Stochastic Character Mapping of Continuous Traits on Phylogenies

B. S. Martin[1–3,*], and M. G. Weber[3]

[1] *Department of Plant Biology, Ecology, Evolution, and Behavior Program, Michigan State University, East Lansing, MI 48824, USA*

[2] *Department of Plant Biology, University of Georgia, Athens, GA 30602, USA*

[3] *Department of Ecology and Evolutionary Biology, University of Michigan, Ann Arbor, MI 48109, USA*

[*] *Correspondence to be sent to: bruce.stagg.martin@gmail.com*

Abstract

Fossilized organisms only represent a small fraction of Earth's evolutionary history, motivating "ancestral state reconstruction" techniques for inferring unobserved phenotypes of evolving lineages based on measurements of their relatives. Stochastic character mapping has emerged as a particularly powerful approach in this regard, allowing researchers to sample histories of discrete variables on phylogenies and better account for the inherent uncertainty of reconstructed ancestral states. Here, we generalize this procedure to work with continuous variables by developing an efficient algorithm for sampling evolutionary histories under Brownian Motion, implemented in a new R package *contsimmap*. To demonstrate potential applications of these "*continuous* stochastic character maps", we develop a pipeline for inferring relationships between rates of continuous trait evolution and continuously-varying factors (e.g., body size, generation time)—a difficult statistical problem for which relatively few methods are available. After verifying this novel pipeline's performance on simulated data, we use it to show that smaller eucalypt trees tend to exhibit higher rates of flower and leaf trait evolution overall, aligning with well-established predictions based on life history theory as well as empirical patterns in other systems. Ultimately, continuous stochastic character mapping is a valuable new tool for analyzing macroevolutionary data, enabling rigorous yet flexible

investigation of complex evolutionary dynamics involving continuous traits and/or continuous variables hypothesized to affect evolutionary processes.

Key words: eucalypts; macroevolution; comparative methods; stochastic character mapping; ancestral state reconstruction; continuous trait evolution; Brownian Motion; rate heterogeneity

Understanding phenotypic evolution over deep time is central to many fundamental questions in evolutionary biology. Accounting for the evolutionary history of particular phenotypes is necessary to comprehensively address questions ranging from "what trait-trait or trait-environment correlations are present across the tree of life?" to "which factors constrain or promote trait and lineage diversity?". Unfortunately, directly observing phenotypic evolution over macroevolutionary timescales is often impossible due to the incompleteness of the fossil record. Accordingly, biologists have developed numerous "indirect" approaches to instead infer these largely unobserved evolutionary processes from comparative phenotypic data. Perhaps the oldest and most explicit approach to this task is "ancestral state reconstruction" (ASR)—that is, estimating the unobserved phenotypes of evolving lineages based on the observed phenotypes of their (typically living) relatives (Dobzhansky and Sturtevant, 1938; Sanger et al., 1955; Pauling et al., 1963; Witmer, 1995; Schultz et al., 1996; Sumrall and Brochu, 2008).

With the development of rigorous statistical methods for performing ASR under various phylogenetic comparative models of trait evolution, ASR has become ubiquitous in macroevolutionary research (Schluter et al., 1997; Groussin et al., 2016; Joy et al., 2016). Stochastic character mapping in particular has become an extremely popular ASR technique, allowing researchers to randomly sample evolutionary histories of characters along phylogenies according to their probability under a given model of trait evolution (Nielsen, 2002; Huelsenbeck et al., 2003; Bollback, 2006). This approach is generally more

useful than other ASR techniques because, while the most likely estimates of ancestral phenotypes under a given trait evolution model are often inaccurate (particularly among older lineages without accompanying fossil data; e.g., Slater et al., 2012; Joy et al., 2016), they are typically so uncertain as to be consistent with a wide variety of reasonable evolutionary trajectories. By sampling hundreds or thousands of stochastic character maps, researchers can flexibly conduct macroevolutionary analyses over distributions of possible evolutionary histories and better account for the inherent uncertainty of reconstructed ancestral states in their inferences. Accordingly, stochastic character mapping has been used across numerous studies for a variety of purposes, from determining the timing/frequency of evolutionary events (e.g., Baliga and Law, 2016; Tornabene et al., 2016; Freyman and Höhna, 2019; Hughes et al., 2021; Landis et al., 2021; Siqueira et al., 2023) to investigating correlations between past life history/environmental factors and evolutionary outcomes (e.g., de Alencar et al., 2017; Borstein et al., 2019; Burns and Bloom, 2020; Fabre et al., 2020; Rincon-Sandoval et al., 2020; Drury et al., 2021; Nations et al., 2021; Friedman and Muñoz, 2023).

While stochastic character mapping has revolutionized statistical approaches to macroevolutionary analysis, all currently available implementations are limited to use on discrete variables, with no analogue for continuous variables. This conspicuous methodological gap constrains approaches for analyzing macroevolutionary data, as many methods for testing whether evolutionary dynamics vary according to some "explanatory factor" (e.g., habitat, reproductive strategy) explicitly rely on fixed estimates of the factor's evolutionary history—often provided in the form of stochastic character maps (e.g., Revell, 2013a; Clavel et al., 2015; Beaulieu and O'Meara, 2023). Consequently, researchers investigating the interplay between evolutionary dynamics and continuously-varying factors are generally forced to arbitrarily divide continuous variation into discrete categories and/or develop custom statistical pipelines tailored to their specific research goals (Cooper and Purvis, 2009; Hansen et al., 2008; Welch and Waxman, 2008;

FitzJohn, 2010; Lartillot and Poujol, 2011; Felsenstein, 2012; Baker et al., 2015; Weir and Lawson, 2015; Clavel and Morlon, 2017; Harvey and Rabosky, 2017; Cooney and Thomas, 2021; Hansen et al., 2022; Uyeda et al., 2021; Boyko et al., 2023a; Tribble et al., 2023). Given that many biologically important variables are continuous in nature (e.g., body size, generation time, climatic niche; Friedman et al., 2019; Gingerich, 2001; Tribble et al., 2023), the lack of stochastic character mapping methods for continuous variables currently presents a limiting hurdle for comparative research.

Here, we introduce an efficient and flexible method for stochastic character mapping of continuous variables under Brownian Motion models of evolution. These "*continuous* stochastic character maps" may be used to better incorporate uncertainty in analyses that aim to characterize the evolutionary dynamics of continuous traits, elucidate the timing/phylogenetic locations of major transitions in continuously-measured phenotypes, and/or explore how continuously-varying factors affect evolutionary dynamics, among other common goals in macroevolutionary research. We first present an algorithm for generating continuous stochastic character maps that works by sampling values of continuous variables across phylogenies while conditioning on any observed data at the phylogeny's nodes/tips. Second, we showcase potential uses of continuous stochastic character mapping by developing a general pipeline for inferring relationships between continuous trait evolution dynamics and continuously-varying factors based on mapped factor histories—focusing on factor-associated evolutionary rate variation in particular. Lastly, after verifying this pipeline's performance on simulated data, we use it to show that lineages of smaller eucalypt trees tend to exhibit higher rates of leaf and flower trait evolution overall, demonstrating the empirical utility of both our pipeline as well as continuous stochastic character mapping in general.

Materials and Methods

Continuous stochastic character mapping is a flexible framework for sampling evolutionary histories of continuous variables along phylogenies under Brownian Motion models. We implement this framework in an R package called *contsimmap*, which supports stochastic character mapping of one or more potentially correlated variables while also accommodating multiple measurements per tip/node with associated intraspecific variation/measurement error (Felsenstein, 2008; Hansen and Bartoszek, 2012; Landis and Schraiber, 2017), missing measurements, and even among-lineage variation in evolutionary trends (Hansen and Martins, 1996) and/or rates. The R package additionally provides tools for transforming, summarizing, and visualizing continuous stochastic character maps (Fig. 1), as well as fitting Brownian Motion models with parameters dependent on mapped continuous variables.

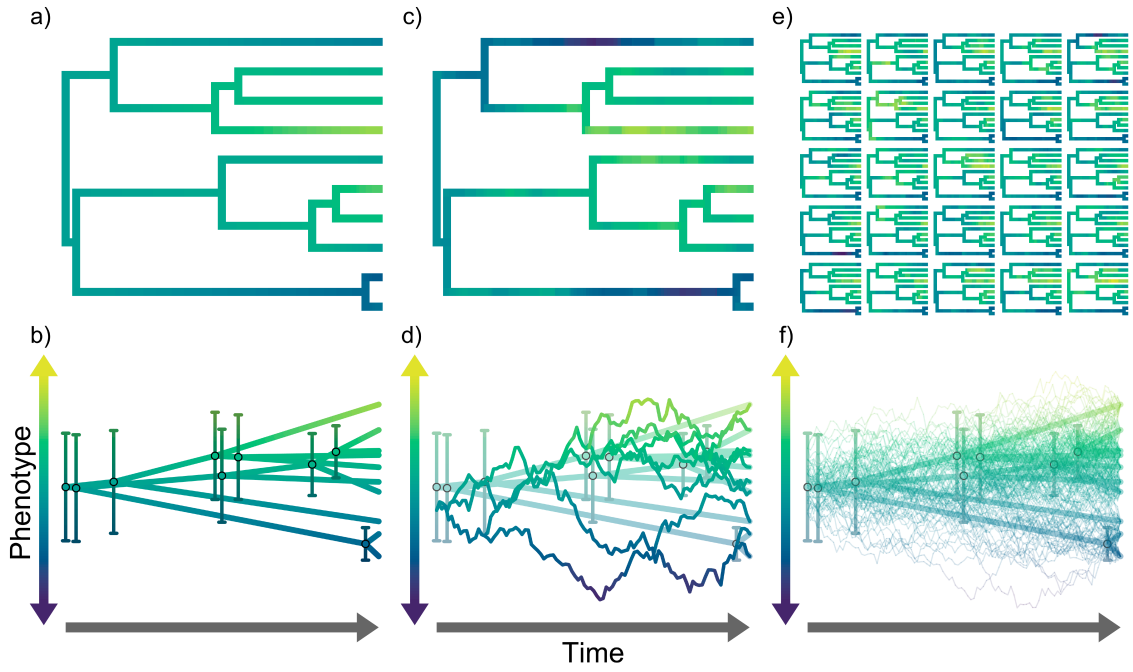

Figure 1. Phylogram and phenogram-based (Revell, 2013b) visualizations of continuous stochastic character maps, colored according to reconstructed phenotypic values. Panels a-b: conventional ancestral state reconstructions of a continuously-varying phenotype under a Brownian Motion model. The phylogram in panel a is colored according to mean phenotypic estimates, which are also depicted by the phenogram in panel b along with error bars representing 95% confidence intervals on the estimates at each node. Panels c-d: a single continuous stochastic character map generated using the same data and Brownian Motion model, again represented as both a colored phylogram and phenogram in panels c and d, respectively. Panels e-f: a sample of 25 continuous stochastic character maps representing the overall distribution of generated maps also depicted as phylograms and phenograms in panels e and f, respectively. Panels c-f also include conventional ancestral state reconstructions for reference.

Overview of Continuous Stochastic Character Mapping

In this section, we provide an overview of our algorithm for generating continuous stochastic character maps. For simplicity, the explanations in this section generally assume we are sampling the evolutionary history of a single continuous variable (hereafter "phenotype") under a homogeneous Brownian Motion model based on exact measurements at the tips of a phylogeny. A more in-depth and technical description of the algorithm—including how it is modified to work with multiple correlated variables, among-lineage variation in evolutionary trends/rates, and "messier" comparative data

involving multiple/missing measurements at both tips and internal nodes—can be found in the Online Appendix (available under Supplementary Material on bioRxiv).

Overall, our algorithm works by sampling phenotypic values at fine-grained time points (including all nodes) across a phylogeny while conditioning on observed measurements at tips. This procedure is analogous to conventional discrete stochastic character mapping, but with two key distinctions. First, discrete stochastic character mapping generally samples evolutionary histories under continuous time Markov chain models, whereas our method uses Brownian Motion. Second, while discrete stochastic character mapping works by directly mapping evolutionary transitions among different states onto a phylogeny, continuous stochastic character mapping instead samples phenotypic values at a finite number of time points evenly distributed across a phylogeny. The overall density of time points is controlled by a user-specified "resolution" parameter corresponding to the approximate number of points along the longest root-to-tip path in the phylogeny. Critically, sampling phenotypic values along edges allows researchers to account for uncertainty in phenotypic evolutionary trajectories between nodes, which increases with edge length. Such uncertainty is completely obscured by conventional methods for ASR of continuous traits, which generally report estimated phenotypes at nodes only.

Our algorithm ensures sampled phenotypic values follow the joint distribution of unobserved phenotypic states across all time points under an assumed Brownian Motion model (i.e., it explicitly accounts for expected correlations among unobserved phenotypes of closely-related lineages). For computational efficiency, we avoid sampling phenotypic values from all time points at once by recursively computing and sampling from distributions of phenotypic values for one point at a time. To do this, our algorithm employs a "backward then forwards" traversal approach, similarly to other ancestral state reconstruction algorithms (e.g., Yang, 2006; Hiscott et al., 2016; Goolsby, 2017; Hassler et al., 2022). This procedure consists of first calculating distributions of possible phenotypic

values at each node during a backwards traversal across the phylogeny from its tips to the root, followed by a forwards (i.e., root to tips) traversal that progressively updates and samples from the distributions at both nodes and time points along edges. We visually illustrate this process in Fig. 2 and describe each of these traversals in more detail below:

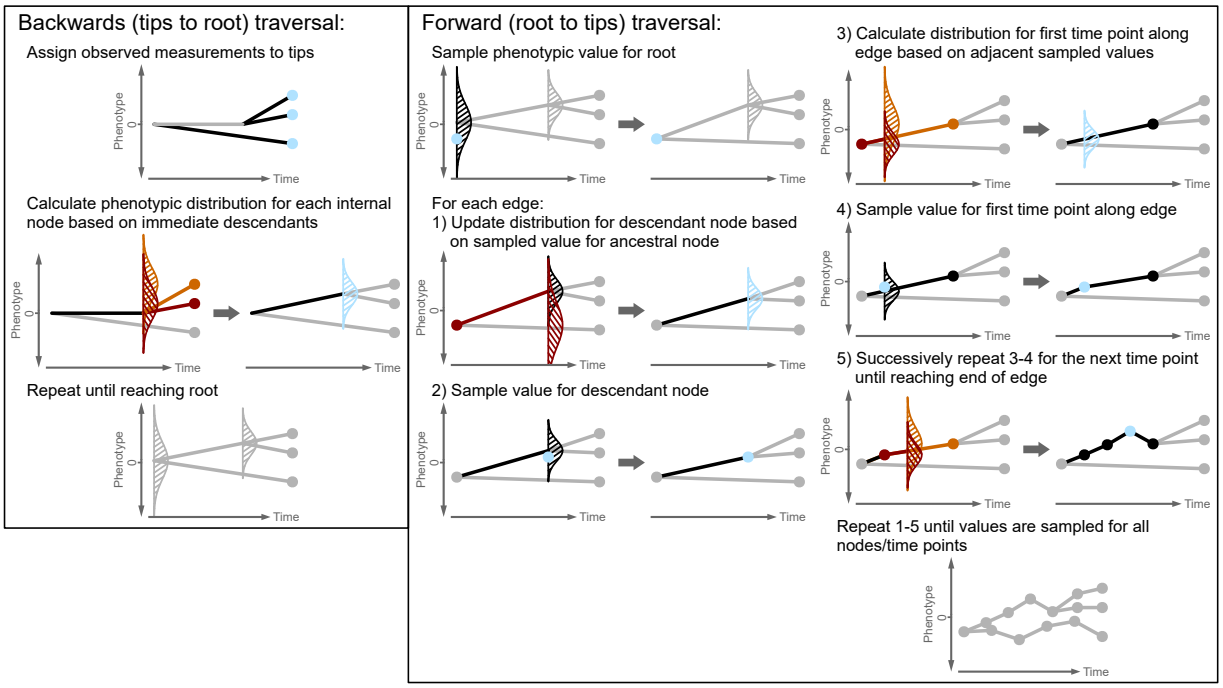

Figure 2. Visual phenogram-based (Revell, 2013b) overview of the major steps in our continuous stochastic character mapping algorithm. Line-shaded, vertically-oriented normal distributions represent distributions of possible phenotypic values at particular points on the phylogeny, while dots indicate exact phenotypic values (either measured or sampled). The algorithm generally focuses on one edge in a given step, which—along with any associated distributions/dots—are drawn in black, though light blue is used to highlight particular distributions/dots on the focal edge that are updated in a given step. Medium-dark red/orange are also used to indicate distributions briefly used in intermediate calculations—these distributions are conditioned on particular nodes/edges drawn in matching colors. Note that, for purely aesthetic purposes, phenotypic distributions at internal nodes are initialized at 0 while values at time points along edges are set to their maximum likelihood expectation before being explicitly sampled (i.e., a straight line connecting adjacent nodes). Line-shaded normal distributions and dots are not drawn at these points to convey that proper phenotypic distributions are yet to be calculated/sampled from.

Backwards traversal: This first traversal calculates phenotypic distributions at all nodes in the phylogeny based only on the measurements associated with each node's descendants. After assigning observed measurements to their corresponding tips, the algorithm visits each internal node in the phylogeny, making sure to only visit a given node

after all of its direct descendants have been visited. By leveraging convenient properties of Brownian Motion and normal distributions (see Online Appendix for specific details), the phenotypic distribution for each node is computed recursively from the distributions associated with the node's immediate descendants. After reaching and calculating the phenotypic distribution at the root of the phylogeny, the algorithm moves on to the subsequent forwards traversal.

Forwards traversal: The second traversal has two main functions: updating phenotypic distributions at each node based on measurements associated with each node's *non*-descendants and sampling exact phenotypic values for all nodes/time points across the phylogeny. The algorithm begins by sampling a phenotypic value for the root from its associated distribution (i.e., the one calculated during the backwards traversal). Next, the algorithm visits each node, first updating the node's phenotypic distribution based on the value sampled for its direct ancestor. Then, the algorithm samples a value for the given node from this updated distribution. Before moving on to another node, the algorithm samples phenotypic values at equally-spaced time points along the given node's preceding edge. The phenotypic distributions at these time points are derived by conditioning the assumed Brownian Motion process to start and end at the sampled phenotypic values for each point's immediate ancestor and the node descending from the given edge (i.e., a "Brownian bridge"; Karatzas and Shreve, 2004; see also Horvilleur and Lartillot, 2014; Lartillot et al., 2016; Quintero and Landis, 2020; Jhwueng, 2021; Quintero, 2025).

Upon completing the forward traversal, the algorithm generates a single continuous stochastic character map. From here, the process can be repeated to sample additional maps in proportion to their probability under the assumed Brownian Motion model.

Validation Study

To verify that our continuous stochastic character mapping algorithm correctly generates samples from the joint distribution of unobserved phenotypic states under an

assumed Brownian Motion model, we conducted a validation study comparing the results of our algorithm to their theoretical expectations. To derive the expected distribution of phenotypic values, we first consider the joint distribution of both observed measurements and phenotypic values at all time points across the phylogeny. Because these values are multivariate normally distributed under Brownian Motion models, we can use well-established formulae to compute the corresponding conditional distribution of phenotypic values given exact values for all observed measurements (Petersen and Pedersen, 2012; Jhwueng, 2021; see Online Appendix for details). We used one-sample Kolmogorov-Smirnov tests (Marsaglia et al., 2003) to determine whether the continuous stochastic character maps generated by our algorithm follow these expected distributions (using Mahalanobis distances and χ^2 distributions as univariate summaries of these multivariate samples/distributions; see Online Appendix).

To comprehensively validate our algorithm, we generated continuous stochastic character maps and calculated expected distributions under a variety of simulated scenarios. In total, we defined 360 distinct simulation conditions that encompassed trait datasets simulated on both ultrametric and non-ultrametric phylogenies ranging from 10-100 tips with or without evolutionary trends, intraspecific variation/measurement error, multiple potentially correlated phenotypes, multiple observations per tip, and missing data (See Online Appendix for specific details). For half of these conditions, we generated continuous stochastic character maps and calculated expected distributions under the same Brownian Motion models used to simulate data. For the other half, we instead generated maps/calculated expected distributions under Brownian Motion models with random parameters (i.e., micmicking parameter misestimation/model misspecificiation). We simulated 10 phylogenies and associated trait datasets per condition for a total of 3,600 simulations, generating 100 continuous stochastic character maps (ranging in resolution from 20–500) per simulated dataset.

Pipeline using Continuous Stochastic Character Maps to Model Factor-Dependent Trait Evolution

Stochastic character maps are particularly useful for "sequential inference" pipelines, whereby the impact of an explanatory factor on an evolutionary process is inferred by first generating stochastic character maps of factor histories, then fitting factor-dependent evolutionary models based on these maps. For discrete stochastic character maps, evolutionary models can be made factor-dependent by simply allowing the model's parameters—such as evolutionary rates or trends in the case of Brownian Motion models—to vary among lineages in different discrete states (e.g., lineages in mountain versus lowland habitats, annual versus perennial plant lineages; see Revell, 2013a for further description of this approach). For continuous stochastic character maps, with an infinite continuum of states, such models must instead use "parameter functions" that convert factor values to parameter values (e.g., the value of a parameter could increase/decrease according to a linear or exponential relationship with the factor; e.g., FitzJohn, 2010).

We implemented a set of tools in our R package for sequential inference of factor-dependent Brownian Motion models based on continuous stochastic character maps of factor histories (Fig. 3). Specifically, our implementation constructs a likelihood function for a trait dataset given a set of continuous stochastic character maps (step 1 in Fig. 3) along with parameter functions that convert these mapped values to the evolutionary rate, trend, and/or intraspecific/measurement error variance parameters of a Brownian Motion model (step 2 in Fig. 3). The parameter functions are themselves controlled by free parameters determining the overall magnitude of rates, trends, etc. and/or how they vary with respect to factor values, analogously to intercepts and slopes in linear regression. Thus, fitting parameter functions to data involves inferring the values of these free parameters via maximum likelihood or Bayesian approaches. Given estimates for free parameters, a likelihood function constructed using our R package automatically

transforms the given set of factor histories into sets of parameter histories—assuming the factor/parameter values at each time point are constant during their preceding time intervals—and uses the pruning algorithm outlined by Hassler et al., 2022 to calculate likelihoods conditional on each continuous stochastic character map (step 3 in Fig. 3). Ultimately, the likelihood function returns a single overall likelihood by averaging all the conditional likelihoods (in practice, we sometimes found it beneficial to account for "outlier" conditional likelihoods before taking these averages; see Online Appendix). This can be viewed as a Monte Carlo approach to estimating the "true" likelihood that would be obtained by integrating over all possible factor histories rather than a finite sample of them (Mayrose and Otto, 2011).

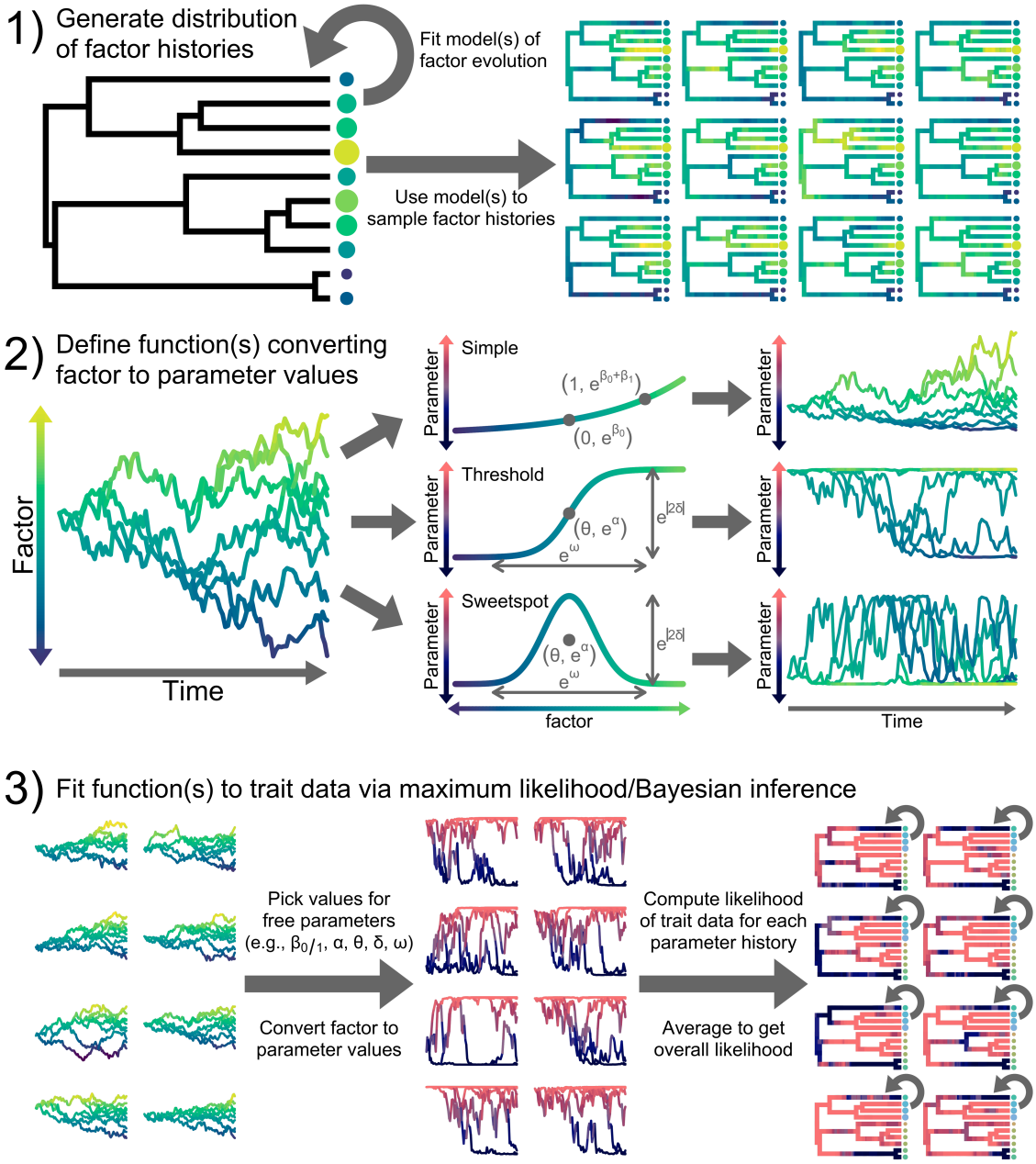

Figure 3. Visual overview of our continuous stochastic character map-based pipeline for inferring relationships between continuously-varying factors and the parameters of a Brownian Motion model of trait evolution. The pipeline consists of three major steps: 1) fitting Brownian Motion models to factor data and using continuous stochastic character mapping to generate a distribution of possible factor histories, 2) defining parameter functions that convert mapped factor values to parameter values, and 3) fitting parameter functions to trait data by using maximum likelihood/Bayesian approaches to infer values of free parameters defining parameter functions. The lighter purple-yellow and darker blue-red color gradients indicate different factor and parameter values, respectively. Differently sized/colored dots arrayed along the tips of phylogenies indicate observed factor and trait data in steps 1 and 3, respectively. Note in step 2 that we illustrate the three main parameter functions (along with their associated free parameters) used to model factor-rate relationships in the current work.

Our R package currently provides tools for using the C++ library NLOPT (Johnson, 2021), interfaced through the R package *nloptr* (Ypma et al., 2022), to infer free parameter values that maximize a given likelihood function. While our implementation allows users to select any optimization algorithm available in the NLOPT library and customize optimizer settings as they see fit, we generally recommend the (currently default) SBPLX (Rowan, 1990) algorithm due to its flexibility and robustness under many conditions.

Pipeline Simulation Study

To assess the performance of our pipeline for inferring relationships between continuous trait evolution dynamics and continuously-varying factors, we tested whether our approach could be used to reliably detect and quantify factor-dependent rates of continuous trait evolution from simulated data. We simulated the evolution of a single continuous trait, Y, with rates depending on a simulated continuous factor—either an observed factor X_o or unobserved factor X_h—and applied our pipeline to the simulated data to infer relationships between rates of Y evolution and X_o. In an empirical context, X_h represents a "hidden" factor that affected rates of trait evolution and may thus mislead hypothesis testing for factor-dependent rates (May and Moore, 2020; see also Beaulieu and O'Meara, 2016; Boyko and Beaulieu, 2023; Boyko et al., 2023b; Tribble et al., 2023). After outlining our simulation procedure below, we outline our analysis procedure, which includes a pragmatic technique for constructing additional null models that help account for hidden factors and thereby mitigate their potential effects on hypothesis testing.

We allowed rates of Y evolution, σ^2, to vary with factor values (denoted X below) according to one of four parameter functions: 1) "simple" functions whereby rates exponentially increase/decrease, 2) "threshold" functions whereby rates shift between some minimum and maximum value, 3) "sweetspot" functions whereby rates peak or dip around a particular factor value, and 4) "constant" functions whereby rates do not vary at all. We

define simple functions as:

$$\sigma^2 = \mathrm{e}^{\beta_0 + \beta_1 X} \tag{1}$$

Where the free parameters β_0 and β_1 represent the intercept and slope of the factor-rate relationship on a logarithmic scale (see step 2 in Fig. 3 for a visual representation). Simple functions are useful in scenarios where rates are hypothesized to increase or decrease in association with a factor like temperature (Clavel and Morlon, 2017; Slater et al., 2017) or generation time (Gingerich, 2001). We use an exponential—as opposed to linear—relationship to ensure estimated rates are always positive while also better reflecting empirical findings that evolutionary rates tend to vary multiplicatively rather than additively (Limpert et al., 2001; Gingerich, 2009).

Threshold functions are useful in cases where rates are hypothesized to follow a logistic relationship with some factor. A benefit of threshold functions is that, unlike simple functions, they are bounded between minimum and maximum rates. Under simple functions, rates of trait evolution can become arbitrarily close to 0 and grow without bound. In reality, various factors like mutation rates, integration among traits, and/or physical limits to trait variation will constrain how slow or fast rates can be. We define threshold functions as:

$$\sigma^2 = \mathrm{e}^{\alpha} \left(\tanh \delta \left(2 \left(1 + \mathrm{e}^{-2\pi\sqrt{3}\frac{X-\theta}{\mathrm{e}^{\omega}}} \right)^{-1} - 1 \right) + 1 \right) \tag{2}$$

Here, α alters the overall scale of rates, specifically corresponding to the natural log of the rate halfway between the minimum and maximum possible rates or "mid rate", while θ determines the factor value or "location" at which rates shift (i.e., the value of X at which $\sigma^2 = \mathrm{e}^{\alpha}$). We denote δ, which controls the direction and magnitude of the rate shift, the "rate deviation". The fold-difference between minimum and maximum rates is given by $\mathrm{e}^{2|\delta|}$, and positive and negative values of δ yield upward and downward shifts with increasing factor values, respectively. Lastly, ω adjusts the "width" of the shift, with rates

roughly reaching their minimum/maximum values at factor values of $\theta \pm \mathrm{e}^{\omega}/2$ (see step 2 in Fig. 3 for a visual representation of these free parameters).

Sweetspot functions can be used when rates are hypothesized to peak or dip at intermediate factor values (Cooper and Purvis, 2009; FitzJohn et al., 2009; Feldman et al., 2016; Amado et al., 2021). For example, Cooper and Purvis (2009) showed that rates of cranial shape evolution are slowest among Paltyrrhine monkeys and Palangeriforme possums of intermediate size, increasing among both smaller and especially larger-bodied lineages. Under sweetspot functions, factor-rate relationships follow positively-valued Gaussian curves of arbitrary height and orientation:

$$\sigma^2 = \mathrm{e}^{\alpha}\left(\tanh\delta\left(2\mathrm{e}^{-18\left(\frac{X-\theta}{\mathrm{e}^{\omega}}\right)^2} - 1\right) + 1\right) \tag{3}$$

Where α, θ, δ, and ω largely have the same effects and interpretations as they do for the threshold function (see step 2 in Fig. 3 for a visual representation). Under the sweetspot function, however, rates peak to their maximum at θ and reach their minimum at factor values around $\theta \pm \mathrm{e}^{\omega}/2$ if δ is positive. Conversely, if δ is negative, rates will instead dip to their minimum at θ and reach their maximum around $\theta \pm \mathrm{e}^{\omega}/2$.

Lastly, we define constant functions, whereby rates do not vary, as $\sigma^2 = \mathrm{e}^{\alpha}$ for consistency with the α parameter under threshold and sweetspot functions.

For each simulation, we used *phytools* (Revell, 2012) to simulate an ultrametric, pure-birth phylogeny of height 1 with either 50, 100, or 200 tips. To simulate the continuous factors X_o and X_h, we generated two densely-sampled continuous factor histories (i.e., resolution of 500) evolving under Brownian Motion processes with root values of 0, no trends, and constant rates of 4. For the trait Y, we simulated an additional Brownian Motion process with root value 0 and no trends, but with rates varying according one of 12 possible factor-rate relationships which differed in type (i.e., constant, simple, threshold, or sweetspot), whether they depended on the observed factor X_o or hidden factor X_h, the overall magnitude of rate variation or "strength", and how gradually

rates change with respect to factor values or "width" (see Table S1 for the specific relationships/free parameter values used). To assess the robustness of our approach to imperfect correspondences between factors and rates, we repeated these simulations while adding random variation or noise around factor-rate relationships in a manner inspired by both uncorrelated relaxed clock models and rate fluctuations under Lévy processes (e.g., Drummond et al., 2006; Landis et al., 2013; Lartillot et al., 2016; see Online Appendix for details on noise simulation). Ultimately, we defined 72 distinct simulation conditions and repeated our simulation procedure 120 times per condition, yielding a grand total of 8,640 simulated phylogenies and associated factor/trait datasets.

For the analysis procedure, we retained the phylogeny and tip factor/trait values of X_o and Y from each simulation. We analyzed the factor/trait datasets by first fitting Brownian Motion models to the X_o data assuming no intraspecific variation/measurement error or evolutionary trends and using the fitted model to generate 100 continuous stochastic character maps of X_o with resolutions of 100 (e.g., step 1 of Fig. 3). We then fit four Brownian Motion models to the Y data conditioned on these maps (again assuming no intraspecific variation/measurement error or evolutionary trends)—three non-null models assuming rates depend on mapped X_o values through either a simple, threshold, or sweetspot function, plus a null model assuming a constant rate (e.g., steps 2 and 3 in Fig. 3). To better account for rate variation due to hidden factors, we fit three additional null models to the Y data assuming rates depend on a mapped "dummy factor" D—which is simulated under the same Brownian Motion model fitted to the X_o data—through either a simple, threshold, or sweetspot relationship. Because D exhibits the same evolutionary dynamics as X_o but with random tip values, support for D-dependent over X_o-dependent models provide evidence that rates of Y evolution vary, but in a way not necessarily related to the observed factor X_o (see also Tribble et al., 2023, which employs a similar approach in testing for relationships between discrete trait evolution and continuous factors). See the Online Appendix for details on the particular conditional likelihood

averaging/optimizer settings used to fit models to simulated data.

We analyzed the results of our simulation study through both model selection and parameter inference-based approaches. For model selection, we used sums of small sample size corrected Akaike Information Criteria (AICc) weights across all X_o-dependent models to gauge whether simulated datasets exhibit "significant" support for a relationship between rates of Y evolution and the observed factor X_o. For parameter inference, we compared model-averaged predictions of X_o-rate relationships based on AICc weights to the true relationships used to simulate data (see Online Appendix for further details on both our model selection and model-averaging procedures).

Empirical Example

We applied our pipeline for inferring factor-dependent rates of continuous trait evolution to test whether rates of phenotypic evolution are associated with size among eucalypts, a diverse clade of ∼800 predominantly Australian woody plant species. Body size is often associated with many aspects of life history, as larger organisms tend to live in smaller populations with slower generational turnover (Calder, 2001; Niklas and Enquist, 2001; White et al., 2007; Sibly and Brown, 2007; Adler et al., 2014; Salguero-Gómez et al., 2016; Bakewell et al., 2020). This pattern has led to the hypothesis that evolutionary rates should generally be slower for larger organisms compared to their smaller relatives. However, how rates of phenotypic evolution scale with body size remains unclear, particularly in plants (Cooper and Purvis, 2009; Baker et al., 2015; Chira et al., 2018; Friedman et al., 2019; Zimova et al., 2023; see also Lanfear et al., 2013; Boucher et al., 2017; Vasconcelos et al., 2022). Eucalypts are a powerful system for investigating relationships between body size and rates of phenotypic evolution in plants, as their growth forms range from shrubby "mallees" only reaching a few meters in height (e.g., *Corymbia setosa*, *Eucalyptus erectifolia*, *E. macrocarpa*, *E. odontocarpa*, *E. vernicosa*) to gargantuan trees that grow nearly 100 meters tall (e.g., *E. obliqua*, *E. regnans*, *E.*

viminalis; Brooker et al., 2015; Nicolle, 2024).

Leveraging a recently-published dataset consisting of 101 nuclear exonic loci (Crisp et al., 2024), we inferred a time-calibrated phylogeny consisting of 388 ingroup tips belonging to 318 unique eucalypt species representing most major subclades of the group (see Online Appendix for details on phylogenetic inference). We then gathered species-level data on maximum reported measurements for 14 continuous traits: height (a proxy for body size which acts as a factor in our analyses), 8 flower traits (peduncle length, inflorescence/infructescence pedicel length, bud length/width, fruit length/width, seed length), and 5 leaf traits (adult petiole length, juvenile/adult leaf length/width). We obtained these data using a regular expression-based approach to extract continuous trait measurements from all species descriptions in the online Australian eucalypt encyclopedia EUCLID (Brooker et al., 2015). Species descriptions yielding missing and/or outlier measurements were flagged and manually checked/corrected as necessary (see Online Appendix for details on trait data processing). We supplemented the EUCLID data with trait measurements for *E. deglupta*—the only species in our phylogeny not native to Australia—from several regional floras (Pool, 2001; Parnell and Chantaranotahi, 2002; Barrie, 2009) and obtained trait data from published descriptions for two recently named eucalypt species in our phylogeny, *E. connexa* and *E. plumula* (Nicolle and French, 2021). Our final dataset consisted of 311 eucalypt species with 0% missing data for height, 0–3.5% for floral traits, and 5.5–9% for leaf traits. All trait values were log-transformed for our comparative analyses. Prior to applying our pipeline, we used Pagel's lambda test via the R package *geiger* (Pennell et al., 2014) to assess the level of phylogenetic signal for each trait.

We generated 400 continuous stochastic character maps of eucalypt height (with a resolution of 100) under an evolving rates or "evorates" model whereby rates of height evolution are allowed to gradually change over time and across lineages (Martin et al., 2023; see appendix for further details). We then generated 400 corresponding dummy factor maps by simulating trait evolution with root values and rates identical to those of

the height maps. Using these continuous stochastic character maps, we fit seven Brownian Motion models to each individual trait—six models assuming rates depend either on height or the dummy factor through a simple, threshold, or sweetspot relationship, plus a constant rates model (see Online Appendix for details on optimizer settings). Despite using maximum reported measurements for our trait data, we still inferred measurement error parameters in all models for two reasons. First, most measurements were not exact but rounded to the nearest mm or cm (or even 5 m increments in the case of height measurements for tall eucalypt species). Second, incorporating measurement error into our models should render rate variation inferences more conservative in the face of general "tip fog" (e.g., assigning trait data to the wrong tip due to taxonomic conflicts; Beaulieu and O'Meara, 2025) and/or deviations from Brownian Motion-like evolution (Landis and Schraiber, 2017).

To explore the effect of uncertain free parameter estimation on our empirical inferences, we used the R package *dentist* (Boyko and O'Meara, 2024) to sample free parameter values around our maximum likelihood estimates. For each height-dependent and constant rate model, we ran Markov chain Monte Carlo under dentist's default settings for 10,000 iterations. We used the resulting samples to construct 95% confidence intervals on predicted size-rate relationships under each height-dependent model by taking the minimum/maximum predicted rates for each height value across all sampled parameter values within their respective 95% confidence regions. We then used AICc weights to model-average the 95% confidence interval bounds for each trait.

Finally, we explored the distribution of marginal rate estimates across the phylogeny under our fitted models ("rate mapping"). To do so, we first calculated the estimated rates at each time point of the phylogeny for each height/dummy factor map under the maximum likelihood estimates for a given model. We then used the conditional likelihoods associated with each continuous stochastic character map under the model to take a weighted average of the estimated rates at each time point, yielding a single overall

rate map for a given model. Lastly, we used AICc weights to model-average the resulting rate maps across all models for a given trait and generate overall rate maps for each trait. In addition to rates maps for individual traits, we also summarized patterns of rate variation across multiple traits by rescaling the rate maps to have a mean rate of 1—thereby accounting for variation in average evolutionary rates across different traits—and averaging these "relative rates" at each time point. To quantitatively describe the size-rate associations implied by these rate maps, we used the *phylolm* R package (Ho and Ané, 2014) to estimate the relationship between mapped rates and maximum heights at the tips of the phylogeny.

Results

Validation Study

Overall, our validation study confirms our continuous stochastic character mapping algorithm correctly samples phenotypic values across a phylogeny according to their expected distribution under a given Brownian Motion model. Kolmgorov-Smirnov (KS) tests comparing observed and expected distributions of Mahalanobis distances for all simulations yielded more-or-less uniformly distributed p-values with only $\sim$5.4% of p-values falling below 0.05, as expected under the null hypothesis (i.e., the observed samples follow their expected distributions; Fig. 4). This result held across different simulation condtions, with the observed proportions of p-values less than 0.05 not significantly differing from 5% across all conditions ($p \approx 0.3$ according to a Monte Carlo χ^2 test).

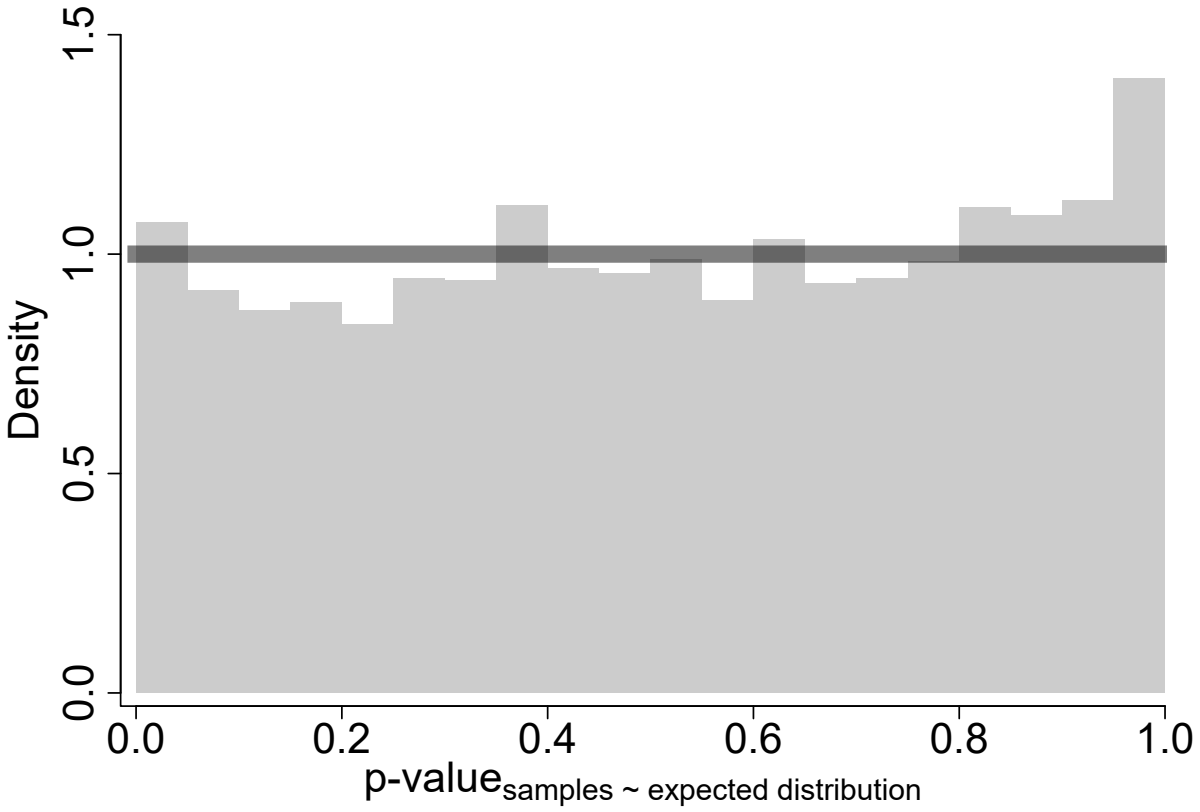

Figure 4. Distribution of p-values for Kolmgorov-Smirnov (KS) tests comparing samples generated by our continuous stochastic character mapping algorithm to their expected multivariate normal distribution under a given Brownian Motion model (using observed Mahalanobis distances and comparing to χ^2 distributions). Assuming the null hypothesis is true—that is, observed samples do follow their expected distributions—this distribution should be uniform. The thick horizontal line indicates the expected density of a uniform distribution for reference.

Closer examination of variation in KS test statistics/p-values based on particular features of different simulation conditions (e.g., number of phenotypes, amount of missing data, etc.) further confirmed a lack of significant differences, with the notable exception that p-values tended to be higher when the same parameters were used to both simulate data and generate continuous stochastic character maps/compute expected distributions. Indeed, an additional KS test suggests the overall distribution of p-values under these conditions deviated from exact uniformity due to a slightly higher-than-expected frequency of p-values around 1 ($p < 5 \times 10^{-7}$; also apparent in overall distribution of p-values in Fig. 4). In contrast, the distribution of p-values for conditions under which maps/expected

distributions were based on randomly sampled parameters appeared exactly uniform ($p \approx 0.75$). Ultimately, we suspect using identical parameters to simulate data and compute expected distributions resembles scenarios whereby expected distributions are derived from observed data, which violates a key assumption of the one-sample KS test and results in high, overly conservative p-values (Capasso et al., 2008). We therefore conclude that our continuous stochastic character mapping algorithm works correctly in general and does not noticeably break down under any of the conditions tested here.

Pipeline Simulation Study

To illustrate potential uses of continuous stochastic character maps, we performed a simulation study to investigate whether continuous stochastic character maps can be used to infer relationships between rates of continuous trait evolution and continuously-varying factors, in analogy with popular approaches for inferring relationships between rates and discretely-varying factors (e.g., Revell, 2013a). Generally, we found model selection using our new pipeline most effective when using AICc weights to summarize the relative goodness-of-fit over multiple models, rather than directly comparing AICc among individual models (see Online Appendix for further details on our model selection procedure). Here, we considered a simulated dataset to exhibit "significant" overall evidence for a factor-rate relationship if the sum of AICc weights for all factor-dependent models (excluding those based on simulated dummy factors) exceeded a cutoff. Averaging across all simulation conditions, a cutoff of 0.9 yields error (i.e., false positive) and power (i.e., true positive) rates of ∼5 and 65%, respectively (Tables 1-2; Fig. 5), while relaxing this cutoff to 0.8–0.7 increases these rates to about 7–11% and 71–76%, respectively (Tables S3-S6).

Table 1. *Proportions of times (based on 120 replicates) our continuous stochastic character map-based method for inferring relationships between rates of trait evolution and continuously-varying factors selected a model assuming a given kind of factor-rate relationship (i.e., null, exponential/simple, logistic/threshold, or Gaussian/sweetspot; corresponding to different rows) as the best-fitting one (based on a summed AICc weight cutoff of 0.9; see Tables S3-S6 for results with lower cutoffs of 0.8 and 0.7). Proportions for simulations with and without random variation around simulated factor-rate relationships (noise) are given within and without parentheses, respectively. Here, we report results for simulations with constant or hidden factor-dependent rates only and consider all models assuming either constant or dummy factor-dependent rates to be null models (see Table S2 for results with each null model separated out).*

	constant	hidden factor-dependent		
		simple	threshold	sweetspot
		50 tips		
null	0.97 (0.95)	0.95 (0.94)	0.98 (0.94)	0.95 (0.94)
simple	0.03 (0.00)	0.03 (0.02)	0.00 (0.02)	0.02 (0.01)
threshold	0.01 (0.03)	0.00 (0.02)	0.01 (0.03)	0.02 (0.02)
sweetspot	0.00 (0.02)	0.02 (0.03)	0.02 (0.02)	0.02 (0.03)
		100 tips		
null	0.99 (0.91)	0.95 (0.95)	0.94 (0.93)	0.94 (0.96)
simple	0.00 (0.01)	0.03 (0.03)	0.00 (0.00)	0.02 (0.00)
threshold	0.01 (0.03)	0.01 (0.00)	0.03 (0.03)	0.01 (0.02)
sweetspot	0.00 (0.06)	0.01 (0.03)	0.03 (0.04)	0.03 (0.03)
		200 tips		
null	0.98 (0.92)	0.96 (0.98)	0.96 (0.91)	0.98 (0.96)
simple	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)
threshold	0.01 (0.01)	0.02 (0.01)	0.01 (0.02)	0.00 (0.02)
sweetspot	0.02 (0.07)	0.03 (0.02)	0.03 (0.07)	0.03 (0.03)

Table 2. *Proportions of times (based on 120 replicates) our continuous stochastic character map-based method for inferring relationships between rates of trait evolution and continuously-varying factors selected a model assuming a given kind of factor-rate relationship (i.e., null, exponential/simple, logistic/threshold, or Gaussian/sweetspot; corresponding to different rows) as the best-fitting one (based on a summed AICc weight cutoff of 0.9; see Tables S3-S6 for results with lower cutoffs of 0.8 and 0.7). Proportions for simulations with and without random variation around factor-rate relationships (noise) are given within and without parentheses, respectively. Here, we report results for simulations with factor-dependent rates only and consider all models assuming either constant or dummy factor-dependent rates to be null models (see Table S2 for results with each null model separated out).*

	simple		threshold			sweetspot		
	strong	weak	strong	weak	wide	strong	weak	wide
				50 tips				
null	0.42 (0.40)	0.76 (0.79)	0.35 (0.45)	0.68 (0.75)	0.47 (0.69)	0.48 (0.48)	0.84 (0.83)	0.72 (0.75)
simple	0.52 (0.42)	0.19 (0.12)	0.39 (0.28)	0.21 (0.10)	0.38 (0.20)	0.04 (0.04)	0.05 (0.03)	0.09 (0.08)
threshold	0.03 (0.07)	0.01 (0.06)	0.12 (0.13)	0.04 (0.08)	0.08 (0.03)	0.04 (0.09)	0.03 (0.05)	0.06 (0.07)
sweetspot	0.03 (0.10)	0.04 (0.03)	0.14 (0.14)	0.07 (0.07)	0.07 (0.08)	0.44 (0.39)	0.08 (0.09)	0.13 (0.10)
				100 tips				
null	0.06 (0.12)	0.35 (0.64)	0.10 (0.16)	0.42 (0.54)	0.24 (0.43)	0.03 (0.14)	0.56 (0.66)	0.39 (0.56)
simple	0.82 (0.62)	0.52 (0.21)	0.19 (0.15)	0.23 (0.10)	0.38 (0.17)	0.04 (0.01)	0.04 (0.00)	0.08 (0.05)
threshold	0.04 (0.14)	0.07 (0.07)	0.40 (0.34)	0.12 (0.18)	0.18 (0.14)	0.10 (0.10)	0.02 (0.08)	0.08 (0.12)
sweetspot	0.07 (0.12)	0.07 (0.08)	0.31 (0.35)	0.22 (0.17)	0.19 (0.25)	0.83 (0.75)	0.38 (0.26)	0.44 (0.27)
				200 tips				
null	0.00 (0.02)	0.09 (0.38)	0.01 (0.03)	0.10 (0.23)	0.03 (0.15)	0.01 (0.00)	0.14 (0.32)	0.07 (0.18)
simple	0.85 (0.58)	0.72 (0.36)	0.05 (0.03)	0.23 (0.19)	0.24 (0.21)	0.00 (0.00)	0.02 (0.00)	0.03 (0.01)
threshold	0.07 (0.17)	0.07 (0.11)	0.65 (0.58)	0.36 (0.25)	0.35 (0.26)	0.05 (0.10)	0.12 (0.10)	0.13 (0.19)
sweetspot	0.08 (0.22)	0.12 (0.16)	0.29 (0.36)	0.31 (0.32)	0.38 (0.38)	0.94 (0.90)	0.72 (0.58)	0.77 (0.62)

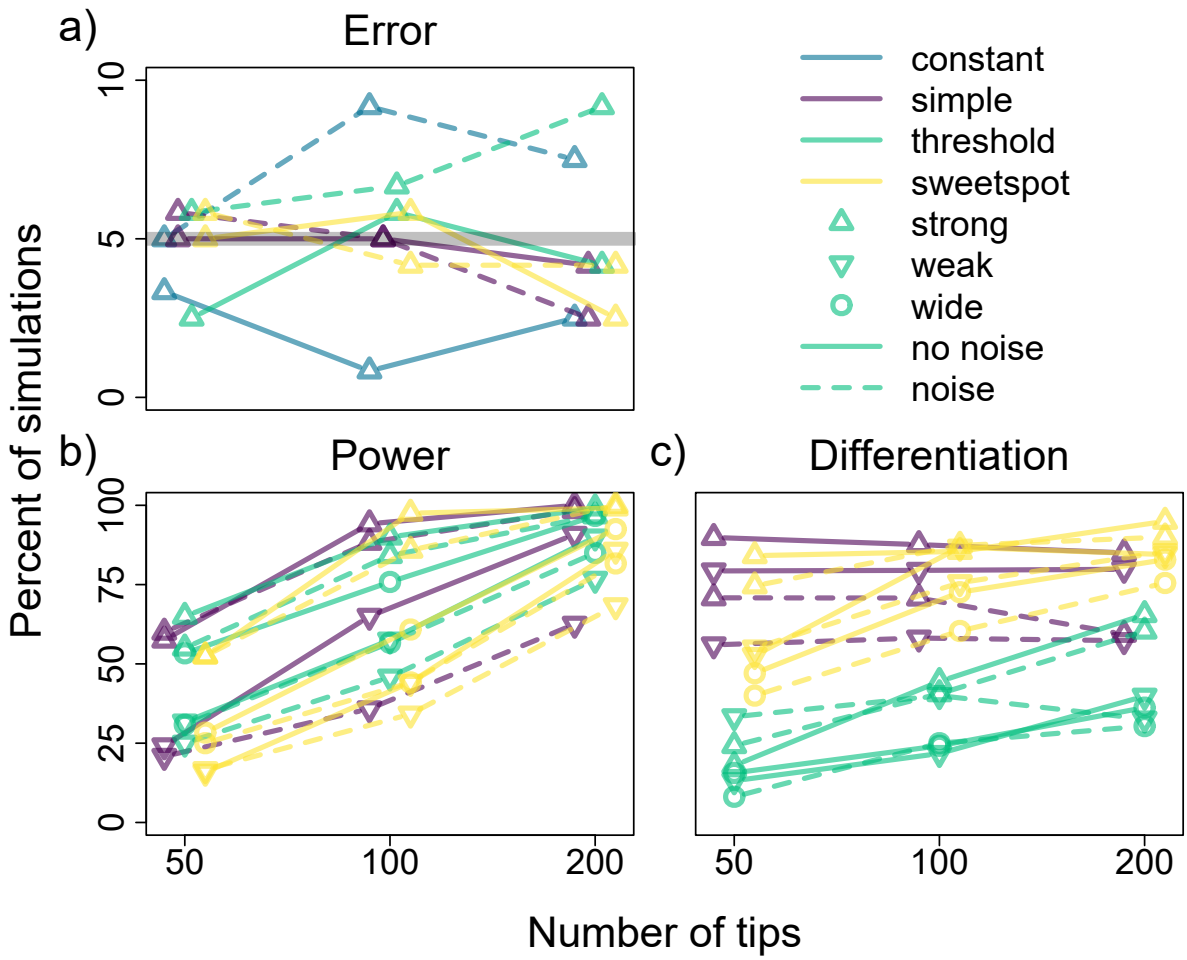

Figure 5. Error, power, and differentiation rates of our continuous stochastic character map-based pipeline for inferring relationships between rates of trait evolution and continuously-varying factors based on a summed AICc weight cutoff of 0.9 (see Figs. S1-S2 for results with lower cutoffs of 0.8 and 0.7). Different colors correspond to simulations with different factor-rate relationships (i.e., constant, exponential/simple, logistic/threshold, and Gaussian/sweetspot), different symbols to simulations with differing relationship strength and—in the case of threshold and sweetspot models—width, and dashed versus solid lines to simulations with versus without random variation in rates (noise) around simulated factor-rate relationships. Panel a: percent of simulations (120 replicates per condition) with either constant or hidden factor-dependent rates for which the best-fitting model is an observed factor-dependent model (i.e., error rates), with a thick gray line indicating a rate of 5%. Panel b: percent of simulations (120 replicates per condition) with observed factor-dependent rates for which the best-fitting model is also an observed factor-dependent model (i.e., power rates). Panel c: percent of observed-factor dependent simulations (varying number of replicates per condtion) exhibiting significant evidence for a factor-rate relationship for which the best fitting model assumed the same kind of factor-rate relationship used to simulate the data (i.e., differentiation rates).

Our pipeline exhibited consistent error rates across most simulation conditions (Fig. 5a), though error tended to decline as sample size increased (i.e., greater numbers of tips), especially when using AICc weight cutoffs lower than 0.9 (Figs. S1-S2a). Additionally,

error rates were elevated under conditions with noisy yet otherwise relatively constant evolutionary rates (e.g., constant and threshold factor-rate relationships)—presumably reflecting dummy factor-dependent models being more effective at modeling rate variation due to hidden factors than noise. Perhaps unsurprisingly, simulation conditions with perfectly constant rates exhibited relatively low error rates.

Power (Fig. 5b) varied drastically across simulation conditions, increasing with both sample size and the magnitude of factor-associated rate variation. Among simulations with strong, narrow factor-rate relationships (i.e., ∼20-fold rate variation over factor values between -2 and 2), power rates ranged from just above 50% for 50-tip simulations to over 90% for 200-tip simulations based on an AICc weight cutoff of 0.9. In contrast, under conditions with wider (i.e., ∼20-fold variation from -4 to 4) or weaker (i.e., ∼5-fold variation from -2 to 2) relationships, power dropped to around 25–50% for 50–100-tip simulations and 60–90% for 200-tip simulations. Relaxing the AICc weight cutoff to lower values increased power rather evenly across all simulation conditions (Figs. S1-S2b). Noise around simulated factor-rate relationships tended to reduce power, though this trend was generally weaker at both small *and* large sample sizes. This pattern likely reflects our pipeline's limited ability to detect rate noise at smaller sample sizes, while signals of factor-rate associations were strong enough to "overcome" any effects of noise at sufficiently large sample sizes. Comparing among different types of simulated factor-rate relationships, our pipeline typically exhibited the most power to detect factor-rate associations under simple (i.e., exponential) relationships, followed by threshold (i.e., logistic) and sweetspot (i.e., Gaussian) relationships. However, noisy and weak simple relationships were associated with exceptionally low power at large sample sizes. Additionally, the power to detect factor-rate associations under sweetspot relationships grew rapidly with sample size.

We also investigated our pipeline's ability to distinguish among different types of factor-rate relationships by comparing AICc among factor-dependent models (i.e., "differentiation rates"; Fig. 5c; note that this metric only considered simulated datasets

already exhibiting significant overall support for a factor-rate relationship based on summed AICc weights). While our approach was effective at correctly inferring simple and sweetspot relationships, it struggled to detect threshold relationships. Instead, threshold relationships were commonly mistaken for simple relationships at small sample sizes and for sweetspot relationships at large sample sizes. Our pipeline also exhibited a slight bias towards inferring more complex factor-rate relationships as sample size increased; while the ability to correctly infer threshold and sweetspot relationships improved with sample size, differentiation rates for simple relationships declined slightly. Noise also caused simple relationships to be mistaken for threshold/sweetspot relationships more frequently. Notably, while power rates were typically lowest for simulations with weak factor-rate relationships, differentiation rates were lowest for those with wide relationships. Thus, while our pipeline's power was mainly affected by the overall magnitude of factor-associated rate variation, differentiation appeared more strongly influenced by how abruptly rates change with respect to factor values.

Beyond AICc-based model selection, we explored how accurately our pipeline can infer the particular shape of factor-rate relationships based on model-averaged predictions (Fig. 6). Overall, predicted factor-rate relationships largely recapitulated simulated ones—particularly as sample size increased. However, predicted rates also exhibited shrinkage whereby low and high rates were over- and underestimated, respectively (Fig. S3). This shrinkage tended to flatten predicted factor-rate relationships, especially under conditions with limited power (e.g., small sample sizes, weak/wide relationships). However, even predictions for 200-tip simulations could exhibit minor shrinkage around factor values associated with abrupt rate shifts as well as extreme factor values in the tails of their simulated distribution (i.e., generally factor values less than -2 or greater than 2 under our simulation conditions). In contrast to predicted rates for extreme factor values, predictions at "typical" factor values towards the center of their distribution (i.e., factors values between -2 and 2) appeared fairly accurate if imprecise (Figs. S4-S5). While predicted rates

at typical factor values often varied widely across simulations with identical factor-rate relationships (∼3–30-fold in general), they only differed from simulated rates by a median 15–50%. Noise around simulated factor-rate relationships decreased the accuracy and precision of predicted rates, though this effect was small relative to that of sample size.

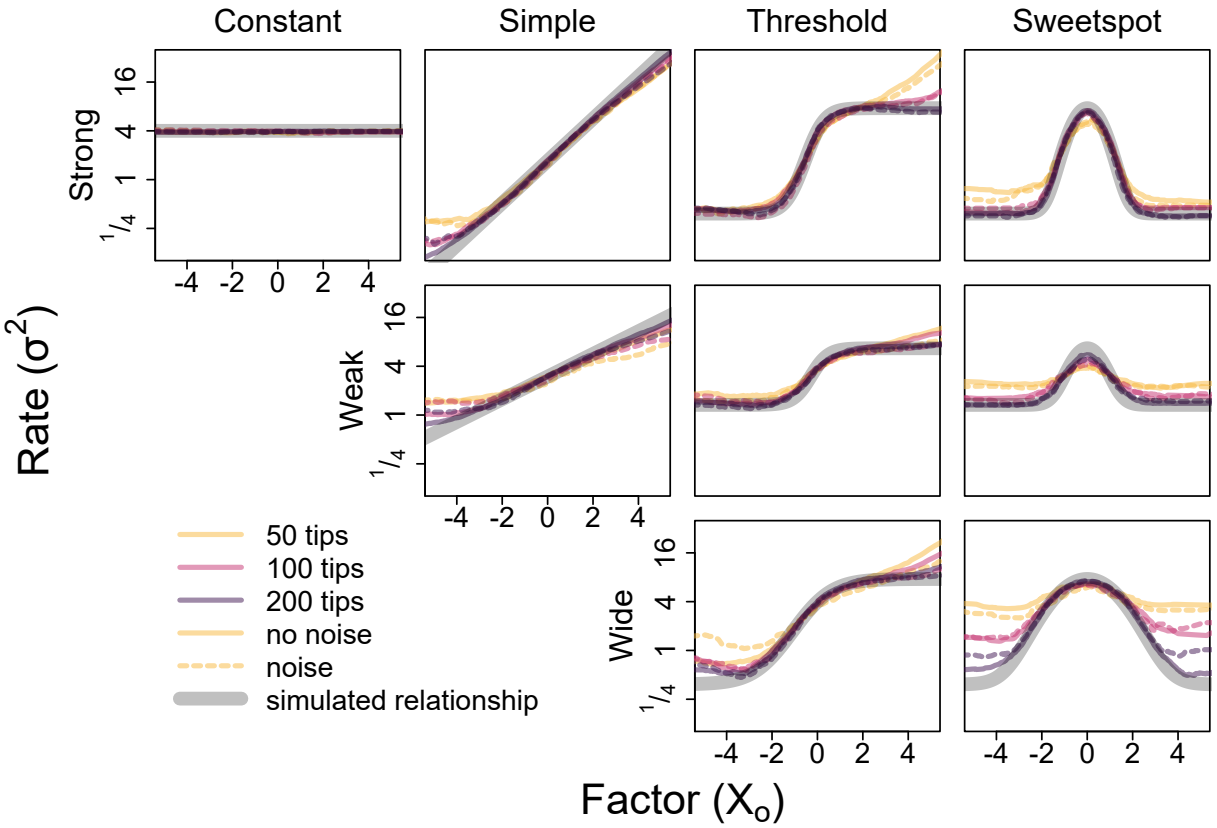

Figure 6. Median model-averaged factor-rate relationships inferred using our continuous stochastic character map-based pipeline for inferring relationships between rates of trait evolution and continuously-varying factors under all simulation conditions with either constant or observed factor-dependent rates. Different colors correspond to different phylogeny sizes (i.e., number of tips) and dashed versus solid lines to simulation conditions with versus without random variation in rates (noise) around simulated factor-rate relationships. Simulated factor-rate relationships are represented by thick gray lines for reference.

Empirical Example

In addition to the simulation study, we applied our continuous stochastic character map-based pipeline for inferring factor-dependent rates of continuous trait evolution to test whether rates of flower and leaf trait evolution scale with body size among eucalypts.

All traits investigated here exhibited variable yet consistently significant phylogenetic signal, with weak to moderate signal among leaf traits ($\lambda \sim$0.3–0.7), moderate to strong signal among flower traits ($\sim$0.6–0.95), and relatively strong signal for height, our proxy for body size ($\sim$0.85). Across all traits, dummy factor-dependent models tended to account for the majority of AICc weight, implying that rates of flower and leaf trait evolution, while variable, are mostly unrelated to size. Only bud width/length (summed AICc weight > 0.85) and adult leaf width (summed AICc weight > 0.7) exhibit notable evidence for associations between size and evolutionary rates (Fig. 7). Bud length and width both show substantial support for sweetspot size-rate relationships in particular, though the specific shape of this relationship differs between the two traits. Model-averaged predictions suggest a $\sim$3-fold peak in rates of bud length evolution among lineages around 5 m tall yet a 10-fold dip in rates of bud width evolution among lineages around 15 m tall (Fig. 8). Nonetheless, these inferences both suggest that rates of bud size evolution decrease with size between roughly 5 and 15 m in height. In contrast to results for bud size, rates of leaf width evolution exhibit a largely simple size-rate relationship, with rates decreasing by $\sim$5% per 10% increase in height.

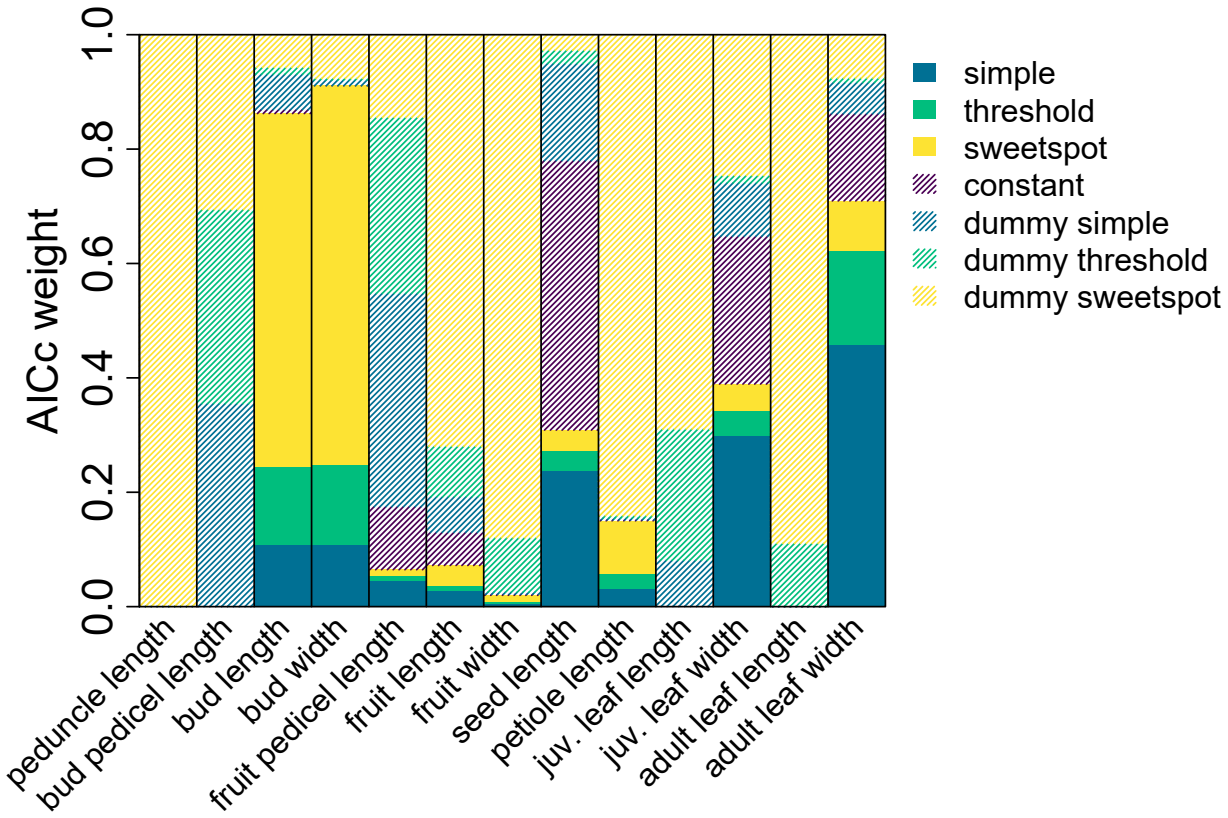

Figure 7. AICc weights of models assuming different factor-rate relationships for each eucalypt trait using our continuous stochastic character map-based pipeline (juv. = juvenile). Colors correspond to different kinds of factor-rate relationships (i.e., constant, exponential/simple, logistic/threshold, and Gaussian/sweetspot). Solid colors correspond to non-null models assuming rates depend on height while line-shaded colors represent null models assuming rates are either constant or depend on simulated dummy factors.

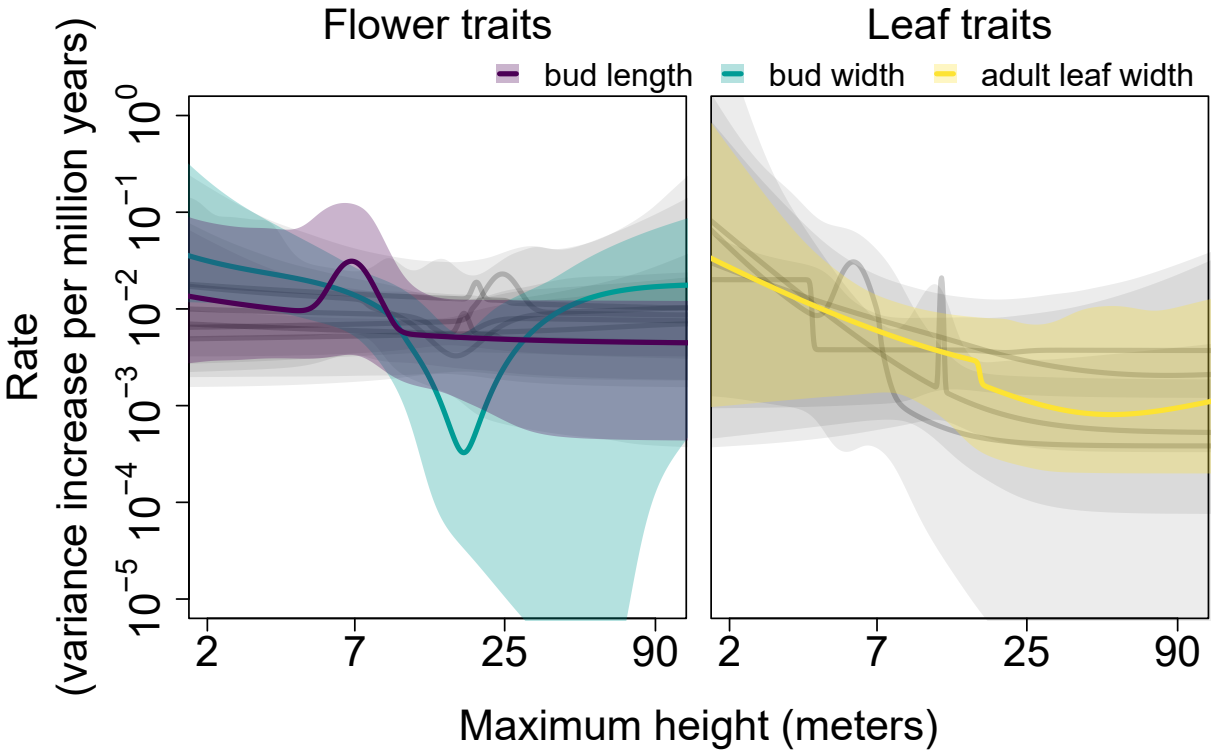

Figure 8. Model-averaged height-rate relationships for all eucalypt traits inferred using our continuous stochastic character map-based pipeline. Lines depict relationships based on maximum likelihood estimates while the sorrounding shaded regions represent 95% confidence intervals calculated using *dentist* (Boyko and O'Meara, 2024). Lines and shaded regions for traits exhibiting substantial evidence of factor-rate associations (i.e., summed AICc weight > 0.7) are drawn in color, while those for other traits are drawn in light gray. Note that all traits were log-transformed, such that rates are measure increases in *geometric* variance per million years.

While evidence for size-rate relationships was weak when considering traits individually, we consistently inferred a negative association between size and rates under simple models for all traits except inflorescence pedicel and seed length (Tables S7-S19), suggesting that rates of flower and leaf trait evolution are higher among smaller-bodied eucalypt lineages overall. We further investigated this trend by calculating fold-differences between average rates predicted for eucalypts greater and less than 15 m tall (close to the median height across all species) based on model-averaged relationships for each trait (Fig. 9). Averaging across all traits, rates of phenotypic evolution are ∼40% lower among eucalypts that grow taller than 15 m compared to those less than 15 m tall, though the

confidence intervals for individual traits overlapped with 1 (i.e., no change) in all cases except bud length/width and petiole length.

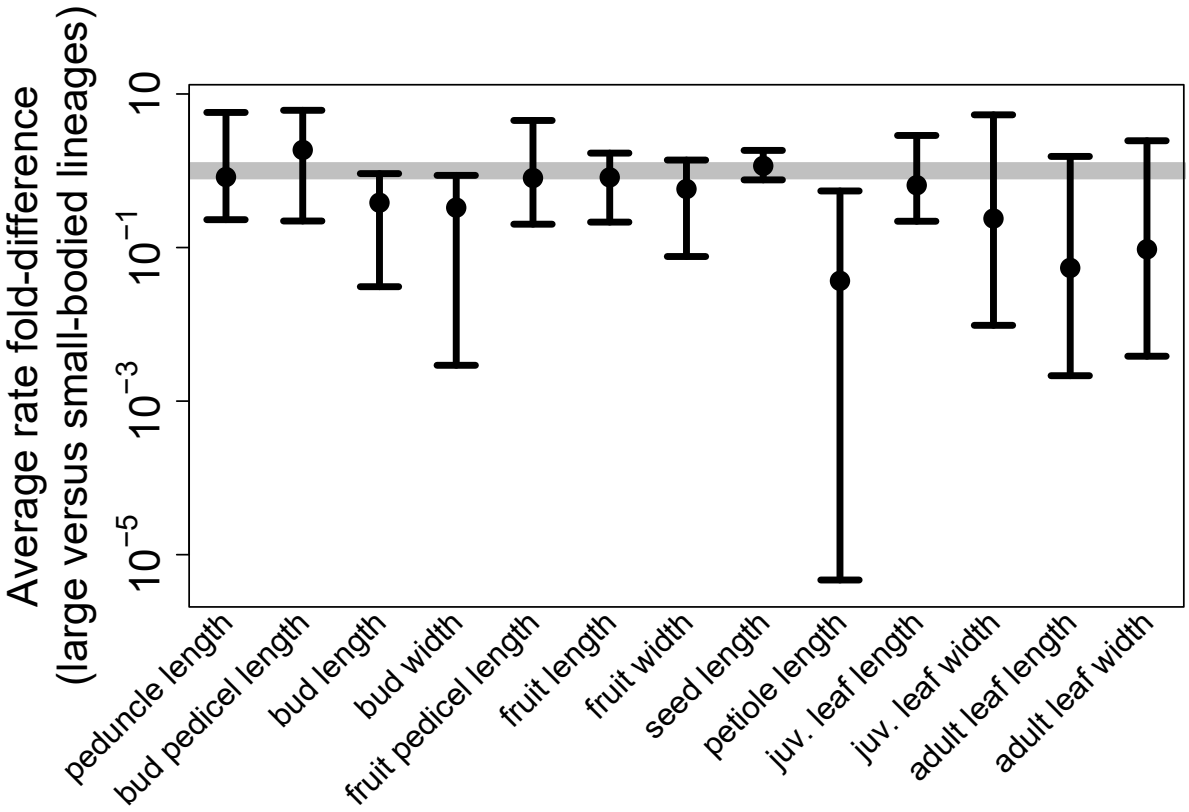

Figure 9. Model-averaged fold-differences in average rates among eucalypts greater and less than 15 meters tall for all traits inferred using our continuous stochastic character map-based pipeline. Points correspond to maximum likelihood estimates while error bars represent 95% confidence intervals on these differences calculated using *dentist* (Boyko and O'Meara, 2024). The thick gray line corresponds to a fold-difference of 1 (i.e., no change) for reference.

Despite substantial size-independent rate variation inferred by dummy factor-dependent models, mapping estimated rates onto the eucalypt phylogeny still revealed a negative association between size and rates of flower and leaf trait evolution overall, especially among smaller eucalypts growing ∼20 m tall or less (Fig. 10; see Figs. S6-S18 for individual traits and Figs. S19-S20 for separate summaries of flower and leaf traits). A phylogenetic generalized least squares regression including both a linear and quadratic term for height best explained variation in relative tip rates averaged across all

traits, suggesting overall rates of phenotypic evolution dip to a minimum around 25 m and increase by about 200% and 50% among lineages around 3 and 75 m tall, respectively. Height explained ∼15% of variation in tip rates according to this model, with much of the residual variation stemming from a relatively small number of outliers characterized by rates some 2 to 10 times higher than expected.

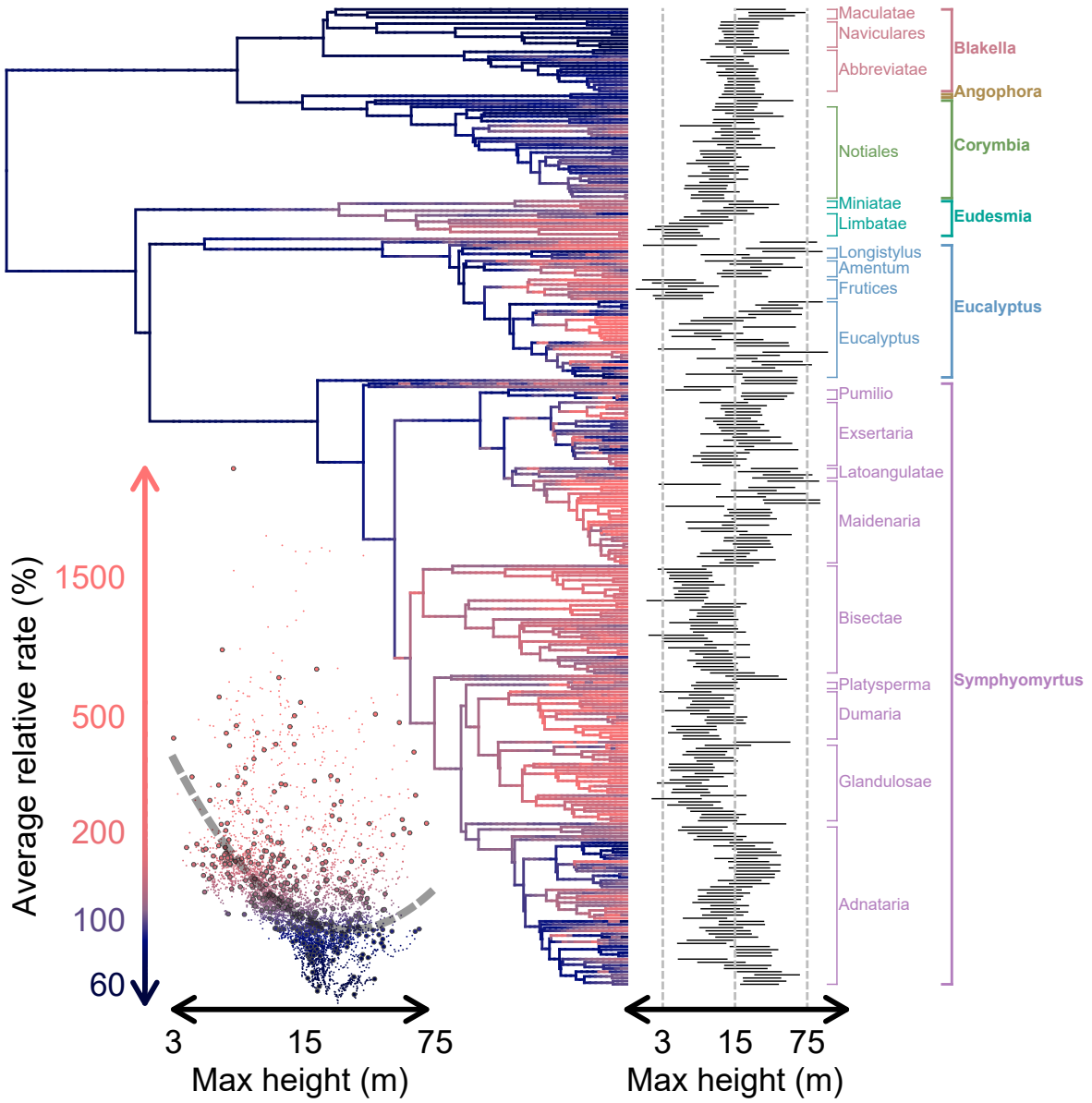

Figure 10. Relative rates of phenotypic evolution, averaged over all flower and leaf traits, mapped onto the eucalypt phylogeny, with dark blue and light red colors indicating relatively low and high rates, respectively (see y-axis of plot on bottom left for color scale; note that the color gradient is percentile-based due to the right-skewed distribution of relative rates on a log scale). Bars depicting 95% credible intervals on maximum height for each tip (inferred under an "evolving rates" or evorates model of height evolution) are arrayed along the right side of the phylogeny, along with clade labels indicating major eucalypt sections and subgenera/genera. The bottom left plot consists of mapped rates for each time point with respect to mean maximum heights (as inferred under the evorates model of height evolution), with larger points indicating tip heights/rates. The gray dashed line running through the plot represents a phylogenetic generalized least squares regression relating tip rates to maximum heights.

Discussion

In this paper we introduce continuous stochastic character mapping, a flexible and efficient tool for sampling evolutionary histories of continuous variables along phylogenies based on Brownian Motion models of trait evolution. Like conventional stochastic character mapping for discrete variables, this technique can be used to generate distributions of probable evolutionary histories for continuous variables. Beyond their use in exploring continuous trait evolutionary dynamics while accounting for uncertainty in reconstructed ancestral states, these distributions enable the development of statistical pipelines for investigating how evolutionary patterns across phylogenies are associated with continuous variables. Such approaches will help researchers address long-standing questions regarding the impacts of continuously-varying factors like body size (e.g., Friedman et al., 2019), generation time (e.g., Gingerich, 2001), and/or climatic niche (e.g., Tribble et al., 2023) on macroevolutionary dynamics. To illustrate these applications, we implemented a pipeline capable of detecting and quantifying relationships between continuously-varying factors and rates of continuous trait evolution. We applied this novel approach to an empirical example, testing whether body size is related to rates of flower and leaf trait evolution in eucalypts. While our analyses revealed substantial size-independent variation in rates of phenotypic evolution across different traits and clades, we found that smaller-bodied eucalypt lineages exhibit elevated rates overall. Ultimately, our results demonstrate how continuous stochastic character maps provide a powerful toolkit for analyzing comparative data on continuous variables.

Continuous stochastic character maps are useful for inferring factor-rate associations

Our simulation study revealed that continuous stochastic character mapping provides an effective and pragmatic approach for inferring relationships between rates of trait evolution and continuously-varying factors. Previous approaches for estimating these relationships were developed for particular empirical studies to circumvent the lack of a

standard method (e.g., Cooper and Purvis, 2009; Uyeda et al., 2021). Accordingly, the general statistical performance of these methods remains largely unknown and their utility is limited by a lack of generalizable code for implementation. The *evorag* method developed by Weir and Lawson (2015) constitutes a notable exception as a well-studied method suitable for a variety of purposes, though it only uses sister pair contrasts to infer factor-rate relationships. By using entire phylogenies, our pipeline more fully utilizes the information available in phylogenetic comparative data to infer factor-rate relationships. More recently, Hansen et al. (2022) developed a regression-based approach for inferring correlations between rates and continuous variables, though this method is only capable of inferring linear or exponential relationships between a single continuous factor and rates for a single continuous trait (or intraspecific variances, which are interpreted as "microevolutionary rates" by the authors). By comparison, our method offers more flexibility, allowing researchers to incorporate explicit reconstructions of both discrete and multiple continuous factors into analyses, fit models to multivariate continuous trait data, and/or simultaneously link multiple parameters (i.e., rates, intraspecific variances, and/or evolutionary trends) to reconstructed factors via arbitrary parameter functions (computational/statistical tractability notwithstanding). Overall, our new method is a substantial generalization of previous ones, expanding the hypothesis-testing toolkit available for macroevolutionary analyses.

While our pipeline was generally accurate and robust, it was limited under certain conditions. For one, the approach struggled to detect relatively weak factor-rate relationships from smaller phylogenetic comparative datasets. Based on the simulation conditions examined here, our pipeline typically requires phylogenies with 100 tips to consistently detect ∼20-fold differences in rates associated with a continuous factor, and 200 tips for 5-fold differences. Nonetheless, our pipeline was able to reliably reject factor-rate associations under most simulation conditions—even when rates varied due to unobserved factors. Thus, while our approach may fail to detect weak factor-rate

associations, it is generally unlikely to infer spurious associations (though, as with any correlative statistical method, associations may reflect unobserved factors confounded with observed factors/rates). Our pipeline also often failed to distinguish threshold relationships from simple or sweetspot relationships—even using large phylogenies with 200 tips. Interestingly, a previously-developed method for inferring relationships between continuously-varying factors and lineage diversification rates exhibited similar difficulties with distinguishing threshold relationships from linear ones (FitzJohn, 2010). Overall, the low power/differentiation rates of our pipeline under these conditions may reflect inherent limits to phylogenetic comparative inference of factor-rate relationships rather than shortcomings particular to our approach. Recent work suggests detecting associations between evolutionary rates and continuously-varying factors from phylogenetic comparative data may be fundamentally challenging (Hansen et al., 2022).

We also observed shrinkage of rates estimated under our pipeline in some scenarios, whereby high and low rates tend to be under- and overestimated, respectively. Shrinkage has also been found in previous work investigating methods for inferring variation in rates of trait evolution (Revell, 2013a; Martin et al., 2023). Here, shrinkage likely reflects a phenomenon first pointed out by Revell (2013a), who noted that uncertainty in factor histories may cause factor values associated with low rates to be inferred in lineages evolving at relatively high rates and vice versa. A joint inference approach, whereby the factor history and its effect on trait evolution processes are inferred simultaneously rather than sequentially, could help mitigate these issues and enable more accurate inference of factor-rate associations in the future. However, such methods entail significant mathematical, statistical, and computational challenges due to their complexity (e.g., FitzJohn, 2010, May and Moore, 2020, Boyko et al., 2023b).

Our work adds to a growing body of literature demonstrating that accounting for "background" or residual heterogeneity in evolutionary processes is often critical for robust phylogenetic comparative inference and hypothesis testing (Maddison and FitzJohn, 2015;

Beaulieu and O'Meara, 2016; Uyeda et al., 2018; May and Moore, 2020; Boyko and Beaulieu, 2023; Boyko et al., 2023b; Tribble et al., 2023). Models based on simulated dummy factors were often crucial for controlling our method's error rates by providing more competitive null models which were able to account for rate variation due to unobserved factors. However, elevated error rates under a few simulation conditions (e.g., noisy constant rates) suggest dummy factor-dependent models were less effective at accounting for primarily random rate variation that does not exhibit phylogenetic autocorrelation (i.e., noise, as opposed to the phylogenetically autocorrelated rate variation driven by observed/hidden factors in our simulations). Despite this, dummy factor-dependent models still reduced error rates quite substantially under such conditions—particularly as sample size increased (see Tables S2, S3, and S5).

Smaller eucalypts exhibit higher rates of phenotypic evolution

Using our continuous stochastic character map-based pipeline for inferring relationships between rates of trait evolution and continuous factors, we show that rates of flower and leaf trait evolution are elevated among lineages of smaller eucalypts overall. As far as we know, this is the first published analysis to investigate the relationship between body size and rates of phenotypic evolution in any plant group. Previous work has been limited to vertebrates and yielded mixed results, suggesting body shape evolution is slower among larger fish and birds (Friedman et al., 2019; Zimova et al., 2023; but see Chira et al., 2018) while body size and cranial shape evolution are generally faster among larger mammals (Cooper and Purvis, 2009; Baker et al., 2015). Nonetheless, our results agree with broad empirical evidence for accelerated molecular evolution and lineage diversification among smaller-bodied lineages (Gillooly et al., 2005; Fontanillas et al., 2007; Bromham, 2011; Wollenberg et al., 2011; Etienne et al., 2012; Lanfear et al., 2013; Weber et al., 2014; Boucher et al., 2017; Igea et al., 2017; Tedesco et al., 2017; Friedman et al., 2019; Zimova et al., 2023). Smaller plants in particular tend to exhibit reduced seed

dispersal capacity and greater microhabitat specialization, presumably leading to more frequent evolutionary divergence at finer spatial scales compared to larger plants (Boucher et al., 2017). These factors may play a particularly important role in shaping size-rate relationships across eucalypts, as their seeds are gravity-dispersed and typically travel no further than the height of their parent tree (Booth, 2017). Accordingly, the geographic ranges of most eucalypt species span 1% or less of Australia's total land area (Hughes et al., 1996). While prior comparative studies suggest both range size (Mathews and Bonser, 2005) and lineage diversification rates (Vasconcelos et al., 2022) are largely unrelated to height variation across eucalypts, such patterns are worth re-investigating given the growing availability of eucalypt trait/occurrence data (e.g., Belbin et al., 2021; Falster et al., 2021) along with increasingly well-resolved phylogenetic resources (e.g., Ferguson et al., 2023; McLay et al., 2023; Crisp et al., 2024).

Despite the overall negative association between size and rates of phenotypic evolution, our results also revealed substantial size-independent rate variation among eucalypts. A fundamental difficulty in inferring factor-rate relationships from phylogenetic comparative data is that evolutionary rates are generally influenced by countless interacting factors, yet macroevolutionary researchers are practically limited to only modeling the effect(s) of one to several factors at a time. Even well-supported factor-rate relationships tend to only account for a small portion of apparent rate variation in any given clade (e.g., Hansen et al., 2022). Using rate maps, however, we can at least gauge which clades generally supported or contradicted the overall negative association between size and rates of phenotypic evolution across eucalypts. Based on rates averaged across all traits, we inferred relatively high rates of phenotypic evolution among several clades predominantly consisting of small eucalypt species (e.g., sections Limbatae, Frutices, Bisectae, Dumaria, and Glandulosae). However, larger eucalypts do not always exhibit low rates by comparison, with particularly high rates occurring among some larger-bodied lineages in the subgenus Eucalyptus and section Maidenaria. Notably, each of these clades

exhibit exceptionally high rates of size evolution (Fig. S21) and encompass both the smallest and largest known eucalypt species (this may partially account for the inferred increase in rates among the largest eucalypt species compared to medium-sized ones). Ultimately, relatively high rates of phenotypic evolution seem to occur among eucalypt lineages of all sizes, though smaller-bodied lineages nonetheless stand out as exhibiting high rates rather consistently based on our analyses. Interestingly, while this trend is apparent across both major eucalypt clades, rates of flower and leaf trait evolution tend to be slower and less heterogeneous among *Corymbia* s.l. (consisting of (sub)genera *Angophora*, *Corymbia*, and *Blakella*) compared to *Eucalyptus* (consisting of the major subgenera Eucalyptus, Eudesmia, and Symphyomyrtus).

While our empirical investigation demonstrated the utility of our new pipeline overall, it also illustrates an important caveat regarding our model selection procedure. Broadly, our current approach appears effective at clarifying when factor-dependent rate variation exceeds factor-independent variation and vice versa. Our approach does not, however, provide results in a regression-like manner (i.e., inferring the amount of residual variation around an overall trend relating rates to a factor of interest), limiting its ability to reliably detect subtler factor-rate relationships. For example, while AICc weights provided weak to equivocal evidence for size-associated variation in rates of petiole length and juvenile leaf width evolution (Fig. 7), rate maps for the same traits revealed clear negative relationships with just a few extreme outliers (Figs. S14, S16). While careful investigation of estimated parameters and rate maps is useful for identifying cases where model selection is being overly conservative in this manner, explicitly modeling residual rate variation around inferred factor-rate relationships would likely yield more robust and powerful inferences which are less cumbersome to interpret. Unfortunately, modeling residual variation in rate parameters under continuous trait evolution models is not straightforward (which is why our pipeline uses continuous stochastic character maps in the first place), though we see this as a natural next step for improving inference of

factor-rate relationships from macroevolutionary data.

Conclusion and Future Extensions

Here, we developed a novel method for sampling histories of continuous variables on phylogenies under Brownian Motion models of trait evolution, thereby generalizing conventional discrete stochastic character mapping methods to work with continuous variables. We show how continuous stochastic character maps may be used to infer relationships between continuous trait evolution dynamics (with a focus on rates) and continuously-varying factors. Finally, we demonstrate the utility of this approach in an empirical example, showing that rates of flower and leaf trait evolution are generally higher among smaller eucalypts. Ultimately, continuous stochastic character mapping can empower researchers with new and innovative strategies for both analyzing the evolutionary dynamics of continuous phenotypes and testing hypotheses that require accounting for how continuous variables have changed across the evolutionary histories of clades.

Looking forward, continuous stochastic character mapping could be extended to sample evolutionary histories under other popular models of continuous trait evolution such as Ornstein-Uhlenbeck and Lévy processes, which are generally interpreted as models of adaptive and pulsed evolution, respectively. Another promising future direction would be the development of methods for fitting lineage diversification or discrete trait evolution models with parameters depending on mapped continuous variables. Such methods would allow macroevolutionary biologists to investigate the impact of continuously-varying factors on evolutionary dynamics with greater flexibility. Lastly, our algorithm for generating continuous stochastic character maps could benefit data augmentation schemes for Bayesian inference of complex phylogenetic comparative models involving continuously-varying factors (e.g., Lartillot and Poujol, 2011; Lartillot and Delsuc, 2012; Lartillot et al., 2016; Quintero and Landis, 2020; Quintero, 2025), as it can directly

transform samples of uncorrelated normal random variables into phenotypic evolutionary histories under Brownian Motion models.

Funding

B.S.M. was partially supported by both a Michigan State University fellowship (Summer Fellowship in Ecology, Evolution, and Behavior) and National Science Foundation fellowship [award number 2410449] in this work, while M.G.W. was partially supported by two National Science Foundation grants [grant numbers DEB-2236747 and DUE-2012014]. Additional support to both B.S.M. and M.G.W. was provided by another National Science Foundation grant [grant number DEB-1831164] awarded to M.G.W.

Acknowledgements

We would like to thank the members of the corresponding author's former dissertation committee—Luke Harmon, Gideon Bradburd, William Wetzel, and Jeffrey Conner—for their invaluable advice throughout the development of this work. We are especially grateful for their writing suggestions, which helped improve the clarity of this manuscript. Past and present members of the Leebens-Mack lab at the University of Georgia as well as the Weber and Bradburd labs at the University of Michigan (both formerly at Michigan State University) provided additional advice which was particularly helpful in refining how we communicated the work to broader audiences. We would also like to thank participants and organizers of the 2019 Nantucket phylogeny developeR workshop—in particular Nicholas Bone, Fabio Machado, and Josef Uyeda—for their helpful feedback on the very first implementation of continuous stochastic character mapping. Lastly, Isabel Sanmartín, Stacey Smith, Mike May, and two anonymous reviewers provided incredibly comprehensive and useful comments on an earlier version of this manuscript, leading to key refinements of our statistical methods, expansion of our empirical case study, and much more concise, accessible writing.

Supplementary Material

Data will be made available from a Dryad Digital Repository upon publication. In the meantime, please reach out B.S.M. (bruce.stagg.martin@gmail.com) to request any of the data and/or scripts used in this work. The Online Appendix can be found under Supplementary Material on bioRxiv, while the current version of the *contsimmap* R package is available at the GitHub repository: `https://github.com/bstaggmartin/contsimmap`.

References

Adler P.B., Salguero-Gómez R., Compagnoni A., Hsu J.S., Ray-Mukherjee J., Mbeau-Ache C., and Franco M. 2014. Functional traits explain variation in plant life history strategies. Proc Natl Acad Sci USA 111:740–745.

Amado T.F., Martinez P.A., Pincheira-Donoso D., and Olalla-Tárraga M.Á. 2021. Body size distributions of anurans are explained by diversification rates and the environment. Glob Ecol Biogeogr 30:154–164.

Baker J., Meade A., Pagel M., and Venditti C. 2015. Adaptive evolution toward larger size in mammals. Proc Natl Acad Sci USA 112:5093–5098.

Bakewell A.T., Davis K.E., Freckleton R.P., Isaac N.J.B., and Mayhew P.J. 2020. Comparing life histories across taxonomic groups in multiple dimensions: How mammal-like are insects? Am Nat 195:70–81.

Baliga V.B. and Law C.J. 2016. Cleaners among wrasses: Phylogenetics and evolutionary patterns of cleaning behavior within Labridae. Mol Phylogenet Evol 94:424–435.

Barrie F.R. 2009. *Eucalyptus deglupta*. Page 80 *in* Flora Mesoamericana (G. Davidse, M. Sousa S, S. Knapp, and F. Chiang, eds.) vol. 4(1). Universidad Nacional Autónoma de México, Missouri Botanical Garden Press, and the Natural History Museum (London), Mexico City, Mexico; St. Louis, MO; and London.

Beaulieu J. and O'Meara B. 2023. OUwie: Analysis of Evolutionary Rates in an OU Framework. R package version 2.10.

Beaulieu J.M. and O'Meara B.C. 2016. Detecting hidden diversification shifts in models of trait-dependent speciation and extinction. Syst Biol 65:583–601.

Beaulieu J.M. and O'Meara B.C. 2025. Navigating "tip fog": embracing uncertainty in tip measurements. Evolution 79:1131–1142.

Belbin L., Wallis E., Hobern D., and Zerger A. 2021. The Atlas of Living Australia: History, current state and future directions. Biodivers Data J 9:e65023.

Bollback J.P. 2006. SIMMAP: Stochastic character mapping of discrete traits on phylogenies. BMC Bioinformatics 7:88.

Booth T.H. 2017. Going nowhere fast: a review of seed dispersal in eucalypts. Aust. J. Bot. 65:401.

Borstein S.R., Fordyce J.A., O'Meara B.C., Wainwright P.C., and McGee M.D. 2019. Reef fish functional traits evolve fastest at trophic extremes. Nat Ecol Evol 3:191–199.

Boucher F.C., Verboom G.A., Musker S., and Ellis A.G. 2017. Plant size: A key determinant of diversification? New Phytol 216:24–31.

Boyko J.D. and Beaulieu J.M. 2023. Reducing the biases in false correlations between discrete characters. Syst Biol 72:476–488.

Boyko J.D., Hagen E.R., Beaulieu J.M., and Vasconcelos T. 2023a. The evolutionary responses of life-history strategies to climatic variability in flowering plants. New Phytol 240:1587–1600.

Boyko J.D. and O'Meara B.C. 2024. dentist: quantifying uncertainty by sampling points around maximum likelihood estimates. Meth Ecol Evol .

Boyko J.D., O'Meara B.C., and Beaulieu J.M. 2023b. A novel method for jointly modeling the evolution of discrete and continuous traits. Evolution 77:836–851.

Bromham L. 2011. The genome as a life-history character: Why rate of molecular evolution varies between mammal species. Philos Trans R Soc B 366:2503–2513.

Brooker I., Slee A., Connor J., Duffy S., and West J. 2015. EUCLID Eucalypts of Australia. 4 ed. Identic Pty Ltd, Brisbane, QLD.

Burns M.D. and Bloom D.D. 2020. Migratory lineages rapidly evolve larger body sizes than non-migratory relatives in ray-finned fishes. Proc R Soc B 287:20192615.

Calder W.A. 2001. Ecological consequences of body size. Pages 1–7 *in* Encyclopedia of Life Sciences. John Wiley & Sons, Ltd, Chichester, UK.

Capasso M., Alessi L., Barigozzi M., and Fagiolo G. 2008. On approximating the distributions of goodness-of-fit test statistics based on the empirical distribution function: The case of unknown parameters .

Chira A.M., Cooney C.R., Bright J.A., Capp E.J.R., Hughes E.C., Moody C.J.A., Nouri L.O., Varley Z.K., and Thomas G.H. 2018. Correlates of rate heterogeneity in avian ecomorphological traits. Ecol Lett 21:1505–1514.

Clavel J., Escarguel G., and Merceron G. 2015. mvMORPH: An R package for fitting multivariate evolutionary models to morphometric data. Methods Ecol Evol 6:1311–1319.

Clavel J. and Morlon H. 2017. Accelerated body size evolution during cold climatic periods in the Cenozoic. Proc Natl Acad Sci USA 114:4183–4188.

Cooney C.R. and Thomas G.H. 2021. Heterogeneous relationships between rates of speciation and body size evolution across vertebrate clades. Nat Ecol Evol 5:101–110.

Cooper N. and Purvis A. 2009. What factors shape rates of phenotypic evolution? A comparative study of cranial morphology of four mammalian clades. J Evol Biol 22:1024–1035.

Crisp M.D., Minh B.Q., Choi B., Edwards R.D., Hereward J., Kulheim C., Lin Y.P., Meusemann K., Thornhill A.H., Toon A., and Cook L.G. 2024. Perianth evolution and implications for generic delimitation in the eucalypts (Myrtaceae), including the description of the new genus, *Blakella*. J. Syst. Evol. 62:942–962.

de Alencar L.R.V., Martins M., Burin G., and Quental T.B. 2017. Arboreality constrains morphological evolution but not species diversification in vipers. Proc R Soc B 284:20171775.

Dobzhansky T. and Sturtevant A.H. 1938. Inversions in the chromosomes of *Drosophila pseudoobscura*. Genetics 23:28–64.

Drummond A.J., Ho S.Y.W., Phillips M.J., and Rambaut A. 2006. Relaxed phylogenetics and dating with confidence. PLoS Biol. 4:e88.

Drury J.P., Clavel J., Tobias J.A., Rolland J., Sheard C., and Morlon H. 2021. Tempo and mode of morphological evolution are decoupled from latitude in birds. PLoS Biol 19:e3001270.

Etienne R.S., de Visser S.N., Janzen T., Olsen J.L., Olff H., and Rosindell J. 2012. Can clade age alone explain the relationship between body size and diversity? Interface Focus 2:170–179.

Fabre A.C., Bardua C., Bon M., Clavel J., Felice R.N., Streicher J.W., Bonnel J., Stanley E.L., Blackburn D.C., and Goswami A. 2020. Metamorphosis shapes cranial diversity and rate of evolution in salamanders. Nat Ecol Evol 4:1129–1140.

Falster D., Gallagher R., Wenk E.H., Wright I.J., Indiarto D., Andrew S.C., Baxter C., Lawson J., Allen S., Fuchs A., Monro A., Kar F., Adams M.A., Ahrens C.W., Alfonzetti M., Angevin T., Apgaua D.M.G., Arndt S., Atkin O.K., Atkinson J., Auld T., Baker A., von Balthazar M., Bean A., Blackman C.J., Bloomfield K., Bowman D.M.J.S., Bragg J., Brodribb T.J., Buckton G., Burrows G., Caldwell E., Camac J., Carpenter R., Catford J.A., Cawthray G.R., Cernusak L.A., Chandler G., Chapman A.R., Cheal D., Cheesman A.W., Chen S.C., Choat B., Clinton B., Clode P.L., Coleman H., Cornwell W.K., Cosgrove M., Crisp M., Cross E., Crous K.Y., Cunningham S., Curran T., Curtis E., Daws M.I., DeGabriel J.L., Denton M.D., Dong N., Du P., Duan H., Duncan D.H.,

Duncan R.P., Duretto M., Dwyer J.M., Edwards C., Esperon-Rodriguez M., Evans J.R., Everingham S.E., Farrell C., Firn J., Fonseca C.R., French B.J., Frood D., Funk J.L., Geange S.R., Ghannoum O., Gleason S.M., Gosper C.R., Gray E., Groom P.K., Grootemaat S., Gross C., Guerin G., Guja L., Hahs A.K., Harrison M.T., Hayes P.E., Henery M., Hochuli D., Howell J., Huang G., Hughes L., Huisman J., Ilic J., Jagdish A., Jin D., Jordan G., Jurado E., Kanowski J., Kasel S., Kellermann J., Kenny B., Kohout M., Kooyman R.M., Kotowska M.M., Lai H.R., Laliberté E., Lambers H., Lamont B.B., Lanfear R., van Langevelde F., Laughlin D.C., Laugier-Kitchener B.A., Laurance S., Lehmann C.E.R., Leigh A., Leishman M.R., Lenz T., Lepschi B., Lewis J.D., Lim F., Liu U., Lord J., Lusk C.H., Macinnis-Ng C., McPherson H., Magallón S., Manea A., López-Martinez A., Mayfield M., McCarthy J.K., Meers T., van der Merwe M., Metcalfe D.J., Milberg P., Mokany K., Moles A.T., Moore B.D., Moore N., Morgan J.W., Morris W., Muir A., Munroe S., Nicholson Á., Nicolle D., Nicotra A.B., Niinemets Ü., North T., O'Reilly-Nugent A., O'Sullivan O.S., Oberle B., Onoda Y., Ooi M.K.J., Osborne C.P., Paczkowska G., Pekin B., Guilherme Pereira C., Pickering C., Pickup M., Pollock L.J., Poot P., Powell J.R., Power S.A., Prentice I.C., Prior L., Prober S.M., Read J., Reynolds V., Richards A.E., Richardson B., Roderick M.L., Rosell J.A., Rossetto M., Rye B., Rymer P.D., Sams M.A., Sanson G., Sauquet H., Schmidt S., Schönenberger J., Schulze E.D., Sendall K., Sinclair S., Smith B., Smith R., Soper F., Sparrow B., Standish R.J., Staples T.L., Stephens R., Szota C., Taseski G., Tasker E., Thomas F., Tissue D.T., Tjoelker M.G., Tng D.Y.P., de Tombeur F., Tomlinson K., Turner N.C., Veneklaas E.J., Venn S., Vesk P., Vlasveld C., Vorontsova M.S., Warren C.A., Warwick N., Weerasinghe L.K., Wells J., Westoby M., White M., Williams N.S.G., Wills J., Wilson P.G., Yates C., Zanne A.E., Zemunik G., and Ziemińska K. 2021. AusTraits, a curated plant trait database for the Australian flora. Sci Data 8:254.

Feldman A., Sabath N., Pyron R.A., Mayrose I., and Meiri S. 2016. Body sizes and diversification rates of lizards, snakes, amphisbaenians and the tuatara. Glob Ecol

Biogeogr 25:187–197.

Felsenstein J. 2008. Comparative methods with sampling error and within-species variation: Contrasts revisited and revised. Am Nat 171:713–725.

Felsenstein J. 2012. A comparative method for both discrete and continuous characters using the threshold model. Am Nat 179:145–156.

Ferguson S., Jones A., Murray K., Andrew R., Schwessinger B., and Borevitz J. 2023. Plant genome evolution in the genus *Eucalyptus* driven by structural rearrangements that promote sequence divergence.

FitzJohn R.G. 2010. Quantitative traits and diversification. Syst Biol 59:619–633.

FitzJohn R.G., Maddison W.P., and Otto S.P. 2009. Estimating trait-dependent speciation and extinction rates from incompletely resolved phylogenies. Syst Biol 58:595–611.

Fontanillas E., Welch J.J., Thomas J.A., and Bromham L. 2007. The influence of body size and net diversification rate on molecular evolution during the radiation of animal phyla. BMC Evol Biol 7:95.

Freyman W.A. and Höhna S. 2019. Stochastic character mapping of state-dependent diversification reveals the tempo of evolutionary decline in self-compatible Onagraceae lineages. Syst Biol 68:505–519.

Friedman S.T., Martinez C.M., Price S.A., and Wainwright P.C. 2019. The influence of size on body shape diversification across Indo-Pacific shore fishes. Evolution 73:1873–1884.

Friedman S.T. and Muñoz M.M. 2023. A latitudinal gradient of deep-sea invasions for marine fishes. Nat Commun 14:773.

Gillooly J.F., Allen A.P., West G.B., and Brown J.H. 2005. The rate of DNA evolution: Effects of body size and temperature on the molecular clock. Proc Natl Acad Sci USA 102:140–145.

Gingerich P.D. 2001. Rates of evolution on the time scale of the evolutionary process. Genetica 112-113:127–144.

Gingerich P.D. 2009. Rates of evolution. Annu Rev Ecol Evol Syst 40:657–675.

Goolsby E.W. 2017. Rapid maximum likelihood ancestral state reconstruction of continuous characters: A rerooting-free algorithm. Ecol Evol 7:2791–2797.

Groussin M., Daubin V., Gouy M., and Tannier E. 2016. Ancestral reconstruction: Theory and practice. Pages 70–77 *in* Encyclopedia of Evolutionary Biology (R. M. Kliman, ed.). Academic Press, Waltham, MA.

Hansen T.F. and Bartoszek K. 2012. Interpreting the evolutionary regression: the interplay between observational and biological errors in phylogenetic comparative studies. Syst. Biol. 61:413–425.

Hansen T.F., Bolstad G.H., and Tsuboi M. 2022. Analyzing disparity and rates of morphological evolution with model-based phylogenetic comparative methods. Syst Biol 71:1054–1072.

Hansen T.F. and Martins E.P. 1996. Translating between microevolutionary process and macroevolutionary patterns: The correlation structure of interspecific data. Evolution 50:1404–1417.

Hansen T.F., Pienaar J., and Orzack S.H. 2008. A comparative method for studying adaptation to a randomly evolving environment. Evolution 62:1965–1977.

Harvey M.G. and Rabosky D.L. 2017. Continuous traits and speciation rates: alternatives to state-dependent diversification models. Methods in Ecology and Evolution 9:984–993.

Hassler G., Tolkoff M.R., Allen W.L., Ho L.S.T., Lemey P., and Suchard M.A. 2022. Inferring phenotypic trait evolution on large trees with many incomplete measurements. J Am Stat Assoc 117:678–692.

Hiscott G., Fox C., Parry M., and Bryant D. 2016. Efficient recycled algorithms for quantitative trait models on phylogenies. Genome Biol. Evol. 8:1338–1350.

Ho L.s.T. and Ané C. 2014. A linear-time algorithm for Gaussian and non-Gaussian trait evolution models. Syst Biol 63:397–408.

Horvilleur B. and Lartillot N. 2014. Monte Carlo algorithms for Brownian phylogenetic models. Bioinformatics 30:3020–3028.

Huelsenbeck J.P., Nielsen R., and Bollback J.P. 2003. Stochastic mapping of morphological characters. Syst Biol 52:131–158.

Hughes J.J., Berv J.S., Chester S.G.B., Sargis E.J., and Field D.J. 2021. Ecological selectivity and the evolution of mammalian substrate preference across the K-Pg boundary. Ecol Evol 11:14540–14554.

Hughes L., Cawsey E.M., and Westoby M. 1996. Geographic and climatic range sizes of australian eucalypts and a test of rapoport's rule. Global Ecol. Biogeogr. Lett. 5:128.

Igea J., Miller E.F., Papadopulos A.S.T., and Tanentzap A.J. 2017. Seed size and its rate of evolution correlate with species diversification across angiosperms. PLoS Biol 15:e2002792.

Jhwueng D.C. 2021. Two Gaussian bridge processes for mapping continuous trait evolution along phylogenetic trees. Mathematics 9:1998.

Johnson S.G. 2021. The NLopt nonlinear-optimization package. Version 2.7.1.

Joy J.B., Liang R.H., McCloskey R.M., Nguyen T., and Poon A.F.Y. 2016. Ancestral reconstruction. PLoS Comput Biol 12:e1004763.

Karatzas I. and Shreve S.E. 2004. Brownian motion and stochastic calculus. Graduate texts in mathematics 2 ed. Springer, New York, NY.

Landis M.J., Eaton D.A.R., Clement W.L., Park B., Spriggs E.L., Sweeney P.W., Edwards E.J., and Donoghue M.J. 2021. Joint phylogenetic estimation of geographic movements and biome shifts during the global diversification of *Viburnum*. Syst Biol 70:67–85.

Landis M.J., Matzke N.J., Moore B.R., and Huelsenbeck J.P. 2013. Bayesian analysis of biogeography when the number of areas is large. Syst Biol 62:789–804.

Landis M.J. and Schraiber J.G. 2017. Pulsed evolution shaped modern vertebrate body sizes. Proc Natl Acad Sci USA 114:13224–13229.

Lanfear R., Ho S.Y.W., Jonathan Davies T., Moles A.T., Aarssen L., Swenson N.G., Warman L., Zanne A.E., and Allen A.P. 2013. Taller plants have lower rates of molecular evolution. Nat Commun 4:1879.

Lartillot N. and Delsuc F. 2012. Joint reconstruction of divergence times and life-history evolution in placental mammals using a phylogenetic covariance model. Evolution 66:1773–1787.

Lartillot N., Phillips M.J., and Ronquist F. 2016. A mixed relaxed clock model. Philos Trans R Soc B 371:20150132.

Lartillot N. and Poujol R. 2011. A phylogenetic model for investigating correlated evolution of substitution rates and continuous phenotypic characters. Mol Biol Evol 28:729–744.

Limpert E., Stahel W.A., and Abbt M. 2001. Log-normal distributions across the sciences: Keys and clues. Bioscience 51:341–352.

Maddison W.P. and FitzJohn R.G. 2015. The unsolved challenge to phylogenetic correlation tests for categorical characters. Syst Biol 64:127–136.

Marsaglia G., Tsang W.W., and Wang J. 2003. Evaluating kolmogorov's distribution. J. Stat. Softw. 8:1–4.

Martin B.S., Bradburd G.S., Harmon L.J., and Weber M.G. 2023. Modeling the evolution of rates of continuous trait evolution. Syst Biol 72:590–605.

Mathews S. and Bonser S.P. 2005. Life histories, ecological tolerance limits, and the evolution of geographic range size in *Eucalyptus* (Myrtaceae). Aust J Bot 53:501–508.

May M.R. and Moore B.R. 2020. A Bayesian approach for inferring the impact of a discrete character on rates of continuous-character evolution in the presence of background-rate variation. Syst Biol 69:530–544.

Mayrose I. and Otto S.P. 2011. A likelihood method for detecting trait-dependent shifts in the rate of molecular evolution. Mol. Biol. Evol. 28:759–770.

McLay T.G.B., Fowler R.M., Fahey P.S., Murphy D.J., Udovicic F., Cantrill D.J., and Bayly M.J. 2023. Phylogenomics reveals extreme gene tree discordance in a lineage of dominant trees: Hybridization, introgression, and incomplete lineage sorting blur deep evolutionary relationships despite clear species groupings in *Eucalyptus* subgenus *Eudesmia*. Mol Phylogenet Evol 187:107869.

Nations J.A., Mount G.G., Morere S.M., Achmadi A.S., Rowe K.C., and Esselstyn J.A. 2021. Locomotory mode transitions alter phenotypic evolution and lineage diversification in an ecologically rich clade of mammals. Evolution 75:376–393.

Nicolle D. 2024. Classification of the eucalypts, genus *Eucalyptus*. version 7.1. `https://www.dn.com.au/Classification-Of-The-Eucalypts.pdf` accessed: 2025-8-10.

Nicolle D. and French M.E. 2021. A taxonomic revision of the semi-cryptic narrow-leaved mallees (*Eucalyptus* series Porantherae, Myrtaceae) from southern Australia. Swainsona 35:73–152.

Nielsen R. 2002. Mapping mutations on phylogenies. Syst Biol 51:729–739.

Niklas K.J. and Enquist B.J. 2001. Invariant scaling relationships for interspecific plant biomass production rates and body size. Proc Natl Acad Sci USA 98:2922–2927.

Parnell J.A.N. and Chantaranotahi P. 2002. *Eucalyptus deglupta*. Page 795 *in* Flora of Thailand (T. Santisuk, K. Larsen, I. Nielsen, K. Chayamarit, I. Hedge, J. Dransfield, B. Hansen, C. Phengkhlai, J. Vidal, K. Iwatsuki, J. Parnell, and M. Roos, eds.) vol. 7(4). The Forest Herbarium, Royal Forest Herbarium, Bangkok, Thailand.

Pauling L., Zuckerkandl E., Henriksen T., and Lövstad R. 1963. Chemical paleogenetics: Molecular "restoration studies" of extinct forms of life. Acta Chem Scand 17:9–16.

Pennell M.W., Eastman J.M., Slater G.J., Brown J.W., Uyeda J.C., FitzJohn R.G., Alfaro M.E., and Harmon L.J. 2014. geiger v2.0: An expanded suite of methods for fitting macroevolutionary models to phylogenetic trees. Bioinformatics 30:2216–2218.

Petersen K.B. and Pedersen M.S. 2012. The matrix cookbook. Version 20121115.

Pool A. 2001. *Eucalyptus*. Pages 1568–1569 *in* Flora de Nicaragua (W. D. Stevens, C. U. Ulloa, A. Pool, and O. M. Montiel, eds.) vol. 85(2). Missouri Botanical Garden Press, Saint Louis, MI.

Quintero I. 2025. The diffused evolutionary dynamics of morphological novelty. Proc Natl Acad Sci USA 122:e2425573122.

Quintero I. and Landis M.J. 2020. Interdependent phenotypic and biogeographic evolution driven by biotic interactions. Syst Biol 69:739–755.

Revell L.J. 2012. phytools: an R package for phylogenetic comparative biology (and other things). Methods Ecol Evol 3:217–223.

Revell L.J. 2013a. A comment on the use of stochastic character maps to estimate evolutionary rate variation in a continuously valued trait. Syst Biol 62:339–345.

Revell L.J. 2013b. Two new graphical methods for mapping trait evolution on phylogenies. Methods Ecol Evol 4:754–759.

Rincon-Sandoval M., Duarte-Ribeiro E., Davis A.M., Santaquiteria A., Hughes L.C., Baldwin C.C., Soto-Torres L., Acero P A., Walker H.J. Jr, Carpenter K.E., Sheaves M., Ortí G., Arcila D., and Betancur-R R. 2020. Evolutionary determinism and convergence associated with water-column transitions in marine fishes. Proc Natl Acad Sci USA 117:33396–33403.

Rowan T.H. 1990. Functional stability analysis of numerical algorithms. Ph.D. thesis Department of Computer Science, University of Texas at Austin, TX.

Salguero-Gómez R., Jones O.R., Jongejans E., Blomberg S.P., Hodgson D.J., Mbeau-Ache C., Zuidema P.A., de Kroon H., and Buckley Y.M. 2016. Fast-slow continuum and reproductive strategies structure plant life-history variation worldwide. Proc Natl Acad Sci USA 113:230–235.

Sanger F., Thompson E.O., and Kitai R. 1955. The amide groups of insulin. Biochem J 59:509–518.

Schluter D., Price T., Mooers A.Ø., and Ludwig D. 1997. Likeliood of ancestor state in adaptive radiation. Evolution 51:1699–1711.

Schultz T.R., Cocroft R.B., and Churchill G.A. 1996. The reconstruction of ancestral character states. Evolution 50:504–511.

Sibly R.M. and Brown J.H. 2007. Effects of body size and lifestyle on evolution of mammal life histories. Proc Natl Acad Sci USA 104:17707–17712.

Siqueira A.C., Muruga P., and Bellwood D.R. 2023. On the evolution of fish-coral interactions. Ecol Lett 26:1348–1358.

Slater G.J., Goldbogen J.A., and Pyenson N.D. 2017. Independent evolution of baleen whale gigantism linked to Plio-Pleistocene ocean dynamics. Proc R Soc B 284:20170546.

Slater G.J., Harmon L.J., and Alfaro M.E. 2012. Integrating fossils with molecular phylogenies improves inference of trait evolution. Evolution 66:3931–3944.

Sumrall C.D. and Brochu C.A. 2008. Viewing paleobiology through the lens of phylogeny. Paleontol Soc Papers 14:165–183.

Tedesco P.A., Paradis E., Lévêque C., and Hugueny B. 2017. Explaining global-scale diversification patterns in actinopterygian fishes. J Biogeogr 44:773–783.

Tornabene L., Van Tassell J.L., Robertson D.R., and Baldwin C.C. 2016. Repeated invasions into the twilight zone: Evolutionary origins of a novel assemblage of fishes from deep Caribbean reefs. Mol Ecol 25:3662–3682.

Tribble C.M., May M.R., Jackson-Gain A., Zenil-Ferguson R., Specht C.D., and Rothfels C.J. 2023. Unearthing modes of climatic adaptation in underground storage organs across Liliales. Syst Biol 72:198–212.

Uyeda J.C., Bone N., McHugh S., Rolland J., and 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., Zenil-Ferguson R., and Pennell M.W. 2018. Rethinking phylogenetic comparative methods. Syst Biol 67:1091–1109.

Vasconcelos T., O'Meara B.C., and Beaulieu J.M. 2022. A flexible method for estimating tip diversification rates across a range of speciation and extinction scenarios. Evolution 76:1420–1433.

Weber C.C., Nabholz B., Romiguier J., and Ellegren H. 2014. Kr/Kc but not dN/dS correlates positively with body mass in birds, raising implications for inferring lineage-specific selection. Genome Biol. 15:542.

Weir J.T. and Lawson A. 2015. Evolutionary rates across gradients. Methods Ecol Evol 6:1278–1286.

Welch J.J. and Waxman D. 2008. Calculating independent contrasts for the comparative study of substitution rates. J Theor Biol 251:667–678.

White E.P., Ernest S.K.M., Kerkhoff A.J., and Enquist B.J. 2007. Relationships between body size and abundance in ecology. Trends Ecol Evol 22:323–330.

Witmer L. 1995. The extant phylogenetic bracket and the importance of reconstructing soft tissues in fossils. Pages 19–33 *in* Functional Morphology in Vertebrate Paleontology (J. J. Thomason, ed.). Cambridge University Press, New York, NY.

Wollenberg K.C., Vieites D.R., Glaw F., and Vences M. 2011. Speciation in little: The role of range and body size in the diversification of Malagasy mantellid frogs. BMC Evol Biol 11:217.

Yang Z. 2006. Computational Molecular Evolution. Oxford University Press, New York, NY.

Ypma J., Johnson S.G., Stamm A., Borchers H.W., Eddelbuettel D., Ripley B., Hornik K., Chiquet J., Adler A., Dai X., and Ooms J. 2022. nlotpr: R Interface to NLOPT. R package version 2.0.3.

Zimova M., Weeks B.C., Willard D.E., Giery S.T., Jirinec V., Burner R.C., and Winger B.M. 2023. Body size predicts the rate of contemporary morphological change in birds. Proc Natl Acad Sci USA 120:e2206971120.