LIKELIHOOD METHODS FOR DETECTING TEMPORAL SHIFTS IN DIVERSIFICATION RATES

Daniel L. Rabosky

Department of Ecology and Evolutionary Biology, Cornell University, Ithaca, New York 14853-2701
E-mail: DLR32@cornell.edu

Abstract.—Maximum likelihood is a potentially powerful approach for investigating the tempo of diversification using molecular phylogenetic data. Likelihood methods distinguish between rate-constant and rate-variable models of diversification by fitting birth-death models to phylogenetic data. Because model selection in this context is a test of the null hypothesis that diversification rates have been constant over time, strategies for selecting best-fit models must minimize Type I error rates while retaining power to detect rate variation when it is present. Here I examine model selection, parameter estimation, and power to reject the null hypothesis using likelihood models based on the birth-death process. The Akaike information criterion (AIC) has often been used to select among diversification models; however, I find that selecting models based on the lowest AIC score leads to a dramatic inflation of the Type I error rate. When appropriately corrected to reduce Type I error rates, the birth-death likelihood approach performs as well or better than the widely used gamma statistic, at least when diversification rates have shifted abruptly over time. Analyses of datasets simulated under a range of rate-variable diversification scenarios indicate that the birth-death likelihood method has much greater power to detect variation in diversification rates when extinction is present. Furthermore, this method appears to be the only approach available that can distinguish between a temporal increase in diversification rates and a rate-constant model with nonzero extinction. I illustrate use of the method by analyzing a published phylogeny for Australian agamid lizards.

Key words.—Diversification rates, macroevolution, maximum likelihood, phylogeny, speciation.

There is great interest in understanding variation in speciation and extinction rates both over time (e.g., Hey 1992; Sanderson and Donoghue 1996; Barraclough and Nee 2001; Zink et al. 2004; Ruber and Zardoya 2005) and among lineages (e.g., Mooers and Heard 1997; Chan and Moore 2002; Moore et al. 2004; Ree 2005). The proliferation of molecular phylogenetic data has provided researchers in this area with a wealth of material for comparative analyses of diversification. Many intriguing and unanswered questions in evolutionary biology concern temporal variation in diversification rates. For example, how have climatic shifts influenced the tempo of diversification (Zink and Slowinski 1995; Kadereit et al. 2004; Weir and Schluter 2004)? How do diversification rates vary over time during major evolutionary radiations (Harmon et al. 2003; Kozak et al. 2006)? Is early, rapid diversification a general feature of adaptive radiation (Lovette and Bermingham 1999; Schluter 2000)?

To address temporal variation in diversification rates, researchers use an increasingly sophisticated statistical toolkit. Typically, branch lengths in a reconstructed phylogeny are used to estimate the timing of speciation events, and observed waiting times between successive speciation events are compared to the expected distribution under a stochastic branching process where rates have been constant over time (Wollenberg et al. 1996). The most widely used null model for diversification rate analyses is the birth-death model (Kendall 1948), where clades grow under constant per-lineage speciation and extinction rates. Both parametric and nonparametric statistics have been developed to test the distribution of internal nodes in a molecular phylogeny for departures from the pure-birth process, a special case of the birth-death process where all lineages have a common per-lineage extinction rate $\mu = 0$ (Zink and Slowinski 1995; Paradis 1998b; Pybus and Harvey 2000).

Although methods based on the pure-birth process can infer temporal decreases in diversification rates, they are unable to detect temporal increases (Pybus and Harvey 2000; Nee 2001). If the per-lineage extinction rate μ is high relative to the speciation rate λ, we will observe an excess of recently diverged lineages relative to our expectation under the pure-birth model (Fig. 1). This phenomenon, termed the "pull of the present" (Nee et al. 1994a), can mislead researchers into concluding that a particular clade has undergone a temporal increase in the net diversification rate (Nee et al. 1994a; Kubo and Iwasa 1995). The distribution of waiting times under the birth-death process is thus contingent on the ratio of extinction and speciation rates. As we typically do not know and cannot reasonably estimate this parameter (Nee et al. 1994a; Paradis 2004), existing methods are limited to testing for temporal declines in diversification rates using the pure-birth null model.

Likelihood methods based on the birth-death model are a potentially powerful approach for reconstructing the history of diversification from phylogenetic data. Likelihood is increasingly used to reconstruct shifts in diversification rates, either through use of the birth-death model (Barraclough and Vogler 2002; Turgeon et al. 2005) or through a related approach involving survival analysis (Paradis 1997; Emerson et al. 2000; Pitra et al. 2004; Paradis 2005). Rather than explicitly testing a specific null hypothesis of diversification, such as the pure-birth process, likelihood methods do not need to assume a particular parameterization of the birth-death null model. It is thus possible to ask, for example, whether phylogenetic data are more likely to have been generated under a general rate-constant birth-death model ($\lambda > 0; \mu \geq 0$), or whether the data are more likely under a model in which rates have varied over time.

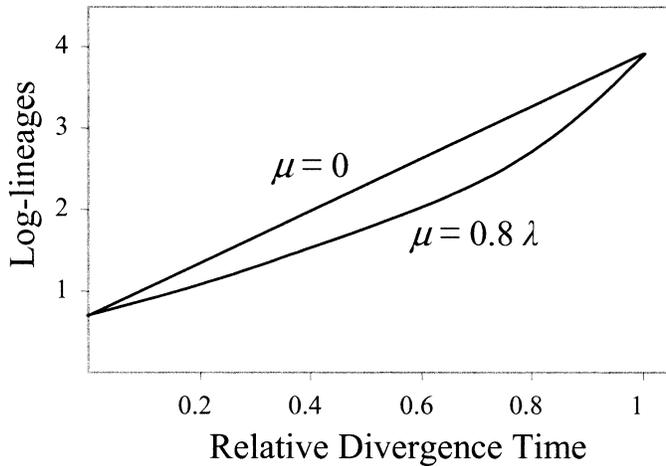Use of a general birth-death null model may afford several

FIG. 1. Average log-lineages-through-time curves for phylogenies of $N = 50$ taxa simulated under a rate-constant birth-death process with $\mu = 0$ or $\mu = 0.8\lambda$. Note the sharp upturn in the number of lineages toward the present for phylogenies simulated under $\mu = 0.8\lambda$.

advantages. Because the null hypothesis of rate-constancy incorporates extinction, it is possible that this approach can separate temporal increases in diversification rates from a rate-constant process with $\mu > 0$. This method may also have greater power to infer temporal decreases in diversification rates when extinction is present, relative to methods based on the pure-birth model. All methods should suffer some loss of power in such situations, due to the tendency of the pull of the present to offset the excess of early branching events, but the effect may be more severe for methods that assume $\mu = 0$. Finally, such model-based approaches can simultaneously assess rate variation over time and provide estimates of relevant diversification parameters (Sanderson 1994; Barraclough and Vogler 2002).

Despite the potential utility of the birth-death likelihood approach, there have been few critical assessments of its performance against datasets simulated under both rate-variable and rate-constant diversification scenarios, and there are additional issues that should be considered by researchers wishing to make use of these methods. Model selection, for example, is the subject of a substantial and growing literature in molecular phylogenetics (Zhang 1999; Posada and Crandall 2001; Pol 2004; Posada and Buckley 2004), but little attention has been given to this subject in the context of diversification rates analysis. Because researchers typically use model selection as a test of the null hypothesis that diversification rates have not varied over time, it is imperative that model selection be based on nonarbitrary criteria that minimize the Type I error rate.

Here, I address model selection and Type I error rates for likelihood-based analyses of diversification, and I evaluate whether this approach can be used to identify temporal increases in diversification rates. I assess the power of likelihood methods to recover the true model of evolution when rates are known to have varied over time, and I compare the birth-death approach to the γ-statistic (Pybus and Harvey 2000), a popular parametric alternative. Because we know little about bias and variances of parameters estimated under

the birth-death model when rates are known to have varied over time, I examine the relative error of parameters inferred from phylogenies simulated under several models of diversification. Finally, I illustrate use of the likelihood approach by analyzing the tempo of diversification in a radiation of Australian lizards.

METHODS

Models for Likelihood-Based Analyses of Diversification

Here I consider the generalized birth-death process as a framework for the analysis of diversification rates. This is a simple null model for the growth of a phylogenetic tree over time: existing lineages give birth to new lineages at a per-lineage rate λ and go extinct at a per-lineage rate μ. The Yule process, or pure-birth model (Yule 1924), is a special case of the birth-death process with $\mu = 0$, where the number of lineages can only increase over time. In contrast, when $\mu > 0$, the number of lineages can decrease over time; a clade diversifying under this model can become extinct even when the speciation rate exceeds the extinction rate.

Kendall (1948) formalized the probability that a stochastic process beginning with n_0 lineages will have n_t progeny after some time t under any constant and time-dependent birth and death rates. Nee et al. (1994b) extended the Kendall model to the case of reconstructed phylogenies; in this case, birth events (speciation) can only be observed if each descendent lineage leaves at least one surviving progeny in the present (Fig. 2). The probability of each waiting time between successive speciation events is equal to the probability that each of the n lineages at the start of the interval has exactly one progeny (itself) after some time t, conditioned on the probability that none of the n lineages go extinct (Nee et al. 1994b; eq. 17):

$$\text{Prob}(t \mid n, \lambda, \mu, T, t_n)$$

$$= n(\lambda - \mu)\exp[-(\lambda - \mu)nt]$$

$$\times \frac{\left\{1 - \dfrac{\mu}{\lambda}\exp[-(\lambda - \mu)(T - t_n - t)]\right\}^{n-1}}{\left\{1 - \dfrac{\mu}{\lambda}\exp[-(\lambda - \mu)(T - t_n)]\right\}^{n}}, \quad (1)$$

where T is the time from the root node to the present, t_n is the birth time of the nth lineage, and t corresponds to the waiting time until the birth of another lineage that can be observed in a reconstructed phylogeny.

From equation (1), we multiply the transition probabilities together to obtain the likelihood that a particular birth-death process has produced the branching times observed in the reconstructed phylogeny (Nee et al. 1994b). Reparameterizing the model such that $r = \lambda - \mu$ (the net diversification rate), $a = \mu/\lambda$ (the extinction fraction), and letting $\mathbf{x}$ be a vector of observed branching times (Fig. 2), we have

$$L(\mathbf{x} \mid a, r) = \prod_{n=2}^{N-1} nr \exp[-nr(x_n - x_{n+1})]$$

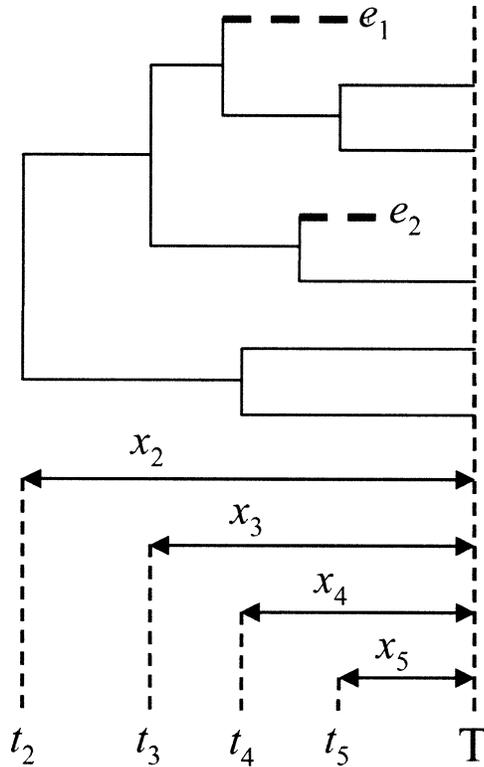$$\times \frac{\{1 - a \exp[-r(x_{n+1})]\}^{n-1}}{\{1 - a \exp[-r(x_n)]\}^{n}}. \quad (2)$$

FIG. 2. A reconstructed phylogenetic tree with five species to illustrate parameters discussed in text. T represents the total time elapsed from root node to the present, t_i is the birth time of the ith lineage, and x_i is the branching time of lineage i ($x_i = T - t_i$). Lineages e_1 and e_2 went extinct at some time before the present, and we are thus unable to see the corresponding speciation events in a reconstructed phylogeny.

Here I consider the special case where a clade diversifies under parameters r_1 and a_1, until some point in time t_s, where rates shift to r_2 and a_2. Due to the computational time required to optimize parameters over all possible values of t_s, I consider only observed branching times as possible shift points. This is a conservative approach, because consideration of shift points other than the observed branching times can only increase the likelihood of the rate-variable model relative to the rate-constant model. This approach has been implemented in several previous studies (Barraclough and Vogler 2002; Turgeon et al. 2005). Throughout the text, I refer to this method as BDL.

An alternative likelihood-based approach to testing for temporal shifts in diversification rate is survival analysis (Paradis 1997, 1998a). Here we think of the time-axis of the phylogeny in reverse: a speciation event in the reconstructed phylogeny becomes a failure event in survival analysis. Each lineage has a probability of $h(t)$ of failure or death; in the case of a reconstructed phylogeny with exponentially distributed branching times, $h(t)$ is equal to the net diversification rate (Cox and Oakes 1984; Paradis 1997). Survival analysis has been used to test for temporal increases in net diversification rates (Near et al. 2003); however, because survival models do not include an extinction term, it is unclear whether they can separate this phenomenon from constant background extinction rates.

I used the birth-death model to analyze phylogenies simulated under a set of rate-variable and rate-constant models of speciation and extinction. Likelihoods of simulated phylogenies were computed under two rate-constant and two rate-variable models: (1) the pure birth model (one parameter, λ, with μ set to zero); (2) a constant-rate birth-death model (two parameters, $\lambda > 0$; $\mu \geq 0$); (3) a pure birth rate-variable model where speciation rate λ_1 shifts to rate λ_2 at some time t_s (three parameters: λ_1, $\lambda_2 > 0$; $0 < t_s < T$); and (4) a rate-variable model with two speciation rates and two extinction rates, but constrained such that the extinction fraction μ/λ remains constant (four parameters: λ_1, $\lambda_2 > 0$; μ_1, $\mu_2 \geq 0$; $0 < t_s < T$; but $\mu_1/\lambda_1 = \mu_2/\lambda_2$). The distribution of branching times in a reconstructed phylogeny is a function of the net diversification rate, $\lambda - \mu$, and the extinction fraction, μ/λ, and throughout the text I denote these parameters by $r = \lambda - \mu$ and $a = \mu/\lambda$. I also computed likelihoods of rate-constant phylogenies using survival models to assess Type I error rates under different values of a.

For each simulated phylogeny, parameters were optimized for each birth-death model under consideration. I then calculated the likelihood of the data given the model for all models using these optimized parameters. For survival analysis, the likelihood of the simulated data was calculated under two models: a rate-constant model, and a rate-variable model where the diversification rate is modeled as a function of time, $r(t) = \alpha\beta(\alpha t)^{\beta - 1}$. In the latter case, the distribution of failure times follows a Weibull distribution, where the diversification rate decreases over time if $\beta > 1$, increases if $\beta < 1$, and is constant if $\beta = 1$. Simulation and optimization details are given in the Appendix.

Model Selection

In the likelihood methods discussed here, model selection is equivalent to hypothesis testing: we are asking, at least initially, whether we can reject a rate-constant model in favor of a model where diversification rates have varied over time.

The Akaike information criterion (AIC; Akaike 1973) has been widely used to select among different models of diversification (e.g., Paradis 1997; Emerson et al. 2000; Barraclough and Vogler 2002; Pitra et al. 2004). Typically, AICs are calculated for a set of models, and the model with the lowest AIC is taken to be the model that best approximates the data. The AIC is a function of both the log likelihood and the number of free parameters in a given model:

$$\text{AIC} = -2 \log L + 2p, \tag{3}$$

where p is the number of parameters that are estimated from the data. Thus, more parameters typically improve the fit of the model to the data, but also increase the penalty term ($2p$).

Some argue that it is inappropriate to use the AIC or any other information-theoretic measure for hypothesis testing, in part because it is believed that the traditional hypothesis testing framework is uninformative and based on arbitrary rejection criteria (Burnham and Anderson 2002). A philosophical discussion of this matter is beyond the scope of the present paper, but several considerations justify the use of the AIC for the present purpose: (1) virtually all research in this area has employed the AIC as a test statistic to distinguish

between rate-constant and rate-variable models of diversification; and (2) as will be shown, the AIC performs well in practice and lends itself easily to Monte Carlo methods used to infer the distributions of other parametric and nonparametric test-statistics.

Is it sufficient to select the model with the lowest AIC, regardless of the difference in AIC scores between the best and second-best models (Paradis 1997)? This issue is essentially one of confidence: if the best rate-variable model represents an improvement in fit of x AIC units over the best rate-constant model, how confident can we be that the rate-variable model is the better approximation of our data?

To address this issue, I simulated 1000 phylogenies each of $N = 50$ taxa under 10 rate-constant diversification scenarios. The net diversification rate r was identical for all 10 scenarios, but the extinction fraction a was varied from $a = 0$ to $a = 0.9$. The likelihood of each simulated phylogeny was calculated under the four variants of the birth-death model and two survival models described previously. AIC scores were computed for each model to address whether the birth-death likelihood approach is susceptible to high Type I error rates and whether objective criteria based on the AIC can be established to select among rate-variable and rate-constant models of diversification such that the Type I error rate is minimized.

For each simulated phylogeny, the difference in AIC score between the best rate-constant and rate-variable models was calculated as

$$\Delta AIC_{RC} = AIC_{RC} - AIC_{RV}, \qquad (4)$$

where AIC_{RV} is the lowest AIC score among rate-variable models (hereafter referred to as the candidate set) and AIC_{RC} is the lowest AIC score among the two rate-constant models. ΔAIC_{RC} is positive if a rate-variable model best fits the data and negative if a rate-constant model is the better fit. Use of ΔAIC_{RC} as a test-statistic permits us to identify a difference in AIC scores between rate-constant and rate-variable models such that $\alpha \leq 0.05$. The ΔAIC_{RC} giving a Type I error rate of $\alpha = 0.05$ corresponds to the 95th percentile of the distribution of ΔAIC_{RC} scores tabulated from phylogenies simulated under the null hypothesis of rate-constancy.

One might predict that the probability of Type I error will increase for larger phylogenies, because of the greater number of likelihood estimates (under each rate-variable model, we have a likelihood of a rate shift for each branching time). If this is the case, AIC criteria for rejecting the null hypothesis of rate-constancy will require adjustment based on the number of taxa in our tree. I simulated 1000 phylogenies of $N = 15, 30, 60,$ and 100 taxa under the pure birth model to determine whether Type I error rates show dependency on sample size.

It is also possible that the Type I error rate will increase as the number of rate-variable models under consideration increases. For each set of simulated phylogenies, I computed the distribution of ΔAIC_{RC} under three different candidate sets of rate-variable models. In addition to the two rate-variable models described previously, I added a five-parameter variant of the pure birth model with three speciation rates and two shift points, all of which were optimized for each simulated phylogeny. This model may be useful in many

situations of interest to biologists; for instance, rapid diversification early in the history of a clade could limit our power to detect recent rate shifts that may have occurred, for example, during the Pleistocene. By decoupling these processes, we may better be able to approximate the true tempo of diversification. The candidate sets of rate-variable models for this analysis consisted of (1) the three-parameter model only; (2) three-and four-parameter models; and (3) three-, four- and five-parameter models.

Finally, I used survival analysis to assess the distribution of ΔAIC_{RC} as a function of the extinction fraction a. If there is a positive relationship between ΔAIC_{RC} and a, this method cannot be used to distinguish temporal increases in diversification from a rate-constant model with $a \geq 0$.

Power to Detect Temporal Variation in Diversification Rates

A good statistical framework for diversification rates analysis cannot simply minimize Type I error: it must have sufficient power to detect temporal variation in diversification rates when it is present. Of primary concern is whether we can ever detect a temporal increase in the net diversification rate using phylogenies of extant taxa only. Constant, nonzero extinction rates produce an apparent excess of recently diverged lineages relative to the pure-birth model (Nee et al. 1994a; Kubo and Iwasa 1995), possibly leading to the erroneous conclusion that diversification rates have increased over time (Nee 2001). The most widely used diversification rates test-statistic, γ, can only be used to detect temporal decreases in diversification (Pybus and Harvey 2000). This fact is not widely appreciated, and positive γ-values have been interpreted to support temporal increases in diversification rates (e.g., Linder et al. 2003; Turgeon et al. 2005).

I assessed the power of the birth-death likelihood approach by simulating 500 phylogenies of $N = 30, 60,$ and 100 taxa under each of eight models of diversification, yielding a total of 24 sets of 500 phylogenies. For each N, phylogenies were simulated with a constant extinction fraction $a = 0$ or $a = 0.5$ under the following scenarios: (1) fivefold decrease in the net diversification rate r; (2) twofold decrease in r; (3) twofold increase in r; and (4) fivefold increase in r. Clades grew under rate r_1 from an initial size of two lineages until the birth of the $N/2$ lineage. Rates then shifted to a second diversification rate r_2, and the clade was permitted to grow until the Nth lineage was born. Thus, a phylogeny of 60 taxa would grow under rate r_1 until the birth of the 30th lineage, after which the clade continued to grow with a new rate r_2 until the clade reached a final size of $N = 60$. Extinction was maintained at a constant level of $a = 0$ or $a = 0.5$ throughout the duration of the simulation.

The power of the likelihood method to recover the true rate-variable model of evolution was simply the percentage of trees for which the null hypothesis of rate-constancy was rejected. The rate constant model was rejected only if ΔAIC_{RC} was less than the critical value of the simulated null distribution. The likelihood of each phylogeny was assessed under two rate-constant and two rate-variable models of diversification.

I contrasted the power of the likelihood approach to that

of the γ-statistic, a test of the distribution of internal nodes in a phylogeny. Because the γ-statistic tests for departures from the pure birth ($a = 0$) model of diversification, it cannot distinguish between an increase in speciation and a constant rate model with $a > 0$ (Pybus and Harvey 2000); both scenarios result in an apparent excess of speciation events near the present relative to the pure birth model. The γ-statistic follows a standard normal distribution under the pure birth process, and we reject the null hypothesis of rate-constancy in favor of a temporal decrease in diversification when $γ < -1.645$.

I further explored the power of these methods to detect temporal decreases in diversification rates when rates have shifted earlier or later than the $N/2$ speciation event. I simulated a two-rate branching process where the net diversification rate decreased 3.5-fold with the birth of the kth taxon ($k = 5, 10, 15, 20, 25, 30, 35, 40, 45, 50,$ or 55) under background extinction rates of $a = 0$ and $a = 0.5$. I generated 500 phylogenies of $N = 60$ taxa for each combination of a and k and applied both BDL and the γ-statistic to each set of simulated phylogenies.

In the above rate-variable scenarios, diversification rates shift instantaneously from rate r_1 to r_2. Because the likelihood models implemented here are designed precisely to detect this type of rate shift, it is possible that the γ-statistic would perform better than likelihood methods under models where diversification rates have changed gradually over time. To address this possibility, I considered a model of density-dependent cladogenesis, where diversification rates are inversely related to the number of lineages surviving at any point in time. I simulated two sets of 500 phylogenies ($N = 60$ taxa; $a = 0$ or $a = 0.5$) under a density-dependent model with $r(n) = r_0 n^{-x}$, with $x = 0.45$. Likelihoods were computed under each model as described above.

Parameter Estimation

A potential advantage of the likelihood approach is that it provides estimates of the timing and magnitude of rate shifts. It is not possible to estimate extinction rates from molecular phylogenies with any reasonable degree of confidence, but estimates of the net diversification rate fare much better (Nee et al. 1994a; Kubo and Iwasa 1995; Paradis 2004). To date, no studies have assessed the performance of the BDL method in parameter estimation when rates are known to have varied over time. For each set of $N = 30, 60,$ and 100 phylogenies simulated under four different rate-variable models, I investigated whether likelihood methods could reasonably approximate the net diversification rates before and after the rate shift, r_1 and r_2, and the timing of the rate shift, t_s.

Because phylogenies were simulated under a specified model of diversification, I calculated the relative error in the estimates of r_1 and r_2 as

$$\frac{\hat{r}_i - r_i}{r_i}, \tag{5}$$

where $\hat{r}_i$ is estimated from the data and r_i is the true value. Thus, positive values indicate overestimates of the net diversification rate, and negative values indicate underesti-

mates. I expressed the error in t_s as a percentage of the total age of each simulated phylogeny,

$$\frac{\hat{t}_s - t_s}{T}, \tag{6}$$

where $\hat{t}_s$ is the inferred time of the rate shift in time units from the start of the simulation ($t = 0$), t_s is the true shift time, and T is age of the simulated phylogeny. Thus, a positive value indicates that the inferred rate shift is later than the true shift point.

Example: Australian Agamid Radiation

To illustrate how these methods might be used, I analyzed a recent molecular phylogeny for Australian lizards in the family Agamidae (Harmon et al. 2003). The agamids (dragon lizards) constitute a substantial component of squamate reptile diversity in Australia and show considerable diversity in morphology and ecology (Pianka 1986; Melville et al. 2001). The analysis of Harmon et al. (2003) was based on 69 extant taxa, including at least 93% of the species known from Australia.

The agamid tree was constructed by maximum likelihood from 1800 bp of mitochondrial DNA using the GTR + I + Γ model of sequence evolution (Harmon et al. 2003), and the tree was made ultrametric using nonparametric rate smoothing (NPRS; Sanderson 1997). I evaluated the tempo of diversification in the Agamidae with BDL, fitting two rate-constant and two rate-variable models of speciation and extinction using maximum likelihood methods described in this paper. To infer diversification parameters, an age of 30 million years was assigned to the basal divergence between Australian and Southeast Asian agamids (Hugall and Lee 2004).

RESULTS

Model Selection

When data are simulated under a model in which rates do not vary over time, selecting the model with the lowest AIC score leads to Type I error rates exceeding 28% in all cases (Fig. 3A); for some values of a, the error rate exceeded 40%. For phylogenies simulated under rate-constant models, the 95th percentile of the distribution of ΔAIC_{RC} corresponds to $α = 0.05$: by definition, 5% of the differences in AIC scores between the best rate-constant and rate-variable models exceed this value. The 95th percentile values are approximately constant or decreasing slightly from $a = 0$ to $a = 0.9$ (Fig. 3B).

There is a high variance in ΔAIC_{RC} values, accounting for the ragged distribution of error rates and $α_{0.05}$ rejection regions. Because error rates do not increase with extinction levels, we can reject the rate-constant model if ΔAIC_{RC} for a test phylogeny is greater than the critical value determined by simulating rate-constant phylogenies under the pure-birth process. The lack of a positive relationship between a and ΔAIC_{RC} suggests that the birth-death likelihood approach may be capable of detecting a temporal increase in the net diversification rate.

The ΔAIC_{RC} required to maintain $α \leq 0.05$ increases with sample size (Fig. 3C): as the number of taxa increases, we
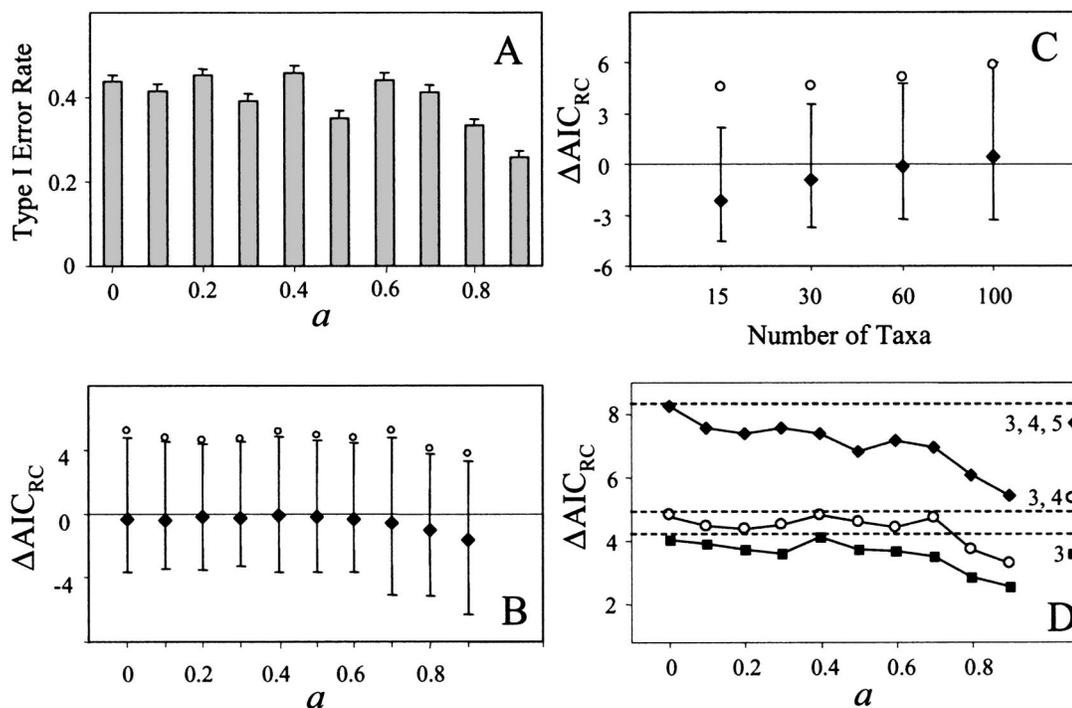
FIG. 3. Type I error rates and ΔAIC_{RC} rejection criteria for rate-constant phylogenies simulated under different values of the extinction fraction a and analyzed with a set of rate-constant and rate-variable birth-death models. (A) Type I error rate as a function of a if the model with the lowest AIC score is selected as that which best approximates the data. Type I error rates range between 28% and 45%, but do not increase substantially with a. (B) Differences in AIC scores between the best rate-constant and rate-variable models as a function of a. Shown are median ΔAIC_{RC} scores (black diamonds), 0–95th percentile range of ΔAIC_{RC} (error bars), and bootstrap standard errors of the 95th percentile ΔAIC_{RC} value (open circles). The 95th percentile of the ΔAIC_{RC} distribution (upper-bound error bar) equals the AIC difference between rate-constant and rate-variable models corresponding to $\alpha = 0.05$ and does not increase with a. Negative values of ΔAIC_{RC} are obtained when AIC scores for the best rate-variable model exceed those of the best rate-constant model. (C) Distribution of ΔAIC_{RC} as a function of the number of taxa in simulated phylogenies: as the number of taxa increases, a greater difference in AIC scores between the best rate-constant and rate-variable models is required to maintain $\alpha = 0.05$. (D) The 95th percentile of the distribution of ΔAIC_{RC} when data are analyzed with multiple rate-variable models. Dashed lines indicate ΔAIC_{RC} rejection criteria required to maintain $\alpha = 0.05$ when analysis is based on different sets of three-, four-, and five-parameter rate-variable models (model set indicated adjacent to each dashed line). As more parameters are added to models in the candidate set, ΔAIC_{RC} values corresponding to $\alpha = 0.05$ increase. All estimates are based on 1000 simulated phylogenies per a.

require greater differences in likelihood scores between the best rate-constant and rate-variable models to reject the null hypothesis. When N is small, there is a high variance in the 95th percentile values of the ΔAIC_{RC} distribution, but this decreases for larger phylogenies. For phylogenies of $N = 60$ and 100 taxa, Type I error rates exceeded 50% if the model with the lowest AIC is selected as that which best approximates the data. We have little confidence that a rate-variable model best approximates the data when the AIC difference between rate-constant and rate-variable models is less than 3. Only when ΔAIC_{RC} approaches 4 for small ($N = 30$) phylogenies and 5.5 for large ($N = 100$) phylogenies can the rate-constant model be rejected with confidence.

There is a positive relationship between ΔAIC_{RC} required to maintain $\alpha \leq 0.05$ and the number of rate-variable models under consideration (Fig. 3D). As more complex models are added to the candidate set, Type I error rates increase unless ΔAIC_{RC} rejection criteria are adjusted accordingly. This is particularly apparent with the addition of the five parameter rate variable model, where the ΔAIC_{RC} corresponding to $\alpha \leq 0.05$ jumps from 4.8 to 8.2.

As predicted, likelihood methods based on survival anal-

ysis show high Type I error rates when the extinction fraction a increases (Fig. 4). Because available survival models do not incorporate the effects of extinction, they cannot distinguish between temporal increases in diversification rates and constant-rate models with $a > 0$.

Power Analyses

For each of the 500 phylogenies simulated under a given diversification model, the rate constant model was rejected in favor of a rate-variable scenario if ΔAIC_{RC} exceeded that required to maintain a Type I error rate of 0.05 under the pure-birth model (Fig. 3C). Power to reject the null hypothesis of rate-constancy is generally high when diversification rates have decreased over time (Fig. 5). Likelihood can separate temporal increases in diversification rates from constant, nonzero extinction, but power to detect this shift is high only when sample sizes are large ($N = 60, 100$) and the magnitude of the rate shift is great. When diversification models included constant background extinction ($a = 0.5$), power to reject the null hypothesis decreased in all cases.

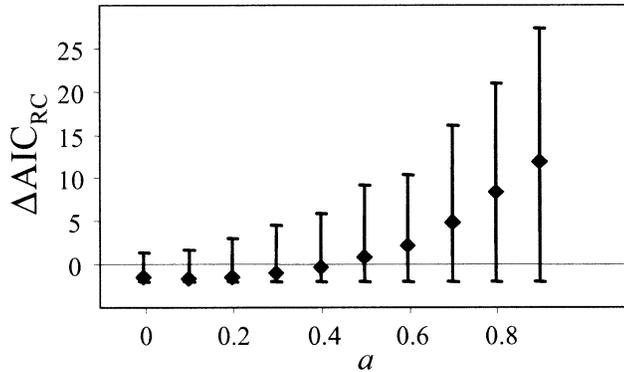Under diversification models with extinction set to zero,

FIG. 4. Relationship between ΔAIC_{RC} and the extinction fraction a when survival analysis is used to compute likelihoods. Shown are median (black diamonds) and the 0–95th percentile range of the distribution of ΔAIC_{RC}. Median values and 95th percentile of the ΔAIC_{RC} distribution are positively correlated with a, indicating that this method cannot separate a temporal increase in diversification from a rate-constant model with $a > 0$. All estimates are based on 1000 simulated phylogenies per a.

BDL and the γ-statistic had approximately equal power to detect temporal decreases in diversification rates. When diversification models included nonzero extinction rates ($a = 0.5$), likelihood methods performed much better than the γ-statistic. BDL showed much greater power to detect temporal decreases in diversification rates when rate shifts occur early in the history of a clade, particularly when extinction is present (Fig. 6). In the case of density-dependent cladogenesis, power was virtually identical for BDL and the γ-statistic when $a = 0$ (Fig. 7). Likelihood performed better under $a =$

0.5, but power to reject the null hypothesis was low for both methods. Despite a nearly five-fold decrease in the net diversification rate over time, power was much lower than in the case where rate shifts occurred instantaneously (Fig. 5).

Parameter Estimation

For models with $a = 0$, estimates of the net diversification rate showed a weak bias for small phylogenies (Fig. 8), but the mean relative errors tended to zero as the number of taxa increased. For weak increases or decreases in diversification rates, variance in estimates of r_1 and r_2 was high. Estimates of the time of the rate shift t_s did not show appreciable bias for most of the diversification scenarios considered. For weak increases in diversification rates, mean relative errors in t_s suggested a consistent underestimate of true shift times; however, median relative errors in this parameter were less than 0.01 for all three sets of simulated phylogenies ($N = 30, 60, 100$). For five-fold increases or decreases in diversification rates, error in estimates of t_s was very low for phylogenies of $N = 60$ and 100 taxa. For models with $a = 0.5$, estimates of r_1, r_2, and t_s generally appeared unbiased and tended to zero as the number of taxa increased.

Australian Agamids

The agamid phylogeny was analyzed under two rate-constant and two rate-variable models of diversification (Table 1). The best rate-variable model had an AIC of 24.72 versus 40.68 for the best rate-constant model ($\Delta AIC_{RC} = 15.96$). The ΔAIC_{RC} value required to maintain α = 0.05 was obtained by simulating 1000 phylogenies of the same size as
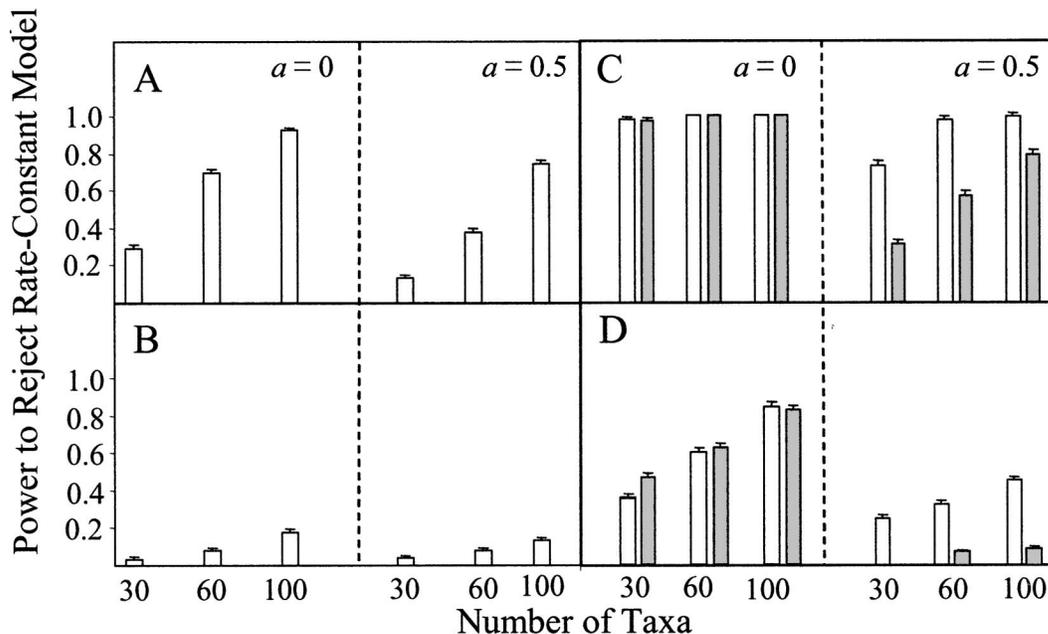

FIG. 5. Power of birth-death likelihood (BDL; open bars) and the γ-statistic (shaded bars) to detect shifts in diversification rates over time. Test phylogenies were simulated under a model where diversification rates shift with the birth of the $N/2$ lineage. Power is the percentage of 500 replicate phylogenies for which the null hypothesis of rate-constancy was rejected for each value of a and N if the ΔAIC_{RC} rejection region is set to maintain α = 0.05 (BDL) or if γ ≤ −1.645. (A) fivefold increase in r over time; (B) twofold increase in r; (C) fivefold decrease in r; (D) twofold decrease in r. Power for BDL exceeded that of the γ-statistic in all but two of the diversification scenarios considered (D; $a = 0$; $N = 30$ and 60). Confidence limits are bootstrap standard errors of each power estimate.
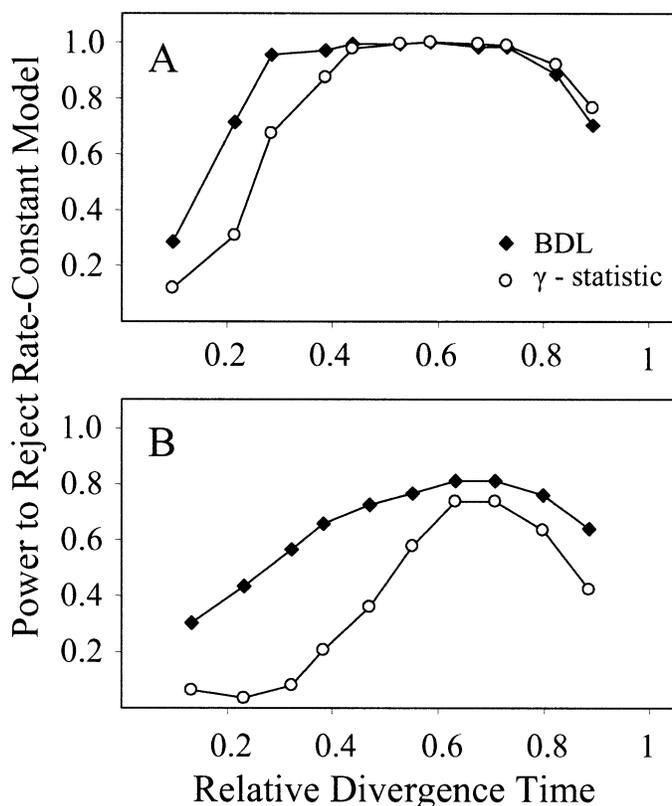
FIG. 6. Power of birth-death likelihood (BDL) and the γ-statistic to detect temporal decreases in diversification rates when rates have shifted at different points in time. Test phylogenies ($N = 60$) were simulated under (A) $a = 0$ and (B) $a = 0.5$, with a 3.5-fold decrease in the net diversification rate occurring with the birth of the 5th, 10th, . . . , 50th, or 55th lineage. Power is the percentage of 500 replicate phylogenies for which the null hypothesis of rate-constancy was rejected. Relative divergence times reflect the mean time of the rate shift for each set of simulated phylogenies, expressed as a fraction of total clade age. Bootstrap standard errors of each power estimate (not shown) were less than 0.022.

FIG. 7. Power of birth-death likelihood (BDL; open bars) and the γ-statistic (shaded bars) when diversification rates have decreased gradually under a model of density-dependent cladogenesis. (A) Simulation model: the net diversification rate r decreases as the number of surviving lineages increases. (B) Percentage of 500 simulated phylogenies ($N = 50$) for which the rate-constant model was rejected for likelihood and the γ-statistic under $a = 0$ and $a = 0.5$.

the test phylogeny under the pure-birth model. For $N = 69$, we reject the null hypothesis of rate-constancy if the observed $\Delta AIC_{RC} \geq 5.0$.

ΔAIC_{RC} for the agamids was much greater than the ΔAIC_{RC} required to reject the null hypothesis at $\alpha = 0.05$, providing strong support for the hypothesis that diversification rates have changed over time. Diversification rates in the Australian agamids appear to have decreased roughly threefold over time (Table 2), with an estimated rate shift occurring 12–14 million years ago. Confidence limits were placed on estimated diversification rates using Moran's variance (Nee 2001).

Log-likelihoods of rate shifts under the best rate-variable model were plotted as a function of time (Fig. 9). Although the maximum likelihood shift point occurs approximately 13 million years ago, two additional peaks in the likelihood plot (Fig. 9B; 6 million and 19 million years ago) suggest that agamid diversification trends may have occurred in a stepped fashion. This pattern of multiple likelihood peaks separated by troughs appears inconsistent with a model of gradually decreasing diversification rates over time, as could occur if
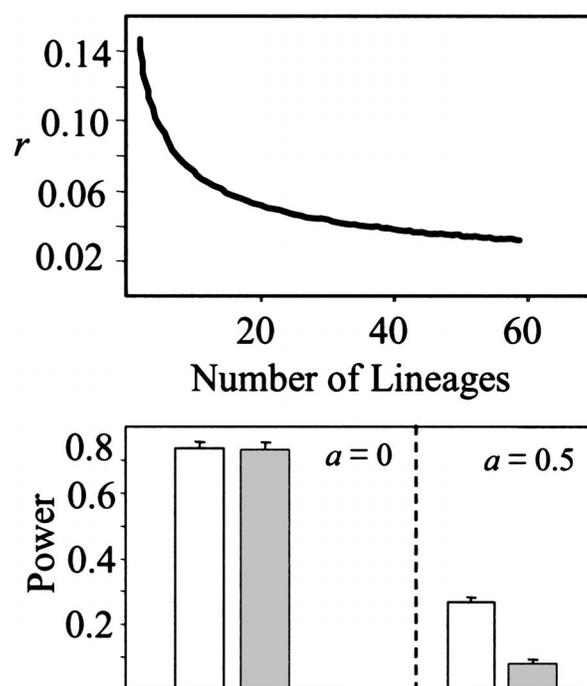
the agamid radiation in Australia was characterized by density-dependent cladogenesis on a continental scale. Additional work is needed to determine whether this pattern can be attributed to stochasticity or a true stepwise decline in diversification.

DISCUSSION

Model Selection

Type I error rates are high when the model with the lowest AIC score is selected as that which best approximates the data, approaching 50% in some trials. This effect appears to be most pronounced for phylogenies generated under $a \leq 0.7$. Because there is no positive relationship between ΔAIC_{RC} and a, we can reject the null hypothesis of rate-constancy if ΔAIC_{RC} is greater than or equal to that required to maintain $\alpha = 0.05$ under the pure-birth model for a given value of N. The dependency of this rejection criterion on the size of the clade (Fig. 3C) and the number of models under consideration (Fig. 3D) suggests that no single criterion can be used to select among rate-constant and rate-variable models. Researchers studying clades larger or smaller than those considered here may wish to obtain ΔAIC_{RC} scores required to maintain an acceptable Type I error rate by simulating phylogenies of the same size as the actual phylogeny under the pure-birth model and examining the resulting distribution of ΔAIC_{RC} scores with respect to the full set of candidate models.
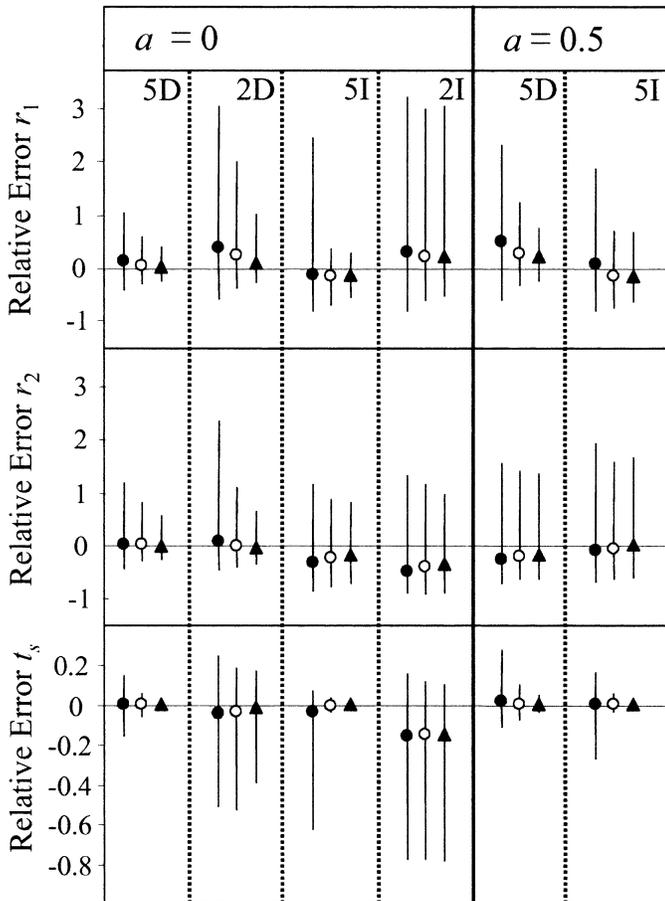
FIG. 8. Relative error in parameters estimated under the best fit rate-variable model for phylogenies simulated under four models of diversification with $a = 0$ and two models with $a = 0.5$. r_1 and r_2 are maximum likelihood estimates of the net diversification rate before and after the inferred rate shift, t_s. Shown are means and 95% confidence limits on the distribution of relative errors for each set of simulated phylogenies. In general, mean relative errors tended to zero as the number of taxa increased.

TABLE 2. Diversification parameters for the Australian Agamidae under the best-fit model before and after the inferred rate shift (my, million years).

	r	95% lower	95% upper
Divergences >13 my ago	0.148/my	0.119/my	0.196/my
Divergences ≤13 my ago	0.048/my	0.039/my	0.063/my

correct the LRT for multiple testing (Paradis 1998a; Pol 2004), and this problem is avoided by use of the AIC.

Power Analyses and Parameter Estimation

Despite the high $\Delta\mathrm{AIC}_{RC}$ values required to maintain Type I error rates of 0.05 or less, BDL retains considerable power to detect variation in diversification rates over time. When extinction rates are equal to zero, differences in power for BDL and the γ-statistic are trivial, at least for the abrupt rate-shift scenarios considered here. However, when clades have grown under a model that includes constant, nonzero extinction rates, BDL has much greater power than the γ-statistic.

Decreased diversification through time results in an excess of early diverging lineages in reconstructed phylogenies, but constant background extinction can reduce this effect by eliminating ancient divergences and favoring recently diverged lineages. By explicitly including extinction, model-based approaches can succeed where other approaches fail.

It is far more difficult to detect a temporal increase in diversification than a temporal decrease. In the birth-death likelihood framework, power to detect weak increases in diversification rates is always low. However, power increases with sample size when the magnitude of the rate shift is large. Furthermore, this may be the only approach available that can detect an increase when it occurs. For most groups, the fossil record does not permit independent estimates of extinction rates, and it is difficult to justify why a particular background extinction rate is less likely than any other; this limits the utility of methods where rejection criteria are dependent on a to cases where diversification rates have decreased over time. Methods that do not explicitly account for extinction, such as the γ-statistic and survival analysis, are unable to distinguish between temporal increases in diversification rates and rate-constant models with $a \geq 0$.

Under both $a = 0$ and $a = 0.5$, parameter estimation using the birth-death model performs well; estimates of both the timing and magnitude of the rate shift were consistent and generally unbiased across a range of rate-variable diversification scenarios. The two main conclusions from previous work on parameter estimation in a birth-death framework are

Alternative model selection criteria, such as the AICc (Burnham and Anderson 2002) and the likelihood ratio test (LRT) also have high Type I error rates (data not shown), and there seems to be little advantage in abandoning the AIC, provided that error rates and rejection criteria are explicitly addressed through simulation. Even if Monte Carlo methods are used to infer the null distribution of the LRT, this method results in a large number of pairwise LRTs, the significance of which are difficult to compute and interpret (Paradis 1998a). When comparing more than two models, one must

TABLE 1. Results of fitting four birth-death models to the agamid data. ΔAIC is the difference in AIC scores between each model and the overall best-fit model.

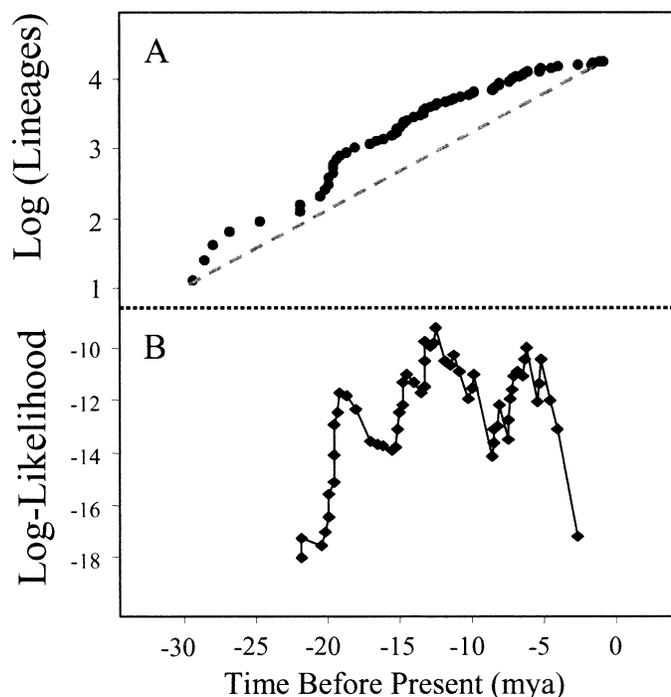| Description of model	Rate-constant $a = 0$	Rate-constant $a \geq 0$	Rate-variable $a = 0$	Rate-variable $a \geq 0$
Parameters in model	r	r, a	r_1, t_s, r_2	r_1, t_s, r_2, a
Log-likelihood	−19.34	−19.34	−9.36	−9.36
AIC	40.68	42.68	24.72	26.72
ΔAIC	15.96	17.96	0	2

FIG. 9. Analysis of diversification in Australian agamid lizards. (A) Log-lineage through time plot using the nonparametric-rate-smoothing tree from Harmon et al. (2003). Dashed line represents expected rate of lineage accumulation under the pure birth model. (B) Log-likelihood of rate shifts at different points in time for the agamid data. The best fit model was the three parameter model with two speciation rates and extinction set to zero, with an estimated rate shift occurring 13 million years ago. Estimated speciation rates are 0.148 per million years and 0.048 per million years before and after the shift point, respectively, assuming a clade age of 30 million years.

that (1) it is always difficult to estimate a in the absence of fossils; and (2) we can infer r with reasonable confidence when a is low (Nee et al. 1994a; Paradis 2004). This study extends previous work in two important respects. It is shown that, at least for the diversification models considered here, estimates of the timing of the rate shift are unbiased and have low variance. The sole exception to this pattern occurs when rates show weak increases over time; this can be explained, at least in part, by the low power of BDL to detect such shifts in diversification rates (Fig. 5B). This consistency implies that the method will be useful for testing specific hypotheses about the causes of variation in diversification rates. To the extent that we can accurately calibrate divergence times in a molecular phylogeny, we should be able to use the method to examine, for example, whether rate shifts coincide with climatic or geological events that may have influenced the tempo of diversification.

A second point is that the relative error in estimates of the magnitude and timing of rate shifts when diversification rates increase over time perform nearly as well as estimates when rates decrease over time. Confidence limits on the distribution of relative errors under scenarios of increasing and decreasing diversification are similar, particularly for phylogenies of $N = 60$ and 100 taxa. Provided that we can detect a temporal increase in the net diversification rate when it occurs, we can

be reasonably confident in the resulting estimates of the magnitude and timing of the rate shift.

Strengths of the Birth-Death Likelihood Approach

Primary strengths of the likelihood-based methods discussed here include power to detect both increases and decreases in diversification over time and parameter estimation. Furthermore, the method is very flexible and could accommodate tests of a wide range of rate variable models; although I considered only simple models, net diversification rates could be modeled as functions of any number of independent variables. For example, rates can be modeled as a function of the number of extant lineages at any point in time to test whether decreasing diversification rates show density dependence. Another application might include testing whether diversification rates are related to climatic oscillations (Kadereit et al. 2004). When we have reason to suspect that a particular event in earth history may have affected diversification, rate-variable models can be adjusted to test a priori hypotheses. Finally, likelihood methods can test for simultaneous rate shifts across multiple, distantly related groups of organisms, provided we can appropriately calibrate divergence times (Turgeon et al. 2005).

Limitations of Likelihood Methods

Conclusions about the temporal nature of diversification are dependent upon the quality of the data themselves. Here I have ignored error in branching time estimates to focus on the analysis of diversification from these data. In reality, many factors can influence or bias estimates of divergence times. Errors in tree topology per se are not necessarily fatal for likelihood-based analyses of diversification; for example, an ancient, rapid radiation may result in a virtual star topology, with all lineages appearing to originate simultaneously. It may never be possible to reconstruct the true phylogeny for these taxa, but it may nonetheless be possible to infer with confidence that diversification rates have decreased over time.

Biased estimates of divergence times, on the other hand, are expected to introduce a similar bias into diversification rate analyses. For example, failure to use an appropriate model of sequence evolution to estimate branch lengths could lead to consistent underestimates of divergence times early in the history of a clade. This is expected to occur when simple models of sequence evolution cannot fully capture the effects of saturation that become apparent when comparing highly diverged lineages (Arbogast et al. 2002). In this case, one might conclude that decreasing diversification rates have prevailed, when the pattern is a simple artifact of saturation (Revell et al. 2005).

All methods developed to infer temporal shifts in diversification rates, including BDL and the γ-statistic, assume that diversification rates have been constant across lineages. The extent to which violations of this assumption influence BDL and other methods is unknown. Revell et al. (2005) found no significant correlation between tree imbalance and the γ-statistic for phylogenies generated under rate-constant models of diversification, but more research is clearly needed on this topic.

Incomplete taxon sampling is a potentially severe problem that has not been addressed for BDL. If a number of taxa in a clade are not sampled, and if sampled lineages are random with respect to divergence times, we will observe a spurious decrease in diversification over time, because missing branching times tend to occur closer to the present than the root of the tree (Nee et al. 1994a). Monte Carlo approaches can be used to generate the null distribution of branching times with incomplete sampling, if the true number of species in the clade is known (Pybus and Harvey 2000; Pybus et al. 2002). A similar approach could be used with BDL to obtain the ΔAIC_{RC} score required to reject a rate-constant model while minimizing Type I error rates.

The BDL approach could be adapted to test a variety of birth-death models other than the simple rate shift models considered here. In the case of density-dependent cladogenesis (Fig. 7), one might predict that power to reject the null hypothesis would increase if a density-dependent model had been included in the candidate set. However, Type I error rates increase with the number of rate-variable models considered (Fig. 3D); a greater difference in AIC scores between the best rate-constant and rate-variable models is required to reject the null hypothesis when the number of fitted models is increased. Because Type I error rates do not increase with a, we can easily determine appropriate rejection criteria for the null hypothesis under any set of candidate models by simulating rate-constant phylogenies and examining the distribution of ΔAIC_{RC} values.

The relationship between hypothesis testing and data exploration can easily become blurred when applying these methods, and this is a particular concern when using a priori hypotheses to test models of diversification. If, before looking at any data, we have reason to believe that diversification rates may have shifted at some point in time, we can justifiably reduce the number of free parameters in rate-variable models by fixing the timing of the rate shift to the hypothesized value (Near et al. 2003; Turgeon et al. 2005). This reduces the AIC penalty term of the rate-variable model relative to the rate-constant model, and we have greater power to reject the null hypothesis of rate-constancy if it is false. However, there is danger that researchers will simply look at a phylogenetic tree or a plot of lineages-through-time, note an apparent increase or decrease in diversification, and then construct an a posteriori hypothesis to explain this pattern. If this a posteriori observation is then treated as an a priori hypothesis in a model-based analysis, the probability of a Type I error may be increased. This is a subtle but important point, and researchers must be careful to justify why a particular a priori hypothesis was used.

Summary and Recommendations

The analyses presented here indicate that the BDL approach is a useful framework for studying temporal variation in diversification rates. In all scenarios considered, BDL showed comparable or greater power than the γ-statistic to detect temporal variation in diversification rates. The advantages afforded by BDL are particularly apparent when extinction is present, suggesting that researchers may wish to revisit datasets for which the γ-statistic was unable to detect a temporal shift in the diversification rate. Furthermore, BDL appears to be unique among available methods in that it can separate a temporal increase in diversification from constant, nonzero extinction rates. Finally, BDL provides parameter estimates that can be used to formulate more specific hypotheses about underlying processes that have influenced the tempo of diversification. However, model overfitting is a potentially serious problem and may lead to a high Type I error rate; researchers must carefully address this issue before concluding that the tempo of diversification has changed over time.

LITERATURE CITED

Akaike, H. 1973. Information theory and an extension of the maximum likelihood principle. Pp. 267–281 *in* Proceedings of the second international symposium on information theory. Akademiai Kiado, Budapest.

Arbogast, B. S., S. V. Edwards, J. Wakeley, P. Beerli, and J. B. Slowinski. 2002. Estimating divergence times from molecular data on phylogenetic and population genetic timescales. Annu. Rev. Ecol. Syst. 33:707–740.

Barraclough, T. G., and S. Nee. 2001. Phylogenetics and speciation. Trends Ecol. Evol. 16:391–399.

Barraclough, T. G., and A. P. Vogler. 2002. Recent diversification rates in North American tiger beetles estimated from a dated mtDNA phylogenetic tree. Mol. Biol. Evol. 19:1706–1716.

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

Chan, K. M. A., and B. R. Moore. 2002. Whole-tree methods for detecting differential diversification rates. Syst. Biol. 51:855–865.

Cox, D. R., and D. Oakes. 1984. Analysis of survival data. Chapman and Hall, London.

Emerson, B. C., P. Oromi, and G. M. Hewitt. 2000. Colonization and diversification of the species *Brachyderes rugatus* (Coleoptera) on the Canary Islands: evidence from mitochondrial DNA COII gene sequences. Evolution 54:911–923.

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.

Hey, J. 1992. Using phylogenetic trees to study speciation and extinction. Evolution 46:627–640.

Hugall, A. F., and M. S. Y. Lee. 2004. Molecular claims of Gondwanan age for Australian agamid lizards are untenable. Mol. Biol. Evol. 21:2102–2110.

Kadereit, J. W., E. M. Griebeler, and H. P. Comes. 2004. Quaternary diversification in European alpine plants: pattern and process. Proc. R. Soc. Lond. Biol. B 359:265–274.

Kendall, D. G. 1948. On the generalized ''birth-and-death'' process. Ann. Math. Stat. 19:1–15.

Kozak, K. H., D. W. Weisrock, and A. Larson. 2006. Rapid lineage accumulation in a non-adaptive radiation: phylogenetic analysis of diversification rates in eastern North American woodland salamanders (Plethodontidae: *Plethodon*). Proc. R. Soc. Lond. Biol. B 273:539–546.

Kubo, T., and Y. Iwasa. 1995. Inferring the rates of branching and extinction from molecular phylogenies. Evolution 49:694–704.

Linder, H. P., P. Eldenas, and B. G. Briggs. 2003. Contrasting patterns of radiation in African and Australian Restionaceae. Evolution 57:2688–2702.

Lovette, I. J., and E. Bermingham. 1999. Explosive speciation in the New World *Dendroica* warblers. Proc. R. Soc. Lond. Biol. B. 266:1629–1636.

Melville, J., J. A. Schulte II, and A. Larson. 2001. A molecular phylogenetic study of ecological diversification in the Australian lizard genus *Ctenophorus*. J. Exp. Zool. 291:339–353.

Mooers, A. O., and S. B. Heard. 1997. Evolutionary process from phylogenetic tree shape. Q. Rev. Biol. 72:31–54.

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.

Near, T. J., T. W. Kassler, J. B. Koppelman, C. B. Dillman, and D. P. Philipp. 2003. Speciation in North American black basses, *Micropterus* (Actinopterygii: Centrarchidae). Evolution 57: 1610–1621.

Nee, S. 2001. Inferring speciation rates from phylogenies. Evolution 55:661–668.

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

Nee, S., R. M. May, and P. H. Harvey. 1994b. The reconstructed evolutionary process. Philos. Trans. R. Soc. Lond. B 344: 305–311.

Paradis, E. 1997. Assessing temporal variations in diversification rates from phylogenies: estimation and hypothesis testing. Proc. R. Soc. Lond. Biol. B 264:1141–1147.

———. 1998a. Detecting shifts in diversification without fossils. Am. Nat. 152:176–187.

———. 1998b. Testing for constant diversification rates using molecular phylogenies: a general approach based on statistical tests for goodness of fit. Mol. Biol. Evol. 15:476–479.

———. 2004. Can extinction rates be estimated without fossils? J. Theor. Biol. 229:19–30.

———. 2005. Statistical analysis of diversification with species traits. Evolution 59:1–12.

Pianka, E. R. 1986. Ecology and natural history of desert lizards. Princeton Univ. Press, Princeton, NJ.

Pitra, C., J. Fickel, E. Meijaard, and P. C. Groves. 2004. Evolution and phylogeny of Old World deer. Mol. Phylogenet. Evol. 33: 880–895.

Pol, D. 2004. Empirical problems of the hierarchical likelihood ratio test for model selection. Syst. Biol. 53:949–962.

Posada, D., and T. R. Buckley. 2004. Model selection and model averaging in phylogenetics: advantages of the Akaike information criterion and Bayesian approaches over likelihood ratio tests. Syst. Biol. 53:793–808.

Posada, D., and K. A. Crandall. 2001. Selecting the best-fit model of nucleotide substitution. Syst. Biol. 50:580–601.

Pybus, O. G., and P. H. Harvey. 2000. Testing macro-evolutionary models using incomplete molecular phylogenies. Proc. R. Soc. Lond. Biol. B 267:2267–2272.

Pybus, O. G., A. Rambaut, E. C. Holmes, and P. H. Harvey. 2002. New inferences from tree shape: number of missing taxa and population growth rates. Syst. Biol. 51:881–888.

Ree, R. H. 2005. Detecting the historical signature of key inno-

vations using stochastic models of character evolution and cladogenesis. Evolution 59:257–265.

Revell, L. J., L. J. Harmon, and R. E. Glor. 2005. Underparameterized model of sequence evolution leads to bias in the estimation of diversification rates from molecular phylogenies. Syst. Biol. 54:973–983.

Ruber, L., and R. Zardoya. 2005. Rapid cladogenesis in marine fishes revisited. Evolution 59:1119–1127.

Sanderson, M. J. 1994. Reconstructing the history of evolutionary processes using maximum likelihood. Pp. 13–26 *in* D. M. Fambrough, ed. Molecular evolution of physiological processes. Rockefeller Univ. Press, New York.

———. 1997. A nonparametric approach to estimating divergence times in the absence of rate constancy. Mol. Biol. Evol. 14: 1218–1231.

Sanderson, M. J., and M. J. Donoghue. 1996. Reconstructing shifts in diversification rates on phylogenetic trees. Trends Ecol. Evol. 11:15–20.

Schluter, D. 2000. The ecology of adaptive radiation. Oxford Univ. Press, New York.

Turgeon, J., R. Stoks, R. A. Thum, J. M. Brown, and M. A. McPeek. 2005. Simultaneous quaternary radiations of three damselfly clades across the Holarctic. Am. Nat. 165:78–107.

Weir, J., and D. Schluter. 2004. Ice sheets promote speciation in boreal birds. Proc. R. Soc. Lond. Biol. B 271:1881–1887.

Wollenberg, K., J. Arnold, and J. C. Avise. 1996. Recognizing the forest for the trees: testing temporal patterns of cladogenesis using a null model of stochastic diversification. Mol. Biol. Evol. 13:833–849.

Yule, G. U. 1924. A mathematical theory of evolution based on the conclusions of Dr. J. C. Willis. Philos. Trans. R. Soc. Lond. B 213:21–87.

Zhang, J. 1999. Performance of likelihood ratio tests of evolutionary hypotheses under inadequate substitution models. Mol. Biol. Evol. 16:868–875.

Zink, R. M., and J. B. Slowinski. 1995. Evidence from molecular systematics for decreased avian diversification in the Pleistocene epoch. Proc. Natl. Acad. Sci. USA 92:5832–5835.

Zink, R. M., J. Klicka, and B. R. Barber. 2004. The tempo of avian diversification during the Quaternary. Philos. Trans. R. Soc. Lond. B 359:215–219.

Corresponding Editor: K. Crandall

APPENDIX

Likelihood Optimization

For the pure-birth model ($a = 0$), maximum likelihood estimates of r were calculated as

$$\hat{r} = \frac{N - 2}{\displaystyle\sum_{i=2}^{N} i b_i} \qquad (A1)$$

after Nee (2001), where b_i is the ith internode distance.

For models incorporating $a \geq 0$, maximum likelihood estimates of a and r were inferred using a steepest descent algorithm. For any trial points r_i and a_i, the gradient of the likelihood function $L(\mathbf{x} \,|\, r, a)$ can be obtained by evaluating the partial derivatives of L with respect to r and a. I used the gradient of the likelihood function to sequentially choose the r_{i+1} and a_{i+1} following the steepest path to the optimum. The optimum was found when 100 successive approximations could not improve the log-likelihood beyond 10^{-5} units. Nonlinear optimization can be difficult due to the possibility of multiple optima in the likelihood surface, and the optimization algorithm employed here is prone to error when multiple optima exist. To guard against this possibility, I repeated each optimization with 10 different initial values of a and r. Estimates of the maximum log-likelihood and r converged to the same values, though a was very slow to converge beyond 0.05 units of precision (which is well within the 95% confidence interval for this parameter). This can be explained by the fact that, at least for phylogenies examined in this study, the likelihood surface for a and r consists

of a long, narrow ridge, with a single optimum, as observed in Nee et al. (1994a). A wide range of a-values give log-likelihoods that are virtually indistinguishable, and none of the analyses presented here require a precise estimate of a.

For rate variable models, parameters were optimized separately for all internode distances occurring before and after a given shift point, t_s. This was repeated for all possible shift points, assumed here to consist only of observed branching times. In the case of survival analysis, likelihood and parameter optimization follows Cox and Oakes (1984) and Paradis (1997).

Phylogenetic Simulation

A discrete time approach was used to simulate phylogenies under rate-variable and rate-constant scenarios. Each simulation begins with two lineages and runs forward in time; at each time step, surviving lineages are evaluated for a speciation event, with probability λ, or an extinction event, with probability μ. Each new lineage contains a reference to the parent lineage, such that a reconstructed phylogeny can be generated at the end of the simulation. When the target number of lineages is reached at some time T, both basal lineages are tested to ensure that at least one descendent lineage has survived to the present; if all of the descendents of one lineage had gone extinct, our simulated phylogeny would not be of age T. For each simulation, $r = \lambda - \mu$ was chosen sufficiently small to generate phylogenies of several thousand time units in total length (mean number of time steps per simulation: 8322; range: 1964–34853). This step was taken to avoid possible artifacts caused by using a discrete simulation procedure to approximate a continuous time process. Software for likelihood analysis, phylogenetic simulation, and model selection were written in ANSI C; source code and Windows-compiled executables are available from the author.