Syst. Biol. 0(0):1–12, 2022

https://doi.org/10.1093/sysbio/syac035
Advance Access Publication 16 May 2022

Fast Likelihood Calculations for Automatic Identification of Macroevolutionary Rate Heterogeneity in Continuous and Discrete Traits

MICHAEL C. GRUNDLER[1,*], DANIEL L. RABOSKY[2], AND FELIPE ZAPATA[1]

[1]*Department of Ecology and Evolutionary Biology, University of California, Los Angeles, CA 90095, USA and* [2]*Museum of Zoology and Department of Ecology and Evolutionary Biology, University of Michigan, Ann Arbor, MI 48109, USA*
[*]*Correspondence to be sent to: Department of Ecology and Evolutionary Biology, University of California, 612 Charles E. Young Dr. South, Los Angeles, CA 90095, USA;*
E-mail: mgru@ucla.edu.

Received 29 September 2021; reviews returned 28 April 2022; accepted 6 May 2022
Associate Editor: Josef Uyeda

Abstract.—Understanding phenotypic disparity across the tree of life requires identifying where and when evolutionary rates change on phylogeny. A primary methodological challenge in macroevolution is therefore to develop methods for accurate inference of among-lineage variation in rates of phenotypic evolution. Here, we describe a method for inferring among-lineage evolutionary rate heterogeneity in both continuous and discrete traits. The method assumes that the present-day distribution of a trait is shaped by a variable-rate process arising from a mixture of constant-rate processes and uses a single-pass tree traversal algorithm to estimate branch-specific evolutionary rates. By employing dynamic programming optimization techniques and approximate maximum likelihood estimators where appropriate, our method permits rapid exploration of the tempo and mode of phenotypic evolution. Simulations indicate that the method reconstructs rates of trait evolution with high accuracy. Application of the method to data sets on squamate reptile reproduction and turtle body size recovers patterns of rate heterogeneity identified by previous studies but with computational costs reduced by many orders of magnitude. Our results expand the set of tools available for detecting macroevolutionary rate heterogeneity and point to the utility of fast, approximate methods for studying large-scale biodiversity dynamics. [Brownian motion; continuous characters; discrete characters; macroevolution; Markov process; rate heterogeneity.]

The uneven distribution of phenotypic diversity among different clades of organisms is one of the most prominent large-scale biodiversity patterns of life on earth. Identifying where and when evolutionary rates change on phylogeny is a key step in understanding the causes underlying this variation, as both ecological and developmental controls on clade-level differences in phenotypic disparity predict that evolutionary tempo and mode vary over phylogeny in response to ecological opportunity and developmental constraints or enablers (Christin et al. 2013; Stroud and Losos 2016). Quantifying among-lineage variation in rates of phenotypic evolution is therefore a primary methodological challenge in macroevolution.

The growth of methods for modeling evolutionary rate heterogeneity has seen substantial progress in both continuous (O'Meara et al. 2006; Revell and Collar 2009; Eastman et al. 2011; Revell et al. 2012; Thomas and Freckleton 2012; Landis et al. 2013; Rabosky et al. 2014; Mitov et al. 2019; Fisher et al. 2021b; Revell 2021) and discrete traits (Lloyd et al. 2012; Marazzi et al. 2012; Beaulieu et al. 2013; Zanne et al. 2014; Davis Rabosky et al. 2016; Fisher et al. 2021a). Broadly, these methods fall into two classes: supervised approaches, which require an a priori hypothesis of phylogenetic rate variation, and unsupervised approaches, which automatically learn the distribution of phylogenetic rate variation from data. Methodological advances in both types of approaches have made describing patterns of among-lineage evolutionary rate heterogeneity relatively routine, and this in turn has enabled researchers to investigate mechanisms underlying broad-scale differences in phenotypic disparity such as ecological opportunity (Alhajeri et al. 2016; Price et al. 2016), phenotypic integration and modularity (Goswami et al. 2014; Watanabe et al. 2019), and the origin of key innovations (Werner et al. 2014; Barua and Mikheyev 2020; Simões et al. 2020).

Despite much methodological progress, however, there are still a number of limitations that hinder investigations of phylogenetic variation in tempo and mode. Supervised approaches are useful for targeted hypothesis testing (e.g., Butler and King 2004; O'Meara et al. 2006; Lloyd et al. 2012), but the large number of potential hypotheses discourages their use in large comparative data sets where phylogenetic variation in tempo and mode is most likely. Unsupervised approaches have the ability to automatically evaluate many possible hypotheses of phylogenetic rate variation (e.g., Eastman et al. 2011; Uyeda and Harmon 2014; Grundler and Rabosky 2020), but the frequent use of Bayesian techniques typically requires long computation times. Moreover, methods suitable for organismal traits that can only be represented as categorical variables have primarily been developed for use in phylogenetic inference, typically in a Bayesian context (e.g., Drummond and Suchard 2010), and have seen limited application in macroevolutionary studies (but see Beaulieu et al. 2013; King and Lee 2015a; Davis Rabosky et al. 2016). Compared with continuous phenotypes, therefore, there is a relative shortage of unsupervised methods available for discrete traits. As a result, a substantial amount organismal trait complexity remains underexplored by studies seeking to identify phylogenetic variation in tempo and mode.

In this article, we describe a likelihood-based, unsupervised method for inferring evolutionary rate heterogeneity that can be applied to both discrete and

continuous organismal traits. The method is similar to many related methods in that the general approach assumes the present-day distribution of trait is shaped by a variable-rate process that arises from a mixture of constant-rate processes. Unlike most related methods, however, our maximum likelihood approach does not require long running, computationally intensive Bayesian techniques to detect rate variation. By employing dynamic programming optimization techniques and closed-form (sometimes approximate) maximum likelihood estimators, our approach is capable of evaluating the phylogenetic distribution of many possible rate-shift configurations with a single tree traversal and permits rapid exploration of data sets for variation in the tempo and mode of phenotypic evolution. Using simulations, we show that the method reconstructs rates of trait evolution with high accuracy, particularly in large phylogenies where the evolutionary process is most likely to vary among lineages in empirical data sets. When we apply the method to empirical data sets on turtle body size and squamate reproductive mode, we achieve order-of-magnitude speedups recovering patterns of rate variation nearly identical to those inferred from much more computationally intensive Bayesian analyses. Overall, our results help to expand the set of tools available for detecting macroevolutionary rate heterogeneity and point to the utility of fast, approximate methods for studying large-scale biodiversity dynamics.

Materials and Methods

Notations

We let T be a rooted binary tree and we write $|T|$ for the number of nodes in T and $\rho(T)$ for the root of T. For all nodes $u \in T$, we write t_u for the length of branch ending at u, a_u for the parent of u, v_u for the left child of u, w_u for the right child of u, and T_u for the subtree of T rooted at u (excluding the branch leading to u). For any subset of nodes $M = \{m_1, m_2, \ldots, m_N\}$, we write $T_{\neg M}$ to denote the induced subtree that results from pruning the set of nodes spanned by the subtrees $T_{m_1}, \ldots, T_{m_N}$. The set of terminal nodes in T is represented by Λ_T and we let X represent a map from Λ_T to the set of real numbers (for a continuous character) or to a finite set of integers (for a discrete character), so that the values of $X(\Lambda_T) = \{X(u) | u \in \Lambda_T\}$ are the character states at the tips of T. Finally, we write $L(\theta_T | X(\Lambda_T))$ for the log likelihood that $X(\Lambda_T)$ was generated by a process with rate θ_T. For the moment we postpone specification of the exact form of L.

General Approach

Our general approach derives from Kapli et al. (2017) and assumes there is a background constant-rate process of character evolution that experiences a series of shifts to a new independent constant-rate process, each such process having a distinct rate. We further assume that each rate-shift occurs at the crown-node of a clade, and we disallow nested rate-shifts. Importantly, this constraint means that the model cannot fully capture the rate heterogeneity that might be present in an empirical phylogeny because it imposes a limit on the number and complexity of rate-shift configurations. However, we will describe a way to mitigate this constraint in our presentation of the algorithm used to fit the model in the next section.

Given a set of internal nodes $S_T = \{s_1, s_2, \ldots, s_J\}$ that represent rate-shift locations, the function we seek to maximize is

$$L\left(\theta_{T_{\neg S_T}} | X\left(\Lambda_{T_{\neg S_T}}\right)\right) + \sum_{j=1}^{J} L\left(\theta_{T_{s_j}} | X\left(\Lambda_{T_{s_j}}\right)\right). \quad (1)$$

The left term represents the log likelihood of the background process whereas the right term is the log likelihood of all rate-shift processes inside T. Although not explicit in equation (1), calculating the likelihood of the background process requires conditioning on the character states of the subset of nodes S_T that are internal nodes of the full phylogeny T but that are terminal nodes in the induced subtree $T_{\neg S_T}$. In effect, we are decomposing the original tree into $J+1$ subtrees by pruning the clades subtended by rate-shifts and treating the crown-node roots of the pruned clades as terminal nodes in the induced subtree governed by the background process. For each rate-shift, we therefore estimate two parameters. In addition to the rate of the process, we also estimate the state at the root of the process. This unobserved state is then treated as known for the purposes of estimating the rate of the background process. Note that these ancestral state estimates are not maximum likelihood estimates but locally parsimonious estimates computed via the methods of Fitch (1971) and Felsenstein (1973) for discrete and continuous characters, respectively.

A Dynamic Programming Algorithm

Because there is no simple solution to directly maximize equation (1), we opt for a greedy dynamic programming approach modified from Kapli et al. (2017). The basic algorithm requires only a single pass over the phylogeny, visiting each internal node of T in a postorder traversal, and is illustrated in Figure 1. For all nodes $u \in T$ we maintain an array of entries $f_{u,i}$ for $0 \leq i \leq |T_u|$ that contain the maximization of equation (1) under the assumption that i branches in T_u belong to the background rate process and using only character state values in $X(\Lambda_{T_u})$. Note that the entries recorded for each node do not depend on the traversal order of the node's subtree. For entry $i = 0$ the clade T_u is part of a rate-shift process and for entry $i = |T_u|$ the clade T_u is part of the background process. The maximization of equation (1) for entries $i > 0$ is performed by considering all combinations of j and k such that $i = j + k + 2$ and using information stored in the corresponding entries for $f_{v_u,j}$ and $f_{w_u,k}$. The plus two

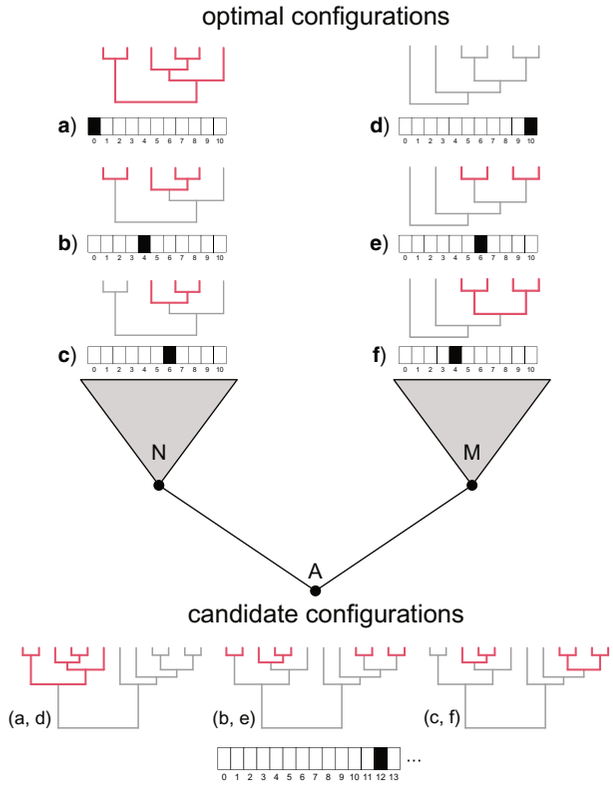

FIGURE 1. Illustration of the dynamic programming algorithm used for fitting multirate processes of character evolution. Each entry in a node's array contains a configuration of rate-shift processes (red branches) on a background process (gray branches). Array indices indicate the number of branches in the node's subtree that belong to the background process. To fill in an array entry for an ancestral node, the algorithm looks at all valid combinations of array entries in the node's immediate descendants and chooses the combination that confers the highest likelihood on the character state data in the node's subtree. A valid combination is any pair of descendant indices whose sum added with two equals the ancestral index. In the current illustration, pairs (a, d), (b, e), and (c, f) are a few of the valid candidate combinations of entries in the (previously optimized) N and M arrays for optimizing entry (12) of the A array. A color version of this figure appears in the online version of this article.

term accounts for the fact that, as we move down T toward the root, we add two branches to the background process for every node visited. Each entry $f_{u,i}$ stores the sufficient statistics needed for maximizing the log likelihood of the background process as well as the log likelihood of all character state data generated by rate-shift processes inside T_u, which is obtained by summing the corresponding log likelihoods in entries $f_{v_u,j}$ and $f_{w_u,k}$.

More concretely, consider a single node u and its left (v_u) and right (w_u) child nodes. To fill in an array entry for an ancestral node, the algorithm looks at all valid combinations of array entries in the node's immediate descendants and chooses the combination that confers the highest likelihood on the character state data in the node's subtree. A valid combination is any pair of descendant indices whose sum added with two equals the ancestral index. Thus, the score for entry $i>0$ is computed as

$$f_{u,i}(Q) = \max_{\substack{j,k: \\ i=j+k+2}} f_{u,i}(A) + f_{v_u,j}(B) + f_{w_u,k}(B),$$

where $f_{u,i}(A)$ is the maximized log likelihood of the background process rooted at u(the left term in equation (1)) and $f_{u,i}(B) = f_{v_u,j}(B) + f_{w_u,k}(B)$ is the maximized log likelihood of all rate-shift processes inside T_u (the right term in equation (1)). For $i=0$ we set $f_{u,i}(Q) = f_{u,i}(A) = f_{u,i}(B)$ because T_u is assumed to evolve under a single process in this case. If u is a terminal node, we set $f_{u,i}(Q) = f_{u,i}(A) = f_{u,i}(B) = 0$.

When the traversal finishes, the algorithm will have recorded $|T|/2$ entries for the root array. One of these entries corresponds to a constant-rate process with no rate-shifts, whereas the others are distinguished by the number and configuration of rate-shifts. The relative likelihood of different configurations can be compared using Akaike weights, and these weights can be used to assign each edge in T a model-averaged rate. Model-averaged branch rates can be used to partially mitigate the constraint that each individual configuration disallows nested rate-shifts because average rates are capable of reconstructing patterns of phylogenetic rate variation consistent with a true pattern of nested rate-shifts when the rate processes leave distinct signatures in comparative data. Nonetheless, the non-nested constraint imposes a limit on the complexity of phylogenetic rate variation that the method is capable of discovering and model-averaged branch rates may miss finer scale phylogenetic rate variation that lies beyond this limit.

Likelihood Equations

So far, the discussion has avoided specifying the forms of the log likelihood function and therefore any indication of how to perform the maximization required in the previous section. When the data are continuous we assume that the evolutionary process is a Brownian motion. In that case, the log likelihood can be computed using Felsenstein's (1973) independent contrasts procedure, which leads to closed-form maximum likelihood parameter estimates reviewed in Freckleton (2012). The sufficient statistics necessary for performing these estimates can be easily tracked while computing the contrasts. When the data are discrete we assume that the evolutionary process is a fully symmetric Markov process. Although the log likelihood can be computed using Felsenstein's (1981) peeling algorithm, this approach does not admit a closed-form maximum likelihood estimate for the rate of the process. However, if we are willing to assume that the rate of the process θ_T is small—that is, that the expected number of character state changes on any given branch in T is typically much less than one—we can derive an approximate log likelihood function that admits a closed-form maximum likelihood estimate for θ_T. We show in the appendix

that this approximation has the form (up to an additive constant)

$$L\left(\theta_T|X\left(\Lambda_T\right)\right)=N_X\log\theta_T-\theta_T\sum_{u\in T}t_u, \quad (2)$$

where N_X is the minimum number of character state changes needed to explain $X(\Lambda_T)$. This leads to the (approximate) maximum likelihood estimate

$$\hat{\theta}_T=\frac{N_X}{\sum_{u\in T}t_u}.$$

The value of N_X is easily computed as we move down the tree toward the root using Fitch's (1971) maximum parsimony algorithm. Although this estimator is expected to have negative bias (especially at high rates) because it assumes all character state changes are visible, it is monotonically related to and highly correlated with true rates and maximum likelihood rate estimates ($r=0.95$; Supplementary Fig. S17 available on Dryad at http://dx.doi.org/10.5061/dryad.wpzgmsbph).

Simulation Study

To assess how well the new method estimates branch-specific rates of evolution, we simulated trait evolution on empirical phylogenies under different rate-shift regimes. To carry out the simulations, we generated 5000 phylogenies by randomly sampling internal nodes from the genes-only ultrametric squamate reptile phylogeny of Tonini et al. (2016). Nodes were assigned weights such that all sizes (measured as the number of living descendants) of extracted clades had an equal probability of being selected. We imposed a minimum clade size of 100 for simulations. We chose to select subsets of a large empirical phylogeny, rather than simulated phylogenies, to introduce more realistic distributions of branch lengths than might be obtained using simple tree simulation models (e.g., Yule or constant-rate birth–death models).

For each phylogeny, we determined the number of rate-shift events to place on the tree by drawing a random integer from a Poisson distribution with a mean of 3. This ensured that most simulations would be variable-rate but that enough would be constant-rate to allow us to assess the method's false-positive rate. This also ensures that the shift-to-tip ratio is relatively low, a point to which we will return in our Discussion section. For variable-rate simulations, we determined the locations of rate-shifts by selecting internal nodes randomly without replacement, again using weights that gave all sizes of subtrees an equal probability of being chosen. Clades chosen as rate-shift locations were constrained to be no smaller than 10% of the total tree size to avoid biasing simulations with numerous rate-shifts that would be undetectable due to their small size. Because rate-shift locations were sampled independently with respect to one another simulations contained both nested and non-nested rate processes.

We used a two-step, double randomization procedure to choose the rate of evolution for each rate-shift. For each rate-shift, we first sampled a lower bound on the phylogenetic signal produced by the rate process, and in a second step, we sampled a rate of evolution that satisfied that lower bound. The two-step approach ensures that all simulations, regardless of differences in underlying phylogenies, are sampling evolutionary rates and the bound on phylogenetic signal produced by those rates from a uniform distribution. For discrete trait simulations, we sampled a lower bound for phylogenetic signal uniformly between 0 and 1, defined as the minimum probability of no character-state change event occurring on any given branch in the clade subtended by the rate-shift event. In other words, all descendant nodes must have a probability of retaining the character state of their ancestral node that is at least as large as this number. This condition defines a maximum allowable rate for the process, and we then chose a number uniformly between 0 and this maximum to serve as the overall transition rate for a particular rate-shift event. Similarly, for continuous trait simulations, we sampled a lower bound for phylogenetic signal uniformly between 0 and 1, defined as the minimum probability *density* of no character-state displacement occurring on any given branch in the clade subtended by the rate-shift event (i.e., the height of the Gaussian transition kernel at 0). This condition again defines a maximum allowable rate for the process, and we then chose a number uniformly between 0 and this maximum to serve as the overall rate for a particular rate-shift event.

We conducted two additional sets of simulations to assess robustness of the method to violations of the assumptions it makes about the evolutionary process. First, because the method assumes discrete characters evolve under a fully symmetric Markov process, we carried out a set of simulations with asymmetric transition rates so that the equilibrium distribution departs from the uniform expectation. Second, because the method assumes continuous characters evolve under a Brownian diffusion process, we carried out a set of simulations where the diffusion is punctuated by discrete (normally distributed) jumps in phenotypic space so that there is excess variance relative to the Brownian expectation. Although these scenarios are not exhaustive, they represent obvious ways that biological data sets may commonly depart from conditions assumed by the model.

Performance Assessment

We used model-averaged branch-specific rate estimates to assess the method's performance in two ways. In the first instance, we calculated the correlation coefficient between the true branch rates and the model-averaged branch rates, which provides a scale-independent metric of the method's ability to detect relative rate changes across phylogeny. In the second instance, we computed the mean proportional rate difference between the

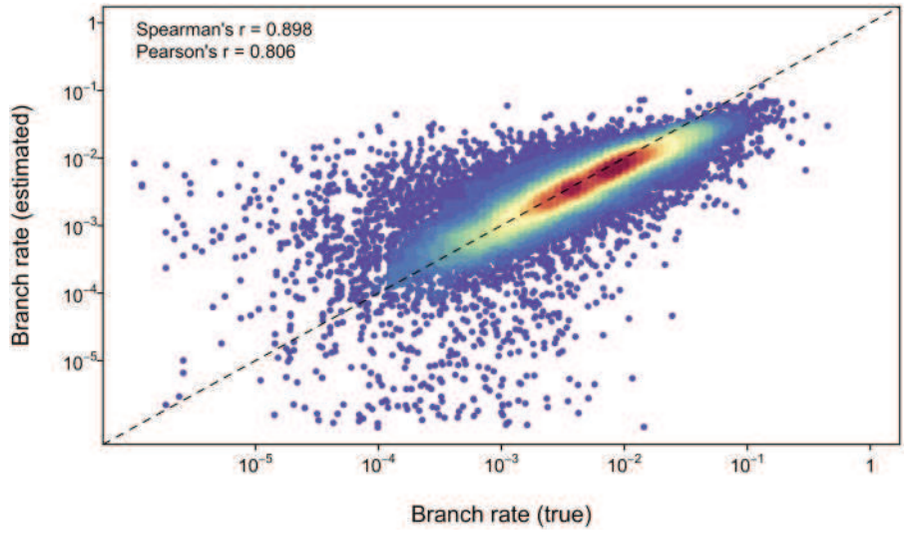

FIGURE 2. Product-moment and rank correlations between true and estimated branch-specific evolutionary rates across all discrete character simulations. Each point represents a single branch and warmer colors indicate a greater density of points. Note the departure of points from isometry toward the right of the plot, which results from underestimation errors when rates are high. Discrete characters lose phylogenetic signal quickly and changes become saturated at high rates, so these underestimates are expected given the approximate rate estimator used in the present study.

true branch rates and the model-averaged branch rates for each simulated rate-shift event, which provides a metric of how accurately the method estimates rates of evolution. The mean proportional rate difference for a rate-shift with true rate equal to r_0 and model-averaged branch rates $\bar{r}_i$ was defined as $\exp(\frac{1}{N}\sum_{i=1}^{N}[\log\bar{r}_i - \log r_0]$. For example, a mean proportional rate difference of one-half indicates that rates are, on average, underestimated by a factor of 2. Because individual rate-shift configurations may also be of interest to researchers, we provided the same performance assessments for the best rate-shift configuration (as determined by AIC) in the supplement (Supplementary Figs. S13–S16 available on Dryad).

Empirical Examples

We applied our new method to two previously analyzed empirical data sets. Eastman et al. (2011) developed a Bayesian method for fitting multirate Brownian models and used it to analyze body size (a continuous trait) evolution for 226 species of turtles. They found strong support for a heterogeneous rate process, with emydid turtles in the genus *Graptemys* showing exceptionally elevated rates of evolution. Additional but more modest rate increases were also observed in tortoises (genus *Geochelone*) and in several geoemydid genera. Pyron and Burbrink (2014) assembled a data set of squamate reptile reproductive modes (a discrete trait) for nearly 4000 species. Analysis of these data with a rate-homogenous state-dependent speciation–extinction model strongly supports a viviparous most recent common ancestor of all squamates, a result that contrasts strongly with parsimony analysis of the same data. King and Lee (2015a) reanalyzed the data set but allowed for among-lineage rate variation using random local clocks (Drummond and Suchard 2010) and found strong evidence of rate variation and strong support for an oviparous ancestor. Thus, this example presents a clear case where biological conclusions depend strongly on assumptions about the phylogenetic distribution of evolutionary rates.

Implementation

Open-source code and documentation for the method is available via R packages mk (for discrete traits; available from https://github.com/blueraleigh/mk) and bm (for continuous traits; available from https://github.com/blueraleigh/bm).

RESULTS

Simulation Study

For discrete traits, the overall correlation across all simulations between model-averaged branch rates and true branch rates was high, with a rank correlation approaching 0.9 (Fig. 2). This was true when using only branches where the true rate decreased over the background rate (rank correlation 0.85) and when using only branches where the true rate increased over the background rate (rank correlation 0.88) (Supplementary Figs. S1 and S2 available on Dryad). Branch rate errors were generally low, with a median proportional error across all simulations of approximately 0.8, corresponding to a slight negative bias (unbiased estimates have a proportional error of 1). The modal behavior of branch-rate correlations and branch-rate errors was largely independent of tree size, but substantially more

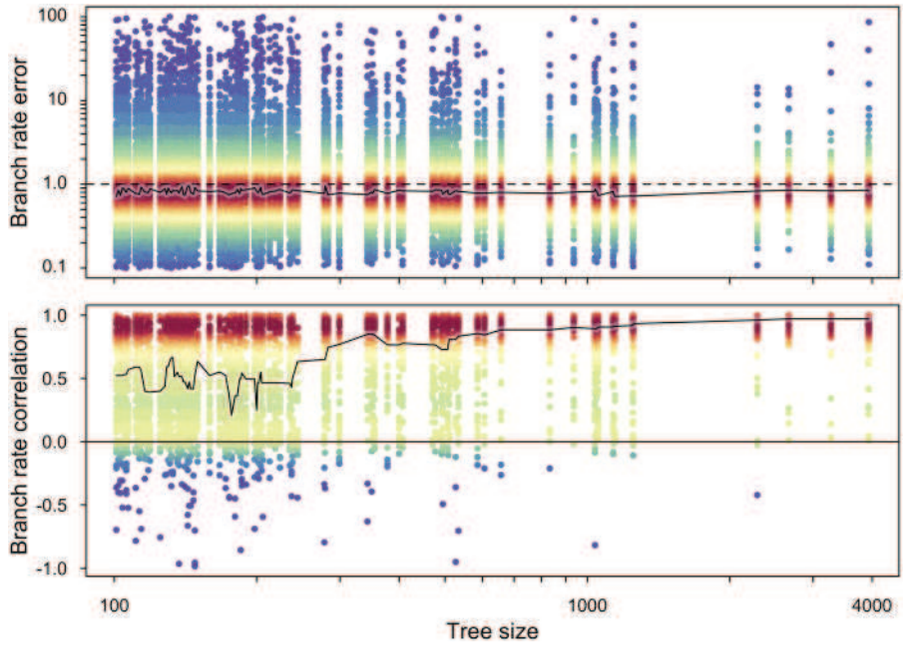

FIGURE 3. Branch rate error (top) and correlation (bottom) between true and estimated branch-specific evolutionary rates as a function of tree size (measured as the number of tips) in discrete character simulations. In the top panel, each point represents the average proportional branch rate error for a single rate-shift event in a single simulation and warmer colors indicate a greater density of points. Unbiased rate estimates have a proportional error of 1, represented by the horizontal dashed line. In the bottom panel, each point depicts the product-moment correlation coefficient between true branch rates and model-averaged branch rate estimates in a single simulation. The solid lines in each panel depict the smoothed median branch rate error or correlation.

variance is observed at smaller tree sizes, which lowers average correlations and raises average errors (Fig. 3). Over the range of values sampled in the simulations, performance of the method appears largely independent of the number of rate-shifts (Supplementary Fig. S3 available on Dryad). Of the 5000 simulated trait data sets, 254 were generated under a constant-rate process. Of these, AIC model selection favored (Δ AIC >0) a variable-rate process over a constant-rate process in 17 cases for an overall false-positive rate of 0.067. However, if we require stronger evidence (ΔAIC >3) to support a variable-rate process over a constant-rate process, the false-positive rate drops to 2 in 254 cases, or 0.008. Allowing for asymmetric transition rates had no evident impact on performance. The overall correlation across all simulations with asymmetric transition rates between model-averaged branch rates and true branch rates was high (rank correlation of 0.9), and the distribution of branch rate errors was essentially unchanged, with a median proportional error across all simulations of approximately 0.8 (Supplementary Figs. S4 and S5 available on Dryad). The false-positive rate remained similarly low (0.08 at ΔAIC >0 and 0.008 at ΔAIC >3).

Similar results were obtained for continuous traits but with generally greater performance overall compared with discrete traits. The rank correlation across all simulations between model-averaged branch rates and true branch rates was 0.98 (Fig. 4) and remained high when using only branches where the true rate decreased over the background rate (rank correlation 0.97) and when using only branches where the true rate increased over the background rate (rank correlation 0.98) (Supplementary Figs. S6 and S7 available on Dryad). The median branch rate error across all simulations was essentially unbiased at approximately 1.02 and showed little variation with tree size, although more variance in estimates is also observed at smaller tree sizes (Fig. 5). Performance of the method was again largely independent of the number of rate-shifts (Supplementary Fig. S8 available on Dryad). Of the 5000 simulated trait data sets, 234 were generated under a constant-rate process. Of these, AIC model selection favored (ΔAIC >0) a variable-rate process over a constant-rate process in 62 cases, for an overall false-positive rate of 0.26. After requiring stronger evidence (ΔAIC >3) to support a variable-rate process over a constant-rate process, the false-positive rate drops to 21 in 234 cases, or 0.09. The impact of allowing for non-Brownian jumps in phenotypic space appears driven by the fraction of phenotypic variance attributable to the jump process over the diffusion process (Supplementary Fig. S9 available on Dryad). When most of the phenotypic variance was attributable to gradual evolution, the method performed well. As the jump process gained in importance the method inferred higher evolutionary rates to accommodate the overdispersion in phenotypic values, leading to generally lower branch rate correlations, higher branch rate errors, and greater false positive rates (Supplementary Figs. S10 and S11 available on Dryad). The rank correlation across all simulations between model-averaged branch rates and true branch rates was 0.89, and the median branch rate

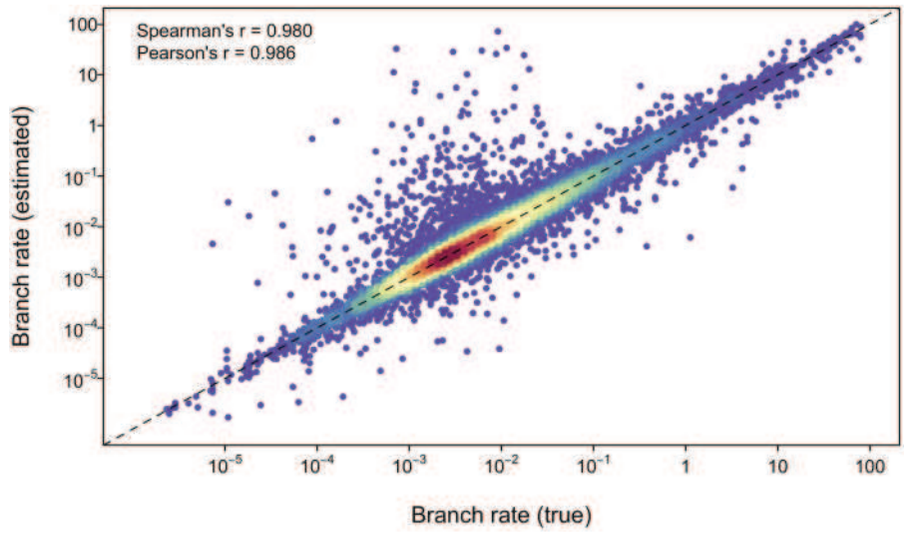

FIGURE 4. Product-moment and rank correlations between true and estimated branch-specific evolutionary rates across all continuous character simulations. Each point represents a single branch and warmer colors indicate a greater density of points. In contrast to discrete characters, continuous characters retain phylogenetic signal when rates are high, and we do not observe systematic underestimates with increasing rate.

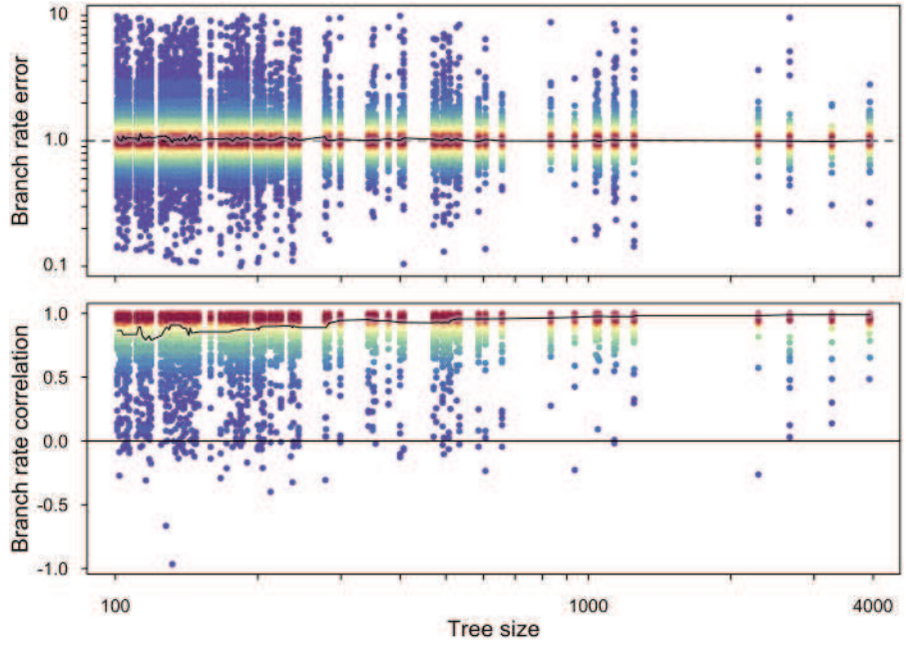

FIGURE 5. Branch rate error (top) and correlation (bottom) between true and estimated branch-specific evolutionary rates as a function of tree size in continuous character simulations. In the top panel, each point represents the average proportional branch rate error for a single rate-shift event in a single simulation and warmer colors indicate a greater density of points. Unbiased rate estimates have a proportional rate error of 1, represented by the horizontal dashed line. In the bottom panel, each point depicts the product-moment correlation coefficient between true branch rates and model-averaged branch rate estimates in a single simulation. The solid lines in each panel depict the smoothed median branch rate error or correlation.

error across all simulations was 1.57. Higher errors were most pronounced at small tree sizes.

Empirical Examples

Analyses of the squamate reproductive mode and turtle body size data sets reveal strong evidence for rate heterogeneity (Figs. 6 and 7). The distribution of phylogenetic rate variation is in broad agreement with results from prior analyses for both squamates (King and Lee 2015a) and turtles (Eastman et al. 2011). In particular, elevated rates of reproductive mode evolution in squamates are observed among skinks, anguimorphs, liolaemids, phrynosomatids, and snakes. Interestingly, the distribution of Akaike weights is strongly bimodal, and individual inspection of these peaks reveals the

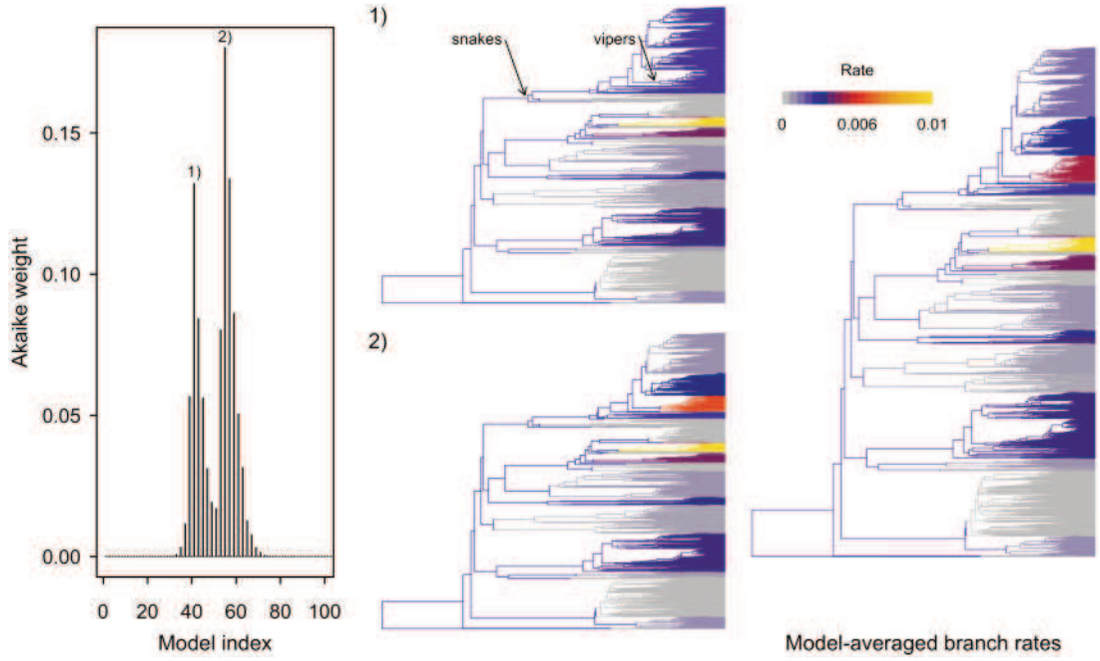

FIGURE 6. Empirical analysis of reproductive mode evolution in squamate reptiles reveals strong evidence of phylogenetic rate variation. The left panel depicts the relative likelihood scores (vertical axis) of different rate-shift configurations (horizontal axis). Rate-shift configurations are indexed by the number of branches belonging to the background process. The middle panels illustrate the distribution of rates on the phylogeny of squamate reptiles for each of the two best rate-shift configurations labeled in the left panel, which differ in the amount of rate heterogeneity inferred among snakes (warm colors denote faster rates). In the right panel, branches are colored by rates that are weighted averages of all rate-shift configurations indexed in the left panel. Substantial increases in the rate of evolution are observed in skinks, some iguanian clades, and snakes, consistent with prior analyses (King and Lee 2015a). This example with nearly 4000 species runs to completion in less than 1 s on a standard laptop computer.

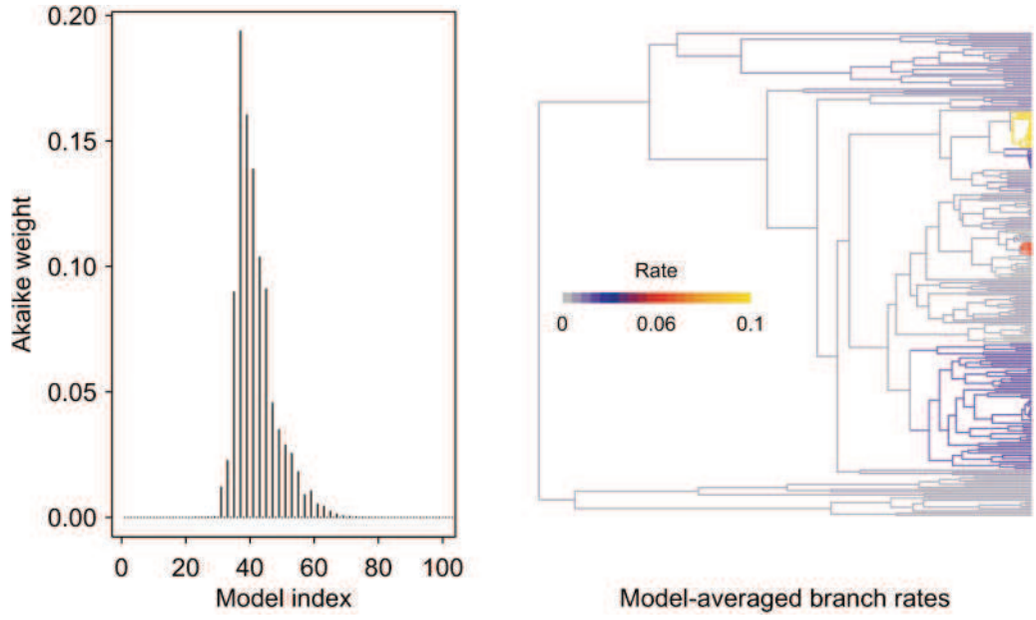

FIGURE 7. Empirical analysis of body size evolution in turtles reveals strong evidence of phylogenetic rate variation. The left panel depicts the relative likelihood scores (vertical axis) of different rate-shift configurations (horizontal axis). The right panel is the phylogeny of turtles with branches colored by model-averaged rate estimates (warm colors denote faster rates). The distribution of phylogenetic rate variation is broadly consistent with prior analyses (Eastman et al. 2011), with notable rate increases observed among emydid turtles and among tortoises.

presence of two distinct, well-supported rate-shift configurations that differ in the amount of rate variation recovered among snakes. In contrast, for turtle body sizes, the distribution of Akaike weights is unimodal with evidence for increased rates of size evolution occurring among the families Emydidae, Geoemydidae, and Testudinidae.

This latter example also helps illustrate the computational efficiency of our approach compared with existing alternatives. Using the recently described maximum likelihood recursive clade partition algorithm (Mitov et al. 2019) to infer rate shifts in turtle body size evolution takes 762 s with a minimum rate-shift size of 6, compared with our approach that takes 0.003 s with a minimum rate-shift size of 2 (2.3 GHz processor with 32 GB RAM). Inferences of phylogenetic rate variation between the two methods are highly similar (Pearson's $r = 0.80$, Supplementary Fig. S12 available on Dryad), but our approach achieves an approximately 254,000-fold speedup in this case.

Discussion

We describe a method for automatic inference of among-lineage evolutionary rate heterogeneity in both continuous and discrete traits that allows researchers to investigate causes underlying the uneven distribution of phenotypic diversity among different clades of organisms. The method assumes that the present-day distribution of a trait is shaped by a multirate Brownian motion in the case of continuous traits or by a multirate fully symmetric Markov process in the case of discrete traits. Unlike many related methods, our new approach automatically evaluates the phylogenetic distribution of many possible multirate processes using a single postorder tree traversal algorithm, thereby permitting rapid exploration of data sets for variable tempo and mode. Moreover, simulation results indicate that the method is able reconstruct branch-specific evolutionary rates with high accuracy, particularly in large phylogenies where the evolutionary process is most likely to vary among lineages in empirical data sets.

Comparison to Existing Methods

Our method is conceptually similar to existing methods that model heterogeneous macroevolutionary trait dynamics as a series of shifts in evolutionary tempo and mode. The majority of these methods have been developed for continuous phenotypes and are implemented in a Bayesian context. By contrast, our approach uses an algorithmic framework that can be applied to both continuous and discrete phenotypes in a maximum likelihood context. Although our approach uses a maximum likelihood formulation, it does not return a single best-fit multirate process. Instead, it evaluates and returns many possible multirate configurations and ranks them according to their relative likelihood. This allows researchers to compute model-averaged branch-specific evolutionary rates or to explore data sets for multiple well-supported rate-shift configurations. In this sense, the method is similar to many existing Bayesian approaches but has the advantage of running orders of magnitude faster. The method's fast likelihood calculations on even the largest phylogenies result from our use of simple symmetric models of character evolution as well as from simplifying assumptions and approximations that allow us to devise a dynamic programming algorithm for fitting multirate models to data. This is an important point of departure from some related methods, which can fit quite complicated and variable evolutionary models to different parts of a phylogeny. Indeed, questions related to variation in evolutionary mode or to contrasting different models of evolution are better addressed by existing software (e.g., Khabbazian et al. 2016; Bastide et al. 2018; Mitov et al. 2019), whereas the current method is best suited for quickly identifying phylogenetic variation in evolutionary tempo. For data sets suspected of harboring complex, fine-scale phylogenetic rate variation, the current method may give only a coarse picture of this variation, and existing methods that formally model nested rate-shifts may be preferred. Importantly, though, our use of model-averaged branch rates can help circumvent some of our simplifying assumptions and approximations, and analyses of simulated data suggest this to be the case: over the range of parameter values examined here model-averaged branch rates were generally quite accurate.

Comparison of Continuous and Discrete Phenotypes

Simulation results indicate that the method generally estimates branch-specific evolutionary rates with greater accuracy for continuous traits than for discrete traits. This is unsurprising given that the likelihood calculations for continuous traits are exact, whereas for discrete traits, they are only approximate. In particular, for discrete traits, we develop an approximation to the likelihood function that uses the number of maximum parsimony implied character state changes to derive an approximate maximum likelihood rate estimate. Explicit in this approximation is the assumption that the rate of evolution is reasonably slow. Nonetheless, our simulation results indicate that relative rate changes can be accurately identified by the method even where this assumption breaks down. The intuition here is that a sufficiently large change in rate from slow to fast or vice versa will still leave a signature in a maximum parsimony reconstruction that is detectable even when the approximate maximum likelihood estimator we use is biased. Even in the absence of approximation error, however, we may still expect rate estimates for continuous traits to be more accurate than their discrete trait counterparts because discrete traits lose phylogenetic signal with increasing rates (Jukes and Cantor 1969; Wagner 2000), although the lower limit to rates where this is actually

an issue may lay beyond values encountered in most empirical data sets.

Our simulation results indicate a higher false-positive rate for continuous traits than for discrete traits. One explanation for this behavior may stem from the greater information content of continuous traits compared with discrete traits. Stochastically large divergences between pairs of terminal taxa separated by short phylogenetic path lengths may (incorrectly) be picked up by the method as evidence for a rate increase. By contrast, a difference in the state of a discrete character between similar such taxon pairs is unlikely to contain enough information to justify a rate increase. We recommend that researchers carefully examine the evidence ratios in favor of variable-rate processes before using model-averaged branch rates in downstream analyses.

Reproductive Mode Evolution in Squamate Reptiles

Squamate reptiles evolved viviparity in approximately 100 independent instances (Blackburn 1992), and numerous hypotheses have been advanced to explain the repeated origins of this trait (Shine 2014; Watson and Cox 2021). The conspicuous phylogenetic clustering of viviparous lineages challenges traditional comparative methods (Pyron and Burbrink 2014), and accounting for among-lineage rate heterogeneity in reproductive mode evolution has been critical in recent attempts to understand its evolution (King and Lee 2015a,b). Our reanalysis of this historically important data set highlights several points. First, our model-averaged results largely recapitulate previous Bayesian analyses of rate variation in this group (King and Lee 2015a), identifying snakes and several distantly related lizard clades as hotspots of viviparity evolution. Second, inspection of individual models reveals two well-supported alternative rate-shift configurations. Whereas one model recovers a homogeneous rate dynamic among snakes, the other, slightly more well-supported, model identifies a substantially elevated rate of evolution among the dangerously venomous viperid snakes. The evolution of viviparity has been cited as a key innovation facilitating the diversification of viperids during the Oligocene, a time of global cooling (Lynch 2009). Elevated rates of viviparity evolution would be consistent with this hypothesis. Our study helps demonstrate how explorations of phylogenetic rate variation can help shed light on macroevolutionary hypotheses but also underscores the position that definitive tests of such hypotheses will need to draw on techniques from outside the standard comparative method statistical toolkit (Uyeda et al. 2018, 2021).

Extensions to the Method

In the current implementation, we estimate the rate of the background process by conditioning on ancestral state estimates at the crown-nodes of rate-shift processes. For this reason, it is important that these ancestral states be both estimated accurately and few in number compared with observed character states. In other words, there should be many more tips in a phylogeny than there are rate-shifts, and this is the most likely explanation for the greater variance in estimation error observed at smaller tree sizes in our simulated data sets. In our simulated data sets, the shift-to-tip ratio never increases beyond approximately 10%. As the shift-to-tip ratio increases, there will be less and less information available for estimating rates, both within rate-shift processes and within the background process. One avenue for future improvements to the method may seek to explore alternative approaches that do not require conditioning on ancestral character states when estimating parameters of the background process, or to devise a two-pass algorithm that more accurately estimates these ancestral states by incorporating information from the whole phylogeny. Currently, no attempt is made to accommodate measurement error or intraspecific phenotypic variation in continuous traits, making the method potentially vulnerable to misleading estimates of large rate increases among close relatives near the present (Ives et al. 2007; Felsenstein 2008). Explicitly accounting for this source of bias is therefore an obvious area for improvement on the method.

In principle, the dynamic programming algorithm used for fitting multirate processes can be extended to more complex models than the Brownian motion and equal-rates Markov processes used in the current study. However, implementing these more complex multiparameter models would offset many of the advantages that come with the current approach. In particular, the lack of closed-form maximum likelihood or approximate maximum likelihood estimates for parameters of more complex models would require additional optimization routines and increase computational runtimes. Furthermore, model-averaging of parameter estimates, whereas straightforward for symmetric, single parameter models, is much more fraught with interpretational difficulties when extended to more complex models (Posada and Buckley 2004; Cade 2015). The widespread availability, at least for continuous phenotypes, of additional methods for modeling heterogeneous tempo and mode under more complex evolutionary processes argues against extending the current approach beyond the simple models currently used.

Finally, it would be straightforward to extend the current approach to accommodate multiple traits simultaneously. However, given the concerns outlined above any attempt to do so would most likely be constrained to work within the confines of the simple models used in the present study, thereby entailing the assumption that a single rate parameter governs the evolution of multiple traits (even if that rate is allowed to vary over phylogeny). The widespread evidence of integration and modularity in phenotypic evolution (Goswami et al. 2014) suggests that the applicability of a multitrait approach would need to be considered carefully for any given data set.

Supplementary Material

Data available from the Dryad Digital Repository: http://dx.doi.org/10.5061/dryad.wpzgmsbph

Funding

This work was supported by the National Science Foundation [DEB-1939128 to F.Z.].

Acknowledgments

We thank members of the Zapata Lab for discussion of ideas and the Associate Editor and two anonymous reviewers for comments that improved the research presented here.

Appendix

Derivation of Equation (2)

Let Υ_T be the set of all nodes in T and let $\hat{X}$ be a map from Υ_T to a finite set of integers such that $\hat{X}(\Lambda_T) = X(\Lambda_T)$. $\hat{X}(\Upsilon_T)$ is called an extension of $X(\Lambda_T)$ and we write P_{θ_T} to represent a probability parameterized by θ_T. The probability of $X(\Lambda_T)$ is computed by summing over all possible extensions as

$$P_{\theta_T}\left(X\left(\Lambda_T\right)\right) = \sum_{\hat{X}(\Upsilon_T)} P_{\theta_T}\left(\hat{X}(\Upsilon_T)\right) = \frac{1}{k}\sum_{\hat{X}(\Upsilon_T)}\prod_{u\neq\rho(T)} P_{\theta_T}(\hat{X}(u)|\hat{X}\left(a_u\right))$$

where k is the number of character states. Ordinarily, this probably is efficiently computed using Felsenstein's (1981) pruning algorithm. This approach does not admit an analytic solution, however, and we therefore seek an approximation that does.

To begin, we write down the lineage transition probability by conditioning on the number of character state changes $N\left(t_u\right)$ on edge u as follows (recalling that these are Poisson distributed)

$$\begin{aligned} &P_{\theta_T}(\hat{X}(u)|\hat{X}\left(a_u\right)) \\ &= \sum_{n=0}^{\infty} P_{\theta_T}(\hat{X}(u)|\hat{X}\left(a_u\right), N\left(t_u\right)=n) P_{\theta_T}(N\left(t_u\right)=n) \\ &= \sum_{n=0}^{\infty} P^n_{\hat{X}(a_u)\to\hat{X}(u)} \frac{\left(\theta_T t_u\right)^n}{n!} e^{-\theta_T t_u} \end{aligned}$$

Here, $P^n_{\hat{X}(a_u)\to\hat{X}(u)}$ is the probability of transitioning from $\hat{X}(a_u)$ to $\hat{X}(u)$ in n steps. Recalling now our assumption that θ_T is small (so that $\theta_T t_u$ is typically much less than 1), we have the approximation

$$\begin{aligned} &P_{\theta_T}(\hat{X}(u)|\hat{X}\left(a_u\right)) \\ &\sim \left(\frac{1}{k-1}\right)^{\mathbb{I}(\hat{X}(a_u)\neq\hat{X}(u))} \left(\theta_T t_u\right)^{\mathbb{I}(\hat{X}(a_u)\neq\hat{X}(u))} e^{-\theta_T t_u} \end{aligned}$$

where $\mathbb{I}(\hat{X}\left(a_u\right)\neq\hat{X}(u))$ is an indicator function that equals 1 when $\hat{X}\left(a_u\right)\neq\hat{X}(u)$ and 0 otherwise. This follows from the fact that when our assumption about the rate of evolution holds, nearly all the probability mass will be concentrated at 0 and 1 events of character state change.

Plugging this approximation in to the equation for the probability of $X\left(\Lambda_T\right)$ gives

$$\begin{aligned} P_{\theta_T}\left(X\left(\Lambda_T\right)\right) &\sim \frac{1}{k}\sum_{\hat{X}(\Upsilon_T)}\prod_{u\neq\rho(T)} \\ &\times \left(\frac{1}{k-1}\right)^{\mathbb{I}\left(\hat{X}(a_u)\neq\hat{X}(u)\right)} \left(\theta_T t_u\right)^{\mathbb{I}\left(\hat{X}(a_u)\neq\hat{X}(u)\right)} e^{-\theta_T t_u} \\ &= \frac{1}{k}\sum_{\hat{X}(\Upsilon_T)} \left(\frac{\theta_T}{k-1}\right)^{N_{\hat{X}}} e^{-\theta_T\sigma_T} \prod_{u\neq\rho(T)} \left(t_u\right)^{\mathbb{I}\left(\hat{X}(a_u)\neq\hat{X}(u)\right)} \end{aligned}$$

where $N_{\hat{X}} = \sum_{u\neq\rho(T)}\mathbb{I}\left(\hat{X}(u)\neq\hat{X}(v)\right)$ is the number of character state changes in $\hat{X}(\Upsilon_T)$ and $\sigma_T = \sum_{u\neq\rho(T)} t_u$ is the sum of all edge lengths in T. When θ_T is small it follows that this likelihood will be almost entirely dominated by those extensions where $N_{\hat{X}}$ is minimal. In the limit as $\theta_T \to 0$ we can drop the summation entirely and consider only the single extension having the fewest changes on the longest branches so that

$$\log P_{\theta_T}\left(X\left(\Lambda_T\right)\right) \sim N_X \log\theta_T - \theta_T\sigma_T + \text{const},$$

where $N_X = \min_{\hat{X}(\Upsilon_T)} N_{\hat{X}}$ and the constant collects terms that do not depend on θ_T.

References

Alhajeri B.H., Schenk J.J., Steppan S.J. 2016. Ecomorphological diversification following continental colonization in muroid rodents (Rodentia: Muroidea). Biol. J. Linn. Soc. 117:463–481.

Barua A., Mikheyev A.S. 2020. Toxin expression in snake venom evolves rapidly with constant shifts in evolutionary rates. Proc. R. Soc. B Biol. Sci. 287:20200613.

Bastide P., Ané C., Robin S., Mariadassou M. 2018. Inference of adaptive shifts for multivariate correlated traits. Syst. Biol. 67:662–680.

Beaulieu J.M., O'Meara B.C., Donoghue M.J. 2013. Identifying hidden rate changes in the evolution of a binary morphological character: the evolution of plant habit in campanulid angiosperms. Syst. Biol. 62:725–737.

Blackburn D.G. 1992. Convergent evolution of viviparity, matrotrophy, and specializations for fetal nutrition in reptiles and other vertebrates. Am. Zool. 32:313–321.

Butler M.A., King A.A. 2004. Phylogenetic comparative analysis: a modeling approach for adaptive evolution. Am. Nat. 164:683–695.

Cade B.S. 2015. Model averaging and muddled multimodel inferences. Ecology 96:2370–2382.

Christin P.-A., Osborne C.P., Chatelet D.S., Columbus J.T., Besnard G., Hodkinson T.R., Garrison L.M., Vorontsova M.S., Edwards E.J.

2013. Anatomical enablers and the evolution of C4 photosynthesis in grasses. Proc. Natl. Acad. Sci. USA 110:1381–1386.
Davis Rabosky A.R., Cox C.L., Rabosky D.L., Title P.O., Holmes I.A., Feldman A., McGuire J.A. 2016. Coral snakes predict the evolution of mimicry across New World snakes. Nat. Commun. 7:11484.
Drummond A.J., Suchard M.A. 2010. Bayesian random local clocks, or one rate to rule them all. BMC Biol. 8:114.
Eastman J.M., Alfaro M.E., Joyce P., Hipp A.L., Harmon L.J. 2011. A novel comparative method for identifying shifts in the rate of character evolution on trees. Evolution 65:3578–3589.
Felsenstein J. 1973. Maximum-likelihood estimation of evolutionary trees from continuous characters. Am. J. Hum. Genet. 25:471–492.
Felsenstein J. 1981. Evolutionary trees from DNA sequences: a maximum likelihood approach. J. Mol. Evol. 17:368–376.
Felsenstein J. 2008. Comparative methods with sampling error and within-species variation: contrasts revisited and revised. Am. Nat. 171:713–725.
Fisher A.A., Ji X., Nishimura A., Lemey P., Suchard M.A. 2021a. Shrinkage-based random local clocks with scalable inference. ArXiv210507119 Q-Bio Stat. Available from: https://doi.org/10.48550/arXiv.2105.07119.
Fisher A.A., Ji X., Zhang Z., Lemey P., Suchard M.A. 2021b. Relaxed random walks at scale. Syst. Biol. 70:258–267.
Fitch W.M. 1971. Toward defining the course of evolution: minimum change for a specific tree topology. Syst. Biol. 20:406–416.
Freckleton R.P. 2012. Fast likelihood calculations for comparative analyses. Methods Ecol. Evol. 3:940–947.
Goswami A., Smaers J.B., Soligo C., Polly P.D. 2014. The macroevolutionary consequences of phenotypic integration: from development to deep time. Philos. Trans. R. Soc. B Biol. Sci. 369:20130254.
Grundler M., Rabosky D.L. 2020. Complex ecological phenotypes on phylogenetic trees: a Markov process model for comparative analysis of multivariate count data. Syst. Biol. 69:1200–1211.
Ives A.R., Midford P.E., Garland T. 2007. Within-species variation and measurement error in phylogenetic comparative methods. Syst. Biol. 56:252–270.
Jukes T.H., Cantor C.R. 1969. Evolution of protein molecules. In: Munro H.N., editor. Mammalian protein metabolism. New York (NY): Academic Press. p. 21–132.
Kapli P., Lutteropp S., Zhang J., Kobert K., Pavlidis P., Stamatakis A., Flouri T. 2017. Multi-rate Poisson tree processes for single-locus species delimitation under maximum likelihood and Markov chain Monte Carlo. Bioinformatics 33:1630–1638.
Khabbazian M., Kriebel R., Rohe K., Ané C. 2016. Fast and accurate detection of evolutionary shifts in Ornstein–Uhlenbeck models. Methods Ecol. Evol. 7:811–824.
King B., Lee M.S.Y. 2015a. Ancestral state reconstruction, rate heterogeneity, and the evolution of reptile viviparity. Syst. Biol. 64:532–544.
King B., Lee M.S.Y. 2015b. Epoch-based likelihood models reveal no evidence for accelerated evolution of viviparity in squamate reptiles in response to cenozoic climate change. J. Exp. Zoolog. B Mol. Dev. Evol. 324:525–531.
Landis M.J., Schraiber J.G., Liang M. 2013. Phylogenetic analysis using Lévy processes: finding jumps in the evolution of continuous traits. Syst. Biol. 62:193–204.
Lloyd G.T., Wang S.C., Brusatte S.L. 2012. Identifying heterogeneity in rates of morphological evolution: discrete character change in the evolution of lungfish (Sarcopterygii; Dipnoi). Evolution 66:330–348.
Lynch V.J. 2009. Live-birth in vipers (Viperidae) is a key innovation and adaptation to global cooling during the Cenozoic. Evolution 63:2457–2465.
Marazzi B., Ané C., Simon M.F., Delgado-Salinas A., Luckow M., Sanderson M.J. 2012. Locating evolutionary precursors on a phylogenetic tree. Evolution 66:3918–3930.
Mitov V., Bartoszek K., Stadler T. 2019. Automatic generation of evolutionary hypotheses using mixed Gaussian phylogenetic models. Proc. Natl. Acad. Sci. USA 116:16921–16926.
O'Meara B.C., Ané C., Sanderson M.J., Wainwright P.C. 2006. Testing for different rates of continuous trait evolution using likelihood. Evolution 60:922–933.
Posada D., Buckley T.R. 2004. Model selection and model averaging in phylogenetics: advantages of Akaike information criterion and Bayesian approaches over likelihood ratio tests. Syst. Biol. 53:793–808.
Price S.L., Etienne R.S., Powell S. 2016. Tightly congruent bursts of lineage and phenotypic diversification identified in a continental ant radiation. Evolution 70:903–912.
Pyron R.A., Burbrink F.T. 2014. Early origin of viviparity and multiple reversions to oviparity in squamate reptiles. Ecol. Lett. 17:13–21.
Rabosky D.L., Donnellan S.C., Grundler M., Lovette I.J. 2014. Analysis and visualization of complex macroevolutionary dynamics: an example from Australian scincid lizards. Syst. Biol. 63:610–627.
Revell L.J. 2021. A variable-rate quantitative trait evolution model using penalized-likelihood. Peer J. 9:e11997.
Revell L.J., Collar D.C. 2009. Phylogenetic analysis of the evolutionary correlation using likelihood. Evolution 63:1090–1100.
Revell L.J., Mahler D.L., Peres-Neto P.R., Redelings B.D. 2012. A new phylogenetic method for identifying exceptional phenotypic diversification. Evolution 66:135–146.
Shine R. 2014. Evolution of an evolutionary hypothesis: a history of changing ideas about the adaptive significance of viviparity in reptiles. J. Herpetol. 48:147–161.
Simões T.R., Vernygora O., Caldwell M.W., Pierce S.E. 2020. Megaevolutionary dynamics and the timing of evolutionary innovation in reptiles. Nat. Commun. 11:3322.
Stroud J.T., Losos J.B. 2016. Ecological opportunity and adaptive radiation. Annu. Rev. Ecol. Evol. Syst. 47:507–532.
Thomas G.H., Freckleton R.P. 2012. MOTMOT: models of trait macroevolution on trees. Methods Ecol. Evol. 3:145–151.
Tonini J., Beard K.H., Barbosa Ferreira R., Jetz W., Pyron R.A. 2016. Fully-sampled phylogenies of squamates reveal evolutionary patterns in threat status. Biol. Conserv. 204:23–31.
Uyeda J.C., Bone N., McHugh S., Rolland J., Pennell M.W. 2021. How should functional relationships be evaluated using phylogenetic comparative methods? A case study using metabolic rate and body temperature. Evolution 75:1097–1105.
Uyeda J.C., Harmon L.J. 2014. A novel Bayesian method for inferring and interpreting the dynamics of adaptive landscapes from phylogenetic comparative data. Syst. Biol. 63:902–918.
Uyeda J.C., Zenil-Ferguson R., Pennell M.W. 2018. Rethinking phylogenetic comparative methods. Syst. Biol. 67:1091–1109.
Wagner P.J. 2000. Exhaustion of morphologic character states among fossil taxa. Evolution 54:365–386.
Watanabe A., Fabre A.-C., Felice R.N., Maisano J.A., Müller J., Herrel A., Goswami A. 2019. Ecomorphological diversification in squamates from conserved pattern of cranial integration. Proc. Natl. Acad. Sci. USA 116:14688–14697.
Watson C.M., Cox C.L. 2021. Elevation, oxygen, and the origins of viviparity. J. Exp. Zoolog. B Mol. Dev. Evol. 336:457–469.
Werner G.D.A., Cornwell W.K., Sprent J.I., Kattge J., Kiers E.T. 2014. A single evolutionary innovation drives the deep evolution of symbiotic N2-fixation in angiosperms. Nat. Commun. 5:4087.
Zanne A.E., Tank D.C., Cornwell W.K., Eastman J.M., Smith S.A., FitzJohn R.G., McGlinn D.J., O'Meara B.C., Moles A.T., Reich P.B., Royer D.L., Soltis D.E., Stevens P.F., Westoby M., Wright I.J., Aarssen L., Bertin R.I., Calaminus A., Govaerts R., Hemmings F., Leishman M.R., Oleksyn J., Soltis P.S., Swenson N.G., Warman L., Beaulieu J.M. 2014. Three keys to the radiation of angiosperms into freezing environments. Nature 506:89–92.