doi:10.1111/j.1558-5646.2012.01791.x

A PARAMETRIC METHOD FOR ASSESSING DIVERSIFICATION-RATE VARIATION IN PHYLOGENETIC TREES

Premal Shah,[1,2,3] Benjamin M. Fitzpatrick,[1,2] and James A. Fordyce[1,2,4]

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

[2]*National Institute for Mathematical and Biological Synthesis, Knoxville, Tennessee 37996*

[3]*Current address: Department of Biology, University of Pennsylvania, Philadelphia, Pennsylvania 19104*

[4]*E-mail: jfordyce@utk.edu*

Received June 24, 2011
Accepted August 7, 2012
Data Archived: Dryad doi:10.5061/dryad.8h350

Phylogenetic hypotheses are frequently used to examine variation in rates of diversification across the history of a group. Patterns of diversification-rate variation can be used to infer underlying ecological and evolutionary processes responsible for patterns of cladogenesis. Most existing methods examine rate variation through time. Methods for examining differences in diversification among groups are more limited. Here, we present a new method, parametric rate comparison (PRC), that explicitly compares diversification rates among lineages in a tree using a variety of standard statistical distributions. PRC can identify subclades of the tree where diversification rates are at variance with the remainder of the tree. A randomization test can be used to evaluate how often such variance would appear by chance alone. The method also allows for comparison of diversification rate among a priori defined groups. Further, the application of the PRC method is not restricted to monophyletic groups. We examined the performance of PRC using simulated data, which showed that PRC has acceptable false-positive rates and statistical power to detect rate variation. We apply the PRC method to the well-studied radiation of North American *Plethodon* salamanders, and support the inference that the large-bodied *Plethodon glutinosus* clade has a higher historical rate of diversification compared to other *Plethodon* salamanders.

KEY WORDS: **Adaptive radiation, chronogram, diversification rate, phylogeny, *Plethodon.***

Phylogenetic hypotheses have become increasingly important and frequently used tools in studies of macroevolutionary patterns. In particular, phylogenetic trees are commonly used to study variation in rates of diversification through time and among groups. Inferences stemming from reconstruction of diversification-rate variation can inform researchers on the role of geological or climatic events, the role of evolutionary novelty and key innovations, and adaptive and nonadaptive radiations in explaining diversification history and extant species richness (e.g., Harmon 2003; Jiggins et al. 2006; Weir 2006; Phillimore and Price 2008; Rabosky and Lovette 2008; Fordyce 2010a). Various approaches have been developed to identify diversification-rate variation evident in phylogenetic hypotheses. Two distinct kinds of questions are addressed by distinct methods; diversification rate variation through time and differences in diversification between groups (Fig. 1).

Variation through time methods directly examine the accumulation of lineages through time based upon a chronogram (i.e., an ultrametric phylogram where branch lengths are scaled as time). That is, they examine the vector of cladogenic events for a monophyletic group through time. This includes methods that explicitly model the process that determines the shape of

EVOLUTION INTERNATIONAL JOURNAL OF ORGANIC EVOLUTION

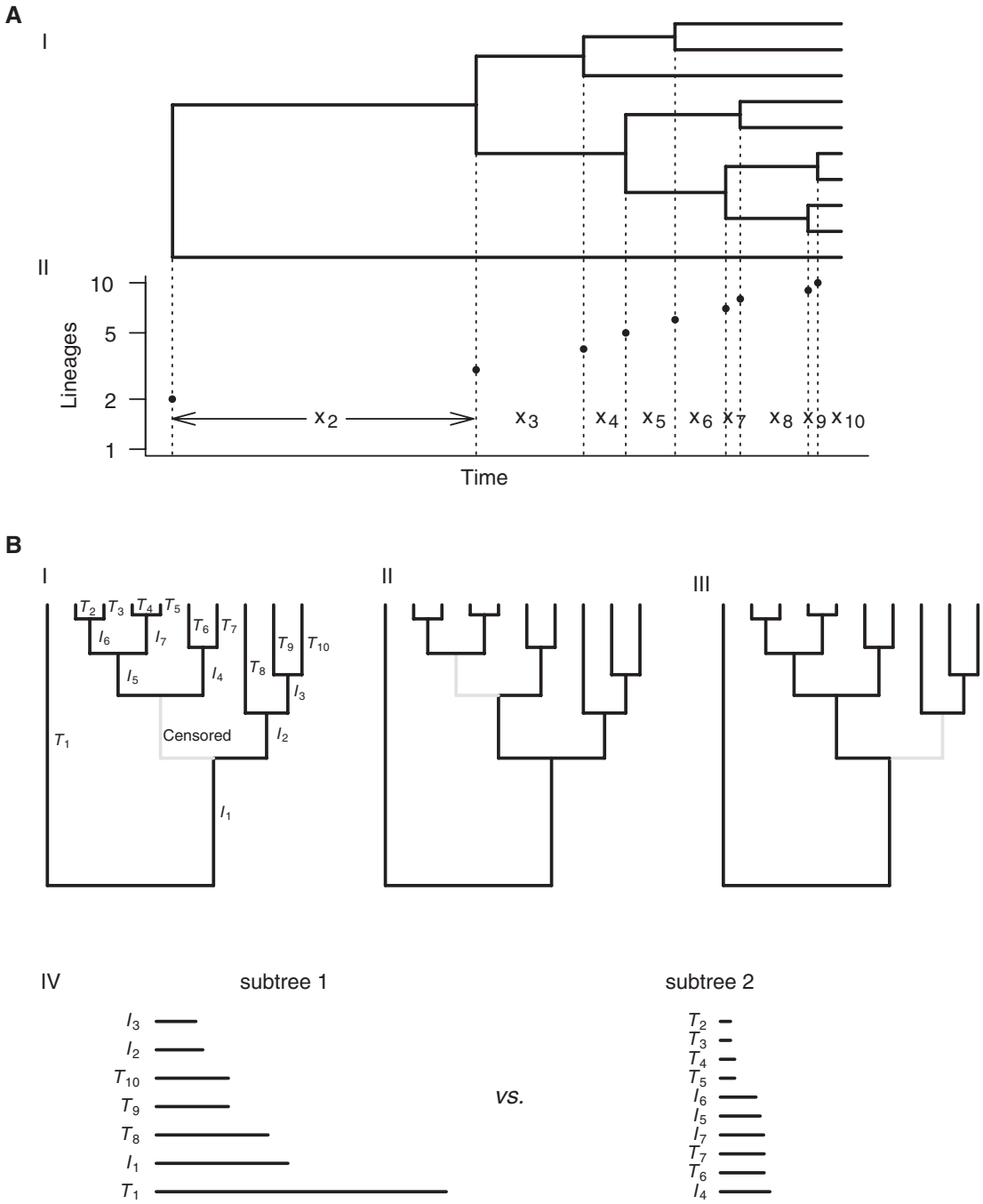

Figure 1. Hypothetical phylogenetic tree illustrating the difference in data used by variation through time methods and difference between group methods. (A_I) Hypothetical tree showing (A_{II}) timing of cladogenic events used by variation through time methods. ($B_{I–III}$) Same hypothetical phylogenetic tree showing subtrees used in difference between group methods (e.g., PRC analysis). ($B_{I–III}$) All possible subtrees for PRC analysis at a threshold of ≥ 3 branches. (B_{IV}) Branches used in PRC for comparison of rates in two subtrees indicated in (B_I).

the log-lineages through time plot, and estimate the parameters that describe the rate. Generally, a pure-birth (Yule) process is used as a null hypothesis. The pure-birth process predicts a linear increase in log-lineages through time. When there is deviation from this pure-birth process, other models, such as birth–death, density-dependence, etc., are applied to the data to find the best model that describes the accumulation of lineages (Nee et al. 1994a,b; Rabosky 2006; Morlon et al. 2010; Stadler 2011).

Other approaches have borrowed from survival analysis where the mean waiting time to cladogensis per lineage is modeled (Paradis 1997). Another commonly used approach, Pybus and Harvey's (2000) constant-rates test (commonly referred to by its test statistic, γ), does not model the rate parameter, but rather examines the shape of the distribution of ordered cladogenic events. The constant-rates test has been used to test for a slowdown in diversification or, conversely, evidence for a burst of diversification early in a group's history, however this method has limitations that compromise its power to detect rate variation early in a clade's history (Fordyce 2010b). Other recently developed approaches search for discrete shifts in patterns of diversification through time (McInnes et al. 2011; Stadler 2011). These approaches examine the entire tree (monophyletic group) as it changes through time, and do not explicitly consider rate variation among lineages within a tree. That is, they assume all lineages/subclades within a tree are characterized by the same time-dependent diversification process.

Difference between groups methods have been developed to examine rate variation within a tree and identify subclades in the tree with higher or lower relative rates. The simplest of these methods compare the species richness of two or more clades to evaluate whether differences in observed diversity are consistent with stochastic variance or are better explained by distinct diversification rates. Different rates of diversification and the processes underlying this variation can also be detected by examining tree balance, or the symmetry of a tree (Shao and Sokal 1990; Rohlf et al. 1990; Mooers and Heard 1997; Chan and Moore 2005). Some methods need not require an ultrametric tree, rather they examine the number of nodes along a path of edges in a tree to determine if punctuated evolution has occurred in a tree's history (Webster et al. 2003; Venditti et al. 2006). The MEDUSA (Modeling Evolutionary Diversification Using Stepwise Akaike Information Criterion [AIC]) approach of Alfaro et al. (2009) uses edge lengths to estimate the parameters of a birth–death process based upon the model described by Rabosky et al. (Rabosky et al. 2007) (but see Rabosky [2010]).

Here, we present a new method aimed at identifying subclades of a tree with relatively higher and lower rates of diversification, the parametric rate comparison test (hereafter, PRC). This approach explicitly examines the distribution of branch lengths, rather than the distribution of cladogenic events across the entire tree (Fig. 1), and does not require that comparisons be made among monophyletic groups. It allows for the detection and comparison of rate variation among both monophyletic and paraphyletic groups. This approach also provides the opportunity to compare diversification rate histories among a priori defined groups. This method is not contingent on a particular evolutionary process; rather, it is phenomenological in nature allowing comparison of various probability distributions. PRC differs from MEDUSA in being a generalized statistical analysis rather than fitting the specific constant birth–death branching model of Kendall (1948) (see also Paradis [2003] and Rabosky [2007]). As such, the PRC method can employ a variety of statistical distributions to characterize the distribution of branch lengths (internodes). This flexibility is the primary advantage of PRC as an addition to the comparative diversification analysis toolbox. We apply the PRC to the well-studied radiation of *Plethodon* salamanders in eastern North America using a few simple distributions as an example of the utility of this method.

Methods

PARAMETRIC RATE COMPARISON (PRC)

To identify shifts in rates of diversification along lines of common ancestry, we iteratively compare each monophyletic subtree to the remainder tree obtained by pruning out the subtree (Fig. 1). We used all of the branch lengths in each partition of the tree as data except the branch linking the subtree to the remainder tree. The linking branch was removed from estimation of diversification rates as we have no way of knowing exactly at what point the rate shift (if any) occurred along the branch length. Obviously, this decision can be made on a case-by-case basis.

Internode distances provide information regarding expected waiting times before a reconstructed split, whereas terminal branch lengths provide information regarding minimum time elapsed before a split along that branch. Therefore, we treat terminal branches as censored at the time of sampling. We jointly model internode distances and the terminal branch lengths as waiting times using the censored form of a given distribution. For an internode of length I, we calculate the probability of cladogenesis at time I as the probability density at I given a distribution. For a terminal branch of length T, we calculate the probability of no cladogenesis by time T as one minus the cumulative probability of cladogenesis by time T given a distribution.

If $\vec{I} = \{I_1, I_2, \ldots, I_{n_I}\}$ is the vector of internode distances of a tree and $\vec{T} = \{T_1, T_2, \ldots, T_{n_T}\}$ the vector of terminal branch lengths, the likelihood of a set of internodes and terminal branches is given by

$$\mathrm{Lik}(\vec{I}, \vec{T}|\mathrm{Model}) = \prod_{i=1}^{n_I} p(I_i|\mathrm{Model}) \prod_{j=1}^{n_T} \times \left(1 - \int_0^{T_j} p(y|\mathrm{Model})dy\right). \quad (1)$$

To illustrate the approach, we implemented it with the four simple distributions used by Venditti et al. (2010): exponential, Weibull, log-normal, and variable rates. These distributions are commonly used in phylogenetic studies because they make different assumptions about waiting times and therefore approximate

Table 1. Likelihood functions.

Model	Likelihood function	Rate
Exponential	$\lambda^{n_I}\exp[-\lambda(\sum_i^{n_I} I_i + \sum_j^{n_T} T_j)]$	$\hat{\lambda}$
Weibull	$k^{n_I}\lambda^{kn_I}\prod_i^{n_I} I_i^{k-1}\exp[-\lambda^k(\sum_i^{n_I} I_i^k + \sum_j^{n_T} T_j^k)]$	$\hat{\lambda}/\Gamma(1+1/\hat{k})$
Log-normal	$\frac{1}{\left(\sigma\sqrt{2\pi}\right)^{n_I}\prod_i^{n_I} I_i}\exp\left[-\frac{\sum_i^{n_I}(ln(I_i)-\mu)^2}{2\sigma^2}\right]\prod_j^{n_T}\left(1-\Phi\left(\frac{ln(T_j)-\mu}{\sigma}\right)\right)$	$\exp\left[-(\hat{\mu}+\hat{\sigma}^2/2)\right]$
Variable rates	$(\alpha\beta)^{n_I}\prod_i^{n_I}(1+I_i\beta)^{-(1+\alpha)}\prod_j^{n_T}\left(1+T_j\beta\right)^{-\alpha}$	$\hat{\beta}(\hat{\alpha}-1)$

distributions of branch lengths expected from different evolutionary processes (Gillespie 1991; Venditti et al. 2010). Table 1 lists the explicit form of the likelihood function under these particular statistical models.

For a given subtree (S_1), we use AIC corrected for sample size (AIC_c) to identify the model that best describes the distribution of branch lengths. We do the same for the remainder set of branch lengths (S_2), and for the pooled set ($\mathbb{S} = S_1 \cup S_2$). This pooled set is different for each subtree partition when the linking branch is left out, so AIC is not comparable among partitions, but ΔAIC_c is comparable among tests. For these analyses, the partition is a random variable, and therefore accounted for as a parameter in calculating AIC_c for two-distribution models (Rabosky et al., 2007; Alfaro et al., 2009).

To identify subtrees where evidence for a diversification rate difference is particularly strong, we calculate a P-value for the null hypothesis that the two sets of branch lengths (S_1, S_2) are drawn from the same statistical distribution. We use a standard likelihood ratio test comparing the joint likelihood of the two subtrees (S_1, S_2) under separate models best describing each subtree to the likelihood of the model that best fits the pooled set of branches, $\mathbb{S}$ (i.e., the model where the subtrees are constrained to have the same distribution is nested within the model where distributions can vary between subtrees). This provides an explicit test of a difference in diversification rate between a given subtree (S_1) and the remainder of the tree (S_2). P-values or ΔAIC_c can be compared across all possible partitions of a tree to identify subtrees; where we have the highest confidence there has been a shift in diversification rate.

K-CLADES PRC

The iterative method described above compares two sets of branch lengths for each partition of a tree. In principle it could be extended to compare multiple sets, but working through all possible pairs or triplets of partitions would be computationally expensive for trees of reasonable size. An alternative is to consider a priori defined set of clades or partitions (K) and ask whether they are best described by a single, or multiple groups ($1 < K' \leq K$). A similar approach is implemented in MEDUSA (Alfaro et al., 2009), except the MEDUSA method estimates the parameters of a birth–death process, whereas the PRC method, being phenomenological in nature, makes no assumptions regarding the underlying evolutionary processes responsible for generating the distribution of edge lengths. Thus, the PRC method is a more general and flexible approach, allowing for the comparison of various probability distributions fit to the sample of edge lengths. The PRC method only assumes that sets of branch lengths can be modeled as if they are drawn from a common distribution. For K-defined clades or groups, there might be up to K distinct distributions. Following the logic above, likelihood or AIC can be used to find the best clustering of K tree partitions into $K' \leq K$ groups, each characterized by a different distribution. Further, alternative values of K' can be compared using the appropriate likelihoods.

INTERPRETING RATES

For most distributions, the reciprocal of the expected waiting time to cladogenesis gives an estimate of the diversification rate. However, comparison of estimated rates for different distributions should be interpreted with caution. An increase or decrease in diversification rate in a given subtree is always relative to the diversification rate in the rest of the tree. Although PRC allows for the estimation of diversification rate of a subtree under various distributions, a direct comparison of diversification rates in different parts of the tree under different distributions might not be ideal. This is primarily due to the fact that different distributions of branch length in different parts of the tree indicate different underlying processes. For instance, if branch lengths are drawn from a Weibull distribution, the probability of observing cladogenesis changes as a branch gets longer. Thus, it might not be appropriate to compare the average rate within that subtree with the rate of a subtree described by an exponential distribution. However, when the analysis is restricted to a particular distribution or if branch lengths in different parts of the tree are drawn from the same distribution, comparing the rates might be informative regarding the magnitude of shifts in diversification rates across the tree.

It is also important to note that different distributions of branch lengths suffer from different biases. For instance, in some cases under the variable rate distribution of branch length, the maximum likelihood estimate of the α parameter might be < 1. Although, a perfectly valid distribution, the expectation of the

branch length and hence the diversification rate remains undefined for such a scenario. Thus, there might be no way of comparing mean diversification rates from different parts of the tree under this scenario.

Finally, high levels of extinction will tend to make the distribution of branch lengths deep in a tree different from the distribution of more recent branches (e.g., Hey 1992; Nee et al. 1994a; Rabosky 2010). For example, PRC, like other methods, will tend to identify recent clades as having relatively low or high diversification rates when the underlying process is actually a constant birth–death model. One way to view this is as a correct detection of different statistical distributions of branch lengths, but it would be incorrect to interpret the result as reflecting distinct patterns of diversification. Thus, if the overall pattern is one of shorter apparent waiting times toward all tips rather than shorter waiting times in one group versus another, then it might be more accurate to infer a birth–death diversification process rather than variation in rates among clades.

Given the potential pitfalls associated with comparing the distribution of branch lengths under various probability density functions, it is not necessarily appropriate to compare the absolute magnitude of rate shifts in various parts of the tree, unless drawn from the same distribution. PRC should primarily be used as an exploratory tool to identify potential regions of shifts in diversification rates on a tree and caution should be exercised in interpreting the magnitude of rate shifts.

FALSE-POSITIVE RATES

To evaluate how often the PRC method incorrectly rejects the null hypothesis of a single diversification model, we applied the method to simulated phylogenetic trees with 25, 50, 75, and 100 terminal taxa, and with extinction rates (death rate/birth rate) of 0.0, 0.1, 0.25, 0.5, 0.75, and 1.0 (the last being equivalent to a coalescent with constant population size). We simulated 1000 trees of each size $\times$ extinction rate for a total of 24,000 trees. For each tree, we ran the PRC test and recorded the number of subtrees that significantly differed from the remainder of the tree at $\alpha = 0.05$. We repeated this once assuming only exponential distributions (the expected distribution of branch lengths under a pure-birth model) and a second time exploring the fit of all four distributions (Table 1). We calculated the false-positive rate (FPR) for each tree as the proportion of *P*-values that were less than 0.05, that is, the number of statistically significant likelihood ratio tests out of the total number of permissible tests given the restriction that each partition of the tree must have ≥ 6 branches. Simulated trees are archived in DRYAD (doi:10.5061/dryad.8h350).

WHOLE-TREE RANDOMIZATION TEST

FPRs are expected to increase with increasing deviation from a pure-birth model. For example, in a constant birth–death model or density-dependent diversification model, the expected internode length of an observed tree changes over time. This violates the i.i.d. assumption when comparing a subtree and paraphyletic remainder tree because they differ in timespan and will tend to make it easier to reject the null hypothesis of identical distributions even when there is no true difference in the diversification process (see results for FPRs). Therefore, to avoid spurious inference of among-lineage rate variation, we use a tree-wide randomization test.

Under the assumption that the same probabilities of speciation and extinction hold across all concurrent lineages in a tree, the set of divergence times contains all the information needed to describe the diversification process (Stadler 2011). That is, given a set of divergence times, all topologies are equally likely under this assumption of homogeneity across lineages. In contrast, diversification is systematically greater in some lineages than others, all topologies are not equally likely. Therefore, as a whole-tree test for among-clade rate variation that will not be biased by deviations from a pure-birth process, we used a randomization test holding the set of branching times constant. Moving backwards through time, for each branching time estimate from the original chronogram, we randomly join two lineages to produce a random sample from all possible topologies. We then run the PRC procedure on each random tree to estimate the distribution of the number of false positives under the null hypothesis that the underlying speciation and extinction probabilities are constant across lineages at any given time. This randomized null distribution makes no other assumptions about the process responsible for the distribution of branching times; it might accelerate or decelerate over time, it might vary idiosyncratically through time. If the number of detections (number of subtrees that significantly differed from the remainder of the tree at $\alpha = 0.05$) or magnitude of the best supported partition (maximum ΔAIC_c) for the observed tree is large relative to this randomized distribution, one can infer that diversification is in fact more concentrated in a subset of lineages than expected.

WHY INTERNODE DISTANCES?

Internode distances are imperfect for summarizing a diversification process other than Yule's pure birth, because their expectations can vary with depth in a tree. For monophyletic groups with no among-lineage rate variation, branching times (distance from each node to the tip) provide superior information (Morlon et al. 2010; McInnes et al. 2011; Stadler 2011). But for paraphyletic groups, the distribution of branching times is distorted and misleading. For the basic birth–death process (Kendall, 1948), this problem was solved by the Rabosky et al. (2007) method implemented in MEDUSA (and the solution requires both branching times and internode lengths). Our approach is completely different: instead of trying to estimate the parameters of a

diversification model, we use the statistical properties of the distribution of internode lengths to test a simple null hypothesis that there is no difference between groups. Internode distances are imperfect, but one can characterize their statistical properties from a paraphyletic group.

DETECTION RATES

To evaluate how the PRC method performed in detecting rate variation, we applied the method to simulated trees with known rate variation. We simulated 2000 pure-birth trees with 35, 50, and 75 terminal taxa, where a monophyletic subtree composed of 25 terminal taxa arises from one constant rate, and the remaining taxa comprise a group arising from a different constant rate. Because the relative rates of these two subtrees was known, we were able to assess the power of the PRC method to detect rate variation. We applied the PRC method both restricting the model of branch lengths to an exponential distribution, as well as exploring the fit of all four distributions simultaneously (Table 1). Successful detection was determined by AIC_c at the node where rate variation was simulated. Simulated trees are archived in DRYAD (doi:10.5061/dryad.8h350).

APPLICATION TO A SALAMANDER RADIATION

To illustrate the method with an empirical dataset, we used a chronogram of eastern North American *Plethodon* salamanders with virtually complete taxon sampling (Kozak et al. 2006, 2009). This group includes what appears to be a relatively recent radiation of large-bodied salamanders, the *Plethodon glutinosus* group and a paraphyletic group of small, slender salamanders. Kozak et al. (2006) used the nonparametric relative cladogenesis (rc) test (Nee et al., 1992) and lineage-through-time analysis to infer that the large-bodied *P. glutinosus* clade had a higher diversification rate than the small-bodied lineages. We applied our PRC method to the eastern *Plethodon* clade using the chronogram of Kozak et al. (2009), kindly supplied by K. Kozak. We also used our *K*-clades algorithm to evaluate whether the three monophyletic groups (*P. glutinosus, P. welleri/wehrlei*, and *P. cinereus* groups) were best described as having one, two, or three distinct diversification models.

Results and Discussion

We present a new method (PRC) for evaluating differences in diversification rates between groups of lineages in a phylogeny. PRC fills a gap left by existing methods because it explicitly tests each monophyletic group for an accelerated or decelerated diversification rate and estimates parameters of statistical distributions describing waiting times between cladogenesis events along lineages in each partition of the tree. The method can also be used to test a priori hypotheses of clade-specific differences in diversification and evaluate how many different distributions best describe a set of clades. The method has a low FPR and good power to detect rate variation.

Table 2. Mean false-positive rates for the PRC test as a function of extinction rate (columns), tree size (rows), and whether the algorithm is restricted to exponential or consider all four models (e or a).

	0	0.1	0.25	0.50	0.75	1
e25	0.010	0.015	0.025	0.046	0.085	0.127
a25	0.042	0.045	0.061	0.086	0.123	0.151
e50	0.041	0.016	0.022	0.042	0.090	0.201
a50	0.044	0.045	0.058	0.078	0.119	0.167
e75	0.012	0.014	0.019	0.040	0.090	0.229
a75	0.042	0.047	0.055	0.077	0.120	0.185
e100	0.012	0.014	0.023	0.039	0.090	0.262
a100	0.044	0.047	0.057	0.076	0.119	0.188

FALSE-POSITIVE RATES

For each simulated tree with no among-lineage rate variation, we calculated a FPR as the proportion of subtrees with apparently significant (likelihood ratio test with $P < 0.05$) differences in diversification rate from the remainder tree. When the algorithm applied all four models, the FPR was below 0.05 for all tree sizes as long as the extinction rate was less than about 50% of the speciation rate (Table 2). When the algorithm was restricted to exponential distributions, the FPR was below 0.05 when the extinction rate was 75% or less. Overall, there was surprisingly little dependence of FPR on tree size, but it did increase to about 0.20–0.25 when extinction was very high (Table 2).

Particularly because the tests within a given tree are not independent, certain trees can have very large numbers of false positives owing to the stochasticity of the branching process. This raises an important general caveat for exploratory analyses of phylogenetic trees. Assessment of FPRs is rare among papers proposing methods for detecting rate variation in phylogenies, but investigators should be aware that the contingent and nonindependent nature of evolutionary branching can cause statistical outliers to be concentrated in particular datasets. The interpretation of a statistical pattern as evidence of an interesting evolutionary process (in this case, a true shift in the probability of cladogenesis) is always strongest when coupled with prior, independent information on natural history.

The whole-tree randomization test had FPRs close to 5% (0.028–0.064) with little effect of tree size, extinction rate, or whether all four distributions were tested (Table 3). Therefore, Type I error rates will be acceptable if the randomization test is used first as a tree-wide evaluation of the null hypothesis of

Table 3. False-positive rates for the whole-tree randomization test as a function of extinction rate (columns), tree size (rows), and whether the algorithm is restricted to exponential or consider all four models (e or a).

	0	0.1	0.25	0.50	0.75	1
e25	0.028	0.037	0.042	0.041	0.041	0.043
a25	0.043	0.037	0.049	0.042	0.051	0.054
e50	0.058	0.038	0.044	0.043	0.044	0.046
a50	0.051	0.041	0.045	0.047	0.046	0.041
e75	0.053	0.041	0.042	0.042	0.046	0.057
a75	0.055	0.051	0.047	0.045	0.046	0.062
e100	0.050	0.049	0.062	0.055	0.042	0.064
a100	0.055	0.047	0.058	0.054	0.046	0.058

no difference in diversification rates among groups. If this null hypothesis is rejected, PRC can be used to identify partitions with strong evidence for faster or slower diversification.

DETECTION RATES

The PRC method was largely able to detect rate variation when the relative diversification rate within the tree differed roughly twofold (Fig. 2). The power to detect rate variation increased as the size of the tree increased. This increase in power as a function of tree size might be expected, as larger trees provide more data (branches) for fitting distributions resulting in lowered error around a given distributions' parameter estimates. Power was decreased by including more possible distributions in the fitting algorithm, presumably because of increased flexibility in fitting a single distribution to full set of internodes.

APPLICATION TO A SALAMANDER RADIATION

We used our lineage-based PRC method to analyze the chronogram of eastern *Plethodon* salamanders (Kozak et al. 2006, 2009). The tree includes 28 tip taxa in the *P. glutinosus* radiation and 16 tip taxa comprising the remainder of the eastern *Plethodon* clade. The tree topology is as found by Kozak et al. (2009) based on mitochondrial and three nuclear sequences. The relative cladogenesis test (without Bonferroni correction) flagged seven nodes within the large-bodied clade as having significantly more descendants than expected (Fig. 3A). MEDUSA, as implemented in the R package geiger (Harmon et al. 2008), found no support for any rate shift if the stem branch was excluded from the ingroup (cutAtStem = FALSE) and weak support ($\Delta AIC_c = 1.27$) for a shift within the *P. glutinosus* group if the stem branch was included (the node marked "m" in Fig. 3A). A whole-tree randomization test with MEDUSA run on each of 1000 randomized topologies never found ΔAIC_c greater than or equal to the observed ($P < 0.001$).

For the PRC, the whole-tree randomization test favored rejecting the null hypothesis of no among-lineage rate variation ($P = 0.034$ based on P-values, and $P = 0.010$ based on ΔAIC_c; if the algorithm was limited to exponential distributions, $P < 0.001$ based on either P-values or ΔAIC_c). PRC suggested that the clade of large-bodied *Plethodon* including *P. ocmulgee* and *P. montanus* has diversified more rapidly than the paraphyletic group including the rest of the *P. glutinosus* group and the small-bodied forms (Fig. 3B). The evidence for this shift was much stronger from the PRC ($\Delta AIC_c = 5.48$) than from MEDUSA ($\Delta AIC_c = 1.27$). If we consider all nodes where a rate difference was favored by AIC_c, there is also an indication that the small-bodied *P. welleri/P.wehrlei* group

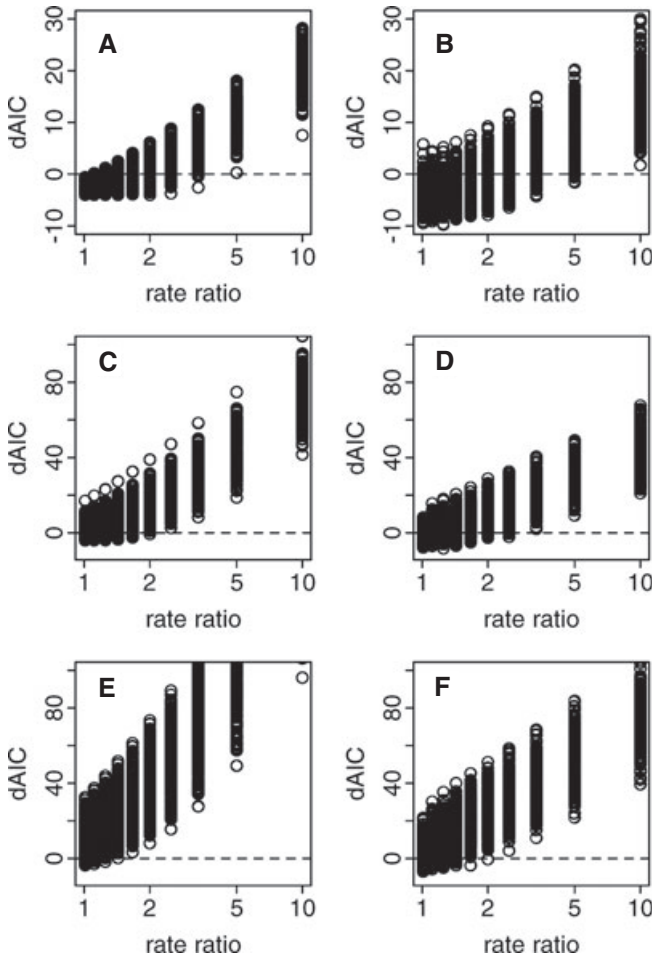

Figure 2. Detection of rate difference between a subtree and remainder tree as a function of known relative diversification rates (rate ratio is the pure-birth rate for the monophyletic subtree of 25 taxa over the base rate). dAIC is the difference in AIC_c between the single-rate and two-rate models. (A, C, and E) show results from fitting only exponential distributions and (B, D, and F) show results from fitting the best of the four distributions. Tree sizes were 35 (A and B), 50 (C and D), and 75 tips (E and F). The dashed horizontal line illustrates the detection threshold (AIC_c lower for the two-rate than one-rate model. A total of 250 simulations were performed for each tree size and rate ratio.

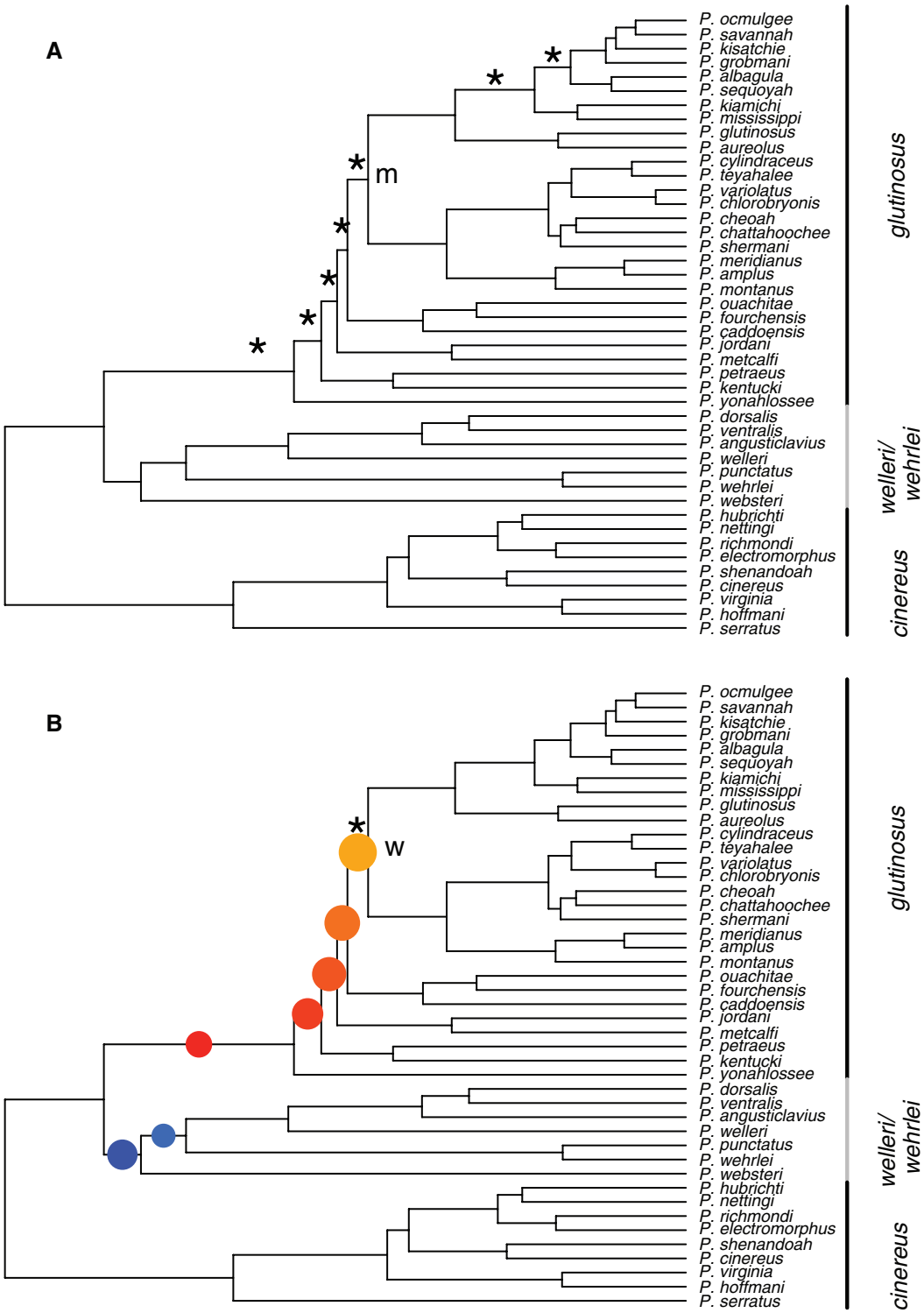

Figure 3. Application of PRC to the radiation of *Plethodon* salamanders in eastern North America. (A) Clades with significantly ($\alpha = 0.05$) higher diversification rates according to the nonparametric relative cladogenesis test (Nee et al. 1992) are indicated by asterisks. (B) Clades with higher or lower diversification rates according to PRC are indicated by hot or cold colors. Asterisks mark statistically significant comparisons. The best fit model was exponential for all subtrees except that marked by "w," which was best fit by a Weibull.

Table 4. ***K*** **-clades PRC test results fitting an exponential distribution to branch lengths. Clades are identified as follows: (1)** ***Plethodon cinereus*** **clade, (2)** ***P. welleri/P. wehrlei*** **clade, and (3)** ***P. glutinosus*** **clade.**

k	Groups	Rate 1	Rate 2	Rate 3	Parameters	LLik	AIC_c	ΔAIC_c
1	1 2 3	0.095			1	−127.328	256.766	2.226
2	1 2 vs 3	0.062	0.127		2	−125.099	254.540	0.000
2	1 3 vs 2	0.112	0.048		2	−125.488	255.320	0.779
2	2 3 vs 1	0.100	0.078		2	−127.128	258.598	4.048
3	1 vs 2 vs 3	0.078	0.048	0.127	3	−124.770	256.246	1.706

diversified more slowly than the composite paraphyletic group including the large-bodied radiation and the *P. cinereus* group (Fig. 3B).

In addition to an increase in diversification rate in the large *Plethodon* group, the strongest pattern is a switch from exponential to Weibull in the large clade including *P. ocmulgee* and *P. montanus* (the node marked "w" in Fig. 3B). Although we stress that the analysis is phenomenological, this qualitative difference suggests a biological difference in mode as well as tempo of diversification in the clade. The estimated shape parameter of the Weibull is 1.46, indicating that longer branches are more likely to split than short ones. This is consistent with the fact that the clade is very balanced, with each long, internal branch ending in a crown group of recently derived lineages. There is no biological mechanism for lineages to "age" or "remember" how long it has been since they originated, so it is unlikely that the diversification process changes as branches go longer without splitting. However, the pattern of long internal branches and short branches toward the tips tends to be generated by a consistently high extinction rate. Even so, a pure-birth model could not be rejected in favor of a constant birth–death model (Kendall, 1948) using the method of Nee et al. (1994b).

The *K*-clades analysis applied to the three groups (*P. glutinosus*, *P. welleri/P.wehrlei*, and *P. cinereus* clades) rejects a single rate model and favors a two-rate model with the large-bodied *P. glutinosus* group diversifying faster than the small bodied clades. However, a two-rate model with a decrease in diversification rate in the *P. welleri/P.wehrlei* group has similar support and a three-rate model including both patterns is also within two units of AIC (Table 4).

Judging by this example our PRC tests are slightly more conservative than the rc test and more powerful than MEDUSA. Moreover, PRC and *K*-clades analyses reveal more about the mode and tempo of diversification than does the nonparametric rc test. In particular, our results imply that variation in diversity among groups of *Plethodon* might be caused by accelerated diversification in the large-bodied clade, decreased diversification in the *P. welleri/P.wehrlei* clade, or both. In addition, the change from exponential to Weibull within the large-bodied clade indicates deviation from a pure-birth model, possibly reflecting an increase in the importance of extinction.

Conclusion

PRC provides a novel approach for exploratory data analysis of rate variation within a phylogenetic tree. We emphasize that the method can be implemented with any distribution appropriate for waiting times, not just those illustrated here. The identification of subclades in a tree showing relative rate heterogeneity can be useful for informing evolutionary hypotheses. For example, it might be interesting to know if cladogenic rates are correlated with rates of morphological evolution. Likewise, if the appropriate data exist, it might be interesting to examine whether rate variation is associated with the origin of evolutionary novelty. PRC provides statistical power similar to the nonparametric relative cladogensis test, but allows for more nuanced interpretation. In the worked example with *Plethodon* salamanders, PRC not only identifies the *P. glutinosus* group as having a high diversification rate, but also indicates a qualitative difference as the branch lengths in the group are better fit by a Weibull distribution.

PRC also provides a framework for explicitly comparing rates among a priori defined clades. Applying the *K*-clades PRC to the *Plethodon* example revealed a competing hypothesis: rate variation in the genus might be explained by a decrease in diversification in the *P. wehrlei* group rather than an increase in the *P. glutinosus* group. This approach might also be useful if there is, for example, a phylogenetic component to variation in a group's habitat preference, life-history, or geographic range. Moreover, it provides a tool to compare rates of cladogenesis among otherwise disparate groups, such as between parasite and host. Unlike variation through time, or lineages-through-time, approaches, the PRC method is not restricted to analyzing the sum total and accumulation of cladogeneic events across the tree. Rather, it models the distribution of branch lengths within subtrees of the tree, and does not require that a given subtree necessarily be monophyletic (i.e., paraphyletic and even polyphyletic groups are appropriate for the analysis). The method is robust to incomplete taxon sampling

provided that taxon sampling is random. As with any methods comparing diversity or rates of diversification, nonrandom sampling can bias inference (Cusimano and Renner, 2010; Brock et al., 2011). The PRC test can be implemented using the package *iteRates* in the R statistical computing environment (R Development Core Team, 2011).

ACKNOWLEDGMENTS

We thank B. O'Meara, C. Hulsey, and T. Near for helpful comments and discussion. K. Kozak kindly provided the *Plethodon* chronogram. This manuscript was improved with the helpful comments of C. Boettiger and L. Harmon. This work was supported in part by the University of Tennessee, a graduate assistantship to PS from the National Institute for Mathematical and Biological Synthesis, and US National Science Foundation (DEB 0614223 and 1050947). PS acknowledges support from a Burroughs Wellcome Fund Career Award and David and Lucille Packard Foundation Fellowship awarded to Joshua B. Plotkin.

LITERATURE CITED

Alfaro, M. E., F. Santini, C. Brock, H. Alamillo, A. Dornburg, D. L. Rabosky, G. Carnevale, and L. J. Harmon, 2009. Nine exceptional radiations plus high turnover explain species diversity in jawed vertebrates. Proc. Natl. Acad. Sci. 106:13410–13414.

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

Chan, K. M. A., and B. R. Moore, 2005. SymmeTREE: whole-tree analysis of differential diversification rates. Bioinformatics 21:1709–1710.

Cusimano, N., and S. S. Renner, 2010. Slowdowns in diversification rates from real phylogenies may not be real. Syst. Biol. 59:458–464.

Fordyce, J. A., 2010a. Host shifts and evolutionary radiations of butterflies. Proc. Roy. Soc. B 277:3735–3743.

———. 2010b. Interpreting the γ statistic in phylogenetic diversification rate studies: a rate decrease does not necessarily indicate an early burst. PLoS ONE 5:e11781.

Gillespie, D. J. H., 1991. The causes of molecular evolution. Oxford Univ. Press, New York.

Harmon, L. J., 2003. Tempo and mode of evolutionary radiation in iguanian lizards. Science 301:961–964.

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

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

Jiggins, C. D., R. Mallarino, K. R. Willmott, and E. Bermingham, 2006. The phylogenetic pattern of speciation and wing pattern change in neotropical Ithomia butterflies (Lepidoptera: Nymphalidae). Evolution 60:1454–1466.

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. Biol. Sci. 273:539–546.

Kozak, K. H., R. W. Mendyk, and J. J. Wiens, 2009. Can parallel diversification occur in sympatry? Repeated patterns of body-size evolution in coexisting clades of north american salamanders. Evolution 63:1769–84.

McInnes, I., C. D. L. Orme, and A. Purvis, 2011. Detecting shifts in diversity limits from molecular phylogenies: what can we know? Proc. Roy. Soc. B 278:3294–3302.

Mooers, A. Ø., and S. B. Heard, 1997. Inferring evolutionary process from phylogenetic tree shape. Quar. Rev. Biol. 72:31–54.

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

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

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

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

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

———. 2003. Analysis of diversification: combining phylogenetic and taxonomic data. Proc. Roy. Soc. B 270:2499–2505.

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

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

R Development Core Team, 2011. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available at http://www.R-project.org. ISBN 3-900051-07-0.

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

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

Rabosky, D., L. S. Donnellan, A. L. Talaba, and I. J. Lovette, 2007. Exceptional among-lineage variation in diversification rates during the radiation of Australia's most diverse vertebrate clade. Proc. Roy. Soc. B 274:2915–2923.

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

Rohlf, F., W. Chang, and R. Sokal, 1990. Accuracy of estimated phylogenies: effects of tree topology and evolutionary model. Evolution 44:1671–1684.

Shao, K., and R. R. Sokal, 1990. Tree balance. Syst. Biol. 39:266–276.

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

Venditti, C., A. Meade, and M. Pagel, 2006. Detecting the node density artifact in phylogeny reconstruction. Syst. Biol. 55:637–643.

Venditti, C., A. Meade, and M. Pagel, 2010. Phylogenies reveal new interpretation of speciation and the Red Queen. Nature 463:349–352.

Webster, A. J., R. J. H. Payne, and M. Pagel, 2003. Molecular phylogenies link rates of evolution and speciation. Science 301:478.

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

Associate Editor: M. Alfaro