

EXTINCTION RATES SHOULD NOT BE ESTIMATED FROM MOLECULAR PHYLOGENIES

Daniel L. Rabosky^{1,2,3,4}

¹*Department of Ecology and Evolutionary Biology, Cornell University, Ithaca, New York 14853*

²*Cornell Laboratory of Ornithology, 159 Sapsucker Woods Road, Ithaca, New York 14850*

³*Department of Integrative Biology, University of California, Berkeley, California 94720*

⁴*E-mail: drabosky@berkeley.edu*

Received February 17, 2009

Accepted December 3, 2009

Molecular phylogenies contain information about the tempo and mode of species diversification through time. Because extinction leaves a characteristic signature in the shape of molecular phylogenetic trees, many studies have used data from extant taxa only to infer extinction rates. This is a promising approach for the large number of taxa for which extinction rates cannot be estimated from the fossil record. Here, I explore the consequences of violating a common assumption made by studies of extinction from phylogenetic data. I show that when diversification rates vary among lineages, simple estimators based on the birth–death process are unable to recover true extinction rates. This is problematic for phylogenetic trees with complete taxon sampling as well as for the simpler case of clades with known age and species richness. Given the ubiquity of variation in diversification rates among lineages and clades, these results suggest that extinction rates should not be estimated in the absence of fossil data.

KEY WORDS: Adaptive radiation, extinction, macroevolution, phylogenetics.

Time-calibrated molecular phylogenies contain information about both the timing and rate of species diversification and thus provide a complementary window into macroevolutionary processes that are often obscured by the incompleteness of the fossil record. Numerous studies have used molecular phylogenies to characterize the pattern of species diversification through time (e.g., Nee et al. 1992; Harmon et al. 2003; Ruber and Zardoya 2005; McPeck 2008) and to quantify differences in diversification rates among lineages (e.g., Mooers and Heard 1997; Moore et al. 2004; Moore and Donoghue 2007; Rabosky et al. 2007).

One of the most intriguing applications of molecular phylogenies involves the inference of extinction rates from data on extant taxa only (Nee et al. 1994a; Paradis 2003; Maddison et al. 2007; Ricklefs 2007). It may seem counterintuitive that living species could provide any information on historical extinction rates, but this is indeed possible, because the shape of phylogenetic trees is influenced both by the net rate of lineage diversification through time as well as the ratio of the extinction rate μ to

the speciation rate λ . This parameter (μ/λ), denoted by ϵ , is also known as the relative extinction rate and is critically important in determining the distribution of speciation times that occur in a molecular phylogeny. Differences in the relative extinction rate can result in different phylogenetic tree shapes, even for clades diversifying under precisely the same net diversification rate (denoted by r , where $r = \lambda - \mu$). This phenomenon occurs because high relative extinction rates lead to high-lineage turnover through time, which changes the “age structure” of nodes in a phylogenetic tree. When ϵ is low, many species will be relatively old, but when ϵ is high, most species will be young, simply as a function of high-lineage turnover. This leads to the appearance of a temporal acceleration in the rate of diversification through time, a phenomenon that has been explored by numerous prior studies (Nee et al. 1994b; Rabosky 2006b).

Methods for analyzing species diversification rates often make assumptions about the constancy of diversification rates through time (Magallon and Sanderson 2001) or among lineages

(Pybus and Harvey 2000; Rabosky 2006b). For example, many studies have inferred diversification rates from clade age and species richness data with the explicit assumption that rates have been constant through time within clades (Hunt et al. 2007; Rabosky et al. 2007; Wiens 2007; Alfaro et al. 2009; Magallon and Castillo 2009). However, Rabosky (2009a) showed that this assumption is often violated, leading to potentially incorrect inferences about the causes of variation in species richness among clades.

In this article, I explore the consequences of among-lineage rate variation for estimates of extinction rates. I demonstrate that, when rates vary across the branches of phylogenetic trees or among clades, estimators that assume rate-constancy among lineages perform poorly. In many cases, the problem is sufficiently severe that extinction should not be estimated at all.

Methods

CLADE AGE AND SPECIES RICHNESS

To assess the effects of among-clade variation in net diversification rates (r) on estimates of extinction, I used a general simulation protocol of (1) drawing diversification rates for a set of clades from a distribution with a specified mean and variance (see below) and under a known relative extinction rate ϵ ; (2) simulating clade diversity under those rates; and (3) estimating ϵ for each set of clades. The central question is thus whether increased variance in net diversification rates among clades can lead to erroneous inference on ϵ .

Given a set of clade ages with associated species richness data, we can use the theory of the birth–death process to estimate both r and ϵ . Under the simple birth–death process (rates constant through time and among lineages), we can compute the probability that a number of ancestral lineages, a , will result in a total of n surviving lineages, after some amount of time t , given a particular speciation and extinction parameterization. This probability can be difficult to compute when a is large, but for the case in which information is available on stem clade ages ($a = 1$), it is simply (Raup 1985)

$$\Pr(n) = (1 - \beta)\beta^{n-1}, \quad (1)$$

where

$$\beta = \frac{e^{rt} - 1}{e^{rt} - \epsilon}. \quad (2)$$

Note that (1) has been conditioned on clade survival to the present. If data are available for many clades, (1) can be used to find the likelihood that a particular birth–death parameterization has generated the observed data (Ricklefs 2007). The likelihood of the full data is given by

$$L(D | r, \epsilon) = \prod_{i=1}^N \Pr(n_i | r, \epsilon), \quad (3)$$

and is maximized with respect to r and ϵ (Bokma 2003).

I conducted simulations using diversification rates and clade ages parameterized from an avian families dataset (Sibley and Ahlquist 1990) that has been the subject of previous analyses of diversification (Paradis 2003; Ricklefs 2003). The data consist of 127 avian families with at least two species and includes their respective species richness and stem-clade ages inferred from DNA hybridization studies. The decision to use this dataset over any other is arbitrary and was made to facilitate meaningful comparisons between the variance in rates used in simulations with the variance in rates observed among real clades. This tree is known to conflict with more recent phylogenetic analyses (e.g., Barker et al. 2004), but is expected to be adequate in the present context, as a rough framework for parameterizing simulations.

I assumed that rates for each family are drawn from a single gamma distribution; gamma is a reasonable choice here, as it is defined on the appropriate interval $(0, \infty)$, has great flexibility in shape, and is widely used to model evolutionary rate variation in molecular evolutionary studies (Yang 1993). Given such a model of rate variation, one possible method for estimating the variance in rates $(\sigma^2)_r$ observed among avian families is to estimate r for each clade independently using stem-clade estimators for r that have been used in previous studies (Raup 1985; Magallon and Sanderson 2001). A gamma distribution can then be fitted to these inferred rates; the shape (k) and scale (θ) parameters of this fitted distribution specify the mean ($k\theta$) and variance ($k\theta^2$) of the distribution of r . This approach assumes that diversification rates within each clade can simply be calculated from the observed species richness and clade age. However, diversification in the birth–death framework is a stochastic process: even if all clades have the same rate, this approach will estimate different rates for most or all clades, simply due to stochastic variation in species richness. This may lead to a problem of overparameterization, because estimation of separate rates for each clade is effectively a model with the same number of parameters as data points. The statistical and biological consequences of assuming this model are poorly known, and model selection procedures are rarely used to test for overparameterization.

A more appropriate framework might be to assume that species richness data are the observed outcomes of a stochastic process in which each clade has a rate r_i drawn from an overall gamma distribution with parameters k and θ . The rates themselves are not directly observed, but we can integrate over all possible rates given a particular gamma(k, θ) parameterization to find the likelihood of the data. In this framework, the likelihood of a particular species richness value n becomes

$$\Pr(n \mid k, \theta, \epsilon, t) = \int \Pr(n \mid r, \epsilon, t) f(r \mid k, \theta) dr,$$
 (4)

where $\Pr(n \mid r, \epsilon)$ corresponds to equation (1) and $f(r \mid k, \theta)$ is a gamma density. The likelihood of the full data D becomes

$$L(D \mid k, \theta, \epsilon) = \prod_{i=1}^N \int_0^\infty \left\{ r_i \frac{e^{-(r_i/\theta)}}{\Gamma(k) \theta^k} \right\} (1 - \beta) \beta^{n_i-1} dr_i,$$
 (5)

where β is given by equation (2) and the expression in curly braces is the probability density of the gamma distribution. By integrating over all possible rates, we can describe rate variation among clades in terms of the two parameters of the gamma distribution (scale and shape), and this fitted model can be directly compared to models that assume homogeneity of rates or that allow rates to vary over time (Rabosky 2009a). I refer to this model as a “relaxed-rate” model of rate variation, analogous to relaxed-clock methods of modeling molecular evolutionary rate variation (Thorne et al. 1998; Drummond et al. 2006).

I estimated the distribution of net diversification rates that best described the variation in species richness among avian families assuming that all clades diversified under a common relative extinction rate, and I considered five relative extinction scenarios: $\epsilon = 0, 0.25, 0.5, 0.75$, and 0.95 . Numerical integration was used to compute the likelihoods in (4), and likelihoods were maximized using the Nelder–Mead algorithm. Results of fitting the relaxed-rate model to the avian families dataset are given in Table 1; maximum log-likelihoods of the data under the constant rate model (eq. 3) are included for comparison and demonstrate a substantial improvement in model fit by accommodating among-lineage variation in r .

Table 1. Parameters and log-likelihoods inferred for the avian families dataset under a relaxed-rate model of rate variation for five relative extinction rates (ϵ). LogL gives the log-likelihood of the data under the relaxed-rate model. For comparison, LogL-CR* gives the log-likelihood under a model with constant rates across all lineages (eq. 3). Parameters are the shape (k) and scale (θ) parameters of the fitted gamma distribution, and S^2_{obs} is the variance of the fitted distribution. Note that the large increase in the log-likelihood of the data under the relaxed-rate model relative to the constant-rate model comes with the addition of just a single parameter.

ϵ	LogL	LogL-CR*	Parameters	Mean rate	S^2_{obs}
0	−656.8	−781.9	$k=4.5, \theta=0.066$	0.298	0.02
0.25	−655	−772.1	$k=4.16, \theta=0.067$	0.276	0.018
0.5	−652.6	−759.2	$k=3.67, \theta=0.067$	0.25	0.016
0.75	−649.1	−739	$k=2.87, \theta=0.069$	0.198	0.014
0.95	−643.5	−705.5	$k=1.69, \theta=0.059$	0.099	0.006

The observed mean and variances (S^2_{obs}) in rates for each ϵ class were used to parameterize simulations. Each simulation entailed drawing a rate for each clade from a gamma distribution with the same mean as the avian families data and variance vS^2_{obs} , where v is a multiplier of the observed variance. One can view v as the variation in diversification rates used for a particular simulation, relative to the variation observed in the avian families dataset. Simulations were conducted under five variance multipliers: $v = 0, 0.5, 1.0, 1.5$, and 2.0 . Thus, a dataset simulated under $v = 1.0$ had rates drawn from a gamma distribution with variance identical to that observed for the avian families, and simulations with $v = 2.0$ had rates drawn from a distribution with twice the observed variance. To rescale the gamma distribution with respect to v , we first compute the scale parameter θ by dividing the target variance (vS^2_{obs}) by the target mean, as θ is the ratio of the mean to the variance. The target mean is then divided by the new θ to give the updated shape parameter k . A scenario in which all clades have the same rate corresponds to $v = 0$. Figure 1 illustrates a fitted gamma distribution of rates (under $\epsilon = 0$) as well as the corresponding curves with half and twice the observed variance ($v = 0.5$ and $v = 2.0$, respectively).

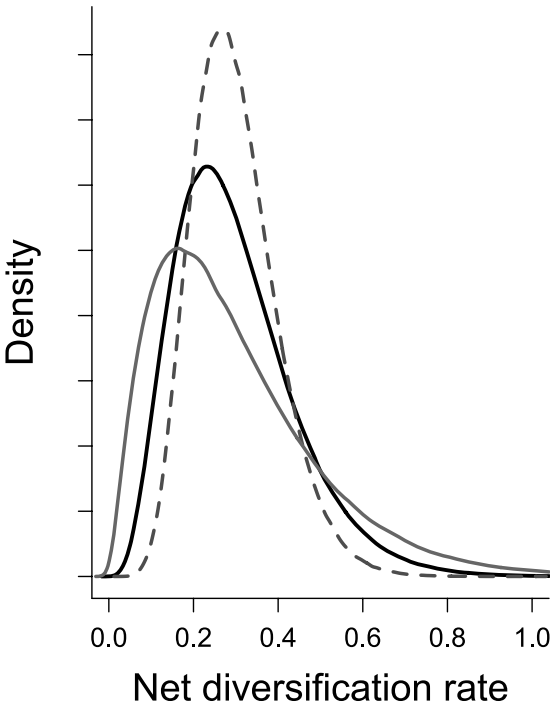


Figure 1. Fitted gamma density of net diversification rates for avian families assuming $\epsilon = 0$ (black curve) and under alternative variance scenarios considered in simulations (gray curves). All curves have an identical mean (see Table 1), but dashed curve corresponds to a distribution of rates with 50% the observed variance ($v = 0.5$) and solid gray curve denotes a distribution with twice the observed variance ($v = 2.0$).

A single simulation consisted of drawing 127 rates from the fitted gamma distribution and pairing each rate uniquely with one of the 127 avian families. The rates were then used to simulate species richness data (see Rabosky 2009a) given the ages of those families. Note that the distribution of species richness for a given diversification process is geometric with parameter $1 - \beta$ (eq. 1). It follows that species richness can be simulated by drawing from this distribution until the first nonzero value is obtained, as we have conditioned on survival to the present. I investigated 25 evolutionary scenarios in all (five ϵ classes, each with five v classes), with 5000 simulations conducted under each scenario.

For each set of clade ages and simulated richness data, ϵ was estimated using the simple constant-rate estimator described above (eq. 3). To guard against the possibility of recovering local rather than global optima, 50 optimizations were performed on each simulated dataset, using random starting parameters. Initial ϵ values for each optimization were drawn uniformly on the interval $[0, 0.99]$. In the Supporting information, I demonstrate that increasing the number of independent optimizations does not change the maximum likelihood estimates of parameters.

DATED PHYLOGENIES WITH COMPLETE TAXON SAMPLING

To investigate the effects of among-lineage variation in diversification rates on estimates of ϵ , I constructed a phylogenetic tree simulation algorithm that allows speciation rates (λ) to evolve along branches of the phylogenetic tree (Heard 1996). Simulations were conducted in continuous time; λ values for each branch were drawn from a lognormal distribution with a mean equal to the value of the parent branch and a standard deviation equal to $T\sigma$, where σ is the standard deviation of the process of rate evolution and T is the length of the parent branch. As σ increases, the magnitude of heterogeneity in rates among branches increases. I conducted simulations under four relative extinction rates: $\epsilon = 0$, $\epsilon = 0.25$, $\epsilon = 0.5$, and $\epsilon = 0.75$. For simulations with nonzero extinction, the relative extinction rate ϵ was kept constant by recalculating μ on each branch once a new speciation rate had been drawn.

Initial speciation rates for each ϵ category were chosen to result in an expected value of $N = 50$ surviving lineages after 50 time units, corresponding to $\lambda = 0.064$, 0.078 , 0.102 , and 0.151 for $\epsilon = 0$, 0.25 , 0.5 , and 0.75 , respectively. The parameter of among-lineage rate variation σ was varied from 0 to 0.06 in increments of 0.005, with 2000 trees generated per σ and ϵ . I estimated ϵ by fitting a constant rate birth–death model to the distribution of speciation times (Nee et al. 1994b). All simulations and analyses were conducted in R, with some code modified from the Ape (Paradis et al. 2004), Geiger (Harmon et al. 2008), and Laser (Rabosky 2006a) packages. Source code (in R) for the simulation of species richness, phylogenetic trees, and model fitting is available as Supporting information.

Results

CLADE AGE AND SPECIES RICHNESS

Among-clade variation in diversification rates exerts a potent effect on inferences about ϵ based on the birth–death model (Fig. 2). For clades simulated under constant r among lineages ($v = 0$), estimates of ϵ are generally consistent with the simulation model (Fig. 2; Table 2). However, as the variance in rates among clades is increased, estimates of ϵ become inconsistent with the simulation model. Any variation in r among clades leads to inaccurate estimates of ϵ , and this problem is especially acute for intermediate relative extinction rates. For example, when the extinction rate is one-half of the speciation rate ($\epsilon = 0.5$), the estimated ϵ values are reasonable for simulations with constant rates among clades (Table 2; Fig. 2, top row). However, these estimates become nonsensical with even minor among-lineage rate variation: median and modal estimates of ϵ are 0.16 and 0.01 for the scenario with $v = 0.5$.

The problem is exacerbated as the variation in rates among clades increases. As v increases, the distribution of estimated ϵ values becomes bimodal, with peaks on both boundaries ($\epsilon = 0$ and $\epsilon = 1$), with few estimates falling between those values. Thus, minor variation in diversification rates among clades results in severely compromised inference. When the data are analyzed with methods that assume constant rates among clades, we are left with the erroneous perception that a birth–death model with extremely high ($\epsilon = 1$) or low ($\epsilon = 0$) relative extinction provides an explanation for observed patterns of species richness.

DATED PHYLOGENIES WITH COMPLETE TAXON SAMPLING

For phylogenetic trees with complete taxon sampling, estimates of ϵ are reasonably unbiased in the absence of among-lineage rate variation (Table 3), although confidence intervals are large. However, a pronounced upward bias in estimates of ϵ is noted when rates vary among branches. Even under a simulation model with $\epsilon = 0$, increased rate heterogeneity among branches inflates estimates of ϵ (Fig. 3). For simulations lacking extinction (Fig. 3A), mean and median estimates of ϵ are 0.56 and 0.60, respectively. A positive bias in estimates of ϵ occurs with increasing heterogeneity in diversification rates across all extinction scenarios considered.

Discussion

Previous studies have shown that confidence intervals on relative extinction rates as inferred from phylogenetic data are generally large (Nee et al. 1994a,b), and that this parameter may be difficult to infer in the absence of fossil data (Paradis 2004; Maddison et al. 2007). My results suggest that the problem is even more severe than typically thought, because among-lineage variation

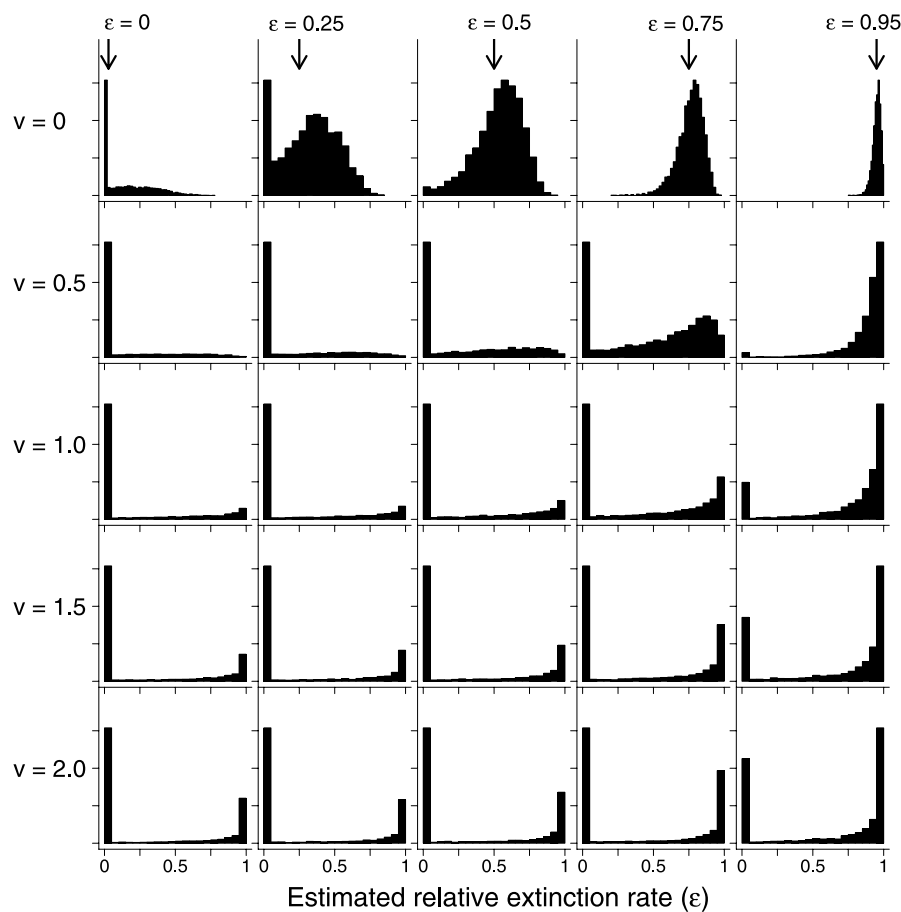


Figure 2. Increased variance in net diversification among clades leads to incorrect estimates of relative extinction, even when extinction is not present. Each histogram represents the distribution of estimated ϵ values. Histograms in the same column were simulated under identical ϵ values (arrows at top), whereas those in the same row were simulated under identical magnitudes of among-clade rate variation. The magnitude of rate variation among clades is given by v . As v increases, estimates of ϵ tend to 0 or 1. Simulation models with $v = 0$ correspond to constant rates for all clades, and simulations with $v = 1.0$ were conducted with among-clade rate variation identical to that estimated for the avian families dataset.

in diversification rates results in erroneous or directionally biased estimates of ϵ . This applies not only to phylogenies with complete taxon sampling (Fig. 3), but also to clade age and species richness

data. Even comparatively minor variation in rates among clades leads to profound error in the estimation of ϵ from clade age and species richness data (Fig. 2).

For clade age and species richness data, the estimator appears to be fundamentally unstable with respect to even minor violations of model assumptions (e.g., $v = 0.5$). It is clear that rate

Table 2. Estimates of relative extinction rates for species richness data simulated under a model with constant rates among clades ($v=0$; Fig. 2, top row) and under five relative extinction ratios. Modes of each distribution were inferred using kernel density estimation with a Gaussian smoothing kernel. Quantiles give the 2.5 and 97.5 percentiles of the distribution of estimates. Note that the distribution of estimated ϵ values is bimodal under $\epsilon=0.25$.

Simulation model (ϵ)	Mean	Median	Mode	Quantiles
0	0.16	0.11	0.004	0–0.55
0.25	0.32	0.33	0.002/0.38	0–0.67
0.5	0.53	0.56	0.6	0.14–0.79
0.75	0.76	0.77	0.79	0.55–0.90
0.95	0.95	0.95	0.95	0.88–0.99

Table 3. Estimates of relative extinction rates for phylogenies with complete taxon sampling simulated under a model with constant rates among lineages ($\sigma=0$; Fig. 3A). Quantiles give the 2.5 and 97.5 percentiles of the distribution of estimates.

Simulation model (ϵ)	Mean	Median	Mode	Quantiles
0	0.21	0.01	0.01	0–0.95
0.25	0.3	0.23	0.01	0–0.99
0.5	0.49	0.51	0.55	0–0.99
0.75	0.68	0.74	0.79	0–1

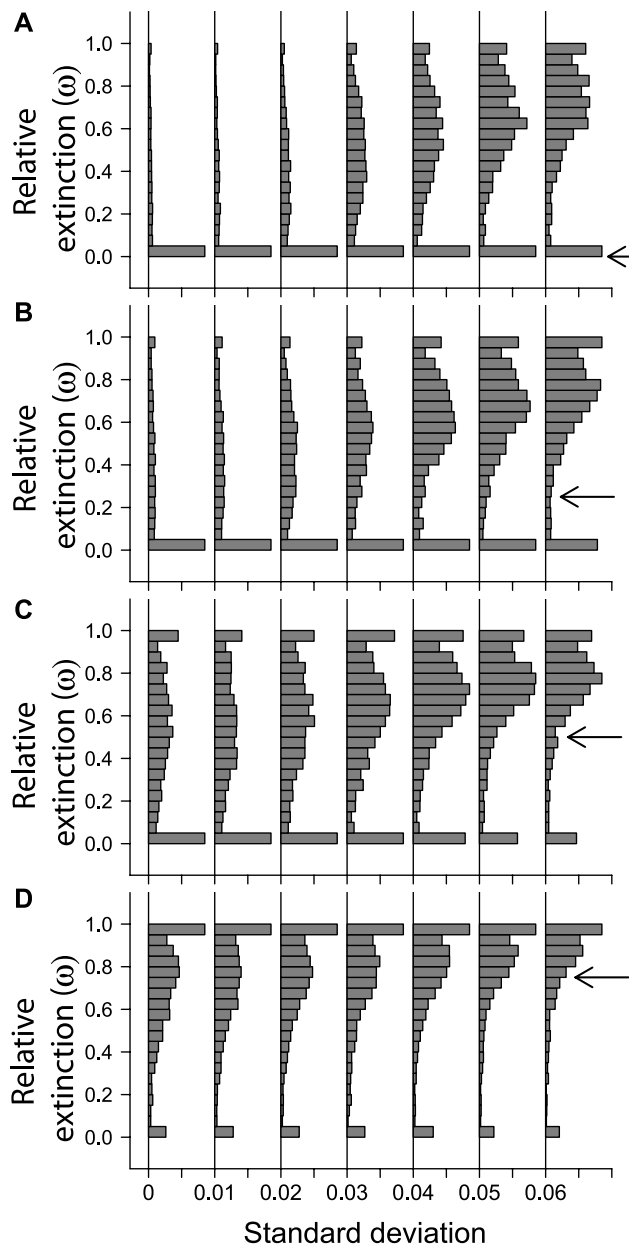


Figure 3. Variation in diversification rates among branches of phylogenetic trees leads to inflated estimates of ϵ . Histograms depict frequency distributions of estimated ϵ values, as a function of σ . Simulations were conducted under (A) $\epsilon = 0$, (B) $\epsilon = 0.25$, (C) $\epsilon = 0.5$, and (D) $\epsilon = 0.75$ (see arrows at right). As σ increases, there is an increase in the number of simulated trees that appear to have high relative extinction rates.

variation among lineages results in estimates of ϵ that approach 0 or 1, regardless of the true relative extinction rate. It is especially interesting that, even within a single parameterization (e.g., $v = 2.0$, $\epsilon = 0.5$), estimates of ϵ are bimodal and centered on the endpoints of the distribution of possible values. In the Supporting information, I show that this bimodality appears to be related, in part, to the ratio of species richness in old clades relative to young

clades (Fig. S5). Because this ratio will vary as a result of stochasticity, simulations conducted with identical parameterizations can show dramatic differences in estimates of relative extinction. This bimodality is not limited to scenarios with high among-lineage rate variation, but is also observed for constant-rate diversification processes with small numbers of clades (Figs. S2 and S3). The latter occurs because, when only few clades are considered, the ratio of richness in young and old clades is driven by stochasticity: in some cases, young clades might even be more diverse than old clades, simply due to chance alone, and this has important consequences for estimated relative extinction rates (Fig. S4).

For phylogenetic trees with complete taxon sampling, the upwards bias in estimates of ϵ with respect to rate variation has to do with the fact that the birth–death model predicts a geometric distribution for species richness with a modal value of $n = 1$. This is best illustrated by an example: suppose a clade is diversifying under a pure-birth process, with $\lambda = 1$ lineage/million years (my). Imagine the effects of a large rate decrease occurring in a lineage at 1 my before present: if this lineage undergoes a complete cessation of diversification in the recent past, such that $\lambda = 0$, we will observe that the lineage will have left only a single descendant in the present—itsself. This is not an especially unlikely event under the pure-birth model; even with $\lambda = 1$, the probability of observing just a single descendant is 0.375. However, suppose that this same lineage has undergone a fivefold increase in the speciation rate at 1 my before present. In this case, the new speciation rate is $\lambda = 5$, leading to a substantial increase in the number of lineages in the past million years. With $\lambda = 5$, the probability of leaving a single lineage is small ($P = 0.006$), and the 2.5% and 97.5% quantiles on the expected number of progeny lineages are 4 and 545, with a mean of 148. One can easily imagine that such increases and decreases would have contrasting effects on the shape of lineage-through-time plots (Fig. 4). A recent rate increase, by generating a pulse of diversification, would suggest a massive rise in total diversity toward the present, but a severe rate decrease would scarcely change the shape of the plot.

In short, decreases in rates along individual branches do not have a large effect on patterns of lineage accumulation through time: assuming the lineage does not go extinct, a decrease in a single lineage, at its most severe, still results in a single descendant surviving to the present. This is expected to be reasonably common even in lineages that do not undergo declines in rates. In contrast, a rate increase that affects only a single branch of a phylogenetic tree can yield a large spike in the number of lineages, because all of the descendants of this lineage will inherit the high rate and thus continue to diversify.

This study does not paint a promising picture of fossil-free extinction estimates, and I have not even considered the additional complications introduced by assuming constant rates through time within clades. This assumption is almost certainly violated in

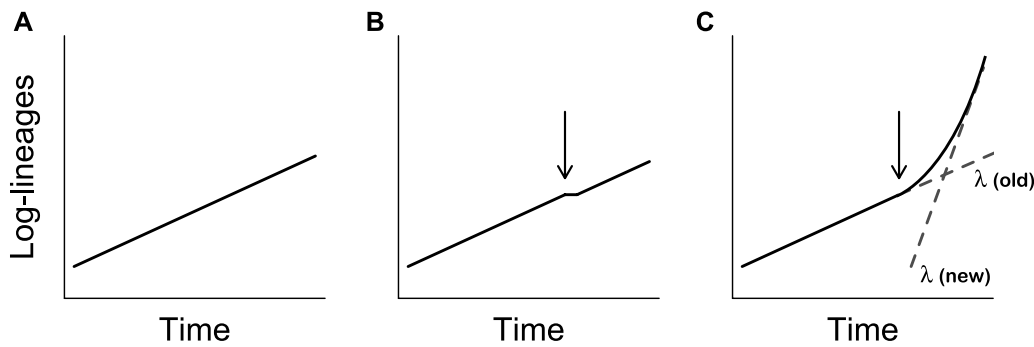


Figure 4. Contrasting effects of lineage-specific increases and decreases in the speciation rate (λ) on the shape of lineage-through-time (LTT) curves. (A) Expected LTT curve under a pure-birth process ($\mu = 0$) with λ constant among lineages. (B) Expected LTT curve when a single lineage undergoes a major decrease in λ . (C) Expected lineage through time plot when a single lineage undergoes a major increase in λ . The rapid rise in the number of lineages toward the present leads to estimates of high ϵ , even when extinction is not present. The effects of a single rate decrease should be statistically indistinguishable from a constant-rate process. However, a rate increase can exert a large effect on the shape of LTT curves, because all descendants of the high-rate lineage inherit an elevated λ and thus contribute disproportionately to clade diversity. The slope of the LTT curve approaches and, given sufficient time, becomes equal to the new (elevated) λ .

many cases (Ricklefs 2007; Ricklefs et al. 2007; Rabosky 2009a) and is likely to pose additional problems for inferences about extinction (Rabosky 2009a). The problem should be especially severe for large clades, because larger clades should contain a more heterogeneous mixture of lineages with different diversification rates.

However, it is impossible to ignore the observation that extinction rates estimated from smaller species-level phylogenies tend toward zero (Nee 2006; Bokma 2008; Purvis 2008). Why these estimates are low remains unclear (Purvis 2008), and even analyses that allow both λ and μ to vary independently through time frequently converge on low estimates for μ (Rabosky and Lovette 2008). Previous authors have found this surprising because the fossil record strongly suggests that relative extinction rates have been high throughout the history of life (Stanley 1979; Gilinsky 1994; Alroy 2000, 2008a,b). The results presented here render this matter somewhat murkier still, because among-lineage variation in rates biases estimates in the opposite direction.

A possible explanation for this was proposed in Rabosky (2009b), which demonstrated that phylogenetically patterned extinction can lead to molecular phylogenies that contain little signature of high background extinction. If extinction events are phylogenetically clustered on phylogenetic trees, then even a high rate of extinction (e.g., $\epsilon = 1$) can fail to leave a signature in the distribution of speciation times for phylogenetic trees of extant species only (Rabosky 2009b). This is likely to be problematic, because a growing body of evidence suggests that extinction rates are phylogenetically conserved (Vamosi and Wilson 2008; Roy et al. 2009). Another possible explanation is implicit in a recent study by Crisp and Cook (2009), who demonstrated that mass extinction events could leave patterns in phylogenetic trees con-

sistent with $\epsilon = 0$. Although their results did not directly address problems in the estimation of ϵ , they showed that mass extinctions could lead to lineage accumulation patterns suggestive of “early burst” diversification; these early burst patterns in turn typically lead to estimates of low ϵ (Rabosky and Lovette 2008; but see Quental and Marshall 2009).

What is the way forward from here? One possibility is to explicitly estimate ϵ after relaxing assumptions about homogeneous diversification rates among lineages. This is the essence of the relaxed-rate model for variation in r developed in this article. However, estimating ϵ with a model similar to that described in equation (4) is a dubious endeavor; minimally, it requires assumptions about: (1) the true distribution of diversification rates among clades; (2) constancy of rate variation through time; and (3) the extent to which extinction is phylogenetically conserved. Results presented here demonstrate that assumptions about the nature of among-lineage rate variation can lead to highly misleading estimates of extinction. Moreover, a previous paper found that estimation of ϵ under a constant-rate birth–death model when rates have varied through time led to high estimates of ϵ with extremely high confidence (Rabosky 2009a). Yet when the data were analyzed under a more appropriate model, there was little ability to discriminate between low and high ϵ .

In summary, the estimation of meaningful extinction rates from data on living species only is a challenging problem that may be insoluble. The results of this study argue strongly for better integration of paleontological data with molecular phylogenetic studies of diversification. At present, molecular phylogenetic studies use the fossil record primarily for dating trees, but combining these complementary sources of information should generate a richer perspective on the dynamics of speciation and extinction.

ACKNOWLEDGMENTS

I thank S. Heard, I. Lovette, A. McCune, R. Ricklefs, J. Vamosi, M. Alfaro, T. Quental, and L. Harmon for discussion of these topics and/or comments on the manuscript. This research was supported by National Science Foundation (NSF-OSIE-0612855), NSF-DEB-0814277, and the Miller Institute for Basic Research in Science at the University of California, Berkeley.

LITERATURE CITED

- Alfaro, M. E., F. Santini, C. Brock, H. Alamillo, A. Dornburg, D. L. Rabosky, G. Carnevale, 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.
- Alroy, J. 2000. New methods for quantifying macroevolutionary patterns and processes. *Paleobiology* 26:707–733.
- . 2008a. Speciation and extinction in the fossil record of North American mammals. Pp. 301–323 in R. Butlin, J. Bridle, and D. Schluter, eds. *Ecology and speciation*. Cambridge Univ. Press, Cambridge, U.K.
- . 2008b. The dynamics of origination and extinction in the marine fossil record. *Proc. Natl. Acad. Sci. USA* 105:11536–11542.
- Barker, F. K., A. Cibois, P. Schikler, J. Feinstein, and J. Cracraft. 2004. Phylogeny and diversification of the largest avian radiation. *Proc. Natl. Acad. Sci. USA* 101:11040–11045.
- Bokma, F. 2003. Testing for equal rates of cladogenesis in diverse taxa. *Evolution* 57:2469–2474.
- . 2008. Problems detecting density dependent diversification on phylogenies. *Proc. R. Soc. Lond. B* 276:993–994.
- Crisp, M. D., and L. G. Cook. 2009. Explosive radiation or cryptic mass extinction? Interpreting signatures in molecular phylogenies. *Evolution* 63:2257–2265.
- Drummond, A. J., S. Y. W. Ho, M. J. Phillips, and A. Rambaut. 2006. Relaxed phylogenetics and dating with confidence. *PLoS Biol.* 4:699–710.
- Gilinsky, N. L. 1994. Volatility and the Phanerozoic decline of background extinction intensity. *Paleobiology* 20:445–458.
- Harmon, L. J., J. A. Schulte, 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. Brock, R. E. Glor, and W. E. 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.
- Hunt, T., J. Bergsten, Z. Levkanicova, A. Papadopoulou, O. S. John, R. Wild, P. M. Hammond, D. Ahrens, M. Balke, M. S. Caterino, et al. 2007. A comprehensive phylogeny of beetles reveals the evolutionary origins of a superradiation. *Science* 318:1913–1916.
- 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.
- Magallon, S., and A. Castillo. 2009. Angiosperm diversification through time. *Am. J. Bot.* 96:349–365.
- Magallon, S., and M. J. Sanderson. 2001. Absolute diversification rates in angiosperm clades. *Evolution* 55:1762–1780.
- McPeck, M. A. 2008. Ecological dynamics of clade diversification and community assembly. *Am. Nat.* 172:E270–E284.
- Mooers, A. O., and S. B. Heard. 1997. Evolutionary process from phylogenetic tree shape. *Q. Rev. Biol.* 72:31–54.
- Moore, B. R., and M. J. Donoghue. 2007. Correlates of diversification in the plant clade dipsacales: geographic movement and evolutionary innovations. *Am. Nat.* 170:S28–S55.
- Moore, B. R., K. M. Chan, and M. J. Donoghue. 2004. Detecting diversification rate variation in supertrees. Pp. 487–533 in O. P. Bininda-Emonds, ed. *Phylogenetic supertrees: combining information to reveal the tree of life*. Kluwer Academic, Dordrecht, Netherlands.
- Nee, S. 2006. Birth-death models in macroevolution. *Annu. Rev. Ecol. Evol. Syst.* 37:1–17.
- Nee, S., A. Mooers, and P. H. Harvey. 1992. 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. 1994a. Extinction rates can be estimated from molecular phylogenies. *Proc. R. Soc. Lond. B* 344:77–82.
- Nee, S., R. M. May, and P. H. Harvey. 1994b. The reconstructed evolutionary process. *Phil. Trans. R. Soc. Lond. B* 344:305–311.
- Paradis, E. 2003. Analysis of diversification: combining phylogenetic and taxonomic data. *Proc. R. Soc. Lond. B* 270:2499–2505.
- . 2004. Can extinction rates be estimated without fossils? *J. Theor. Biol.* 229:19–30.
- Paradis, E., J. Claude, and K. Strimmer. 2004. APE: analyses of phylogenetics and evolution in R language. *Bioinformatics* 20:289–290.
- Purvis, A. 2008. Phylogenetic approaches to the study of extinction. *Annu. Rev. Ecol. Evol. Syst.* 39:301–319.
- Pybus, O. G., and P. H. Harvey. 2000. Testing macro-evolutionary models using incomplete molecular phylogenies. *Proc. R. Soc. Lond. B* 267:2267–2272.
- Quental, T. B., and C. R. Marshall. 2009. Extinction during evolutionary radiations: reconciling the fossil record with molecular phylogenies. *Evolution* 63:3158–3167.
- Rabosky, D. L. 2006a. LASER: a maximum likelihood toolkit for detecting temporal shifts in diversification rates. *Evol. Bioinform. Online* 2:247–250.
- . 2006b. Likelihood methods for detecting temporal shifts in diversification rates. *Evolution* 60:1152–1164.
- . 2009a. Ecological limits on clade diversification in higher taxa. *Am. Nat.* 173:662–674.
- . 2009b. Heritability of extinction rates links diversification patterns in molecular phylogenies and fossils. *Syst. Biol.* 58:629–640.
- Rabosky, D. L., and I. J. Lovette. 2008. Explosive evolutionary radiations: decreasing speciation or increasing extinction through time? *Evolution* 62:1866–1875.
- Rabosky, D. L., S. C. Donnellan, A. L. Talaba, and I. J. Lovette. 2007. Exceptional among-lineage variation in diversification rates during the radiation of Australia's most diverse vertebrate clade. *Proc. R. Soc. Lond. B* 274:2915–2923.
- Raup, D. M. 1985. Mathematical models of cladogenesis. *Paleobiology* 11:42–52.
- Ricklefs, R. E. 2003. Global diversification rates of passerine birds. *Proc. R. Soc. Lond. B* 270:2285–2291.
- . 2007. Estimating diversification rates from phylogenetic information. *Trends Ecol. Evol.* 22:601–610.
- Ricklefs, R. E., J. B. Losos, and T. M. Townsend. 2007. Evolutionary diversification of clades of squamate reptiles. *J. Evol. Biol.* 20:1751–1762.
- Roy, K., G. Hunt, and D. Jablonski. 2009. Phylogenetic conservatism of extinctions in marine bivalves. *Science* 325:733–737.
- Ruber, L., and R. Zardoya. 2005. Rapid cladogenesis in marine fishes revisited. *Evolution* 59:1119–1127.
- Sibley, C. G., and J. E. Ahlquist. 1990. *Phylogeny and classification of the birds of the world*. Yale Univ. Press, New Haven, CT.
- Stanley, S. M. 1979. *Macroevolution: pattern and process*. Freeman, San Francisco, CA.

- Thorne, J. L., H. Kishino, and I. S. Painter. 1998. Estimating the rate of evolution of the rate of molecular evolution. *Mol. Biol. Evol.* 15:1647–1657.
- Vamosi, J. C., and J. R. U. Wilson. 2008. Nonrandom extinction leads to elevated loss of angiosperm evolutionary history. *Ecol. Lett.* 11:1047–1053.
- Wiens, J. J. 2007. Global patterns of diversification and species richness in amphibians. *Am. Nat.* 170:S86–S106.
- Yang, Z. 1993. Maximum-likelihood estimation of phylogeny from DNA sequences when substitution rates differ over sites. *Mol. Biol. Evol.* 10:1396–1401.

Associate Editor: J. Vamosi

Supporting Information

The following supporting information is available for this article:

Figure S1. With the exception of very small numbers of optimizations ($k = 1$), there is no change in the estimated ϵ value with increased numbers of optimizations.

Figure S2. (Left) Relationship between clade age and species richness for clades of age 10, 15, and 20, where species richness in each clade is the expected richness under a diversification process with $r = 0.25$ and $\epsilon = 0.5$.

Figure S3. (A) Bimodality of relative extinction estimates for 3-clade dataset when species richness is simulated under a constant-rate process with $r = 0.25$ and $\epsilon = 0.5$.

Figure S4. Varying species richness in any of the focal clades leads to predictable effects on relative extinction estimates.

Figure S5. (Top) Relationship between estimated relative extinction rates and the R ratio (ratio of richness of old clades to young clades).

Supporting Information may be found in the online version of this article.

Please note: Wiley-Blackwell is not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.