# Slowdowns in Diversification Rates from Real Phylogenies May Not be Real

NATALIE CUSIMANO AND SUSANNE S. RENNER*

* Correspondence to be sent to: Botanische Staatsanstalten, Menzinger-Str. 67, D-80638 Munich, Germany; E-mail: renner@lrz.uni-muenchen.de.

Abstract.—Studies of diversification patterns often find a slowing in lineage accumulation toward the present. This seemingly pervasive pattern of rate downturns has been taken as evidence for adaptive radiations, density-dependent regulation, and metacommunity species interactions. The significance of rate downturns is evaluated with statistical tests (the $\gamma$ statistic and Monte Carlo constant rates (MCCR) test; birth–death likelihood models and Akaike Information Criterion [AIC] scores) that rely on null distributions, which assume that the included species are a random sample of the entire clade. Sampling in real phylogenies, however, often is nonrandom because systematists try to include early-diverging species or representatives of previous intrataxon classifications. We studied the effects of biased sampling, structured sampling, and random sampling by experimentally pruning simulated trees (60 and 150 species) as well as a completely sampled empirical tree (58 species) and then applying the $\gamma$ statistic/MCCR test and birth–death likelihood models/AIC scores to assess rate changes. For trees with random species sampling, the true model (i.e., the one fitting the complete phylogenies) could be inferred in most cases. Oversampling deep nodes, however, strongly biases inferences toward downturns, with simulations of structured and biased sampling suggesting that this occurs when sampling percentages drop below 80%. The magnitude of the effect and the sensitivity of diversification rate models is such that a useful rule of thumb may be not to infer rate downturns from real trees unless they have >80% species sampling. [Diversification rate; diversification slowdown; gamma statistic and MCCR test; species sampling; overconfidence.]

Numerous recent studies have used chronograms from molecular dating to infer the distribution of clado-genetic events across a tree (e.g., Weir and Schluter 2007; Linder 2008; Phillimore and Price 2008, 2009). In animal clades as diverse as birds, reptiles, beetles, and fishes, such studies have often revealed apparent slowing in lineage accumulation toward the present (Nee et al. 1992; Zink and Slowinski 1995; Lovette and Bermingham 1999; Price et al. 2000; Harmon et al. 2003; Rüber and Zardoya 2005; Kozak et al. 2006; McKenna and Farrell 2006; Weir 2006; Phillimore and Price 2008; Rabosky and Lovette 2008a,b). A predominant pattern of rate slowdowns was also found in an analysis of 245 phylogenies, including 39 angiosperm phylogenies (McPeek 2008). Other plant phylogenies (not included in McPeek 2008) also exhibited slowing diversification rates (Kadereit et al. 2004; Good-Avila et al. 2006; Merckx et al. 2008; Smith et al. 2008). This seemingly pervasive pattern of diversification downturns across major groups of animals and plants has been taken as evidence for adaptive radiations (Harmon et al. 2003; Weir 2006), density-(diversity-)dependent regulation (Weir 2006; Phillimore and Price 2008, 2009; Rabosky and Lovette 2008a,b), and metacommunity species interactions (McPeek 2008).

The significance of rate downturns typically is evaluated with the $\gamma$ statistic and the Monte Carlo constant-rates (MCCR) test (Pybus and Harvey 2000) or with Akaike Information Criterion (AIC) scores in birth–death likelihood analyses (Rabosky 2006a,b). To create the probability distribution of the null hypothesis, one simulates phylogenies for the complete number of species in the focal clade (Pybus and Harvey 2000). If

the molecular phylogeny for the focal clade is incomplete, the simulated tree sets are then pruned to the number of species actually sequenced. These null distributions are based on the assumption that the species in a tree represent a random sample of all species in a clade. Overdispersed sampling will raise the Type I error of the MCCR test, whereas underdispersed sampling will raise the Type II error (Pybus and Harvey 2000). Random sampling is a critical assumption because incompletely sampled clades may be typical in molecular phylogenetics, and the included species may not have been chosen at random.

A review of plant diversification studies shows that many have inferred diversification rates from phylogenies with <60% species sampling, down to as little as 11% (e.g., von Hagen and Kadereit 2001, 2003; Malcomber 2002; Becerra 2005; Hughes and Eastwood 2006; Good-Avila et al. 2006; Smith et al. 2008; Sauquet et al. 2009; Janssens et al. 2009). Sampling completeness here refers to the total numbers of species in a focal clade according to the most recent taxonomic treatment, relative to the species included in the phylogeny. Similarly, of the 245 phylogenies included in meta-analysis of diversification rates of McPeek (2008), 56 had less than 80% species sampling, and some of the least complete phylogenies yielded the strongest downturns. An example is the extreme diversification slow-down in Limonium, a genus of approximately 350 species for which "at least one species from each of the sections" (Lledó et al. 2005: 1191) was sampled for a 19% sampling density (not 100% as scored in McPeek 2008: OSM table 1). To guard against incomplete phylogenies possibly biasing his metaanalysis, McPeek (2008: Figure 1) tested
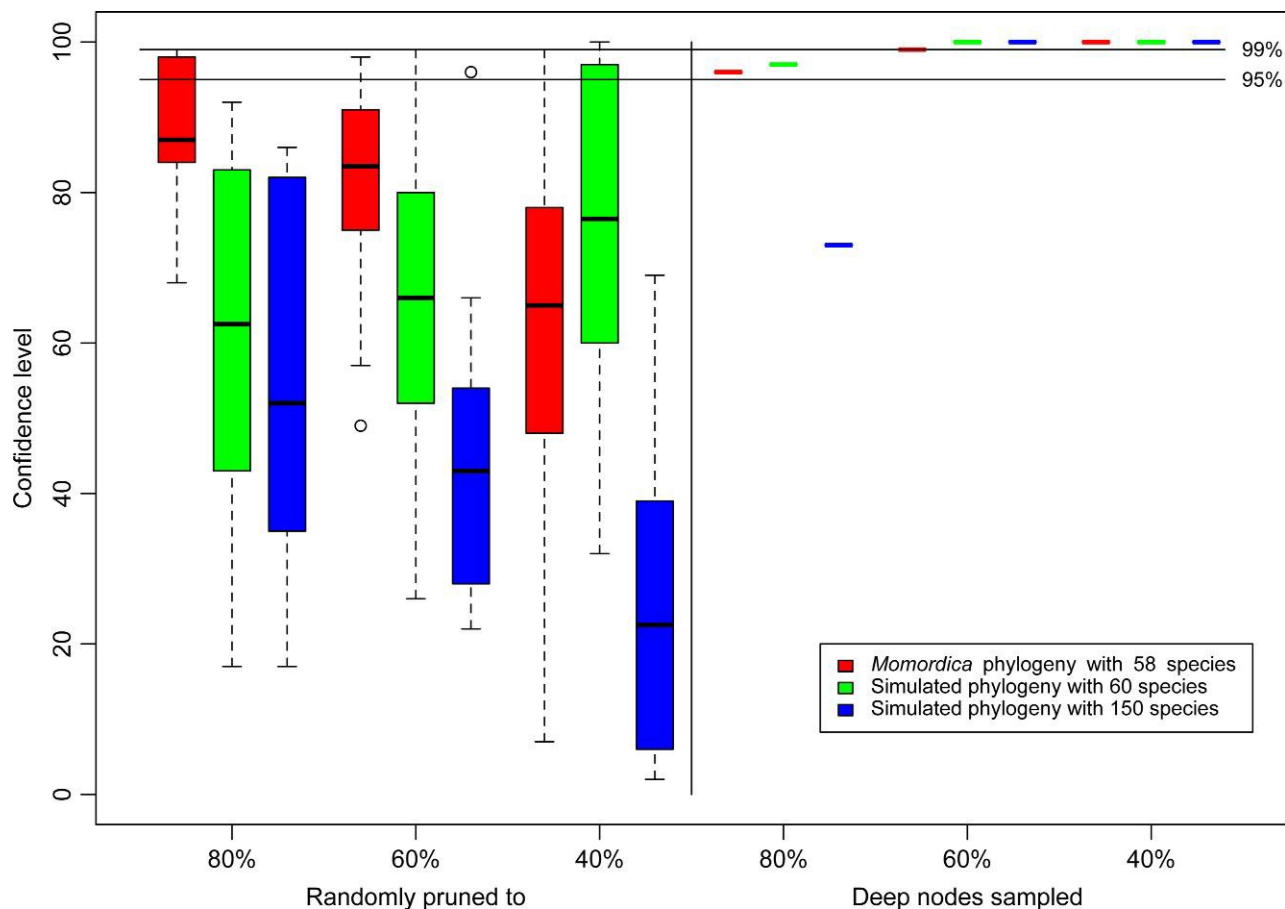
FIGURE 1. Statistical confidence in one's ability to reject an underlying diversification model for real and simulated phylogenies of different sizes, randomly pruned to 80%, 60%, or 40% of their initial species (left panel) and confidence levels for phylogenies manually pruned so as to oversample deep nodes (right panel). Different colors indicate the three starting phylogenies (one empirical, two simulated). Horizontal lines indicate the 95% and 99% confidence levels.

for a correlation between γ and the percent of species sampled for a clade. No correlation was found, a result that we take up in the Discussion section.

As the preceding examples show, many published diversification rates come from incomplete phylogenies. This raises the question whether species sampling in these incomplete phylogenies was random. In many cases, this does not seem to be the case. Rather, studies with incomplete species sampling usually state that species were included to cover a taxon's geographic ranges and morphological diversity, often also previously recognized subtaxa (e.g., Linder et al. 2003; Lledó et al. 2005; Good-Avila et al. 2006; Merckx et al. 2008; Smith et al. 2008; Sauquet et al. 2009; Janssens et al. 2009; Crisp and Cook 2009). Indeed, nonrandom sampling is a sensible strategy as systematists try to strike a balance between the expense of time and funds on sequencing and the return in terms of insights into phylogenetic relationships. Such insights can be gained even with sparse sampling as long as each morphologically or geographically defined group is represented. The matter is different when the goal is inferring diversification rates. Here, biased, structured, or random species sampling have different outcomes, and if real phylogenies

involve overrepresentation of deep nodes, this will introduce a bias into any results concerning clade diversification.

Here, we quantify the magnitude of the bias introduced by phylogenetically informed sampling, that is, sampling such that deep nodes are overrepresented, as well as effects of structured sampling, such that nodes are represented proportionally to the number of nodes in each time slice. Our experiments involve simulated trees with 60 and 150 tips and a real phylogeny with 58 tips. We chose these clade sizes as fairly typical for groups used in diversification studies (see Discussion section). Effects were measured by using birth–death likelihood analyses and the γ statistic and MCCR test.

## MATERIAL AND METHODS

We simulated ultrametric trees with either 60 or 150 tips under the Yule model with a diversification rate of 0.5 and slightly positive γ values ($\gamma = 0.076$ and $\gamma = 0.07$, respectively). Trees were then randomly pruned 10 times to 80, 60, and 40% of their original species numbers. To test for the effect of oversampling deep

nodes, we manually pruned 20, 40, or 60% of the species in shallow nodes, leaving 80, 60, and 40% of the species representing deeper nodes. These experiments address both, the effects of biased species sampling and the effects of structured sampling (because of the specific percentages of species left in the pruned trees). Manual pruning, different from random pruning, was only performed once per phylogeny because it makes no difference which of two sister clades of young clades are cut to end up with the required percentage of deep nodes. Put differently, there is a limited number of ways in which a clade of 60 tips can be oversampled so as to leave 80, 60, and 40% of its deep nodes.

A third phylogeny on which we carried out the same kinds of experiments was an ultrametric phylogeny of the Cucurbitaceae genus Momordica, which includes 58 of the clade's 59 species (Schaefer and Renner 2010; the species number is based on monographic work by H. Schaefer).

All resulting trees were analyzed with the γ statistic and the MCCR test (1000 replicates), which test whether splitting events across the phylogeny are evenly distributed or accumulated toward the root or the tips of the phylogeny (Pybus and Harvey 2000). These tests as well as birth–death likelihood analysis (BDL) were carried out in Laser 2.2 (Rabosky 2006a). In the BDL analysis, five models were fitted to the 34 Momordica chronograms (see Results section), two constant-rate models of diversification (a pure-birth (Yule) model with constant speciation rate; a birth–death model with constant speciation and extinction rates) and three variable-rate models (logistic density dependence model [DDL]; exponential density dependence model [DDX]; a two-rates variant of the pure-birth model with a rate shift at a certain time point [Yule-2-rates]). The ΔAICrc is calculated for every phylogeny as the difference in AIC scores of the best-fit rate-constant and the best-fit rate-variable model.

To obtain a null distribution of AIC scores, we simulated 1000 trees for each of the 34 chronograms independently, with the number of tips corresponding to the complete number of species (58) under a pure-birth model with a diversification rate obtained by fitting the Yule model to the respective phylogeny. The simulated trees were then pruned (if necessary) to the real number of species occurring in the variously pruned experimental phylogenies. Fitting each of the diversification models to every simulated tree yielded a null distribution of AIC scores under the null model against which to compare the scores obtained with the respective "true" phylogeny. Tree simulations were carried out in Cass (Gernhard 2008). To obtain a null distribution of γ values against which to test the empirical (and experimental tree) γ values, we used the MCCR test.

## RESULTS

### Simulated Data

For the simulated phylogenies of initially 60 species, random pruning resulted in inferred downturns in

seven cases and in inferred constant diversification in the remaining 23 (Table 1). Statistical confidence in the downturns grew the more severely pruned a phylogeny (20% pruned: confidence level (CL) = 90%, 40% pruned: CL = 95/99%; 60% pruned: CL = 95/99%). For the simulated phylogenies of initially 150 species, random pruning resulted in an inferred downturn in just one case (CL = 95%; Table 2).

Manual pruning of the 150-species tree leaving 80% of its species resulted in the acceptance of the (true) constant rate model (Fig. 1). Manual pruning of the smaller phylogenies (58 or 60 tips) led to a downturn being inferred with CLs of 95% even when 80% of their species were sampled (Fig. 1). When phylogenies oversampled deep nodes more severely (60% or 40% of the species manually pruned in either the 60- or the 150-species tree), downturns were inferred with high confidence (99.9% CL; Fig. 1 and Table 1). The γ values of all but two pruned phylogenies were negative, in contrast to the initial phylogeny (Table 1).

### Real Data

The best model for the Momordica phylogeny is the Yule-2-rates model, with a rate decrease at 2 Ma. However, a simple Yule model could not be rejected, nor did the γ statistic reject a constant-rates model. The diversification rate is 0.08 species per million years (sp./Myr), and the lineage-through-time plot for the complete Momordica tree is shown in Figure 2 (original data), together with plots for the 33 pruned trees.

Random pruning of 20% of the species (leaving 80% in the tree) already affected the diversification rates such that they became lower than they had been in the complete phylogeny (0.07–0.08 sp./Myr; Table 2). Randomly pruning 40% of the species had nearly the same effects as pruning 20% (Fig. 2), but diversification rates under the Yule model became even lower (0.06–0.07 sp./Myr; Table 2). Of the variable-rate models, the DDL model became the best-fitting model, in one case with a 99.9% CL (Table 2). Randomly pruning 60% of the species (leaving 40%) resulted in the Yule model becoming the preferred model, with inferred diversification rates ranging from 0.05 to 0.06 sp./Myr (Table 2). The constant-rate model could be rejected only in one randomly pruned phylogeny.

Manually pruning trees such that deep nodes were oversampled resulted in the Yule-2-rates model becoming the preferred model when 80% of the species were still in the tree, and the DDL model when 60% or 40% of species were left (Table 2, lowest three rows). Confidence levels were 95% for the 80%-sampled tree and 99.9% for the other two sampling schemes (Table 2). The extreme downturns are also seen in the lineage-through-time plots in Figure 2. All the inferred diversification rates and breakpoint times differed strongly from those obtained when fitting the DDL or Yule-2-rates model to the complete phylogeny (Table 2), with the effects becoming larger the more pruned the phylogeny, an

TABLE 1. Results from the experimental pruning of two simulated trees of 60 and 150 tips; oversampling of deep nodes was simulated by manual pruning (see Materials and Methods section). Resulting trees were analyzed with the γ statistic and the MCCR test (1000 replicates)

| | | 60 species | | | 150 species | | |
|---|---|---|---|---|---|---|---|
| | | γ value | cr. value | p value | γ value | cr. value | p value |
| Complete phylogeny | | 0.08 | — | 0.53 | 0.07 | — | 0.53 |
| Pruned randomly to | 80% | −0.46 | −1.63 | 0.39 | −0.55 | −2.33 | 0.61 |
| | | −0.37 | −2.32 | 0.68 | −0.93 | −1.59 | 0.18 |
| | | 0.64 | −1.66 | 0.83 | −0.20 | −2.34 | 0.83 |
| | | −0.47 | −1.73 | 0.36 | −0.32 | −2.08 | 0.65 |
| | | −0.55 | −1.61 | 0.33 | −1.17 | −1.79 | 0.17 |
| | | −0.86 | −1.09 | 0.08 | −0.46 | −1.70 | 0.41 |
| | | −0.40 | −0.94 | 0.17 | −0.81 | −1.77 | 0.27 |
| | | 0.19 | −1.30 | 0.53 | −0.19 | −2.10 | 0.65 |
| | | −1.14 | −1.34 | 0.08 | −1.09 | −1.52 | 0.14 |
| | | −0.05 | −1.77 | 0.57 | −0.51 | −2.13 | 0.55 |
| Pruned randomly to | 60% | −2.33 | −1.62 | **0.01** | −1.20 | −2.65 | 0.55 |
| | | −0.56 | −2.37 | 0.74 | −1.92 | −1.86 | **0.04** |
| | | −1.71 | −1.49 | **0.03** | −1.08 | −2.60 | 0.59 |
| | | −0.45 | −2.57 | 0.73 | −1.17 | −3.18 | 0.78 |
| | | −0.92 | −2.09 | 0.39 | −0.92 | −2.21 | 0.48 |
| | | −1.26 | −2.33 | 0.35 | −0.98 | −2.29 | 0.46 |
| | | −0.02 | −1.42 | 0.48 | −1.64 | −2.56 | 0.34 |
| | | −0.78 | −1.64 | 0.29 | −0.65 | −2.57 | 0.72 |
| | | −1.91 | −2.59 | 0.20 | −1.05 | −2.76 | 0.67 |
| | | −1.72 | −2.67 | 0.33 | −0.90 | −2.75 | 0.72 |
| Pruned randomly to | 40% | −2.17 | −1.91 | **0.03** | −1.01 | −4.03 | 0.98 |
| | | −2.92 | −1.66 | **0.00** | −0.96 | −2.78 | 0.75 |
| | | −0.92 | −1.91 | 0.40 | −1.47 | −3.22 | 0.72 |
| | | −2.07 | −1.59 | **0.01** | −1.37 | −3.28 | 0.80 |
| | | −0.84 | −2.26 | 0.57 | −1.80 | −3.26 | 0.61 |
| | | −1.96 | −2.53 | 0.17 | −2.20 | −3.14 | 0.36 |
| | | −1.78 | −2.57 | 0.26 | −0.92 | −3.30 | 0.94 |
| | | −1.24 | −2.91 | 0.68 | −1.21 | −2.06 | 0.31 |
| | | −1.56 | −2.17 | 0.21 | −1.51 | −3.71 | 0.89 |
| | | −1.43 | −2.29 | 0.28 | −0.79 | −3.41 | 0.95 |
| Deep nodes sampled | 80% | −1.01 | −0.88 | −0.03 | −1.02 | −1.86 | 0.27 |
| | 60% | −3.12 | −1.43 | **0.00** | −2.49 | −1.62 | **0.00** |
| | 40% | −4.53 | −1.84 | **0.00** | −4.58 | −2.00 | **0.00** |

Note: cr. value = the critical value of the MCCR test. Bold indicates p < 0.05.

effect seen earlier in the two smaller phylogenies than in the larger one. The structured sampling of 80%, 60%, or 40% of the species in the manually pruned phylogenies showed that any sampling below 80% will bias results toward diversification downturns. The most striking result of our experiments may be the high confidence levels obtained with the MCCR test for rate downturns (Fig. 1 and Table 2).

## DISCUSSION

These experiments with simulated phylogenies of 60 or 150 species, plus a real phylogeny of 58 species, reveal a strong and consistent effect of phylogenetically informed species sampling (i.e., biased toward deeper nodes) on the diversification patterns likely to be inferred from incomplete phylogenies. Trees in which deep nodes are oversampled will bias results toward rate slowdowns, that is, apparent density-dependent diversification. Experiments with structured manual pruning (leaving 80%, 60%, or 40% of the species) suggest that this occurs already when <80% of a clade's species are sampled such that deeper nodes are overrepresented (Table 2). Preferred inferred models always implied diversification slowdowns; a rate upswing was never inferred. For trees with random species sampling, the true model (i.e., the one fitting the full phylogenies) could be inferred in all but one case using BDL or in 86.67% of the cases using the γ statistic.

The magnitude of the downturn bias, at least in smaller phylogenies (Table 2: 60-species phylogenies vs. 150-species phylogenies), suggests great caution in making generalizations about diversification slowdowns across clades of animals or plants. A compounding effect may be the sensitivity of γ toward clade size and age demonstrated by Phillimore and Price (2008, 2009). Clades larger than 50 species and younger than 20 Myr tend to show more negative γ values (i.e., stronger slowdowns) than do smaller or older clades. For the Momordica phylogeny, with its 58 species and approximately 35 Myr age, this should not have been a problem. However, γ is also biased toward negative values if molecular clocks underestimate deep branches (Revell et al. 2005), a problem of unknown magnitude. Here, we have focused only on the effects of oversampling deep nodes in incompletely sampled, medium-sized phylogenies. Methods that attempt to correct for missing species in diversification analyses all assume that the sampled species are a random sample of a clade's total species (Pybus and Harvey 2000; Rabosky 2006b), and this assumption may be valid for some phylogenies.

TABLE 2. Results from fitting five diversification models to the complete Momordica phylogeny (58 species) and variously pruned subsets of it (see Materials and Methods section)

| Pruning Scheme | % of species remaining | Birth-death likelihood analysis | | | | | | | | γ statistic | MCCR test | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Best RC model | r1 | Best RV model | Model parameters | | | dAICrc | p value | γ value | cr. value | p value |
| | | | | | r1 | xp/k/r2 | st | | | | | |
| Complete Phylogeny | 100% | pure birth | 0.08 | yule-2-rates | 0.09 | 0.03 | 2.02 | 2.31 | 0.13 | −1.25 | | 0.11 |
| | | | | DDX | 0.20 | 0.28 | | | | | | |
| | | | | DDL | 0.11 | 116.00 | | | | | | |
| Pruned randomly to | 80% | pure birth | 0.07 | yule-2-rates | | | | 1.15 | 0.20 | −0.99 | −1.95 | 0.32 |
| | | pure birth | 0.07 | yule-2-rates | | | | 2.08 | 0.13 | −1.49 | −2.26 | 0.24 |
| | | pure birth | 0.07 | DDx | | | | 1.77 | 0.14 | −1.51 | −1.10 | 0.02 |
| | | pure birth | 0.07 | DDL | 0.12 | 66.14 | | 2.62 | 0.10 | −1.88 | −1.28 | 0.01 |
| | | pure birth | 0.07 | DDX | | | | 2.06 | 0.12 | −1.57 | −2.09 | 0.15 |
| | | pure birth | 0.07 | DDX | 0.27 | 0.45 | | 2.77 | 0.09 | −1.87 | −2.36 | 0.14 |
| | | pure birth | 0.07 | DDX | 0.28 | 0.46 | | 3.04 | 0.07 | −1.88 | −1.32 | 0.01 |
| | | pure birth | 0.07 | yule-2-rates | 0.09 | 0.03 | 3.27 | 2.69 | 0.09 | −1.53 | −1.50 | 0.05 |
| | | pure birth | 0.08 | yule-2-rates | | | | 1.36 | 0.19 | −1.16 | −1.70 | 0.16 |
| | | pure birth | 0.07 | yule-2-rates | | | | 1.31 | 0.17 | −1.40 | −1.92 | 0.12 |
| Pruned randomly to | 60% | pure birth | 0.07 | DDX | | | | 1.03 | 0.30 | −1.36 | −2.10 | 0.25 |
| | | pure birth | 0.06 | DDX | 0.38 | 0.66 | | 4.64 | 0.07 | −2.14 | −2.88 | 0.21 |
| | | pure birth | 0.07 | DDX | | | | 0.09 | 0.44 | −0.88 | −2.06 | 0.43 |
| | | pure birth | 0.06 | DDL | 0.12 | 42.62 | | 3.71 | 0.08 | −2.06 | −2.61 | 0.16 |
| | | pure birth | 0.06 | DDL | 0.13 | 41.82 | | 4.24 | 0.06 | −2.17 | −2.39 | 0.09 |
| | | pure birth | 0.07 | DDL | | | | 1.10 | 0.31 | −1.44 | −1.93 | 0.16 |
| | | pure birth | 0.07 | yule-2-rates | | | | 0.76 | 0.31 | −1.00 | −2.27 | 0.51 |
| | | pure birth | 0.07 | DDX | | | | 0.77 | 0.33 | −1.27 | −1.76 | 0.17 |
| | | pure birth | 0.06 | yule-2-rates | | | | 3.01 | 0.12 | −1.73 | −1.97 | 0.08 |
| | | pure birth | 0.06 | DDL | 0.15 | 37.50 | | 9.44 | 0.00 | −3.01 | −2.74 | 0.02 |
| Pruned randomly to | 40% | pure birth | 0.05 | DDX | | | | 2.01 | 0.32 | −1.52 | −2.53 | 0.38 |
| | | pure birth | 0.05 | DDL | | | | 3.21 | 0.16 | −1.86 | −1.49 | 0.01 |
| | | pure birth | 0.05 | DDX | | | | 2.17 | 0.30 | −1.33 | −2.10 | 0.28 |
| | | pure birth | 0.06 | DDX | | | | −0.91 | 0.82 | −0.41 | −2.82 | 0.93 |
| | | pure birth | 0.06 | DDX | | | | −0.23 | 0.69 | −0.93 | −2.87 | 0.81 |
| | | pure birth | 0.05 | DDL | | | | 2.84 | 0.22 | −1.78 | −2.63 | 0.32 |
| | | pure birth | 0.05 | DDX | 0.51 | 0.95 | | 4.57 | 0.10 | −1.78 | −2.39 | 0.18 |
| | | pure birth | 0.05 | yule-2-rates | | | | 3.89 | 0.13 | −1.73 | −2.99 | 0.43 |
| | | pure birth | 0.05 | DDL | 0.13 | 24.45 | | 4.29 | 0.10 | −2.08 | −3.37 | 0.52 |
| | | pure birth | 0.05 | DDX | | | | 3.02 | 0.21 | −1.44 | −2.08 | 0.22 |
| Deep nodes sampled | 80% | pure birth | 0.07 | yule2rate | 0.08 | 0.02 | 3.27 | 3.81 | 0.04 | −2.01 | −1.92 | 0.04 |
| | 60% | pure birth | 0.06 | DDL | 0.14 | 40.00 | | 7.44 | 0.01 | −2.72 | −2.09 | 0.01 |
| | 40% | pure birth | 0.04 | DDL | 0.20 | 22.70 | | 18.49 | 0.00 | −3.99 | −1.22 | 0.00 |

Note: Best RC = best rate-constant model; r1 = net diversification rate; Best RV = best rate-variable model; xp/k/r2 refer to the parameters associated with the DDX, DDL, and Yule-2-rate models; xp = rate change parameter in the DDX model, k refers to the carrying capacity parameter in the DDL model, and r2 is the second net diversification rate after the breakpoint time st (in Myr) in the Yule-2-rates model; st = breakpoint in the Yule-2-rates model; $\Delta$AICrc is the difference between the best-fitting rate-variable and the best-fitting rate-constant model; cr. value = the critical value of the MCCR test.
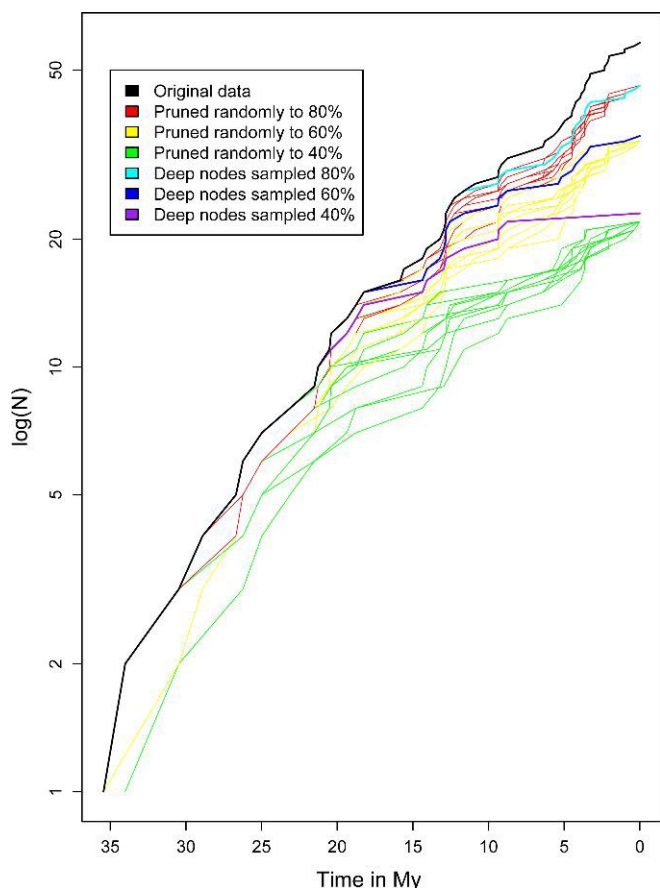
FIGURE 2. Lineage-through-time plots obtained from the complete Momordica phylogeny (58 species) and variously pruned subsets of it (Table 2).

under assumed diversification rates that themselves are obtained from more or less incomplete phylogenies (due to incomplete sampling as well as extinction). Our results, which show that the corrections for missing species by standard tree simulation and pruning (MCCR test) commonly lead to wrong models being preferred with confidence, stress the importance of correcting for missing species in ways that will reduce Type I error. Such correction may be possible if the missing splits (i.e., node ages) were simulated under some biologically realistic distribution and then added to the available splits, permitting calculation of confidence levels.

Of course, this paper highlights only one problematic aspect of inferring diversification rates from empirical phylogenies, and the bias quantified here may only affect medium-sized phylogenies that sample less than 80% of a clade's extant species and do so in a nonrandom way. Nevertheless, phylogenies that oversample deep nodes may be the rule rather than the exception, and being aware of the biases introduced by such sampling is therefore important. If one's main goal is to infer diversification, phylogenies that include at least 80% of a clade's known species may be necessary.

A test that related γ to species sampling density in 245 real phylogenies of arthropods, cordates, flowering plants, and mollusks (McPeek 2008), revealed no correlation, suggesting that some incomplete real phylogenies may indeed have randomly sampled species (thus eliminating the bias introduced by over-sampling deep nodes). Of the 245 phylogenies, 134 had complete species sampling and 56 had over 80% species sampling. However, half of the 245 trees contained fewer than 15 species, whereas 39 had 31 to 116 species (116 being the largest included phylogeny). This large sample likely includes clades with real diversification rate down turns as well as clades with oversampled deep nodes and (erroneously) inferred negative γ values. Therefore, the absence of a correlation between γ and species sampling density found in this metaanalysis has no bearing on the strong statistical bias introduced by oversampling deep nodes that we quantified in this study. Most researchers are inferring diversification rate changes using individual phylogenies (not a metaanalysis), and our findings now call from increased caution when using incomplete phylogenies with possibly nonrandom species sampling for such inferences.

An additional problem in diversification inference may be that the trees for the statistical tests are simulated

## REFERENCES

Becerra J.X. 2005. Timing the origin and expansion of the Mexican tropical dry forest. Proc. Natl. Acad. Sci. USA 102:10919–10923.

Crisp M.D., Cook L.G. 2009. Explosive radiation or cryptic mass extinction? Interpreting signatures in molecular phylogenies. Evolution 63:2257–2265.

Gernhard, T. 2008. Cass: A Python tool for analyzing and manipulating phylogenetic trees. http://www-m9.ma.tum.de/twiki/bin /view/Allgemeines/TanjaGernhard.

Good-Avila S.V., Souza V., Gaut B.S., Eguiarte L.E. 2006. Timing and rate of speciation in Agave (Agavaceae). Proc. Natl. Acad. Sci. USA 103:9124–9129.

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

Hughes C., Eastwood R. 2006. Island radiation on a continental scale: Exceptional rates of plant diversification after uplift of the Andes. Proc. Natl. Acad. Sci. USA 103:10334–10339.

Janssens S.B., Knox E.B., Huysmans S., Smets E.F., Merckx V.S.F.T. 2009. Rapid radiation of Impatiens (Balsaminaceae) during Pliocene and Pleistocene: result of a global climate change. Mol. Phyl. Evol. 52:806–824.

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

Kozak K.H., Weisrock D.W., Larson A. 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, B 273:539–546.

Linder, H.P. 2008. Plant species radiation: where, when, why? Phil. Trans. R. Soc. Lond., B 363:3097–3105.

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

Lledó, M.D., Crespo M.B., Fay M.F., Chase M.W. 2005. Molecular phylogenetics of Limonium and related genera (Plumbaginaceae): biogeographical and systematic implications. Am. J. Bot. 92:1189–1198.

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

Malcomber S.T. 2002. Phylogeny of Gaertnera Lam. (Rubiaceae) based on multiple DNA markers: Evidence of a rapid radiation in a widespread, morphologically diverse genus. Evolution 56:42–57.

McKenna D.D., Farrell B.D. 2006. Tropical forests are both evolutionary cradles and museums of leaf beetle diversity. Proc. Natl. Acad. Sci. USA 103:10947–10951.

McPeek M.A. 2008. The ecological dynamics of clade diversification and community assembly. Am. Nat. 172:E270–E284.

McPeek M.A., Brown J.M. 2007. Clade age and not diversification rate explains species richness among animal taxa. Am. Nat. 169:E97–E106.

Merckx V., Chatrou L.W., Lemaire B., Sainge M.N., Huysmans S., Smets E.F. 2008. Diversification of myco-heterotrophic angiosperms: Evidence from Burmanniaceae. BMC Evol. Biol. 8:178–193.

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

Phillimore A.B., Price T.D. 2008. Density dependent cladogenesis in birds. PLoS Biol. 6:e71.

Phillimore A.B., Price T.D. 2009. Ecological influences on the temporal pattern of speciation. In: Butlin R., Bridle J., D. Schluter, editors. Speciation and Patterns of Diversity. Cambridge: Cambridge University Press. p. 240–256

Price T.D., Lovette I.J., Bermingham E., Gibbs H., Richman A. 2000. The imprint of history on communities of North American and Asian warblers. Am. Nat. 156:354–367.

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

Rabosky D.L. 2006a. LASER: a maximum likelihood toolkit for inferring temporal shifts in diversification rates. Evol. Bioinf. Online 2:257–260.

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

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

Rabosky D.L., Lovette I.J. 2008b Explosive evolutionary radiations: decreasing speciation or increasing extinction through time? Evolution 62:1866–1875.

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

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

Sauquet H., Weston P.H., Anderson C.L., Barker N.P., Cantrill D., Mast A.R., Savolainen V. 2009. Contrasted patterns of hyperdiversification in Mediterranean hotspots. Proc. Natl. Acad. Sci. USA 106:221–225.

Schaefer H., Renner S.S. 2010. A three-genome phylogeny of Momordica (Cucurbitaceae) suggests seven returns from dioecy to monoecy and recent long-distance dispersal to Asia. Mol. Phyl. Evol. 54:553–560.

Smith C.I., Pellmyr O., Althoff D.M., M. Balcazar-Lara, J. Leebens-Mack, and Segraves K.A. 2008. Pattern and timing of diversification in Yucca (Agavaceae): specialized pollination does not escalate rates of diversification. Proc. R. Soc. Lond., B 275:249–258.

von Hagen K.B., Kadereit J.W. 2001. The phylogeny of Gentianella (Gentianaceae) and its colonization of the southern hemisphere as revealed by nuclear and chloroplast DNA sequence variation. Org. Divers. Evol. 1:61–79.

von Hagen K.B., Kadereit J. W. 2003. The diversification of Halenia (Gentianaceae): Ecological opportunity versus key innovation. Evolution 57:2507–2518.

Weir J.T. 2006. Divergent timing and patterns of species accumulation in lowland and highland neotropical birds. Evolution 60:842–855.

Weir J.T., Schluter D. 2007. The latitudinal gradient in recent speciation and extinction rates of birds and mammals. Science 315:1574–1576.

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