

Detecting Environment-Dependent Diversification From Phylogenies: A Simulation Study and Some Empirical Illustrations

ERIC LEWITUS* and HÉLÈNE MORLON

IBENS, Département de Biologie, Ecole Normale Supérieure, CNRS, Inserm, PSL Research University, F-75005 Paris, France

*Correspondence to be sent to: IBENS, Département de Biologie, Ecole Normale Supérieure, CNRS,

Inserm, PSL Research University, F-75005 Paris, France

E-mail: lewitus@biologie.ens.fr.

Received 26 January 2017; reviews returned 6 December 2017; accepted 11 December 2017

Associate Editor: Richard Ree

Abstract.—Understanding the relative influence of various abiotic and biotic variables on diversification dynamics is a major goal of macroevolutionary studies. Recently, phylogenetic approaches have been developed that make it possible to estimate the role of various environmental variables on diversification using time-calibrated species trees, paleoenvironmental data, and maximum-likelihood techniques. These approaches have been effectively employed to estimate how speciation and extinction rates vary with key abiotic variables, such as temperature and sea level, and we can anticipate that they will be increasingly used in the future. Here we compile a series of biotic and abiotic paleodatasets that can be used as explanatory variables in these models and use simulations to assess the statistical properties of the approach when applied to these paleodatasets. We demonstrate that environment-dependent models perform well in recovering environment-dependent speciation and extinction parameters, as well as in correctly identifying the simulated environmental model when speciation is environment-dependent. We explore how the strength of the environment-dependency, tree size, missing taxa, and characteristics of the paleoenvironmental curves influence the performance of the models. Finally, using these models, we infer environment-dependent diversification in two empirical phylogenies: temperature-dependence in Cetacea and $\delta^{13}\text{C}$ -dependence in Ruminantia. We illustrate how to evaluate the relative importance of abiotic and biotic variables in these two clades and interpret these results in light of macroevolutionary hypotheses. Given the important role paleoenvironments are presumed to have played in species evolution, our statistical assessment of how environment-dependent models behave is crucial for their utility in macroevolutionary analysis. [Biodiversity; diversification; macroevolution; paleoenvironment; phylogenetics.]

For decades, evolutionary biologists and paleontologists have debated the relative role of major biotic and abiotic drivers of macroevolutionary dynamics (Smith 1989; Vrba 1993). The view that links biodiversity to abiotic change, which has been apparent to biologists since the 19th century (von Humboldt 1860; Wallace 1878) and is generally referred to as the Court Jester hypothesis (Barnosky 2001), sees shifts in diversification as timed with climatic or geologic events. The broad-scale synchrony between global biodiversity and temperature throughout the Phanerozoic (Sepkoski 1984; Mayhew et al. 2008) supports this view. More focused observations, such as the rise of angiosperms during the early Cretaceous thermal maximum (Spicer and Chapman 1990; Coiffard et al. 2012), also find support for close ties between abiotic change and species diversification. The view that links biodiversity shifts to biotic interactions, on the other hand, as outlined, for example, in Leigh Van Valen's Red Queen hypothesis (Van Valen 1973) and Ehrlich and Raven's Escape and Radiate hypothesis (Ehrlich and Raven 1964), suggests that interspecific interactions (e.g., predation, mutualism, and species recognition) drive species diversification (Liow et al. 2011). While these various abiotic and biotic factors offer fundamentally different interpretations of how species evolve over geological time, evaluating their relative importance in practice is not trivial and has been inhibited by methodological shortcomings.

Investigations into the drivers of species diversification have been conducted using fossil

specimens and phylogenetic comparative methods. Time-series of fossil specimens have been used to study speciation and extinction rates through time (Roy et al. 2009; Silvestro et al. 2014), the timing and effect of mass extinction events on species richness (Raup and Sepkoski 1982; Erwin 2008; Chen and Benton 2012; Liow et al. 2015), and correlations between environmental changes and shifts in biodiversity (Benton 2009; Ezard et al. 2011; Hannisdal and Peters 2011). Fossil studies remain key to our understanding of macroevolutionary dynamics, but they are hindered by incomplete sampling (Smith 2007), clumping of specimens near mass extinction events (Holland and Patzkowsky 2015), and high sensitivity to taxonomic and spatial scale (Mittelbach et al. 2007; Erwin 2009). Importantly, fossil studies are usually restricted to the few lineages for which sufficient fossil information is available (Benton 1995). Phylogenetic comparative methods, which tender statistical or probabilistic information from species trees, provide an alternative approach to the same type of questions, applicable to a more diverse set of organisms. They have been used to study patterns of speciation and extinction in clades (Nee et al. 1994; Morlon 2014), as well as how these patterns of diversification are related to paleoenvironmental drivers (Steeiman et al. 2009; Winkler et al. 2010; Condamine et al. 2012; Cantalapiedra et al. 2014; Kergoat et al. 2014; Claramunt and Cracraft 2015; Condamine et al. 2015; Pérez-Escobar et al. 2017).

Until recently, phylogenetic approaches for analyzing the role of paleoenvironmental drivers relied on

correlative techniques (Steeman et al. 2009; Winkler et al. 2010; Condamine et al. 2012), often assuming that environmental changes are punctuated events. This has been improved in the last couple of years, with the development of “environment-dependent” likelihood-based approaches (Condamine et al. 2013), implemented in the R package RPANDA (Morlon et al. 2016), that allow testing whether gradual changes in paleoenvironments had a significant influence on speciation and extinction rates, as well as quantifying the direction and magnitude of this potential influence. Environment-dependent models take the form of time-dependent birth–death models (Morlon et al. 2011), but where the speciation and extinction rates, $\lambda(t)$ and $\mu(t)$, respectively, at a given time, t , can depend on measured time-variable palaeodata, $E(t)$. In order to test the influence of the paleoenvironment on diversification, hypotheses are made about the functional form of the dependency. Two main dependencies have been used: a linear dependency, of the form:

$$\lambda(t) = \lambda_0 + \alpha_\lambda t \quad (1)$$

and

$$\mu(t) = \mu_0 + \alpha_\mu t, \quad (2)$$

and an exponential dependency of the form

$$\lambda(t) = \lambda_0 e^{\alpha_\lambda t} \quad (3)$$

and

$$\mu(t) = \mu_0 e^{\alpha_\mu t} \quad (4)$$

where λ_0 , μ_0 , α_λ , and α_μ are free parameters. The probability density of observing a set of branching times in a phylogeny with a specific environmental dependence can then be computed using the likelihood expressions derived for time-dependent models, adapted to accommodate a time-dependent environmental variable (Morlon et al. 2011; Condamine et al. 2013). The significance of the environmental dependence, its form (e.g. linear or exponential), its sign (positive or negative), and its strength can thus be assessed using maximum likelihood techniques.

Likelihood-based phylogenetic comparative methods have been essential aids to uncovering patterns of macroevolutionary change (Morlon 2014), and we can thus anticipate that the environment-dependent models will be of great use for understanding the relative role of major biotic and abiotic drivers on deep-time diversity dynamics. These models have already identified, for example, a positive relationship between speciation rates and global temperature in ruminants in the Cenozoic (Cantalapiedra et al. 2014), an inverse relationship between net diversification rates and temperature in bird species since the K-Pg boundary (Claramunt and Cracraft 2015), a positive relationship between extinction rates and sea level in birdwing butterflies (Condamine et al. 2015), and a positive relationship between speciation rates and Andean uplift in neotropical orchids (Pérez-Escobar et al. 2017). However, we can

only have a relative confidence in these empirical results, because the performances of the environment-dependent models have not been formally assessed. We can be confident that the likelihood expressions are correct, because the environment-dependent models are straightforward extensions of time-dependent models that have themselves been thoroughly tested using simulations (Morlon et al. 2011; FitzJohn 2012); but this does not guarantee that parameter estimation on reasonably sized phylogenies is not biased, that there is enough statistical power in the data to distinguish between alternative evolutionary scenarios (such as alternative environmental dependences), or that environmental dependence is not artifactually inferred when it did not occur. So, for example, temperature-dependency of birds diversification (Claramunt and Cracraft 2015) was inferred on a highly undersampled phylogeny ($\sim 0.02\%$); even though the likelihoods accommodate missing taxa, can such results be trusted? Environment-dependent extinction was detected in birds (Claramunt and Cracraft 2015) and birdwing butterflies (Condamine et al. 2015). Can we trust these results given the well-known difficulty in inferring extinction rates from reconstructed phylogenies (Morlon 2014)? We cannot answer these questions, or similar questions that will undoubtedly arise in future use of environment-dependent models, unless the statistical properties of these models are thoroughly assessed using simulations.

In this article, we compile four biotic and five abiotic paleodatasets that can potentially be used as explanatory variables in environment-dependent models and use simulated trees to evaluate the statistical performance of the corresponding models. We first focus on temperature-dependent models: we test the ability to recover speciation rates, extinction rates, and their temperature-dependencies; we also test the ability to detect temperature-dependent diversification when it occurs and check that such dependency is not wrongly detected when it does not occur. Next, because a particularly promising use of these models is to find which environmental variable, among a set, best explains diversification dynamics, we test the ability to correctly identify the environmental variable (of the nine) that actually influenced diversification rates and assess whether specific characteristics of the paleoenvironmental curves influence this ability. Finally, we offer best-practice guidelines for using the environment-dependent model and implement them on empirical phylogenies.

MATERIALS AND METHODS

Paleodata

We collected paleodata on nine variables that have been associated with macroevolutionary hypotheses of diversification. These paleodata represent only a subset of the paleoenvironments that could potentially

be used in environment-dependent diversification models and are focused primarily on the marine environment. However, most of them are also relevant to the terrestrial environment; and they cover a wide range of biotic and abiotic hypotheses with minimal redundancy. Most importantly, they cover a wide range of resolution and temporal trends; we expect the statistical properties of the environment-dependent models on these paleodata to provide a good representation of their statistical properties in general. The paleodata can roughly be classified into three main categories. (i) Paleoclimatic variables: temperature data (inferred from $\delta^{18}\text{O}$ measurements) taken from (Zachos et al. 2008); atmospheric CO_2 data taken from (Hansen et al. 2013); benthic $\delta^{13}\text{C}$ data taken from (Lazarus et al. 2014). Temperature is the canonical indicator of climate change. Atmospheric CO_2 varies as a result of biotic changes, such as the rise of photosynthetic plankton (Park et al. 2015), or tectonic activity, which causes either a reduction in CO_2 caused by the subduction of carbonate-rich ocean crust or the emission of CO_2 through volcanic eruptions (Beerling and Royer 2011); levels of atmospheric CO_2 are thought to impact, in particular, photosynthetic organisms. $\delta^{13}\text{C}$ reflects global changes in organic carbon sequestration and organic isotope fractionation ratios. In the Cenozoic, $\delta^{13}\text{C}$ can represent changes in the proportion of C_4 and C_3 grasses dominating the planet and is therefore likely to be crucial to the evolution of grasses and the herbivores that feed on them; $\delta^{13}\text{C}$ can also represent increases in ^{13}C -rich marine species, such as diatoms, which contribute considerably to the overall carbon present in the planktonic food web (Katz et al. 2005; Oehlert and Swart 2014). (ii) Variables reflecting changes in oceanic composition: silica weathering ratio (Cermeño et al. 2015) and sea level estimates (Miller et al. 2005). The silica weathering ratio accounts for influxes of silicic acid in the ocean, which is needed for silica-precipitating microalgae (e.g., diatoms and radiolaria), which are responsible for about a quarter of global primary production (Buchan et al. 2014). Sea level has several effects on both the terrestrial and the marine biome, through, for example, different overall land surfaces, different weathering levels, changes in stoichiometry, and currents. (iii) Variables reflecting interspecific interactions (primarily in the ocean): we constructed diversity curves for fossil archaeplastida (green algae, red algae, and land plants), radiolaria, foraminifera, and ostracods. Fossil data were compiled from the Neptune Database (Lazarus 1994) and Paleobiology Database (<https://paleobiodb.org/>) and diversity curves were estimated at the genus level using shareholder quorum subsampling (Alroy 2010) at 2-million-year bins. These diversity curves are useful for testing hypotheses of direct species interactions, such as grazing of archaeplastida by herbivores, or indirect interactions, such as competition for the same feeding resources; they may also be useful, for example, as proxies for global productivity (archaeplastida,

(Raven and Giordano 2014)). Additionally, these fossil species were chosen because they are abundant enough in the fossil record to generate detailed diversity curves.

All paleoenvironmental curves were stored in the R package RPANDA freely available on CRAN (Morlon et al. 2016). For the purpose of our study, to avoid biases in model selection, curves were truncated at 52 million years ago (the earliest date of the silica curve) and values were scaled to between 0 and 1.

Performance of Temperature-Dependent Diversification Models

We used simulations to test the statistical properties of the temperature-dependent diversification models. All our simulations were conducted over the time span of the environmental curves (i.e., 52 Myrs). We discarded trees with fewer than 50 tips, which always corresponded to less than 10% of the simulated trees. Throughout, speciation and extinction rates are in units of events per lineage per million years.

Testing the ability to recover simulated parameter values.— We tested the ability of temperature-dependent models to recover accurate parameter values. We simulated environment-dependent birth–death trees with speciation rates, λ , and/or extinction rates, μ , that varied as an exponential function of temperature T , such that at each given point in time, t , $\lambda(t) = \lambda_0 e^{\alpha_\lambda T(t)}$ or $\lambda(t) = \lambda_0$ and $\mu(t) = \mu_0 e^{\alpha_\mu T(t)}$ or $\mu(t) = \mu_0$. For trees with a λ -dependency on T (hereafter referred to as temperature-dependent λ trees), we simulated with $\lambda_0 = 0.1, 0.15, 0.2$, fixing $\alpha_\lambda = 0.02$ and $\mu_0 = 0.01$; $\alpha_\lambda = -0.5, -0.2, 0.4, 0.8, 1.6$ fixing $\lambda_0 = 0.1$ and $\mu_0 = 0.01$; and $\mu_0 = 0.01, 0.03, 0.06, 0.09$ fixing $\lambda_0 = 0.1$ and $\alpha_\lambda = 0.02$. We furthermore simulated temperature-dependent λ trees with a considerably higher μ ($\mu_0 = 0.7$) to reflect estimates of μ inferred from the fossil record (Foote et al. 2007). To do so, we increased λ too ($\lambda_0 = 0.7$, $\alpha_\lambda = 0.35$), otherwise it was difficult to simulate surviving trees. For trees with a μ -dependency on T (hereafter referred to as temperature-dependent μ trees), we simulated with $\lambda_0 = 0.2, 0.25, 0.3$, fixing $\mu_0 = 0.02$ and $\alpha_\mu = 0.02$; $\mu_0 = 0.02, 0.06, 0.1$, fixing $\lambda_0 = 0.3$ and $\alpha_\mu = 0.01$; and $\alpha_\mu = -1, -0.6, 0.2, 0.4$, fixing $\lambda_0 = 0.5$ and $\mu_0 = 0.1$. Additionally, we simulated trees with both a λ - and μ -dependency on T , with (i) $\lambda_0 = 0.05$, $\alpha_\lambda = -0.15, 0.15$, $\mu_0 = 0.05$, and $\alpha_\mu = 0.1$; (ii) $\lambda_0 = 0.25$, $\alpha_\lambda = 0.4$, $\mu_0 = 0.05$, and $\alpha_\mu = 0.1$; (iii) $\lambda_0 = 0.02$, $\alpha_\lambda = 0.5$, $\mu_0 = 0.05$, and $\alpha_\mu = 0.1$; (iv) $\lambda_0 = 0.075$, $\alpha_\lambda = 0.5$, $\mu_0 = 0.05$, and $\alpha_\mu = -0.4, -0.15, 0.35, 0.5$. Values were chosen to test the most amount of variation in the λ and μ dependency, while reliably simulating surviving trees. For each combination of parameter values, we simulated 5000 trees using the RPANDA function `sim_env_bd`, stopping the simulations when the equivalent of 52 Myrs was achieved. The median tree size (of trees

with more than 50 tips) was 280 for trees with a λ -dependency on T , 435 for trees with a μ -dependency on T , and 812 for trees with both a λ - and μ -dependency on T . Finally, we fitted the generating model (exponential dependence of λ or μ on T) by maximum-likelihood using the RPANDA function `fit_env` (Morlon et al. 2016) and recovered parameter estimates for λ_0 , α_λ , μ_0 , and α_μ . The likelihood implemented in this function is the one described in (Morlon et al. 2011) and (Condamine et al. 2013); here we conditioned the likelihood on stem age. Throughout, we used a Nelder–Mead optimization algorithm (Nelder and Mead 1965) to fit models.

Testing the ability to recover the proper model.—We tested the ability to detect temperature-dependent diversification when it occurs and to not detect it when it does not occur. Many paleoenvironmental curves, despite fluctuations in their trends, sport a general tendency to increase or decrease over time. Therefore, in addition to testing the ability to distinguish between temperature-dependent and null, constant rate models, we tested the ability to distinguish between temperature-dependent and time-dependent models. We simulated birth–death trees with a positive or negative exponential dependency of λ on time ($\lambda = \lambda_0 e^{\alpha \cdot t}$ for $\alpha = -0.075, -0.05, 0.05, 0.075$ and $\mu(t) = 0.05$) and then fit constant-rate, time-dependent, and temperature-dependent models to those trees. We calculated the percentage of best-fit models for each set of simulated trees. We similarly simulated birth–death trees with a positive or negative exponential dependency of λ on temperature ($\lambda = \lambda_0 e^{\alpha \cdot T}$ for $\alpha = -1, -0.5, -0.1, 0.1, 0.5, 1$ and $\mu(t) = 0.05$) and fit constant-rate, time-dependent, and temperature-dependent models to them. We simulated 1000 trees for each α value. The median tree size (of trees with more than 50 tips) was 309 tips. Additionally, because unaccounted-for background variation in diversification rates across lineages can potentially lead to biased inferences (Rabosky 2010; Morlon et al. 2011; Rabosky and Goldberg 2015), we tested whether rates that vary among lineages can be misleadingly interpreted as environmental dependence. We therefore simulated trees with a 10% probability of a rate-shift in λ at each branching event, with an initial $\lambda_0 = 0.2$; at each event where a shift occurred, the new λ was drawn from a normal distribution of values (with a standard deviation of 0.1) centered on the ancestral value; μ was held constant at 0.02. For each scenario we simulated 5000 trees; the median tree size (of trees with more than 50 tips) was 455 tips. Finally, to test whether high relative extinction rates affect our ability to properly infer temperature-dependence on speciation rates, we simulated 1000 temperature-dependent λ trees (median tree size of 190 tips) with $\lambda = 0.75 e^{0.35 \cdot T(t)}$ and $\mu = 0.7$. Time-dependent trees were simulated using `tess.sim.age` (Höhnna 2013) and trees with rate-shifts were simulated with our own code. We fit all three models (exponential dependency on temperature, exponential dependency on time, and

constant rates) to each tree and compared support using the corrected (based on number of tips) Akaike Information Criterion (AICc) (Burnham and Anderson 2004) and Akaike weights. Fits of constant and time-dependent models were performed using the `fit_bd` function in RPANDA. We likewise simulated birth–death trees with a fixed λ and exponential dependencies of μ on time ($-0.075, -0.05, 0.05, 0.075$) or temperature ($-1, -0.5, -0.1, 0.1, 0.5$) and calculated the percentage of trees under each simulated scenario best fit by a constant-rate, or time-dependent or temperature-dependent extinction model. For the temperature-dependent extinction trees, in addition to fitting the corresponding models, we also fit models with an exponential dependency of λ on temperature. This was done to see whether a μ -dependency on temperature left a detectable signal in λ . We simulated 1000 trees for each α value. The median tree size (of trees with more than 50 tips) was 401 tips.

Testing the effects of undersampling on parameter estimation and model selection.—Because many empirical phylogenies are undersampled, we tested the effect of undersampling on inferring parameters in temperature-dependent trees. We simulated 1000 temperature-dependent λ trees ($\lambda(t) = 0.2 e^{0.4 \cdot T(t)}$) and constant extinction ($\mu = 0.05$) and 1000 temperature-dependent μ trees ($\mu(t) = 0.05 e^{0.1 \cdot T(t)}$) and constant speciation ($\lambda = 0.5$). We jackknifed the trees (i.e., sampled without replacement) at 10% intervals from 90–50% and inferred parameter estimates by fitting a temperature-dependent speciation model to each set of 1000 trees using the function `fit_env`, while accounting for undersampling by specifying the corresponding sampling fraction (Morlon et al. 2011; Condamine et al. 2013). To test the effect of undersampling on model selection, we fit constant-rate, time-dependent, and temperature-dependent models to trees simulated with an exponential dependence on temperature ($\lambda(t) = 0.2 e^{0.9 \cdot T(t)}$) and constant extinction ($\mu = 0.05$). Here, we jackknifed simulated trees as above, but from 90–20%. Because model selection is poor for fully sampled temperature-dependent μ trees, we only conducted this analysis for temperature-dependent λ trees. The median tree size (of trees with more than 50 tips) of the original temperature-dependent λ trees was 1209 tips and of the original temperature-dependent μ trees was 1119 tips.

Testing the effects of tree size on parameter estimation and model selection.—In order to have an idea of the effect of tree size on parameter estimation and the ability to recover temperature dependency, we simulated 5000 temperature-dependent λ trees ($\lambda = 0.2 e^{\pm 0.05 \cdot T(t)}$) and $\mu = 0.05$. We fitted the corresponding temperature-dependent model and recovered parameter estimates for each tree as above. We also fitted the exponential time-dependent and constant rate models and computed the

Akaike weights of these models on each tree. Finally, we reported the results by tree size bin. This procedure can potentially introduce a bias, because trees falling in a particular size-bin are trees that, by chance, even though they share the same simulated parameter values, diversified more (or less) than others. We did not observe such a bias, however. We similarly tested the effect of tree size on parameter estimation and model recovery for temperature-dependent μ trees ($\mu = 0.2e^{0.1 \cdot T(t)}$); in these simulations, we varied λ_0 across simulations so as to avoid the bias mentioned above. λ_0 values were 0.1, 0.2, 0.3, 0.4, and 0.5.

Testing the effects of loss of resolution in the temperature curve on parameter estimation and model selection.—There will necessarily be a loss of resolution between the actual fluctuations in an environmental variable over time and the reconstruction of those fluctuations with sampled data. It is therefore important to understand how sensitive the environment-dependent model is to the reconstructed fluctuations. To evaluate this, we tested the effect of discrepancies in resolution between the temperature curve used to simulate trees and the one used to fit the temperature-dependent model. We simulated pure-birth ($\mu = 0$) trees with an exponential dependency between λ and temperature, where the smoothing function for the temperature curve was determined using generalized cross-validation (Golub et al. 1979). We simulated 5000 trees with positive temperature-dependency ($\lambda = 0.2e^{0.05 \cdot T(t)}$) and 5000 trees with negative temperature-dependency ($\lambda = 0.2e^{-0.05 \cdot T(t)}$). The median tree size (of trees with more than 50 tips) was 610 tips. We then fitted temperature-dependent models to the simulated trees, using increasingly smoothed temperature curves. Smoothed curves are meant to represent degraded environmental data and were obtained by reducing the number of degrees of freedom (DOFs) applied to the smoothing function. For each tree, we compared parameter estimates, as well as statistical support in comparison with time-dependent models (with exponentially trending λ) obtained when fitting temperature-dependent models with increasingly smoothed temperature curves.

Performance of model selection on trees simulated with various environmental dependencies

Testing the ability to recover environment-dependent speciation.—Generalizing the analyses performed for temperature-dependency, we assessed whether environmental dependence can be detected on trees simulated with λ dependencies on specific paleoenvironmental curves. Because we did not see any effect of extinction on the ability to recover the environment-dependent model in the case of temperature-dependency, we simulated pure-birth trees ($\mu = 0$) with exponential dependencies between

λ and each paleoenvironmental curve X (see *Paleodata* above). We simulated trees at various values of α_λ (0.15, 0.3, 0.45, 0.6, 0.75, 0.9, 1.05, 1.3, 1.6, 1.9, 2.2) to test how the strength of the simulated dependency influenced the recoverability of the environmental model. For each value of α_λ and for each paleoenvironmental curve we simulated 1000 trees with a median tree size (for trees with more than 50 tips) of 310 tips. λ_0 was identical across paleoenvironmental curves, but varied between α_λ values, in order to maximize the number of surviving, reasonably sized simulated trees. For each tree, we fitted the same environment-dependent model that was used to simulate that particular tree, as well as a constant-rate λ model and a time-dependent λ model, and compared their statistical support with AICc. We selected the best model as the one with the lowest AICc support. We also compared the relative support of the models using Akaike weights.

Testing the effects of characteristics of the paleoenvironmental curves on the ability to recover environment-dependent speciation.—Environmental curves differ widely in their shape and characteristics (Fig. 1). In particular, the biotic curves that we considered here (constructed from fossil data) appeared coarser than the abiotic ones. To assess whether environmental dependency was more easily detected for certain types of curves than others, we characterized the curves based on defining attributes, such as biotic/abiotic, linearity/non-linearity, and six further metrics, two capturing temporal autocorrelation, two capturing global trend, and two capturing the composition of the data. Next, we assessed the effect, if any, these features had on the rate of recovery of the models at different values of α_λ . To test for linearity, we tested for the presence of one or two breakpoints in the data using the threshold test (Hansen 1999). We used two metrics reflecting temporal autocorrelation; first we measured the correlation between values in the original time-series and a new series that lags behind by some amount of time, for a series of time lags. We then fit regressions between the various time lags and the corresponding correlation value. We used the intercept of the regression to measure short-term autocorrelation, and the slope to measure the rate at which autocorrelation decreases. A global trend analysis was conducted by fitting a linear model using generalized least squares to each curve and computing the correlation, slope, and DOFs for the fitted model. Finally, we estimated the average rate of change of the curves by calculating the mean over all time periods of the slope between values within each time period.

Testing the ability to recover the proper environmental dependency on speciation.—For trees simulated under each paleoenvironmental curve with an exponential dependency on λ at different values of α_λ , we compared fits for models using all the paleoenvironmental curves. For each set of simulated trees, we calculated the percentage of trees best fit (i.e., lowest AICc) by each

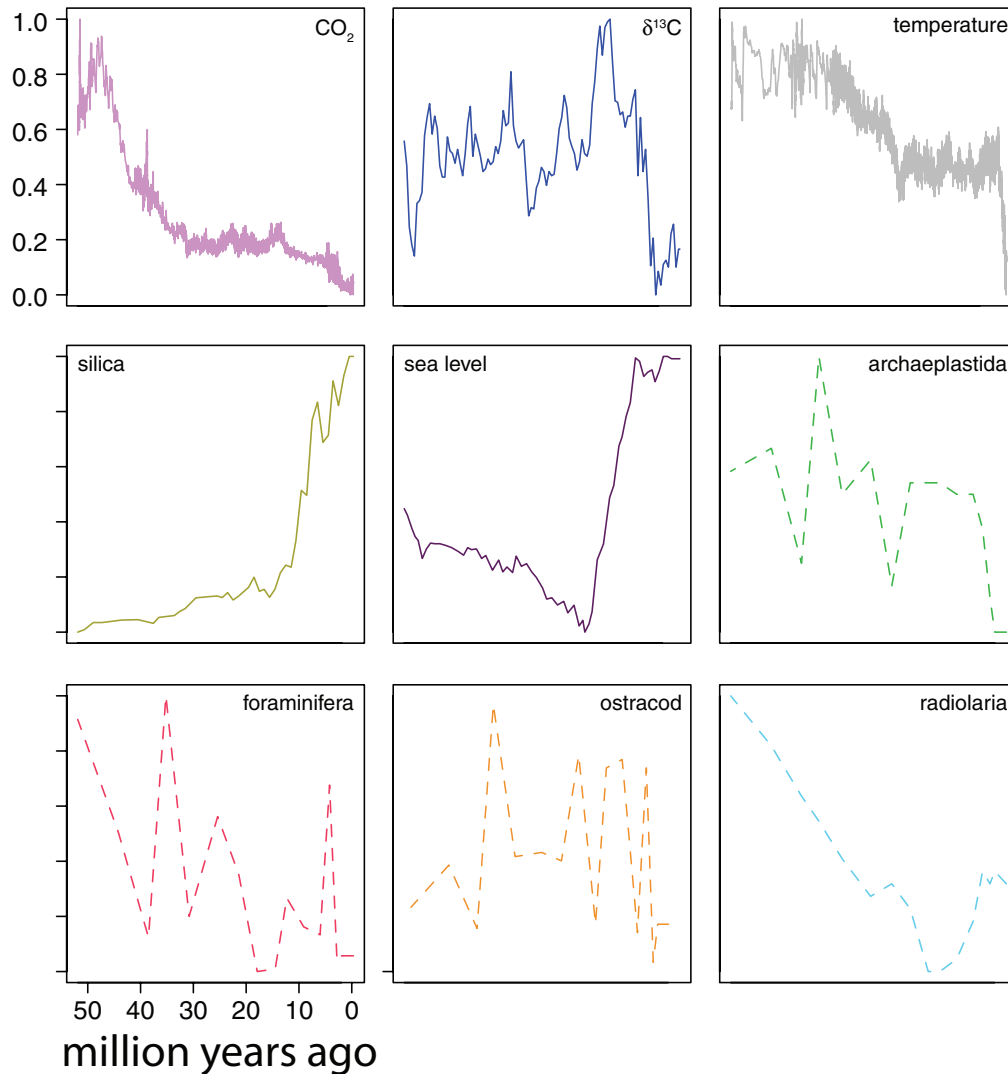


FIGURE 1. Scaled paleoenvironmental curves. Dashed lines correspond to biotic variables.

paleoenvironmental model and reported this percentage as a function of α_λ .

Testing the ability to recover environment-dependent extinction.—We carried out the same type of analyses to assess whether an environmental dependence on μ can properly be inferred. We simulated trees with an exponential dependency between μ and a paleoenvironmental curve X , holding λ and μ constant ($\lambda_0 = 0.15$, $\mu_0 = 0.02$), for different values of α_μ as above. For each value of α_μ and for each paleoenvironmental curve we simulated 1000 trees; the median tree size (of trees with more than 50 tips) was 320. For each tree, we fitted the same environment-dependent model that was used to simulate that particular tree ($\mu = \mu_0 e^{\alpha_\mu X(t)}$, $\lambda = \lambda_0$) as well as a constant-rate birth–death model and a model with μ exponentially dependent on time and constant λ . We also fitted a model where the environment influences speciation rather

than extinction rates ($\lambda = \lambda_0 e^{\alpha_\lambda X(t)}$, $\mu = \mu_0$). We did this under the hypothesis that trees simulated under a μ dependency on a paleoenvironmental curve could manifest a signature of that dependency in their λ .

Empirical illustrations

We used time-calibrated molecular trees for Cetacea and Ruminantia. The Cetacean phylogeny was 98% sampled and constructed from six mitochondrial and nine nuclear genes in a Bayesian framework; it was time-calibrated using fossil data; and divergence dates were estimated using a relaxed molecular clock (Steeman et al. 2009). For Ruminantia we used two different trees: one was constructed using full mitochondrial genomes and 16 fossil calibration points in a Bayesian framework, but was only 65% sampled (Bibi 2013); the other one was constructed using a supermatrix of 124 previously published trees and was 100% sampled

(Cantalapiedra et al. 2014). The two phylogenies differed markedly in their datation; in particular, the phylogeny of (Bibi 2013) hypothesized significantly more recent crown ages for many families and for Ruminantia than that of (Cantalapiedra et al. 2014). We conducted analyses on Bayesian posterior trees for Cetacea and both Ruminantia trees. We chose the above trees because diversification in each has previously been associated with an abiotic process (Steeman et al. 2009; Cantalapiedra et al. 2014).

We fitted the nine paleoenvironment models plus a constant-rate birth–death model, a time-dependent speciation model (with and without constant extinction) and a time-dependent extinction model (with constant speciation) to the empirical phylogenies, accounting for missing taxa by applying the relevant sampling fraction (Morlon et al. 2011; Condamine et al. 2013). For each paleoenvironment model, we tested λ and μ as constant, linear, and exponential functions of the paleoenvironmental curves, as well as $\mu=0$. The best-fit model for each phylogeny was determined by $\Delta AICc$ as above, and the corresponding speciation and extinction parameter estimates were recorded. All palaeoenvironmental curves were tested. While some variables (temperature and sea level) are directly biologically relevant for both clades, other variables (CO_2 , $\delta^{13}C$, silica, archaeplastida diversity, ostracod diversity) are directly relevant for only the Cetacea, and others still (foraminifera and radiolarian diversity) are not directly biologically relevant to any of the two clades. When environmental variables were not directly relevant, we still included them as negative controls. We also note that while significant support for environment-dependent models is typically interpreted by an effect of the environment on clade dynamics, such support could also occur if the rise or demise of a given clade impacts the targeted environmental variable (e.g., silica levels in the ocean are often attributed to the expansion of land plants). Hence, environmental variables that are biologically relevant for a given clade are those that can either influence, or be influenced by, that clade.

Both clades can potentially be influenced by temperature and sea level changes. For cetaceans, aside from these two variables, we may expect a dependency on silica levels, archaeplastida, and ostracods. Indeed, the evolution of cetaceans has been influenced by diatoms, which are phytoplankton that require silica (Marx and Uhen 2010). Archaeplastida include green algae that can have a negative impact on diatom blooming (Keating 1978), and thus potentially also affect cetaceans. Finally, ostracods may also be expected to influence cetacean diversification, either directly as a food source or indirectly as diatom predators. For ruminants, changes in plant life across the globe have been crucial: reductions in CO_2 levels have been linked to radiations of angiosperms during the Cretaceous and the rise of C_4 plants (Gerhart and Ward 2010), which is also reflected in global levels of $\delta^{13}C$; and Archaeplastida, which includes land plants, can directly

influence diversification in Ruminantia via change in their feeding regime.

We quantified false detection rates among the different environmental models using simulations above, however these rates may be sensitive to tree age. In order to assess false detection rates associated to our specific empirical phylogenies, we simulated 500 trees for each model (including the constant-rate and time-dependent models) with the parameters inferred from fitting those models to each of the empirical phylogenies. We then fit all the models to the simulated trees to estimate the percentage of trees in each set that are best supported by each model.

In order to illustrate how one could test the relative importance of abiotic vs. biotic variables, we selected the best supported abiotic and biotic variables and computed the relative probability of these two models based on their Akaike weights; we also computed the Akaike weight of each model among the set of all models and summed these Akaike weights over all the abiotic variables (or biotic variables) to obtain an estimate of the overall support for abiotic (or biotic) variables. The first approach has the advantage to not be biased by the total number of variables, or biologically relevant variables, in each category (abiotic and biotic). However, it measures the relative importance of the two most supported abiotic and biotic variables rather than the relative importance of abiotic and biotic variables as a whole. Intermediate approaches could be envisioned where the relative support for abiotic and biotic variables is computed based on a fixed, homogenized number of biologically relevant biotic and abiotic variables.

Additionally, for each phylogeny, we assessed the support of the best-fit paleoenvironmental model with smoothing splines computed with increasingly smaller DOFs. This was done in order to test at what resolution of the paleoenvironmental curve support for environment-dependent diversification started to be lost. We compared all fits using $AICc$ as above.

RESULTS

Performance of Temperature-Dependent Diversification Models

Parameter estimation.—Temperature-dependent models were able to accurately recover simulated parameter values. Maximum-likelihood estimates of the parameters of the temperature-dependent λ models were unbiased (Fig. 2), even when extinction rates were high (Fig. 2d). Maximum-likelihood estimates of the parameters of the temperature-dependent μ models were likewise unbiased, but they were more variable (Supplementary Fig. S1 available on Dryad at <http://dx.doi.org/10.5061/dryad.m96r3>). When both λ and μ were exponentially dependent on temperature, parameter recovery remained reliable (Supplementary Fig. S2 available on Dryad).

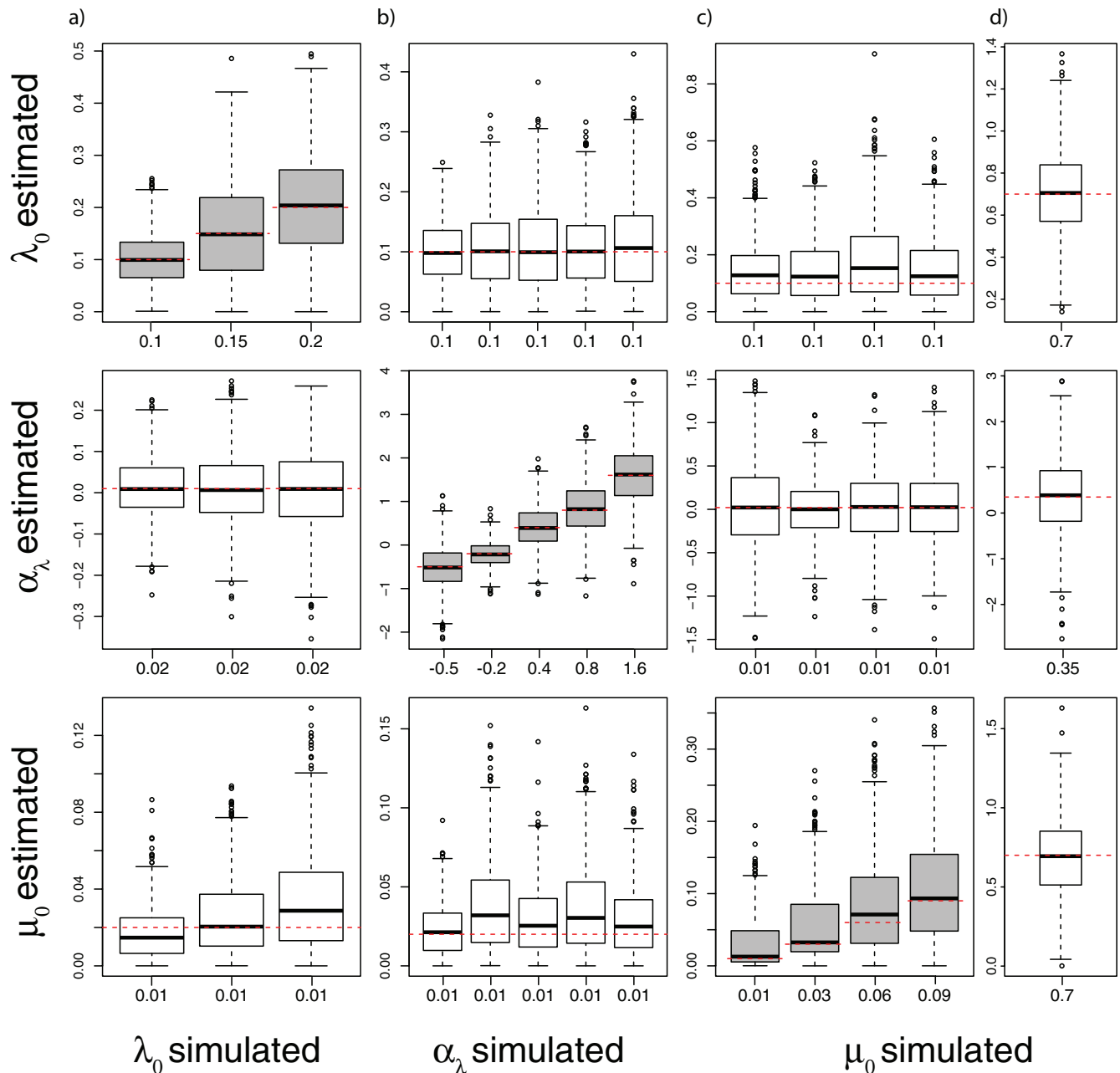


FIGURE 2. Recovered parameter estimates for trees simulated with an exponential dependency of λ on temperature. Simulations with a) varying λ_0 and constant α_λ and μ_0 , b) varying α_λ and constant λ_0 and μ_0 , and c,d) varying μ_0 and constant λ_0 and α_λ . Dashed red lines mark the simulated parameter value.

Model selection.—Temperature-dependent speciation models can be correctly distinguished from other models. When time-dependent speciation trees were simulated, they were falsely recovered by a temperature-dependent model <25% of the time for all values of α (Fig. 3a) and the mean Akaike weight of the temperature-dependent model was always below 0.3 (Fig. 3c). The correct recovery of temperature-dependence was more contingent on the strength of the temperature-dependence (i.e., α). For $\alpha \geq 0.5$, temperature-dependence was best supported for

58–80% of trees, while for $\alpha < 0.5$, it was best supported for only 11–14% of trees (Fig. 2b). Likewise, for $\alpha < 0.5$, the mean Akaike weight of the temperature-dependent model was 50%, while for $\alpha \geq 0.5$ the mean Akaike weight was 0.64 (Fig. 3d). For high α under high relative extinction, the temperature-dependent model was also overwhelmingly supported (with mean Akaike weights for constant-rate (0.07), time-dependence (0.22), and temperature-dependence (0.71)). Heterogeneity in speciation rates across lineages did not lead to a false support for temperature dependency: trees with

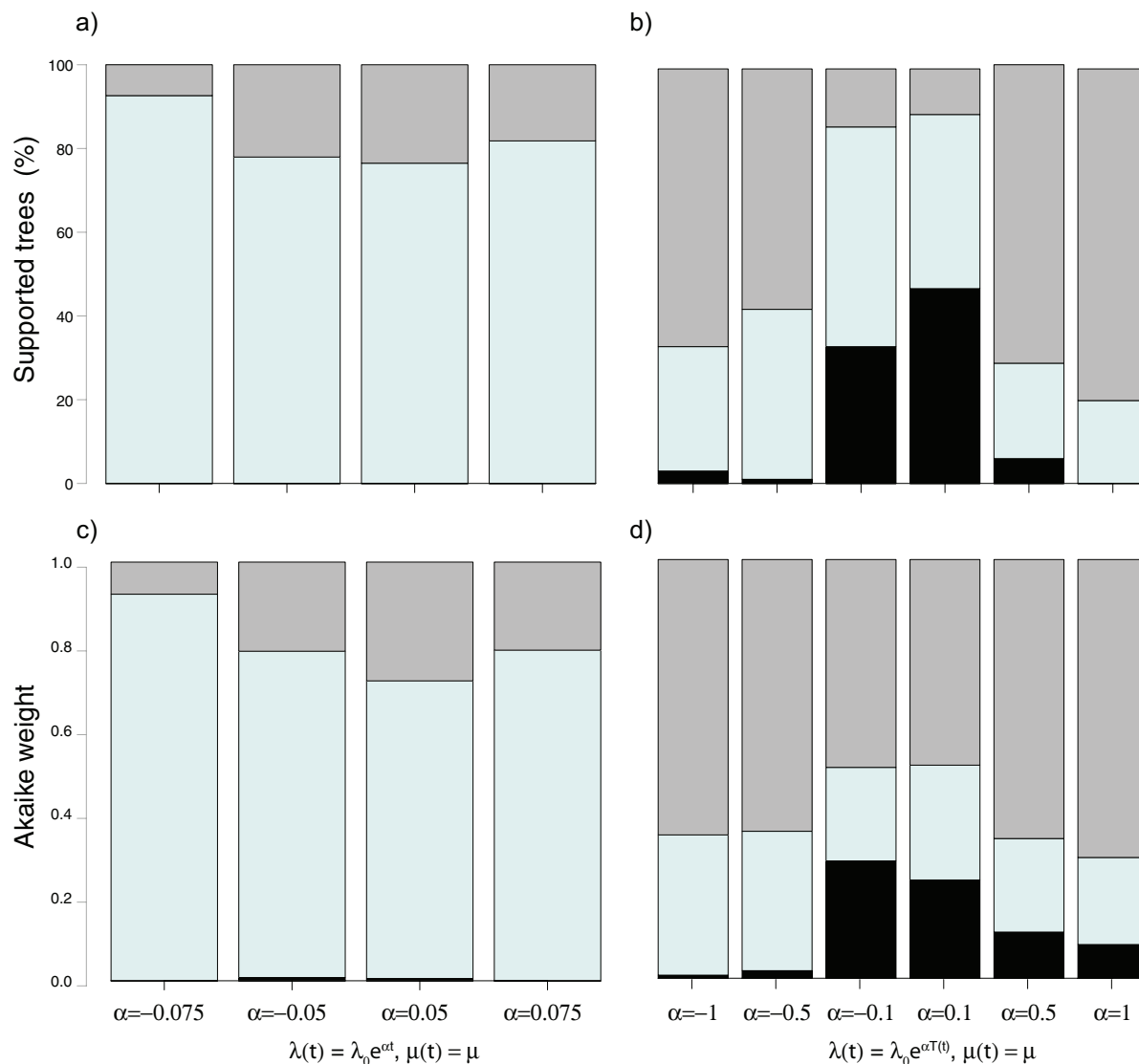


FIGURE 3. The ability to correctly recover trees simulated under a time-dependent model or temperature-dependent model. Columns in each plot show the percentage of trees recovered (a,b) or Akaike weights (c,d) for constant-rate (black), time-dependent (light blue), or temperature-dependent (gray) models for a set of trees simulated under (a,c) time-dependence or (b,d) temperature-dependence. The x-axis shows the strength of the dependencies (i.e., α value) used to simulate each set of trees. Akaike weights are averaged over all trees simulated with the same α .

speciation rate shifts were recovered by constant-rate models 32% of the time, models with increasing speciation rate through time 62% of the time, and temperature-dependent models only 6% of the time. The support for increasing speciation rate models can be explained by the fact that lineages that by chance have high speciation rates leave more descendants, such that the average speciation rate over lineages increases through time. When trees with increasing extinction rates through time were simulated, they were falsely recovered by a temperature-dependent model <5% of the time for all values of α (Supplementary Fig. S4A available on Dryad) and the mean Akaike weight of temperature-dependent model was always below 0.03 (Supplementary Fig. S4c available on Dryad). The correct recovery of temperature-dependence on

extinction was relatively poor, irrespective of α : the correct model was recovered <25% of the time with a mean Akaike weight of 0.21; whereas the temperature-dependent speciation model, time-dependent extinction model, and constant-rate model were recovered >20%, >40%, and >10% of the time, respectively, with mean Akaike weights of 0.22, 0.45, and 0.12 (Supplementary Fig. S4c,d available on Dryad).

Effects of undersampling.—Undersampling did not bias parameter estimates, but affected the uncertainty around the estimates (Supplementary Fig. S3 available on Dryad). While the median estimates did not deviate significantly from the simulated estimates ($P < 0.01$), the variance increased with decreasing sampling

fraction. Model selection was consistently good for undersampled trees, such that at a sampling fraction of 20%, the correct model was selected (by lowest AICc) for 65% of trees with an average Akaike weight of 0.72.

Effects of tree size.—As expected, the ability to recover parameter estimates and the temperature-dependent model improved with tree size. For temperature-dependent λ models, median parameter estimates approximated simulated values precisely as soon as trees had more than 100 tips, and the variance of recovered estimates decreased considerably with tree size (Supplementary Fig. S5 available on Dryad). With the α_λ values considered in these simulations ($\alpha = \pm 0.05$), the Akaike weight was well above 0.5 for trees with more than 200 tips, and steadily increased to 0.9 for trees with more than 900 tips (Supplementary Fig. S5 available on Dryad). For temperature-dependent μ models, median estimates for α_μ approximated simulated values for trees with more than 400 tips and the range of estimates decreased considerably for trees with more than 600 tips (Supplementary Fig. S6 available on Dryad). Median estimates for μ_0 were slightly higher than simulated values for all tree sizes, which is consistent with our previous estimates of μ_0 at $\alpha \geq 0.1$ (see Supplementary Fig. S1 available on Dryad). The Akaike weight of the temperature-dependent μ model increased with tree size, but in agreement with the results above it remained low, below 0.4 even for trees with more than 800 tips (Supplementary Fig. S6 available on Dryad).

Effects of loss of resolution in the temperature curve.—When fitting models with degraded temperature curves, we found that accurate parameter estimates could still be properly recovered, unless the environmental data were very highly degraded (i.e., for a smoothed spline constructed with $\leq 15\%$ of the total DOFs available in generalized cross-validated model (Supplementary Fig. S7 available on Dryad). Fitting models with a degraded temperature curve slightly reduced the ability to distinguish temperature-dependent models from time-dependent ones, but this ability remained good, with $<15\%$ of the phylogenies simulated under a temperature-dependent model with the most degraded curve (DOFs = 5% of the generalized cross-validated value) finding support for time-dependent rather than temperature-dependent models (Supplementary Fig. S7 available on Dryad).

Performance of Model Selection on Trees Simulated with Various Environmental Dependencies

Model selection on environment-dependent speciation.—The ability to detect environment-dependence when it exists does not drastically change with the paleoenvironmental variable considered. For all nine paleoenvironmental variables, environment-dependent speciation can be

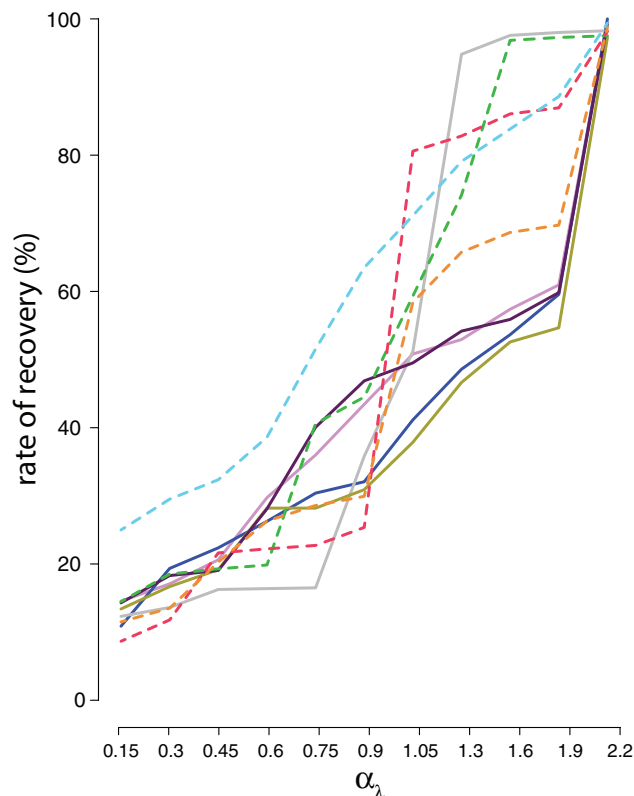


FIGURE 4. Model identifiability for trees simulated with an exponential dependency of λ on different paleoenvironments, X , ($\lambda = \lambda_0 e^{\alpha_\lambda X(t)}$ and $\mu = 0$) for varying values of α_λ . Rate of recovery is defined as the percentage of trees with better AICc support for the simulating environment-dependency model vs. an exponential time-dependent model and constant rate model. Colors correspond to Fig. 1.

detected and distinguished from a constant rate and time-dependent effect, as soon as the environmental effect is relatively strong (Fig. 4). The ability to distinguish between the simulating environment-dependent model and a pure-birth or time-dependent model increases roughly linearly with α_λ , from approximately 20% at $\alpha_\lambda = 0.15$ to 100% at $\alpha_\lambda = 2.2$. This result is unlikely to be an effect of tree size, as the median tree size does not vary much across α_λ values (185 ± 36). As α_λ increases, so does the support for the simulated environmental model, as defined by the Akaike weight support for the environment-dependent, time-dependent, and constant-rate models (Supplementary Fig. S8 available on Dryad), consistent with expectations.

Effects of characteristics of the paleoenvironmental curve.—For $\alpha_\lambda > 1$, curves for biotic palaeodata, which tend to be coarser than abiotic curves, show, on average, higher recovery rates, while for $\alpha_\lambda < 1$, curves for abiotic palaeodata show, on average, higher recovery rates (Supplementary Fig. S9a available on Dryad). There is no such noticeable effect on Akaike weights, though (Supplementary Fig. S8 available on Dryad). There is no apparent effect of linearity, the level of

autocorrelation, the direction or magnitude of the best-fit regression slope, DOFs, or the average rate of change of the environmental curves on our ability to recover how the corresponding environment affected speciation (Supplementary Fig. S9b,c available on Dryad).

Ability to recover the proper environmental dependence on speciation.—The ability to distinguish between λ -dependence on different environmental variables is good as soon as the environmental dependence is strong enough: while a considerable percentage of trees were recovered by the wrong environment-dependent model for low α_λ (a condition under which the environmental model would likely not be selected when compared to a time-dependent model (Fig. 4), the correct model was recovered in the majority of trees for $\alpha_\lambda > 1$ (Fig. 5). For several paleoenvironmental dependencies—specifically, $\delta^{13}\text{C}$, sea level, and temperature—the simulating model was recovered the majority of the time, even at low α_λ values. These three environmental variables were also those that were most often misleadingly detected for trees generated under other environmental dependencies (Fig. 5). As above, the median tree size did not vary much across α_λ values (249 ± 34).

Model selection on environment-dependent extinction.—Environment-dependent extinction is generally not detected, regardless of the strength of the environmental effect: the ability to recover environment-dependent μ is $< 10\%$ across all α_μ values (Supplementary Fig. S10 available on Dryad). Even when only trees with more than 400 tips are considered (median tree size = 1183 tips), the ability to recover the correct model does not exceed 20%. Environment-dependent extinction is more often detected as an environmental dependence on λ than on μ (Supplementary Fig. S10 available on Dryad), but this still happens in less than 20% of the trees.

Empirical Illustrations

Temperature-dependence in Cetacea.—The Cetacean phylogeny was best fit by a pure-birth exponential speciation model with a positive dependency on temperature across all posterior distributions ($\lambda = 0.097e^{0.02 \cdot T(t)}$ for the consensus phylogeny) (Fig. 6). We show that there is a rather low false-positive rate for the temperature-dependent model, with a maximum false-positive rate of 17% for trees where the generating model is foraminifera diversity. The probability that the abiotic model is the best, as calculated by the Akaike weight for the best-fit abiotic model (temperature) when compared with the best-fit biotic model (ostracod diversity) was 0.72. The sum of Akaike weights across abiotic variables was 0.73 ± 0.12 and across biotic variables 0.27 ± 0.15 . We furthermore compared fits of the temperature-dependent λ model with splines smoothed by a range of DOFs (5–190). Using the

maximum DOFs (208), as defined by generalized cross-validation, the temperature-dependent model was significantly better supported than a time-dependent model ($\Delta\text{AICc} = 6.9$) and other environmental models ($\Delta\text{AICc} > 1.3$). The time-dependent model only showed a comparable fit to the tree when the temperature curve was smoothed with DOFs below 10. This suggests that diversification in Cetaceans is best supported by the global trend in the temperature curve, although the decrease in AICc support with reduced DOFs is evidence that diversification is also supported by specific fluctuations within the temperature curve.

Temperature-dependence in Ruminantia.—The Ruminantia phylogeny of (Bibi 2013) was best fit by a pure-birth exponential speciation model with a positive dependency on $\delta^{13}\text{C}$ across all posterior distributions ($\lambda = 0.112e^{0.005 \cdot \delta^{13}\text{C}}$ for the consensus phylogeny) (Fig. 7). Here, too, there is a rather low false-positive rate for the $\delta^{13}\text{C}$ -dependent model, with a maximum false-positive rate of 8% for trees where the generating model is foraminifera diversity. The probability that the abiotic model is the best, as calculated by the Akaike weight for the best-fit abiotic model ($\delta^{13}\text{C}$) when compared with the best-fit biotic model (ostracod diversity) was 0.94. The sum of akaike weights across abiotic variables was 0.79 ± 0.06 and across biotic variables 0.11 ± 0.06 . The Ruminantia supertree of (Cantalapiedra et al. 2014), however, was best supported by a constant-rate birth–death model ($\lambda = 0.11, \mu = 1.20e^{-7}$) (Supplementary Fig. S11 available on Dryad). The analysis by (Cantalapiedra et al. 2014) separated dietary groups and found maximum support for a model with a linear dependency of temperature on λ ; in our analysis (which did not separate dietary groups), a model with linear dependency on temperature found considerably poorer support than the constant-rate model ($\Delta\text{AICc} = 142.329$). When the curve of the best-fit models for Ruminantia (using the phylogeny of (Bibi 2013)) was smoothed with increasingly fewer DOFs and refit to the data, there was little effect on the computed AICc, suggesting that diversification in Ruminantia is supported by the general trend of the $\delta^{13}\text{C}$ curve. AICc support for the best-fit environment-dependent model became poorer than for a time-dependent model only when the DOFs dropped to two, which effectively made the $\delta^{13}\text{C}$ -dependent model equivalent to a time-dependent models.

DISCUSSION

The history of Earth has been marked by shifting land masses, volcanic eruptions, and continuous global atmospheric change, as well as a widespread and often dramatic turnover and dispersal of life. There is general recognition that changes to biodiversity over geological time are driven by the environment

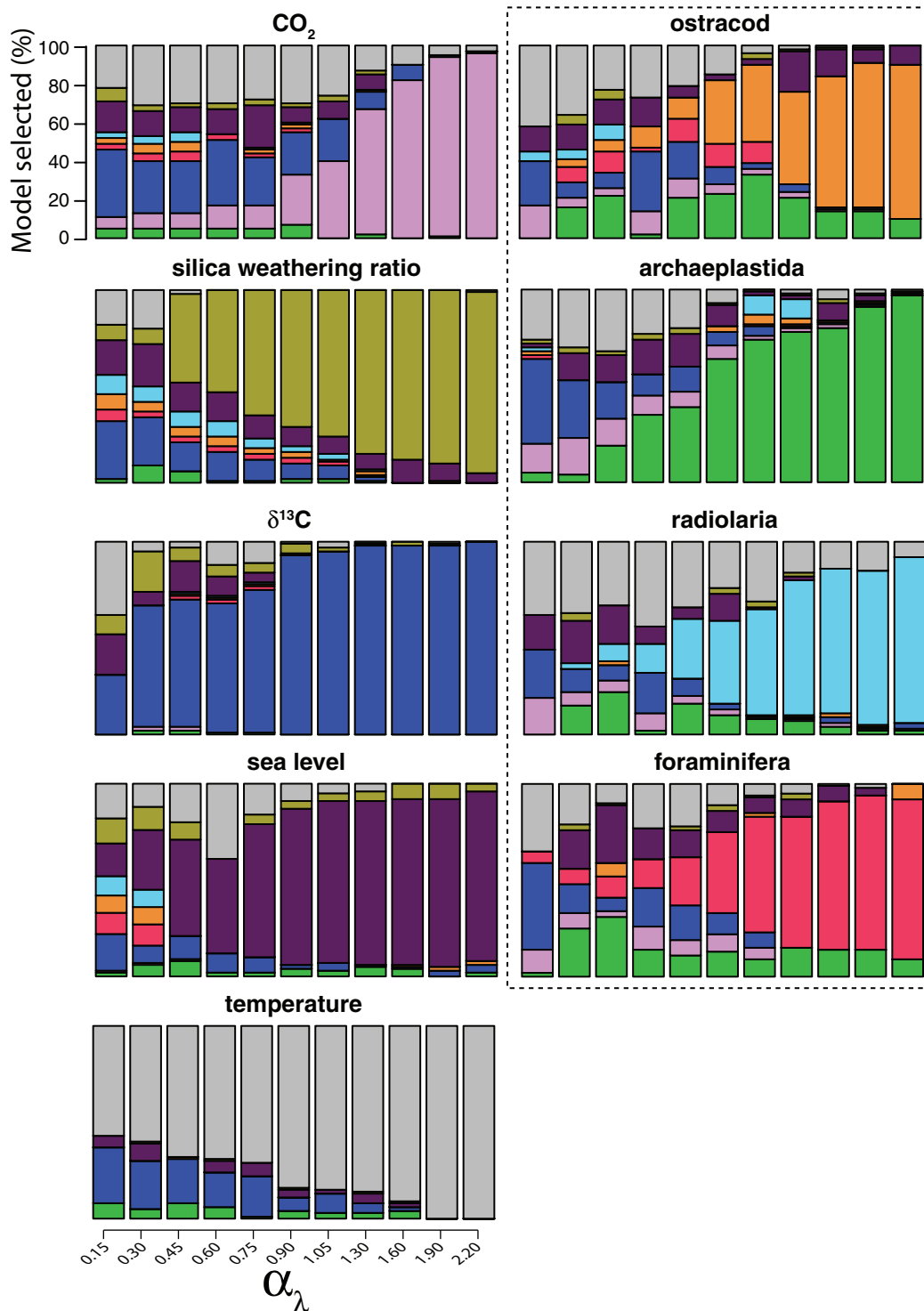


FIGURE 5. Rate of model identifiability between environment-dependent models. For each set of trees simulated with an exponential dependency of λ on a particular paleoenvironmental curve, $X (\lambda = \lambda_0 e^{\alpha X(t)}, \mu = 0)$, the percentage of trees recovered by each paleoenvironmental model for different values of α_λ . The simulated model is listed above each barplot. Colors correspond to Fig. 1; biotic variables are circumscribed by the dashed box.

(Erwin 2009) and species interactions (Voje et al. 2015). The proportional contribution of various abiotic and biotic drivers within individual clades and across the tree of life, however, is still debated (Ezard et al. 2016).

This is due, in part, to the challenges of estimating the effects of these abiotic and biotic factors on clade diversification. We have tested a phylogenetic approach to address these challenges under a model-based

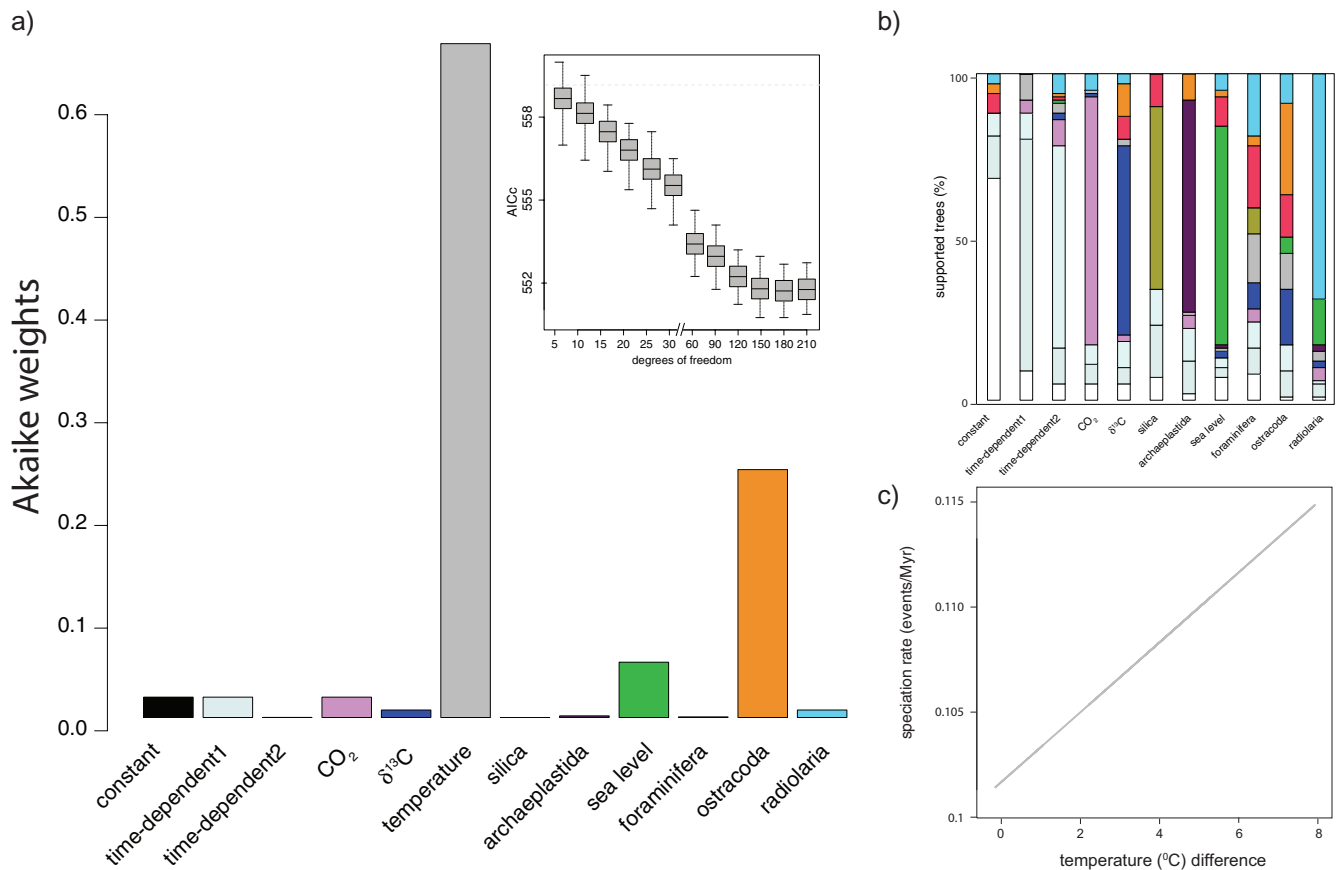


FIGURE 6. Environment-dependency in Cetaceans. a) Akaike weights for different environment-dependent models, a constant-rate birth-death model, a pure-birth exponential time-dependent model (time-dependent1), and an exponential time-dependence on extinction (time-dependent2) on a distribution of 100 posteriorly sampled probabilities of the Cetacean phylogeny. Colors and line type correspond to Fig. 1. (a, inset) AICc support for temperature-dependent models fitted with smoothed splines defined by varying DOFs. AICc value for the time-dependent1 model is shown (dashed line). b) The percentage of trees best fit by each model for trees simulated under the generating model described on the x -axis. c) Speciation as a function of temperature as inferred by fitting the temperature-dependent model.

framework that allows us to directly relate speciation and extinction rates in a clade to biotic and abiotic factors in the paleoenvironment. Applying this modeling framework to empirical phylogenies, we identify specific environmental variables that likely played a major role in shaping the evolution of the corresponding clades.

Statistical Behavior of Environment-Dependent Models and Best-Practice Guidelines for Users

Our simulation results point to aspects of the use of environment-dependent models that should be considered.

(i) Environment-dependent models show very low type I error when compared to constant-rate and time-dependent models, irrespective of the strength of the time-dependency. Therefore, when an environmental dependent model is supported, whether on speciation or extinction, it can safely be concluded that the given environment influenced diversification. (ii) Type II error in environment-dependent models is contingent on the strength of the dependency. Environment-dependent speciation trees that are weakly dependent

on an environmental variable ($\alpha_\lambda \leq 0.9$) are commonly incorrectly supported by a time-dependent or constant-rate model. This contingency is even stronger for environment-dependent extinction trees. Therefore, an absence of support for environment-dependent models should definitively not be interpreted as evidence that the given environment did not influence diversification. Similarly, if environmental dependency is more often supported on speciation than on extinction, this should be interpreted with caution (in particular if the fossil record suggests otherwise), as such distinct support may be of statistical rather than biological origin. (iii) Similarly to what is observed with other diversification models, undersampling—so long as it is uniform and accounted for in model-fitting—increases uncertainty around, but does not bias speciation and extinction rates estimates; likewise, model selection is robust against undersampling. In practice however, as sampling is rarely uniform, we recommend running analyses on trees that are a least 70% sampled, as is also typically recommended for other models. (iv) Finally, although the power to infer environment dependent speciation decreases with decreasing tree

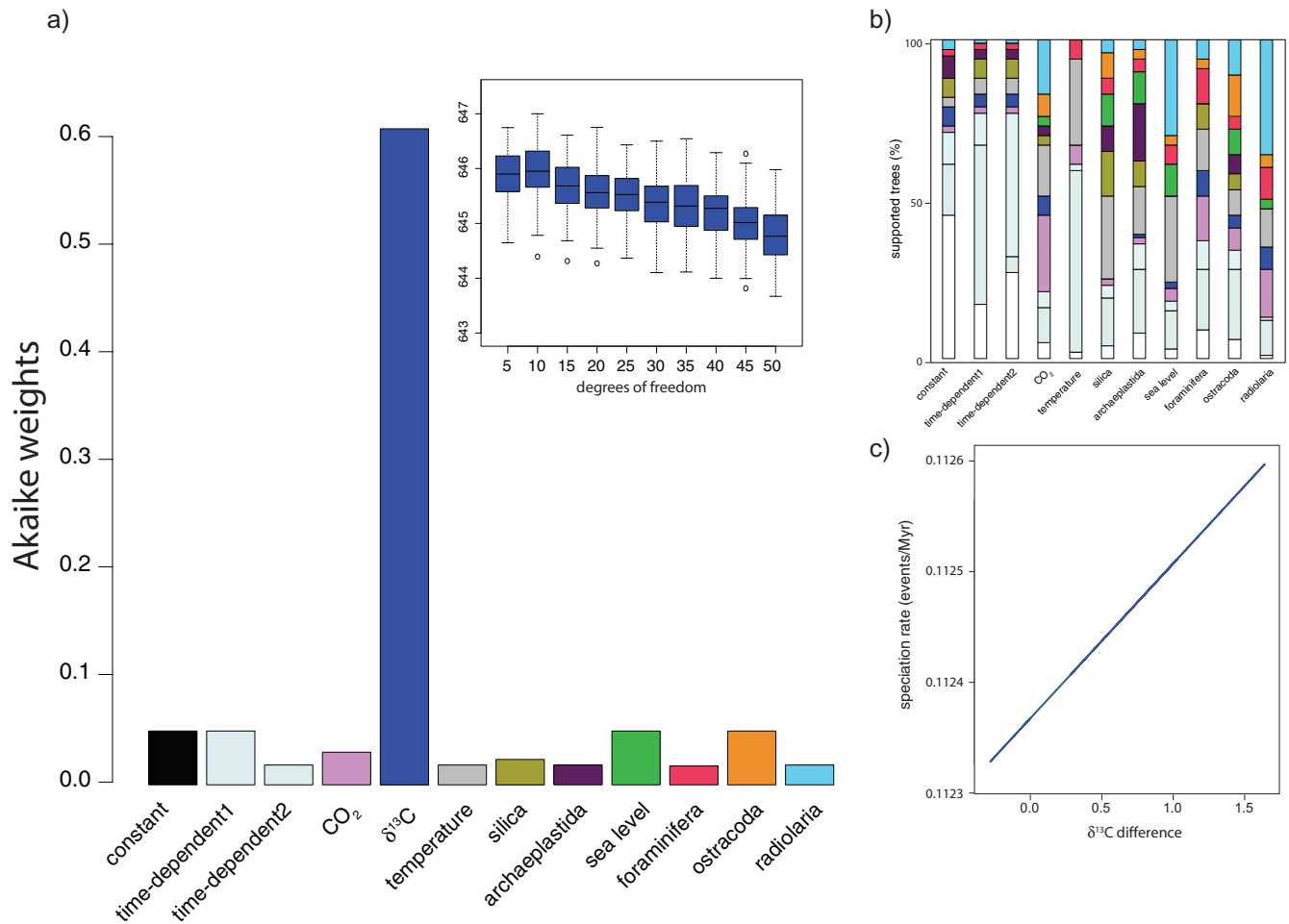


FIGURE 7. Diversification dynamics in Ruminantia. a) Akaike weights for different environment-dependent models, a constant-rate birth-death model, and a pure-birth exponential time-dependent model (time-dependent1), and an exponential time-dependence on extinction (time-dependent2). (a, inset) AICc support for the best supported model fitted with smoothed splines defined by varying DOFs. b) The percentage of trees best fit by each model for trees simulated under the generating model described on the *x*-axis. c) Speciation as a function of temperature as inferred by fitting the $\delta_{13}\text{C}$ -dependent model.

size, parameter estimation is only affected in terms of variability, not in terms of bias. Hence, while we nevertheless recommend avoiding running analyses on trees with less than 30 tips, there is in principal no issue with small trees other than a potential lack of statistical power.

We have two additional recommendations for users: (i) We recommend including environmental variables that are not obviously biologically relevant to the clade of interest, as these can be useful as negative controls or to detect potential hidden variables. In our analyses, for example, we do not find strong support for any biologically irrelevant variable, which gives us confidence in the dependencies for which we do find strong support. (ii) When a specific environmental dependency is detected as the best fit among a set of other environmental dependencies, in order to be confident that the support is not in fact due to the effect of another environmental variable, we recommend evaluating false-positive rates specific to the

environment and tree of interest, as we did here in our empirical illustrations. This can be done by simulating trees under each environmental model (with parameters obtained from the model fit), fitting all models to these trees, and recording the percentage of simulations falsely supporting the best-fit environmental dependency.

A recent work that used the temperature-dependent model (Condamine et al. 2013) to infer diversification in modern birds found rates of speciation and extinction to be negatively dependent on temperature (Claramunt and Cracraft 2015). We can assess how robust this result is in light of our simulation analyses. The temperature curve was smoothed with ~25% of the DOFs inferred by generalized cross-validation. Given our observation that the temperature-dependent model is robust to smoothing, we expect this smoothed curve to be sufficient to generate accurate parameter estimates. The tree has 230 tips, which is sufficient for accurately inferring parameter values for λ , but not necessarily for μ . The result that extinction is temperature-dependent,

therefore, should be interpreted with caution. The strength of the temperature-dependence is not reported, so we cannot judge whether it is likely to be robust against type I error. The tree is severely undersampled, with a sampling fraction of $\sim 0.02\%$, which might produce estimates for both speciation and extinction considerably deviated from their true values. Finally, the temperature-dependent model-fitting was conducted on only one tree (the maximum clade credibility tree), which is not ideal. Dating discrepancies can be sources of potential discordance between results, which is something we account for in the Cetacean and (both) Ruminantia phylogenies by analyzing posteriorly sampled trees.

Environment-Dependence in Cetaceans and Ruminants

In agreement with previous macroevolutionary (Marx and Uhen 2010; Condamine et al. 2013) and macroecological (Whitehead et al. 2008) work, we find that diversification in Cetaceans was positively associated with temperature. The best-fit model is a pure-birth model with exponential dependence of λ on temperature. We show, in addition, that temperature-dependence is better supported than dependence on other paleoenvironmental variables. It has been suggested that Plio-Pleistocene fluctuations in oceanic temperatures created ecological opportunities for allopatric speciation in delphinids (Steeiman et al. 2009) and this may be why we find diversification in Cetaceans to be dependent on temperature. Notably, there is a low type I error for selecting the temperature-dependent model at the inferred α_λ values for the other environment-dependent models and for a phylogeny the same age as the Cetacean phylogeny, which gives us confidence that the temperature-dependent result is not artifactual. The support we find for a positive dependency of ostracod diversity on diversification in Cetaceans ($\lambda = 0.089 * e^{0.193 * \text{ostracod diversity}}$) further suggests that the role of ostracods as a food source may have positively affected Cetacean speciation.

In Ruminantia, the results depended on the phylogeny we used, which is not surprising given the markedly different dates hypothesized by these two phylogenies. With the phylogeny of (Bibi 2013), we found diversification to be best supported by a pure-birth exponential speciation model with a positive dependency on $\delta^{13}\text{C}$. This substantiates a major role for C_4 grasses, which became dominant during the Miocene and have been hypothesized to have spurred diversification in ruminant mammals (Cantalapiedra et al. 2014). The $\delta^{13}\text{C}$ data we used were collected from marine samples rather than from terrestrial samples (e.g., bovid tooth enamel), and so are not a direct measurement of the relative abundance of C_4 and C_3 grasses. However, there are notable similarities in these data, particularly during the Neogene (Cerling et al. 1997). Again, there is a low type I error for selecting the $\delta^{13}\text{C}$ -dependent model at the inferred α_λ ,

values for the other environment-dependent models and for a phylogeny the same age as the Ruminantia phylogeny, which gives us confidence that the this result is not artifactual. For the Ruminantia supertree of (Cantalapiedra et al. 2014), we find support for a constant-rate birth–death model rather than any environment-dependent model, which is contrary to what was found in (Cantalapiedra et al. 2014). This may be because that study conducted separate analyses on different dietary groups.

For both phylogenies, we estimated the support of all biotic and abiotic models, regardless of whether they were biologically relevant or not. In both clades, we found strong support for the best abiotic model when compared with the best biotic one, as well as stronger support for all abiotic variables combined compared with all biotic variables combined. This result may be biologically true, in which case it emphasizes the overarching importance of the global abiotic environment on species diversification in vastly different organisms. It may instead reflect the lack of relevant biotic variables we tested.

We note that the absence of extinction inferred in Cetacea and Ruminantia is not realistic, and is inconsistent with what we know from the fossil record (Quental and Marshall 2010; Cantalapiedra et al. 2014). It has been noted before that, although possible in theory, inferring extinction rates from molecular phylogenies is difficult in practice (see (Rabosky 2010; Morlon 2014; Beaulieu and O’Meara 2015; Rabosky 2016)). Accounting for environmental dependence in speciation rates does not seem to help recovering more meaningful extinction rate estimates. What would more likely help would be to account for heterogeneity in diversification rates across lineages (Morlon et al. 2011). This could be done by applying the same type of tree partitioning that was used by (Morlon et al. 2011), while fitting temperature-dependent models. Even though extinction seems to be poorly estimated in Cetacea and Ruminantia, it is unlikely to spuriously drive the dependency of speciation rates to environmental changes, as simulations suggest that environment-dependent extinction is not frequently mistaken for environment-dependent speciation.

Future Directions in the Inference of Environment-Dependent Diversification

It would be important to gain a better understanding of how differences in the shape of the paleoenvironmental curves influence our ability to make inferences. Indeed, while there were notable differences in the ability to recover some palaeoenvironmental curves compared with others (temperature and fossil diversity curves are more likely to be correctly recovered than the other variables), we could not characterize which features of these curves were influencing their differential recovery rates. Doing so will require simulating environmental curves with distinct and controlled characteristics.

There are many potential extensions of the environment-dependent models presented here. First, paleoenvironmental variables not included in this work could be considered, such as planetary albedo or the origination rate of marine guilds. Second, models accounting for the simultaneous effect of several paleoenvironmental variables as well as their interaction may provide more resolution on how various abiotic and/or biotic factors combine to influence species diversification. While fitting such models is already possible with the current implementation (any kind of dependence of speciation and extinction rates with environmental variables can be imputed, including dependencies on several environmental variables and their interactions), one would need to test the performance of such models using simulations before trusting the associated empirical results. Implementing the potential influence of hidden variables as well as a variable selection procedure would also be important directions for future work.

The way we have accounted for effects of the biotic environment is by considering interclade biotic interactions (e.g., the effect of ostracod diversity through time). However, there is a wide range of intraclade processes that are thought to influence diversification dynamics. For example, speciation may slow down as ecological opportunity recedes and interspecific competition intensifies following an adaptive radiation, resulting in a diversity-dependent effect (Phillimore and Price 2008; Etienne et al. 2012). Alternatively, speciation may accelerate if intra-clade species interactions (such as competition) promote reproductive isolation through phenotypic divergence (Seddon et al. 2013; Bacquet et al. 2015). It would already be possible, with currently available models, to test these alternative hypotheses: likelihoods are comparable, and a larger set of models, including for example diversity-dependent models but also models reflecting other types of heterogeneous diversification, could be included in the model comparison. However these hypotheses are not mutually exclusive, and future forms of the environment-dependent model, therefore, could be developed to additively or synergistically account for intraclade and interclade effects. Similar models were recently developed for analyzing fossils (Liow et al. 2015; Ezard and Purvis 2016). Models directly incorporating intraclade effects on phenotypic divergence have also started to be developed (Nuismer and Harmon 2015; Drury et al. 2016). Accommodating such factors in environment-dependent models would be useful for evaluating the relative effect of the environment, intra- and interclade species interactions, as well as their combined effect, on clade diversification.

We can also imagine adapting the model to analyse population genetic models on genealogical trees. Here, the environment-dependent model could be used to analyse how demographics (birth and death rates) have varied with time-dependent environmental variables (e.g., the rate of influx of immigrants in a region, the density of air pollution in a city). In traditional

population genetic models, population structure is measured as a function of gene flow (Wright 1931). Changes in gene flow can be the result of recent increases in population size or rates of dispersal; and so the factors that promote or limit gene flow cannot be easily inferred (Slatkin 1987). The environment-dependent model, adapted to analyse the effects of, for example, habitat range contraction, environmental disturbance, or precipitous geographic isolation on birth and death rates within a population may therefore help resolve the contribution of those factors to changes in the population. As above, one would need to test the performance of such models using simulations before trusting the associated empirical results.

The phylogenetic comparative toolbox is being extended to better account for the effect of environmental changes. Similar to its use here, where we have shown the effect of different environmental variables on diversification in clades, we have seen a growing body of work relating, for example, decreases in temperature to accelerations in diversification in birds (Claramunt and Cracraft 2015), increases in sea level to increased extinction in butterflies (Condamine et al. 2015), and even orogeny to Andean plant diversification (Pérez-Escobar et al. 2017). Models of phenotypic evolution accounting for environmental variations have similarly been developed recently, and have found decreases in temperature to drive increases in the rate of body mass evolution in birds and mammals (Clavel and Morlon 2017). Future analyses with these two types of models should allow us to better understand whether diversification and phenotypic evolution respond similarly (or differently) to past climatic changes, as well as the link between environmental variations, phenotypic divergence, and diversification.

Environment-dependent diversification models provide a statistical forum for testing a range of macroevolutionary hypotheses on species trees, implemented in user-friendly software. They provide a means for identifying which paleoenvironmental variable has most impacted diversification in a clade and how specifically that variable has impacted diversification. We have analyzed the statistical behavior of and provided guidelines for an environment-dependent model of species diversification. We think this will help researchers test fundamental macroevolutionary hypotheses on the effects of various abiotic and biotic factors on species diversification.

SUPPLEMENTARY MATERIAL

Data available from the Dryad Digital Repository: <http://dx.doi.org/10.5061/dryad.m96r3>.

FUNDING

Funding was provided through a European Research Council grant (ERC-CoG-PANDA) attributed to H.M.

ACKNOWLEDGMENTS

We would like to thank Julien Clavel, Jonathan Drury, Guilhem Sommeria Klein, Marc Manceau, Cesar Martinez, and Olivier Missa for helpful comments on the manuscript, Odile Maliet for comments and her code to simulate rate-shifted phylogenies, and Juan Cantalapiedra for supplying the distribution of resolved Ruminantia supertrees. E.L. would like to thank Evan Charles for helpful discussion.

REFERENCES

- Alroy J. 2010. Geographical, environmental and intrinsic biotic controls on Phanerozoic marine diversification: controls on phanerozoic marine diversification. *Palaeontology* 53:1211–1235.
- Bacquet P.M.B., Brattström O., Wang H.-L., Allen C.E., Löfstedt C., Brakefield P.M., Nieberding C.M. 2015. Selection on male sex pheromone composition contributes to butterfly reproductive isolation. *Proc. Biol. Sci.* 282:20142734.
- Barnosky A.D. 2001. Distinguishing the effects of the red queen and court jester on miocene mammal evolution in the northern rocky mountains. *J. Vertebr. Paleontol.* 21:172–185.
- Beaulieu J.M., O'Meara B.C. 2015. Extinction can be estimated from moderately sized molecular phylogenies. *Evolution* 69:1036–1043.
- Beerling D.J., Royer D.L. 2011. Convergent cenozoic CO₂ history. *Nat. Geosci.* 4:418–420.
- Benton M. 1995. Diversification and extinction in the history of life. *Science* 268:52.
- Benton M.J. 2009. The Red Queen and the Court Jester: species diversity and the role of biotic and abiotic factors through time. *Science* 323:728–732.
- Bibi F. 2013. A multi-calibrated mitochondrial phylogeny of extant Bovidae (Artiodactyla, Ruminantia) and the importance of the fossil record to systematics. *BMC Evol. Biol.* 13:166.
- Buchan A., LeCleir G.R., Gulvik C.A., González J.M. 2014. Master recyclers: features and functions of bacteria associated with phytoplankton blooms. *Nat. Rev. Microbiol.* 12:686–698.
- Burnham K.P., Anderson D.R. eds. 2004. Model selection and multimodel inference. New York (NY): Springer.
- Cantalapiedra J.L., Fitzjohn R.G., Kuhn T.S., Fernández M.H., DeMiguel D., Azanza B., Morales J., Mooers A.Ø. 2014. Dietary innovations spurred the diversification of ruminants during the Cenozoic. *Proc. Biol. Sci.* 281:20132746.
- Cerling T.E., Harris J.M., MacFadden B.J., Leakey M.G., Quade J., Eisenmann V., Ehleringer J.R. 1997. Global vegetation change through the Miocene/Pliocene boundary. *Nature* 389:153–158.
- Cermeño P., Falkowski P.G., Romero O.E., Schaller M.F., Vallina S.M. 2015. Continental erosion and the Cenozoic rise of marine diatoms. *Proc. Natl. Acad. Sci. U.S.A.* 112:4239–4244.
- Chen Z.-Q., Benton M.J. 2012. The timing and pattern of biotic recovery following the end-Permian mass extinction. *Nat. Geosci.* 5:375–383.
- Claramunt S., Cracraft J. 2015. A new time tree reveals Earth history's imprint on the evolution of modern birds. *Sci. Adv.* 1:e1501005.
- Clavel J., Morlon H. 2017. Accelerated body size evolution during cold climatic periods in the Cenozoic. *Proc. Natl. Acad. Sci. U.S.A.* 114:4183–4188.
- Coiffard C., Gomez B., Daviero-Gomez V., Dilcher D.L. 2012. Rise to dominance of angiosperm pioneers in European Cretaceous environments. *Proc. Natl. Acad. Sci. U.S.A.* 109:20955–20959.
- Condamine F.L., Rolland J., Morlon H. 2013. Macroevolutionary perspectives to environmental change. *Ecol. Lett.* 16(1 Suppl):72–85.
- Condamine F.L., Sperling F.A.H., Wahlberg N., Rasplus J.-Y., Kergoat G.J. 2012. What causes latitudinal gradients in species diversity? Evolutionary processes and ecological constraints on swallowtail biodiversity. *Ecol. Lett.* 15:267–277.
- Condamine F.L., Toussaint E.F.A., Clamens A.-L., Genson G., Sperling F.A.H., Kergoat G.J. 2015. Deciphering the evolution of birdwing butterflies 150 years after Alfred Russel Wallace. *Sci. Rep.* 5:11860.
- Drury J., Clavel J., Manceau M., Morlon H. 2016. Estimating the effect of competition on trait evolution using maximum likelihood inference. *Syst. Biol.* 65:700–710.
- Ehrlich P.R., Raven P.H. 1964. Butterflies and plants: a study in coevolution. *Evolution* 18:586.
- Erwin D.H. 2008. Macroevolution of ecosystem engineering, niche construction and diversity. *Trends Ecol. Evol.* 23:304–310.
- Erwin D.H. 2009. Climate as a driver of evolutionary change. *Curr. Biol.* 19:R575–R583.
- Etienne R.S., Haegeman B., Stadler T., Aze T., Pearson P.N., Purvis A., Phillimore A.B. 2012. Diversity-dependence brings molecular phylogenies closer to agreement with the fossil record. *Proc. Biol. Sci.* 279:1300–1309.
- Ezard T.H.G., Aze T., Pearson P.N., Purvis A. 2011. Interplay between changing climate and species' ecology drives macroevolutionary dynamics. *Science* 332:349–351.
- Ezard T.H.G., Purvis A. 2016. Environmental changes define ecological limits to species richness and reveal the mode of macroevolutionary competition. *Ecol. Lett.* 19:899–906.
- Ezard T.H.G., Quental T.B., Benton M.J. 2016. The challenges to inferring the regulators of biodiversity in deep time. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 371:20150216.
- FitzJohn R.G. 2012. Diversitree: comparative phylogenetic analyses of diversification in R: *Diversitree*. *Methods Ecol. Evol.* 3:1084–1092.
- Foote M., Miller A., Raup D., Stanley S. 2007. Principles of paleontology. New York, (NY): W. H. Freeman.
- Gerhart L.M., Ward J.K. 2010. Plant responses to low [CO₂] of the past. *New Phytol.* 188:674–695.
- Golub G.H., Heath M., Wahba G. 1979. Generalised cross-validation as a method for choosing a good ridge parameter. *Technometrics* 21:215–223.
- Hannisdal B., Peters S.E. 2011. Phanerozoic Earth system evolution and marine biodiversity. *Science* 334:1121–1124.
- Hansen B. 1999. Testing for linearity. *J. Econ. Surv.* 13:551–576.
- Hansen J., Sato M., Russell G., Kharecha P. 2013. Climate sensitivity, sea level and atmospheric carbon dioxide. *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.* 371:20120294–20120294.
- Höhna S. 2013. Fast simulation of reconstructed phylogenies under global time-dependent birth-death processes. *Bioinformatics* 29:1367–1374.
- Holland S.M., Patzkowsky M.E. 2015. The stratigraphy of mass extinction. *Palaeontology* 58:903–924.
- Katz M.E., Wright J.D., Miller K.G., Cramer B.S., Fennel K., Falkowski P.G. 2005. Biological overprint of the geological carbon cycle. *Mar. Geol.* 217:323–338.
- Keating K.I. 1978. Blue-Green algal inhibition of diatom growth: transition from mesotrophic to eutrophic community structure. *Science* 199:971–973.
- Kergoat G.J., Soldati L., Clamens A.-L., Jourdan H., Jabbour-Zahab R., Genson G., Bouchard P., Condamine F.L. 2014. Higher level molecular phylogeny of darkling beetles (Coleoptera: Tenebrionidae): darkling beetle phylogeny. *Syst. Entomol.* 39:486–499.
- Lazarus D. 1994. Neptune: a marine micropaleontology database. *Math. Geol.* 26:817–832.
- Lazarus D., Barron J., Renaudie J., Diver P., Türke A. 2014. Cenozoic planktonic marine diatom diversity and correlation to climate change. *PLoS One* 9:e84857.
- Liow L.H., Reitan T., Harnik P.G. 2015. Ecological interactions on macroevolutionary time scales: clams and brachiopods are more than ships that pass in the night. *Ecol. Lett.* 18:1030–1039.
- Liow L.H., Van Valen L., Stenseth N.C. 2011. Red Queen: from populations to taxa and communities. *Trends Ecol. Evol.* 26:349–358.
- Marx F.G., Uhen M.D. 2010. Climate, critters, and cetaceans: cenozoic drivers of the evolution of modern whales. *Science* 327:993–996.
- Mayhew P.J., Jenkins G.B., Benton T.G. 2008. A long-term association between global temperature and biodiversity, origination and extinction in the fossil record. *Proc. Biol. Sci.* 275:47–53.
- Miller K.G., Kominz M.A., Browning J.V., Wright J.D., Mountain G.S., Katz M.E., Sugarman P.J., Cramer B.S., Christie-Blick N., Pekar S.F. 2005. The Phanerozoic record of global sea-level change. *Science* 310:1293–1298.

- Mittelbach G.G., Schemske D.W., Cornell H.V., Allen A.P., Brown J.M., Bush M.B., Harrison S.P., Hurlbert A.H., Knowlton N., Lessios H.A., McCain C.M., McCune A.R., McDade L.A., McPeck M.A., Near T.J., Price T.D., Ricklefs R.E., Roy K., Sax D.F., Schluter D., Sobel J.M., Turelli M. 2007. Evolution and the latitudinal diversity gradient: speciation, extinction and biogeography. *Ecol. Lett.* 10: 315–331.
- Morlon H. 2014. Phylogenetic approaches for studying diversification. *Ecol. Lett.* 17:508–525.
- Morlon H., Lewitus E., Condamine F.L., Manceau M., Clavel J., Drury J. 2016. RPANDA: an R package for macroevolutionary analyses on phylogenetic trees. *Methods Ecol. Evol.* 7:589–597.
- Morlon H., Parsons T.L., Plotkin J.B. 2011. Reconciling molecular phylogenies with the fossil record. *Proc. Natl. Acad. Sci. U.S.A.* 108:16327–16332.
- Nee S., Holmes E.C., May R.M., Harvey P.H. 1994. Extinction rates can be estimated from molecular phylogenies. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 344:77–82.
- Nelder J.A., Mead R. 1965. A simplex method for function minimization. *Comput. J.* 7:308–313.
- Nuismer S.L., Harmon L.J. 2015. Predicting rates of interspecific interaction from phylogenetic trees. *Ecol. Lett.* 18:17–27.
- Oehlert A.M., Swart P.K. 2014. Interpreting carbonate and organic carbon isotope covariance in the sedimentary record. *Nat. Commun.* 5:4672.
- Park J.-Y., Kug J.-S., Bader J., Rolph R., Kwon M. 2015. Amplified Arctic warming by phytoplankton under greenhouse warming. *Proc. Natl. Acad. Sci.* 112:5921–5926.
- Pérez-Escobar O.A., Chomicki G., Condamine F.L., Karremans A.P., Bogarín D., Matzke N.J., Silvestro D., Antonelli A. 2017. Recent origin and rapid speciation of neotropical orchids in the world's richest plant biodiversity hotspot. *New Phytol.* 215:891–905.
- Phillimore A.B., Price T.D. 2008. Density-dependent cladogenesis in birds. *PLoS Biol.* 6:e71.
- Quental T.B., Marshall C.R. 2010. Diversity dynamics: molecular phylogenies need the fossil record. *Trends Ecol. Evol.* 25: 434–441.
- Rabosky D.L. 2010. Extinction rates should not be estimated from molecular phylogenies. *Evolution* 64:1816–1824.
- Rabosky D.L. 2016. Challenges in the estimation of extinction from molecular phylogenies: a response to Beaulieu and O'Meara. *Evolution* 70:218–228.
- Rabosky D.L., Goldberg E.E. 2015. Model inadequacy and mistaken inferences of trait-dependent speciation. *Syst. Biol.* 64:340–355.
- Raup D.M., Sepkoski J.J. 1982. Mass extinctions in the marine fossil record. *Science* 215:1501–1503.
- Raven J.A., Giordano M. 2014. *Algae*. *Curr. Biol.* 24:R590–R595.
- Roy K., Hunt G., Jablonski D. 2009. Phylogenetic conservatism of extinctions in marine bivalves. *Science* 325:733–737.
- Seddon N., Botero C.A., Tobias J.A., Dunn P.O., Macgregor H.E.A., Rubenstein D.R., Uy J.A.C., Weir J.T., Whittingham L.A., Safran R.J. 2013. Sexual selection accelerates signal evolution during speciation in birds. *Proc. Biol. Sci.* 280:20131065.
- Sepkoski J.J. 1984. A kinetic model of Phanerozoic taxonomic diversity. III. Post-Paleozoic families and mass extinctions. *Paleobiology* 10:246–267.
- Silvestro D., Schnitzler J., Liow L.H., Antonelli A., Salamin N. 2014. Bayesian estimation of speciation and extinction from incomplete fossil occurrence data. *Syst. Biol.* 63:349–367.
- Slatkin M. 1987. Gene flow and the geographic structure of natural populations. *Science* 236:787–792.
- Smith A.B. 2007. Marine diversity through the Phanerozoic: problems and prospects. *J. Geol. Soc.* 164:731–745.
- Smith J.M. 1989. The causes of extinction. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 325:241.
- Spicer R.A., Chapman J.L. 1990. Climate change and the evolution of high-latitude terrestrial vegetation and floras. *Trends Ecol. Evol.* 5:279–284.
- Steeman M.E., Hebsgaard M.B., Fordyce R.E., Ho S.Y.W., Rabosky D.L., Nielsen R., Rahbek C., Glenner H., Sorensen M.V., Willerslev E. 2009. Radiation of extant cetaceans driven by restructuring of the oceans. *Syst. Biol.* 58:573–585.
- Van Valen L. 1973. A new evolutionary law. *Evol. Theory* 1:1–30.
- Voje K.L., Holen Ø.H., Liow L.H., Stenseth N.C. 2015. The role of biotic forces in driving macroevolution: beyond the Red Queen. *Proc. R. Soc. B Biol. Sci.* 282: 20150186.
- von Humboldt A. 1860. *Ansichten der Natur: mit wissenschaftlichen Erläuterungen* vol. 1. Leipzig, Germany: JG Cotta.
- Vrba E.S. 1993. Turnover-pulses, the Red Queen, and related topics. *Am. J. Sci.* 293:418–452.
- Wallace A.R. 1878. *Tropical nature, and other essays*. London, UK: Macmillan and Company.
- Whitehead H., McGill B., Worm B. 2008. Diversity of deep-water cetaceans in relation to temperature: implications for ocean warming. *Ecol. Lett.* 11:1198–1207.
- Winkler A.M., Kochunov P., Blangero J., Almasy L., Zilles K., Fox P.T., Duggirala R., Glahn D.C. 2010. Cortical thickness or grey matter volume? The importance of selecting the phenotype for imaging genetics studies. *NeuroImage* 53:1135–1146.
- Wright S. 1931. Evolution in Mendelian populations. *Genetics* 16: 97–159.
- Zachos J.C., Dickens G.R., Zeebe R.E. 2008. An early Cenozoic perspective on greenhouse warming and carbon-cycle dynamics. *Nature* 451:279–283.