When Can Decreasing Diversification Rates Be Detected with Molecular Phylogenies and the Fossil Record?

LEE HSIAH LIOW1,4, TIAGO B. QUENTAL2,3,4, AND CHARLES R. MARSHALL2,3,4

1Centre for Ecological and Evolutionary Synthesis, Department of Biology, University of Oslo, N-0316 Oslo, Norway; 2Department of Organismic and Evolutionary Biology and 3Department of Invertebrate Paleontology, Museum of Comparative Zoology, Harvard University, Cambridge, MA 02138, USA; and 4Present address: Department of Integrative Biology, University of California Museum of Paleontology, University of California, Berkeley, CA 94720, USA;

Correspondence to be sent to: Centre for Ecological and Evolutionary Synthesis, Department of Biology, University of Oslo, Oslo, Norway; E-mail: l.h.liow@bio.uio.no.

Lee Hsiang Liow and Tiago B. Quental contributed equally to this article.

Received 8 December 2009; reviews returned 16 February 2010; accepted 13 August 2010

Guest Associate Editor: Daniel L. Rabosky

Abstract.—Traditionally, patterns and processes of diversification could only be inferred from the fossil record. However, there are an increasing number of tools that enable diversification dynamics to be inferred from molecular phylogenies. The application of these tools to new data sets has renewed interest in the question of the prevalence of diversity-dependent diversification. However, there is growing recognition that the absence of extinct species in molecular phylogenies may prevent accurate inferences about the underlying diversification dynamics. On the other hand, even though the fossil record provides direct data on extinct species, its incompleteness can also mask true diversification processes. Here, using computer-generated diversity-dependent phylogenies, we mimicked molecular phylogenies by eliminating extinct lineages. We also simulated the fossil record by converting the temporal axis into discrete intervals and imposing a variety of preservation processes on the lineages. Given the lack of reliable phylogenies for many fossil marine taxa, we also stripped away phylogenetic information from the computer-generated phylogenies. For the simulated molecular phylogenies, we examined the efficacy of the standard metric (the $\gamma$ statistic) for identifying decreasing rates of diversification. We find that the underlying decreasing rate of diversification is detected only when the rate of change in the diversification rate is high, and if the molecular phylogeny happens to capture the diversification process as the equilibrium diversity is first reached or shortly thereafter. In contrast, estimating rates of diversification from the simulated fossil record captures the expected zero rate of diversification after equilibrium is reached under a wide range of preservation scenarios. The ability to detect the initial decreasing rate of diversification is lost as the temporal resolution of the fossil record drops and with a decreased quality of preservation. When the rate of change of the diversification rate is low, the $\gamma$ statistic will typically fail to detect the decreasing rate of diversification, as will the fossil record, although the fossil record still retains the signature of the diversity dependence in yielding approximately zero diversification rates. Thus, although a significantly negative $\gamma$ value for a molecular phylogeny indicates a decreasing rate of diversification, a nonsignificantly negative or positive $\gamma$ value might mean exponential diversification, or a slowly decreasing rate of diversification, or simply species turnover at a constant diversity. The fossil record can be of assistance in helping choose among these possibilities. [Computer simulations; diversity dependence; extinction; incomplete sampling; speciation; stratigraphic resolution.]

One of the main goals of biology is to understand how diversity arises and is maintained and/or altered through space and time. As with other complex manifestations of biological processes, there exist a multitude of approaches for studying diversity dynamics. The number of species in a natural group is the result of the processes of speciation and extinction (Ricklefs 2007). Diversification rates, and their component speciation and extinction rates, can be estimated from phylogenetic hypotheses constructed using extant species (Nee et al. 1992; Havey et al. 1994; Nee, Holmes, et al. 1994; Nee, May, et al. 1994; Ricklefs 2006; Bokma 2008; Rabosky and Lovette 2008). However, there are limitations to such approaches, including questions that arise over the appropriateness of the models used (Ricklefs 2007; Bokma 2009). In contrast, the fossil record, by including extinct taxa, allows speciation, extinction, and diversification rates to be estimated directly (Foote 2003; Alroy 2008; Etienne and Apol 2009). As is also the case when using only extant taxa for inferring diversification rates, incomplete sampling and biases in the fossil record can lead to misleading results, but many methods have been developed (Alroy et al. 2001, 2008; Foote 2003; Liow et al. 2008; Peters and Ausich 2008) to attenuate these problems.

Here, we compare how well the fossil record performs against approaches that use phylogenies of extant organisms to recover the true underlying processes of diversification. Specifically, we focus on how well these two approaches correctly identify diversity-dependent diversification (sometimes termed density-dependent diversification) given the renewed interest in the detection and implications of this mode of diversification (Weir 2006; Ricklefs 2007; Phillimore and Price 2008).

Many paleontologists have argued for the existence of carrying capacities that cap both global and individual clade diversities, and some have used logistic models to quantify the corresponding diversification process (Rauf 1972; Sepkoski 1978, 1979, 1984; Walker and Valentine 1984). These models assume diversity-dependent diversification, where competition for limited resources is thought to retard diversification as species numbers increase (MacArthur 1969; Wilson 1969). However, other paleontologists propose a radically different class of models, namely exponential models, which suggest that diversity is either not limited by
a carrying capacity or the biosphere is so far from its theoretical equilibrium diversity that we can effectively ignore the effects of diversity dependence in macroevolution (Benton and Emerson 2007; Stanley 2007).

The tempo of diversification can also be estimated from molecular phylogenies (Pybus and Harvey 2000; Phillimore and Price 2008; Rabosky and Lovette 2008). Even though decreases in net diversification are commonly observed in the molecular phylogenies analyzed to date, about half of the clades examined are better described by models of constant diversification rates (Ruber and Zardoya 2005; Haron, Melville, et al. 2008; McPeek 2008; Phillimore and Price 2008; Elias et al. 2009).

Here, we investigate the abilities of standard methods for analyzing the fossil record and molecular phylogenies to correctly identify an underlying diversity-dependent diversification process. We have chosen this class of model because of its venerable history (MacArthur 1969; Wilson 1969) and because of the renewed interest it is receiving in the paleontological and neontological literatures (see above). Specifically, we investigate the behavior of the γ statistic (Pybus and Harvey 2000), an increasingly common method used to explore the tempo of diversification from molecular phylogenies. Using computer simulations, we first assess the ability of the γ statistic, calculated from phylogenies using only the extant taxa, to recover the diversity-dependent diversification used to generate the synthetic phylogenies. Unlike previous simulation studies, we evaluate the efficacy of the γ statistic to recover the diversity-dependent diversification throughout the diversification process, not just when the simulated clades are close to the time they first reached their equilibrium diversity. We then ask, for the same simulations, whether synthetic fossil records are able to recover the true underlying diversity-dependent diversity dynamics. We generated the fossil records by degrading the temporal ranges of the lineages in the true phylogenies to mimic incomplete preservation; reducing the timescale to a series of the coarse temporal intervals to mimic the temporal resolution of the geologic timescale; and, removing the phylogenetic information, given the lack of well-resolved phylogenies for many fossil groups (especially marine invertebrates that are among the best preserved groups in the fossil record). We assessed the diversity dynamics from the simulated fossil data using origination and extinction rates estimated directly from the simulated fossil records.

METHODS

Simulating the Diversification Process

We assumed a simple diversity-dependent model of diversification, although more complicated models could be used. Specifically, we assumed that the speciation rate changes linearly as a function of the number of species, whereas extinction rate remains constant (Fig. 1, the same as the SP_VAR model described in Rabosky and Lovette 2008, and used by Quental and Marshall 2009b). In this model, the speciation rate initially exceeds the extinction rate and the clade begins to diversify. However, as diversity increases, the speciation rate drops until the clade first reaches its equilibrium diversity (50 species in our simulations), the first time that the speciation rate equals the extinction rate (yielding a diversification rate of zero). After this time, diversity will fluctuate around the equilibrium diversity. Stochastic variation in the number of speciation and extinction events will sometimes lead to the diversity rising above the carrying capacity, in which case the speciation rate will drop below the extinction rate, allowing extinction to bring the number of species back down to the equilibrium value. Alternatively, when excess extinction causes the number of species to drop below the equilibrium diversity value, the speciation rate increases, which brings the number of species back to the equilibrium diversity once again (Fig. 1).

It has been shown previously that extinction can erase the signature of diversity dependence from molecular phylogenies (Rabosky and Lovette 2008; Quental and Marshall 2009b). More specifically, our ability to detect decreases in diversification rates from molecular phylogenies depends on the ratio between initial speciation rate (Lambda initial) and extinction rate at equilibrium (Mu at equilibrium), called the LiMe ratio (Quental

![Figure 1](https://academic.oup.com/sysbio/article-abstract/59/6/646/1711105/fig1)

**Figure 1.** Three diversity-dependent diversification scenarios. In all diversification scenarios, the speciation rate (λ) decreases in a linear diversity-dependent fashion, whereas the extinction rate (μ) is constant. a) Scenario where the initial speciation rate (λ0) is twice the value of the extinction rate (LiMe 2). b) Scenario where the initial speciation rate is 5 times the value of the extinction rate (LiMe 5). c) Scenario where the initial speciation rate is 10 times the value of the extinction rate (LiMe 10). For all scenarios, the equilibrium diversity is set to 50 species at which point the speciation rate equals the equilibrium extinction rate (μ).
and Marshall 2009b), and not on the absolute values of the initial speciation and extinction rates. However, these studies only analyzed phylogenies that were close to (Rabosky and Lovette 2008), or had just reached (Quental and Marshall 2009b), the equilibrium diversity. Here, we examine how the diversification process will appear from its initiation until well after the equilibrium diversity was first reached. In particular, we are interested in how soon after the beginning of the radiation the diversity dependence becomes evident and for those simulations where the diversity dependence can be inferred from the corresponding molecular phylogeny as the equilibrium diversity is first reached, how quickly that signature of diversity-dependence is erased as the clade experiences turnover at its equilibrium diversity.

Given the importance of the LiMe ratio in determining the shape of a molecular phylogeny, we simulated 3 different scenarios: LiMe equals 2, 5, and 10 (Fig. 1, henceforth shortened to LiMe 2, LiMe 5, and LiMe 10, respectively). For all simulations, we assumed an equilibrium diversity of 50 species and a constant extinction rate, $\mu$, equal to 0.1 L/Lmyr (lineage per lineage million years). These simulations were also used to determine the extent to which diversity-dependent diversification can be inferred from the fossil record. Each simulation was run for 5 times the expected species duration (ESD = 1/$\mu$). Given an extinction rate of $\mu = 0.1$ L/Lmyr, the ESD was 10 myr and thus, we stopped the simulations after 50 time units (where each time unit was 1 myr). Finally, although we know that the appearance of a molecular phylogeny at the time the equilibrium diversity is first reached is determined by the LiMe ratio and not the absolute extinction rate (Quental and Marshall 2009b), we tested to see that this also holds throughout the diversification process by rerunning simulations with the same LiMe values but with an extinction rate that was 10-fold higher. As expected, it is indeed the LiMe ratio (as well as the equilibrium diversity), and not the absolute value of the speciation or extinction rates, that controls the appearance of a molecular phylogeny, regardless of where in the logistic diversification process the clade happens to be.

A total of 1000 trees were simulated for each of the LiMe values. LiMe 2 had 734 trees that survived for 50 time units; LiMe 5 had 949, and LiMe 10 had 990. Even though each tree is different, trees generated within a given LiMe ratio have undergone radiation with same underlying process and hence have species durations generated from the same statistical distribution. For each simulation, the molecular phylogeny that would have been observed every 2.5 myr (i.e., 0.25 times the ESD) was used to calculate the standing diversity and the $\gamma$ statistic, enabling us to monitor our ability to infer the diversity-dependent dynamics as the dynamics unfolded.

Additional simulations were also run for 10 and 20 times the ESD to confirm that the expected behaviors of the $\gamma$ statistic and of the expected diversity were in fact achieved (results not shown).

**Describing the Diversification Process Using Molecular Phylogenies: The $\gamma$ Statistic at Different Times in a Clade’s History**

The $\gamma$ statistic is a simple way of detecting whether there have been changes in the diversification rate during the diversification of a living clade. The statistic is measured using lineage-through-time plots: if $g_2, g_3, \ldots, g_n$ are the internode distances of a phylogeny with $N$ lineages, then

$$\gamma = \frac{\left(1 - \frac{1}{N-2} \sum_{i=2}^{N-1} \sum_{j=2}^{i} g_{ij} \right) - \left(\frac{\tau}{2}\right)}{\sqrt{\frac{1}{12(N-2)}}},$$

following Pybus and Harvey (2000).

The $\gamma$ statistic assesses where the center of gravity of a given tree lies with respect to where it is expected to lie under a pure birth process. A significantly negative $\gamma$ value indicates that the nodes are concentrated toward the root of the tree, indicating that the diversification rate has decreased over time. For any given complete phylogeny (i.e., with no missing taxa), the null hypothesis of a constant rate of diversification (i.e., the null model of a constant pure birth process) is rejected if its $\gamma$ value is more negative than $-1.645$ (type I error probability of 0.05, one tailed) (Pybus and Harvey 2000). Thus, one of our major goals is to determine how effectively the $\gamma$ statistic enables detection of diversity-dependent decreases in diversification rates (see below).

To determine how well the diversification process is accurately reflected in the $\gamma$ value at different stages of the process (e.g., during its initial radiation, the first time it reaches the equilibrium diversity, after the equilibrium diversity has been reached), we evaluated each of the diversity-dependent simulations at different points in time (Fig. 2). At each of these times, we calculated the $\gamma$ statistic and tabulated the species richness using the “then” extant species. The phylogeny was sampled 20 times over the $5$ species durations the simulations were run for (4 times per ESD). The trees were manipulated and analyzed using the R packages Geiger (Harmon, Weir, et al. 2008) and paleoPhylo (Ezard and Purvis 2009). Some functions in Geiger and paleoPhylo were modified for the specific purposes of our study.

**Describing the Diversification Process Using an Incomplete Fossil Record**

In the computer simulations described above, time was treated as a continuous variable. However, in the fossil record, our general inability to date fossils directly and the limitations of the techniques of temporal correlation needed to tie fossil localities to the geological timescale means that time is usually divided into discrete intervals. That is, the majority of fossil localities are only dated to a given epoch (e.g., the Miocene), or, if they have been defined, a subepoch (e.g., the Upper Miocene), or a specific age (e.g., the Tortonian, the
lower Upper Miocene), etc. Thus, to mimic the fossil record, the timescale for the phylogenies generated by the computer simulations described above was discretized before estimating their diversification parameters. Temporal resolution varies across the timescale, but for the Phanerozoic, the epochs are on average ~10 myr in duration, through the Cenozoic, the stages are ~5 myr in duration, whereas in the Neogene (the Miocene to the Recent), the ages are ~2.5 myr in duration (Gradstein et al. 2004). Thus, we analyzed our simulated phylogenies within discrete time intervals of 2.5, 5, and 10 myr (henceforth termed observation windows [OWs]).

However, it is not just the size of the OW that determines our ability to discern diversity dynamics in the fossil record but also how rapidly the diversity dynamics unfold. A simple metric for characterizing the rate at which the dynamics unfold is the ESD, the reciprocal of the extinction rate (ESD = 1/μ = λ_0/LiMe). Thus, we used the ratio of the OW to the ESD (OW/ESD) as a measure of the effective temporal resolution of the fossil
For both the true and the degraded phylogenies, diversification rates ($\hat{d}$) are given by

$$\hat{d} = \hat{p} - \hat{q},$$

where $N_b$ is the number of taxa that cross the earlier time boundary of the OW, $N_b$ is the number of taxa that cross the later time boundary, $N_b$ is the number that cross both the time boundaries of the given OW, and $\Delta t$ is the duration of the interval. Note that singletons (taxa that occur only within the given time interval) do not figure in these calculations. In addition, note that the rates in the first and last OWs cannot be calculated because the desired rates can only be calculated if the given OW is flanked on both the sides by OWs.

### Simulating Preservation Processes

Although we examined many preservation scenarios, we only present results from four here (Fig. 3), given the overall consistency of the results across the various scenarios (not shown). In preservation scenario R and P (Fig. 3b,c), we removed a random proportion of all the species present in each time interval. The proportion of species we retained ranged from 50, 30 to only 10% per interval, but we presented only the results from 50% (RandP50%) and 10% (RandP10%).

To mimic the fact that fossil recovery potential often increases in progressively younger rocks, we removed a smaller proportion of occurrences as we moved forward in time in preservation scenario IncP (Fig. 3d). In other words, a greater proportion of taxa were removed at the base of the tree compared with the proportion removed further up the tree (Fig. 3d,e). We simulated multiple variants of this scenario, but we present the representative case where the per interval preservation rate was set to 10% at the inception of the clad and increased to 50% in the last temporal bin. We conservatively selected these values in part because preservation rates for well-preserved (marine) groups at the generic level (the most common level used for paleontological diversity analyses) are in the order of 60–90% per stage (Foote and Raup 1996; Foote and Sepkoski, 1999), whereas those for Cenozoic mammal species are about 25–37% (Foote 1997; Foote et al. 1999). Some groups have a much poorer fossil record, for example, it is estimated that less than 7% of Cenozoic primates have so far been found in the fossil record (Tavaré et al. 2002; Soligo et al. 2007).

To mimic the observation that most species will, after their inception, slowly increase in abundance and geographic range with time, plateau, and then undergo a decline before finally going extinct (Foote et al. 2007; Liow and Stenseth 2007; Liow et al. 2010), we developed preservation scenario called HatP (Fig. 3f) because the preservation potential is hat shaped. Here, the preservation potential of each species is modeled as a beta distribution with a zero probability of sampling it at the beginning and end of its duration and with the sampling probability increasing smoothly from both ends of its temporal range to a maximum (sampling probability = 1) in the middle of its duration (Fig. 3f,g). This was achieved by setting the parameters for the beta distribution $\alpha = \beta = 4$. In practice, we discretized the true presence of a given species in each time interval and then subject this to probabilistic sampling based on the density of the beta distribution.

In preservation scenario HatP, the peak probability of sampling is unrealistically high, and so in the last preservation scenario, HatP + IncP, we reduced the overall sampling probabilities in HatP according to the IncP model, where the probability of preservation increases linearly with time, with an initial preservation rate of 10%, culminating in a final preservation rate of 50%. We feel that this is probably the most realistic sampling regime in terms of biological and geological factors. Species are less likely to be sampled during the early and late phases of their existence due to limited geographic spread, lower abundances, and perhaps, at least at the inception, morphological similarity to their ancestors (Liow et al. 2010). In addition, the older fossil record has had more time to be subject to erosion, metamorphosis, and other physical, chemical, and biochemical processes that increasingly limit our ability to sample it. This scenario also provides the most severe test of the underlying logistic diversity pattern because the lower preservation probabilities early in the clad’s history mitigate against observing the initial drop in the diversification rate at the beginning of the radiation and against observing the plateauing of diversity after the initial radiation.

For all preservation scenarios, each species was treated as only being “present” once in each time interval, and after the application of the sampling, the new stratigraphic durations were calculated for the remaining species, which were then used to calculate speciation ($\hat{p}$), extinction ($\hat{q}$), and diversification ($\hat{d}$) rates. Note that if a species was sampled in the time interval before and after a reference time interval but not within it, it is considered as being alive in that reference time interval (i.e., we assumed it “ranged through” the gap) (see Fig. 3e).
Finally, we note that the morphological differences are used to distinguish between species in the fossil record: cryptic species are impossible to distinguish and incomplete preservation means that species-level differences in the phenotype may also be invisible to the paleontologist. Thus, for many groups, paleontologists effectively work with a subset of the species that a neontologist would have available for study. Hence, in modeling the incompleteness of the fossil record, our preservation scenarios encompass both the failure of lineages to make it into the fossil record, as well as the inability to recognize cryptic species, and sometimes closely related species. There are also other differences between paleontological and neontological practices. For example, a single lineage in a molecular phylogeny might have had sufficient anagenetic change such that several species might be recognized in the fossil record. However, for this study, this is not a problem, given that we have no anagenesis in our simulations. Despite the differences between the way taxa are defined and recognized by paleontologists and neontologists, we do not know of any biases that would invalidate the conclusions made here, although we note that for some groups, it is possible that more effort is put into finding basal members of a clade and that for the youngest fossils, there is still a tendency to assign them to living taxa, potentially decreasing the number of young extinct taxa that might otherwise be described. This latter bias has long been recognized and is one component of the “Pull of the Recent” (Raup 1972, 1979). We do not have quantitative data on the extent and magnitude of these potential biases in the context of our simulations, but we suspect that they are small compared with large-scale biases we have tried to capture with our preservation scenarios.

**RESULTS**

\( \gamma \) and Diversity-Dependent Diversification Viewed through Molecular Phylogenies

When the initial speciation rate is low compared with the extinction rate (LiMe 2), the number of species increases to the equilibrium value of 50 species sufficiently slowly that the equilibrium diversity is not reached for most trees after 5 times the ESD (Fig. 4a). Allowing the simulations to run beyond 5 times, the ESD showed that equilibrium is reached after approximately 10 times the ESD (not shown). As the initial rate of speciation increases from LiMe 2 to LiMe 10, the...
equilibrium diversity is reached progressively more quickly (at about 2 times ESD for LiMe 5 and around 1 ESD for LiMe 10, Fig. 4b,c). Note also that as LiMe increases, that is, when the initial speciation rate is high compared with the extinction rate, there is less variation in the number of species at equilibrium (cf. Fig. 4b,c). This difference is the result of differing severities of the diversity dependence in the different scenarios. The higher the LiMe value, the greater the change in the diversification rate as the diversity fluctuates about the equilibrium value (reflected in the steeper slope of the speciation rate line in Fig. 1). Thus, with high LiMe values, even small changes in diversity lead to relatively large increases or decreases in the speciation rate, which in turn lead to strongly restoring rates of diversification (e.g., if the diversity reaches 56 species for LiMe 10, the speciation rate drops to zero, Fig. 1c), and the constant extinction rate will strongly direct the diversity back toward the equilibrium value. In contrast, the realized diversity is more susceptible to stochastic fluctuations given low LiMe values.

The γ values for the lowest LiMe value we simulated (LiMe 2) consistently lie above the 5% cutoff for rejecting constant rates of diversification (Fig. 4d), even though diversification is diversity dependent, and, on average, the nodes on the resulting phylogenies should be concentrated deeper in the trees. Thus, for low values of LiMe, diversity-dependent diversification produces phylogenies that are indistinguishable from those expected under constant rates of diversification throughout the diversification process, not just at or close to the point where the equilibrium diversity is first reached, as has been previously reported (Rabosky and Lovette 2008; Quental and Marshall 2009b).

For both LiMe 5 and 10, γ becomes progressively more negative as the radiations proceed and reach their most negative values as the clades first reach their equilibrium diversities (Fig. 4e,f). γ reaches its most negative value at about 2 times the ESD for LiMe 5 and at about 1 time the ESD for LiMe 10. Although the LiMe 5 and 10 scenarios both produce a dip in the γ values centered around the first time equilibrium diversity is reached, for LiMe 5 the average value of γ only barely dips below the 5% cutoff for rejecting the null hypothesis of a constant diversification rate (Fig. 4e). γ then becomes progressively more positive after the equilibrium diversity is attained, and, approximately 0.5 and 1.5 times the ESD after achieving equilibrium for LiMe 5 and 10, respectively, the average γ becomes indistinguishable from the null model of a constant diversification rate.
Thus, phylogenies driven by diversity-dependent diversification will only be identified as such with the γ statistic if the LiMe values were large and if the phylogenies happen to have been caught at about the first time their equilibrium diversities were reached or within a relatively short period of time thereafter.

**Diversification Rates Viewed through an Imperfectly Known Fossil Record**

We explored the effects of the discretization of the timescale, and of the different preservational scenarios, on our ability to correctly infer the diversity-dependent diversification process from the simulated fossil records. Given that we removed the phylogenetic information from our computer simulations to imitate the typical fossil record (at least for many marine invertebrates), we could not assess the diversification dynamics using the γ statistic. So, instead, we inferred the dynamics by directly estimating diversification rate as a function of time. Under logistic growth, the rate rapidly decays until the equilibrium diversity is reached, and then the diversification rates should fluctuate around zero, with the amplitude of the fluctuation being controlled by intensity of the dependence of the speciation rate on the diversity—the higher the LiMe value the greater that intensity.

Although we applied OWs of two sizes, the impact of the discretization also depends on the ESD and on the rapidity at which the diversification process unfolds (controlled by LiMe). We here define the effective temporal resolution by the ratio OW/ESD, the size of the OW measured as a proportion of the average species duration (smaller numbers mean finer resolving power).

The first step in our analysis was to determine how the diversification process appears with the discretized timescale before we imposed the preservational scenarios. We first explored the simulations with the highest effective temporal resolution, those with an OW of 2.5 myr and a species duration of 10 myr, yielding an effective temporal resolution of 0.25. At this resolution, we wanted to first know how the diversification process appears with the most rapid diversification process (LiMe 10), where we might expect to miss the drop in diversification rate that characterizes the growth phase of diversity-dependent diversification. However, despite the discretization of the timescale, both the expected decrease in the diversification rate and the following plateau at a rate of zero are clearly seen (Fig. 5a); at the highest (reasonable) level of resolution, the diversity dynamics can be discerned.

We next wanted to know how seriously the different preservational scenarios compromise our ability to detect diversity-dependent diversification for the high-resolution (OW/ESD = 0.25) rapid diversification (LiMe 10) scenario discussed immediately above. Surprisingly, the effects were quite mild. The initial decrease in the diversification rate can be observed in all scenarios except for RandP10% (Fig. 5c), whereas the zero rate of diversification at equilibrium is observed for all scenarios (Fig. 5b-f). The RandP10% scenario also had a much wider confidence band than the other scenarios (Fig. 5c). Given the overall similarity of the RandP10%, IncP, HatP, and IncP + HatP scenarios, we only present results from the IncP + HatP scenario, the noisiest of the scenarios, in the analyses that follow. We also present data from the RandP10% scenario given that the signal of diversity-dependent diversification is almost dissipated under that preservational scheme. Note that all the preservation scenarios produced a slightly negative diversification rate in the last OW analyzed (Fig. 5), although slightly less so in IncP (Fig. 5d). This edge effect derives from those lineages that actually persisted through to the last time interval but whose ranges were truncated such that they were judged to have become extinct. Finally, although we conclude that the effects of the preservational scenarios on our ability to discern both the initial and equilibrium phases of the diversification process are mild, a substantial number of simulations were sufficiently degraded by the preservational scenarios that they do not contain enough information for the estimation of diversification rates. Thus, the data presented in Figure 5 overrepresent the strength of the diversification signal captured by the simulated fossil record.

Finally, we wished to know how well the fossil record retains the signal of the diversity-dependent diversification signal when the temporal resolution is lowered and when the diversification process proceeds more slowly. When the temporal resolution was lowered so that the OW was twice the size compared with the scenario analyzed above (OW/ESD = 0.5 while LiMe remained at 10), the initial decline in the diversification rate can only rarely be detected with a perfect fossil record, under the IncP + HatP scenario and not at all in the RandP10% scenario (Fig. 6d-f). However, for both the perfect fossil record and under the two preservational scenarios, the equilibrium diversification rate is still well constrained at zero. Note that the confidence interval for the LiMe 10, OW = 0.5 ESD (Fig. 6d-f) is smaller than those for the LiMe 10, OW = 0.25 ESD case (Fig. 6a-c) because larger OWs contributed to decreased stochasticity in the ratio of various boundary crossers (see Methods section) needed for diversification rate estimation.

For a high effective temporal resolution (OW = 0.25 ESD), but where the diversification process is slow (LiMe 2), a perfect fossil record only barely detects the slow drop in the mean diversification rate, whereas the two preservational scenarios show no sign of the expected gradual drop in the diversification rate (Fig. 6g-i). The drop in the diversification rate can be seen indirectly through the contraction of the confidence interval around the mean with time (Fig. 6g), but this is of little use if one is only analyzing single groups at a time. For the IncP + HatP preservational scenario, the increasing preservation rate with time actually produces a subtle bias, leading to a slightly increasing rate of diversification when it should be fluctuating about zero (Fig. 6h). This is due to the IncP component of...
the scenario, but the confidence interval is very large (Fig. 6h), so for any given clade, the increase is unlikely to be significant.

**DISCUSSION**

**Diversity Dependence and γ**

Even though more sophisticated methods for studying diversification using molecular phylogenies have been developed, for example, based on explicit modeling of speciation and extinction rates through time using likelihood (Rabosky 2006) or approximate Bayesian methods (Rabosky 2009), the γ statistic has become a widely used tool for rejecting the null model of a constant diversification rate (Weir 2006; McPeek 2008; Phillimore and Price 2008; Rabosky and Lovette 2008). A sufficiently negative γ (≤ −1.645) indicates that the diversification rate has, on average, slowed over the time span encompassed by the given phylogeny. Phylogenies with significantly negative γ values are typically interpreted as being the result of diversity-dependent diversification (e.g., Phillimore and Price 2008), although other mechanisms, such as a model that incorporates a pulsed turnover dynamics with heritable extinction (Rabosky 2009), can also generate phylogenies with significantly negative γ values. Conversely, γ values consistent with a constant diversification rate could, at face value, be interpreted as exponential growth. However, we have shown here (see also Rabosky and Lovette 2008 and Quental and Marshall 2009b for the special case where the clade happens to be first at, or just reaching, its equilibrium diversity) that a true underlying diversity-dependent process of diversification can also lead γ values that are in accord with constant rates of diversification. This will only happen if the extinction rate changes (Rabosky and Lovette 2008); if the diversification rate reaches an equilibrium state slowly, that is, for low LiMe values (Quental and Marshall 2009b); or, as we shown here, once enough time has passed because the equilibrium diversity was first reached (i.e., 0.5–1.5 average species durations [ESDs] for the diversification scenarios studied here).

However, the observation that γ values can be generated with diversity dependence similar to those generated under a model of constant diversification does not negate the meaning of the γ statistic as described by Pybus and Harvey (2000). As these authors noted, γ reflects not the diversification rate, per se, but whether or not there is “decrease” in the diversification rate with time. For a clade diversifying under a diversity-dependent model, the clade enters a period of dynamic...
The number of species will then fluctuate around the equilibrium diversity value. Under this scenario, \( \gamma \) becomes progressively more positive and after some time (about 0.5–1.5 times ESD for the specific conditions of our simulations with high LiMe values), its value is appropriately consistent with a null model of constant diversification rate: in our case, the diversification rate is zero. Thus, a phylogeny that results from species turnover at a constant diversity due to diversity dependence is indistinguishable from a phylogeny that reflects constant diversification due to exponential growth when using the \( \gamma \) statistic alone; both have constant rates of diversification.

The value of \( \gamma \) is affected not only by intrinsic diversification rates (as captured by the LiMe ratio) but also by clade size. For example, for a simulation where LiMe is equal to 10 and the equilibrium diversity is set to 100 species, the average \( \gamma \) value at the first time equilibrium diversity reached is \(-5.05\) (Quental and Marshall 2009b). For the simulations presented here, with LiMe 10 and equilibrium diversity set to 50 species, the average value of \( \gamma \) is \(-4.14\) at the first time equilibrium diversity is reached.

More formally, McPeek (2008) has shown that the maximum value of \( \gamma \) is determined by the number of species in a phylogeny. The most negative value that \( \gamma \) can obtain is for a star phylogeny and its most positive...
value is for a phylogeny with one basal split with the remaining branches having zero length (McPeek 2008). Thus, the greater the number of species the wider the range of possible \( \gamma \) values. Additionally, Phillimore and Price (2008) suggest, based on the analysis of empirical and simulated phylogenies, that bigger clades might have more negative \( \gamma \) values simply because of an over-representation of clades that have experienced stochastically driven, above-average diversification rates early in their history. Hence, diversity-dependent diversification will be most readily detected using the \( \gamma \) when the clades in question are large and when they are close to the point in time when diversity equilibrium is first reached.

That about half of the empirical clades studied to date (McPeek 2008) show that the negative \( \gamma \) values poses a dilemma because even if all diversification were diversity-dependent, we would not expect to see so many phylogenies with negative \( \gamma \) values: not all clades of interest are large today, nor do we expect most clades studied to be at a point in time where the signature of their diversity-dependent diversification would be preserved in their molecular phylogenies. In fact, molecular phylogenies are essentially blind to where a clade is in its evolutionary trajectory, that is, whether it is still radiating, just reaching equilibrium, sitting at a constant diversity, or even declining.

One possible explanation for the apparent overabundance of phylogenies with negative \( \gamma \) values is that the current tools used to generate chronograms are biased in some way and tend to artificially concentrate nodes deep in the tree (i.e., perhaps, they artificially shorten basal branches and/or increase the length of terminal branches). For instance, underparameterization of models of DNA evolution can lead to a negative bias in the estimation of the \( \gamma \) statistic (Revell et al. 2005). Alternatively, clades that are actually declining in diversity (i.e., have negative diversification rates) can also produce phylogenies with negative \( \gamma \) values even when no diversity dependence is operating (Quental and Marshall 2009a). Finally, a tendency to oversample deeper nodes might also lead to the predominance of negative \( \gamma \) values (Renner and Cusimano 2010). Thus, although we have shown under what conditions diversity dependence can be detected in a molecular phylogeny generated by diversity-dependent processes, a negative \( \gamma \) value might not mean that the clade has actually diversified via such a process.

**Preservation and the Fossil Record**

The fossil record gives a direct window into how diversity has changed through time, but the observed diversity cannot be interpreted at face value because of preservational issues. The issue of incomplete sampling is well studied and increasingly understood (Marshall 1994; Solow and Smith 1997; Alroy et al. 2001, 2008; Connelly and Miller 2001; Foote 2003). Phylogenetic hypotheses can also be used to infer gaps in the fossil record (Wills 2007), although this approach is biased in that range, extensions can only be made to the bottoms of stratigraphic ranges but not to their tops (Wagner 2000). The distribution of fossil occurrences is often implicitly or explicitly assumed to be random with respect to traits and time (Madin et al. 2006; Alroy et al. 2008), although covariates, such as body size, have been used to account for heterogeneity in the preservation of different taxa (Liow et al. 2008). On a broad temporal scale, sampling rates may decline as we go back in geologic history (Raup 1979; Kidwell and Holland 2002) while many individual taxa have hat-shaped temporal occurrence distributions that need to be accounted for (Foote et al. 2007; Liow et al. 2010). In this contribution, we explicitly explored both the general temporal and lineage temporal aspects of preservation to understand how they affect our direct observations of diversification rates in the fossil record.

There are two components to diversity-dependent diversification, the initial phase where the diversity increases to the equilibrium diversity and the equilibrium phase where there is species turnover at a stochastically constant diversity. During the initial phase, the rate of diversification steadily drops to zero as the diversity first reaches the equilibrium diversity. During the equilibrium phase, the rate of diversification fluctuates but is constrained to oscillate around zero. At a relatively high temporal resolution in the synthetic fossil records (\( OW = 0.25 \) ESD) and where the difference between the initial and equilibrium phases is large (e.g., LiMe 10), the deceleration of diversification is generally detectable and once the equilibrium diversity is reached, the rate is well constrained at zero (Figs. 5, 6a–c). However, the initial phase becomes harder to detect as the temporal resolution is lowered (\( OW = 0.5 \) ESD) even for large values of LiMe, but the signal of a zero diversification rate during the equilibrium phase is still preserved (see, e.g., Fig. 6d–f). When LiMe is small (Fig. 6g–h), the average diversification rate does not reach zero for many average species durations, and even with a perfect fossil record (Fig. 6g), it is not possible to discern the slowly decreasing rate of diversification. With an incomplete fossil record, it is essentially impossible to detect the decreasing diversification rate, and while the average rate over many simulations is around zero, the estimated rate fluctuates widely. Unfortunately, we cannot see at present how to rigorously discriminate between rates that are fluctuating around zero (for clades that have already reached their equilibrium diversity) and those that are still diversifying at slow rates (i.e., at low LiMe values). Nonetheless, although the fossil record fails to pick up the decreasing rate of diversification for low values of LiMe, this is not as serious a problem as it might seem. The reason is that for low values of LiMe even the initial diversification rate is small (Fig. 1a) and is likely to be quite close to zero. In fact, the same argument could be used to explain why the \( \gamma \) statistic has a hard time in detecting decreases in diversification rates when LiMe is low.

Despite the degradation of the temporal data, we find that the fossil record correctly gives, on average,
an indication of constant and zero diversification regardless of the preservation scenarios, except perhaps in the case where preservation rates are as low as only 10%, which for the better preserved groups is an overly pessimistic number. However, the fossil record will not always pick up the signal of decreasing diversification at the beginning of the clade’s diversification. Thus, if the fossil record shows a constant diversification rate, it will be difficult to determine whether the underlying diversification dynamics were controlled by a low LiMe value (and thus that the clade is still diversifying) or whether the dynamics involved an undetected higher value of LiMe (and thus that the clade is at equilibrium diversity).

Finally, we note that the observed diversification rates unexpectedly sometimes dip below zero (Figs. 5b,c,e,f and 6b,h). This is an artifact caused by edge effects: older lineages are likely to have their time of extinction estimated earlier within the time frame of analysis, and younger lineages are likely to have their time of speciation estimated later within the same frame such that close to the end of the time frame, extinction rates are inflated, whereas speciation rates are deflated. This bias is “rescued” in part by increasing preservation forward in time (Fig. 5d), but other preservation scenarios tend to emphasize this effect (Figs. 5 and 6).

In summary, our ability to observe diversity-dependent diversity dynamics in the fossil record is more seriously compromised by the loss of temporal resolution than by incomplete preservation per se. We also found that the estimated diversification rates are more robust than the estimates of the speciation or extinction rates alone (results not shown), most likely due to the fact that the preservation processes affect both speciation and extinction rates in similar ways, so that any biases and noise introduced when calculating those rates alone are canceled out when the difference between them is calculated.

We note here that we have modeled very simple diversification scenarios. In reality, in cases where diversity-dependent diversification might have occurred, a clade’s process of diversification may have been perturbed for substantial periods of time by environmental change and its carrying capacity might also have changed with time. Thus, for real fossil data, it will be difficult to determine whether deviations from the pattern of diversification rates expected under diversity-dependence are due to the vagaries of the preservation process and the dampening of the true signal due to low temporal resolution (see above) or whether the deviations might be due to real departures from a simple model of diversity-dependent diversification.

Temporal Resolution and Natural Units for Measuring Time

The size of the OW has a large impact on our ability to discern patterns of interest in the fossil record. For instance, given the case where the ESD is 10 myr and the OW is 2.5 myr, the fossil record easily recovers the pattern for diversity dependence, given a large enough LiMe value (e.g., 10). However, as the ratio OW/ESD increases information is quickly lost from the fossil record. We also ran analyses with OW/ESD = 1, and as expected, little information remained (results not presented). An average marine invertebrate species has an average duration of 4 myr and the median duration of a geological stage is 5.5 myr (Gradstein et al. 2004), hence the OW/ESD ratio is close to 1. However, most global analyses are done with genus data (Foote et al. 2007; Alroy et al. 2008), where the OW/ESD ratio is much lower. In cases where smaller geographical subsets of fossilized organisms are studied, temporal resolution is often better and hence the OW/ESD ratio decreases too (Jackson et al. 1993; O’Dea et al. 2007). We note, however, that the ability to detect diversity dependence, as well as other modes of diversification, is also controlled by LiMe (Quental and Marshall 2009b), which also involves the ESD. Hence, the dimensionless ratio OW/(LiMe × ESD) is crucial to determining whether the fossil record can be used to unravel a clade’s diversity dynamics. However, at present, we do not have reliable estimates of the range of LiMe values in the real world, although molecular phylogenies in conjunction with the fossil record may be able to help determine LiMe values (Quental and Marshall 2009b).

Although ratios such as LiMe, OW/ESD, etc., help us understand diversification processes, the absolute time it takes to reach a given point in a clade’s history (e.g., the time to first reach the equilibrium diversity) is determined by the absolute values of its diversification rate and species longevities. For example, in our simulations with LiMe 10, with the extinction rate equal to 0.1 L/Lmyr, it takes about 10 myr to first reach the equilibrium diversity, whereas for the simulations with the extinction rate equal to 1.0 L/Lmyr, it takes about 1 myr to reach that point. In both cases, this time represents about 1 ESD. Despite the fact that there is an extra step required to translate our temporal scales into absolute time, we have used the ESD as a substitute for absolute time in our plots because for the analyses of diversification dynamics, the ESD is a natural unit for scaling evolutionary processes, just as generation times are useful timescales for analyzing evolutionary change (Gingerich 2001).

Combining the Best of Both Worlds

Computer simulations have shown that molecular phylogenies with significantly negative $\gamma$ values are consistent with diversity-dependent diversification (Rabosky and Lovette 2008; Quental and Marshall 2009b; results presented here). However, for clades that have not yet reached but are close to their equilibrium diversity, relatively high extinction rates can erode the diversity-dependent signal from molecular phylogenies (Rabosky and Lovette 2008; Quental and Marshall 2009b). Here, we show how far the diversification process has to proceed before that the diversity-dependent...
process can be seen in the first place. More significantly, we also show that if enough time has elapsed after the equilibrium diversity is first reached, the signature of diversity-dependent diversification is lost even for the cases where one would expect to see it (e.g., when extinction rates are relatively low). The loss of the signature of the diversity dependence is due to both the erosion of the evidence of the initial radiation by extinction and the steady turnover of species that dilutes the concentration of deep nodes established during the initial radiation. Hence, when $\gamma$ values are not significantly negative, it is difficult to determine whether diversity-dependent dynamics were present or not. In other words, the type I error rate is low (given the null hypothesis of exponential diversification), whereas the type II error rate is quite high. In cases where $\gamma$ values are not indicative of change in the diversification rate, the fossil record may be used to help determine whether the clade is really diversifying exponentially or is sitting at its carrying capacity. We also note that the type II error rate decrease as the clade sizes increase, given that $\gamma$ values become more negative for a given LiMe value for higher carrying capacities (McPeek 2008; Phillimore and Price 2008; Quental and Marshall 2009b). Finally, the type II error rate also drops with increasing LiMe values.

Despite the emphasis on the imperfectness of the fossil record, our results show that it can be used to robustly detect the equilibrium phase of a diversity-dependent diversification model, even when the fossil record is quite incomplete. On the other hand, the fossil record has a much harder time capturing the initial phases of the radiation. Here, molecular phylogenies can be extremely useful given that significantly negative $\gamma$ values indicate decreasing rates of diversification, the hallmark of the initial phase of the radiation. Moreover, both types of data can be used in combination to provide more reliable estimates the absolute values of initial speciation rates (Quental and Marshall 2009b). In summary, our analyses show that the molecular phylogenies and the fossil record each have their strengths and weaknesses when it comes to identifying diversity-dependent diversity dynamics, but when combined, may be used to more fully characterize those dynamics.

ACKNOWLEDGMENTS

L.H.L. thanks D. Rabosky for the invitation to give a talk at the 2009 SSBE Symposium held at the Society for the Study of Evolution meeting in Moscow Idaho, and all the other invitees for inspiring this work. T.B.Q. and C.R.M. thank N. Nagalingum and Mark Clements for valuable discussion. We also thank D. Rabosky and 2 anonymous reviewers for their comments and suggestions.

REFERENCES


