986)—an indication that ε must be nonzero. We have shown that even with realistic values of heritable rate, heterogeneity ε can be estimated fairly accurately with large enough trees. However, as Rabosky (2010) pointed out, why, then, do empirical results often show ε near or at zero? One clear possibility, which we have demonstrated, is that estimates of ε will often be low on small trees with or without rate heterogeneity (Figs. 1 and 2). Other potential explanations, not explored here, might involve mass extinction scenarios, or scenarios where extinctions are phylogenetically patterned (Rabosky 2010).

We suspect that in many cases low estimates of ε are also due simply to the use of trees that are subsampled from a much larger whole. However there are corrections for dealing with subsampled trees when inferring diversification parameters (reviewed by Stadler 2013), some popular implementations of birth–death estimation methods (e.g., see *laser*, Rabosky and Schliep 2013; *ape*, Paradis et al. 2004) have not incorporated them, although this trend seems to be changing (e.g., see *diversitree*, Fitzjohn 2012; *BAMM*, Rabosky 2014; *geiger*, Harmon et al. 2008). Nevertheless, estimating diversification parameters with methods that do not incorporate the sampling fraction can have a severe impact on estimates of ε. We illustrate this point in Figure 4. Here, we use a heterogeneity model (initial $\lambda = 0.376$, $\sigma = 0.06$) with $\varepsilon = 0.75$ to obtain a set of trees of 500 or more taxa, then subsample, and estimate ε with and without incorporating a subsampling correction. With empirical data coming from subsampled trees, if subsampling is not included in the model the reconstructed rates are clearly biased toward underestimating ε. In fact, estimates begin to approach a Yule model ($\varepsilon = 0$), even though the most severely subsampled trees had at least 100 taxa remaining. Once the sampling fraction is included, however, even trees with less than 20% of true diversity, generated with high heterogeneity, result in somewhat biased upwards, but still reasonable, estimates of ε.

A reviewer pointed out that the issue of nonrandom sampling is likely to be far more pervasive than is currently recognized—we agree. Looking at the age distribution of terminal taxa in the simulated trees from above, the average median tip branch length is relatively short, about 1 MA. This raises an important question: would these young species be recognized or treated as valid taxa in a phylogeny? In many cases, the answer is probably not, and this will have the effect of lengthening terminal branch lengths at the cost of decreasing estimates of extinction. There are methods to deal with this particular problem (see Brock et al. 2011; Cusimano et al. 2012), which are important steps forward, but have yet to be broadly incorporated in many of the popular implementations of birth–death models.

It may be tempting to interpret our study as advocating the use of rather simple birth–death models to molecular phylogenies. This would be a dramatic overstatement of our objective. Our goal here is to simply show how a particular model violation does not necessitate the abandonment of an entire field. Thus, we hope that our much more positive message is taken to heart by developers and reviewers of new methods for examining diversification who may be unduly hesitant to create or approve such methods, based on an influential but problematic study. As the possible size of molecular phylogenies continues to grow, we are provided with an unprecedented amount of data and power to develop and fit very complex models that contain a large number of parameters. However, at least for the forms of heterogeneity examined here, extinction can be estimated from trees as long as they are of moderate size, even if only a subset of taxa is included. Given the importance of extinction throughout evolution, it is important to recognize that in many cases it can still be estimated from molecular phylogenies.

ACKNOWLEDGMENTS

We thank D. Rabosky, an anonymous reviewer, J. Fordyce, B. Fitzpatrick, and members of the HOFF reading group for helpful discussions. We also thank D. Rabosky for providing code used in his original 2010 study. Support for JMB has been provided by the National Institute for Mathematical and Biological Synthesis, an Institute sponsored by the National Science Foundation, the U.S. Department of Homeland Security, and the U.S. Department of Agriculture through NSF Award #EF-0832858, with additional support from The University of Tennessee, Knoxville.

DATA ARCHIVING

The doi for our data is doi:10.5061/dryad.33p91.

LITERATURE CITED

Bokma, F. 2008. Bayesian estimation of speciation and extinction probabilities from (in)complete phylogenies. Evolution 62:2441–2445.

Brock, C. D., L. J. Harmon, and M. E. Alfarao. 2011. Testing for temporal variation in diversification rates when sampling is incomplete and nonrandom. Syst. Biol. 60:410–419.

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

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

Etienne, R.S., B. Haegeman, T. Stadler, T. Aze, P. Pearson, A. Purvis, and A. Phillimore. 2012. Diversity-dependence brings molecular phylogenies closer to agreement with the fossil record. Proc. R. Soc. B 279:1300–1309.

FitzJohn, R. G. 2012. Diversitree: comparative phylogenetic analyses of diversification in R. Methods Ecol. Evol. 3:1084–1092.

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

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

Harmon, L. J., J. A. Schulte II, A. Larson, and J. B. Losos. 2003. Tempo and mode of evolutionary radiation in iguanian lizards. Science 301:961–964.

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

Heard, S. B. 1996. Phylogenetic tree balance with variable and evolving speciation rates. Evolution 50:2141–2148.

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

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

Magallón, S., and A. Castillo. 2009. Angiosperm diversification through time. Am. J. Bot. 96:349–365.

Magallón, S., and M. J. Sanderson. 2001. Absolute diversification rate in angiosperm clades. Evolution 55:1762–1780.

Moore, B. R., K. M. A. Chan, and M. J. Donoghue. 2004. Detecting diversification rate variation in supertrees. Pp. 487–533 *in* O. R. P. Bininda-Emonds, ed. Phylogenetic supertrees: combining information to reveal the tree of life. Kluwer Academic, Dordrecht, The Netherlands.

Morlon, H., M. D. Potts, and J. B. Plotkin. 2010. Inferring dynamics of diversification: a coalescent approach. PLoS Biol. 8:e1000493.

Morlon, H., T. L. Parsons, and J. B. Plotkin. 2011. Reconciling molecular phylogenies with the fossil record. Proc. Natl. Acad. Sci. 108:16327–16332.

Nee, S., A. O. Mooers, and P. H. Harvey. 1992. The tempo and mode of evolution revealed from molecular phylogenies. Proc. Natl. Acad. Sci. USA 89:8322–8326.

Nee, S., E. C. Holmes, R. M. May, and P. H. Harvey. 1994. Extinction rates can be estimated from molecular phylogenies. Philos. Trans. R. Soc. B 344:77–82.

Paradis E., J. Claude, and K. Strimmer. 2004. APE: analyses of phylogenetics and evolution in R language. Bioinformatics 20:289–290.

Rabosky, D. L. 2006. Likelihood methods for detecting temporal shifts in diversification rates. Evolution 60:1152–1164.

———. 2009. Heritability of extinction rates links diversification patterns in molecular phylogenies and fossils. Syst. Biol. 58:629–640.

———. 2010. Extinction rate should not be estimated from molecular phylogenies. Evolution 64:1816–1824.

———. 2014. Automatic detection of key innovations, rate shifts, and diversity-dependence on phylogenetic trees. Plos ONE. 9:e89543.

Rabosky, D. L., and I. J. Lovette. 2008. Density-dependent diversification in North American wood warblers. Proc. R. Soc. B 275:2363–2371.

Rabosky, D. L., and K. Schliep. 2013. Laser: likelihood analysis of speciation/extinction rates from phylogenies. R package version 2.4–1. Available at http://CRAN.R-project.org/package=laser

Raup, D. M. 1986. Biological extinction in Earth history. Science 231:1528–1533.

Stadler, T. 2011. Mammalian phylogeny reveal recent diversification rate shifts. Proc. Natl. Acad. Sci. 108:6187–6192.

———. 2013. How can we improve accuracy of macroevolutionary rate estimates? Syst. Biol. 62:321–329.

Associate Editor: T. Near
Handling Editor: T. Lenormand

Supporting Information

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

Figure S1. Histograms are the same as in Figure 1, but here we use the Nelder–Mead algorithm with the parameter search of ε bounded on the interval (0,1).