

How macroecology affects macroevolution: the interplay between extinction intensity and trait-dependent extinction in brachiopods.

Peter D. Smits^{1,*}

1. University of California – Berkeley, Berkeley, California 94720;

* Corresponding author; e-mail: psmits@berkeley.edu

Manuscript elements: Figure 1, figure 2, table 1, online appendices A and B (including figure A1 and figure A2). Figure 2 is to print in color.

Keywords: species selection, geographic range, environmental preference, paleobiology, Bayesian

Manuscript type: Article.

Prepared using the suggested L^AT_EX template for *Am. Nat.*

Abstract

As average species fitness decreases, do trait-based differences in species' fitness (selection) increase or decrease? Similarly, as taxonomic extinction intensity increases, how do the effects of species traits on taxonomic survival change? Do the selective differences associated with some traits change more than others? Using a hierarchical Bayesian approach, I develop a model of how the effects of biological traits on extinction risk can vary with respect to extinction intensity, origination cohort (i.e. time of origination), and in relation to each other. The emergent traits I analyze in relation to their patterns of Paleozoic brachiopod genus durations are geographic range, affinity for epicontinental seas versus open ocean environments, and body size. Additionally, I estimate the effects of environmental generalization versus specialization on taxonomic survival by allowing environmental preference to have a nonlinear effect on duration. My analytical framework eschews the traditional distinction between background and mass extinction, and instead considers extinction intensity as a continuum. I find that the cohort-specific effects of geographic range and environmental preference are negatively correlated with baseline extinction intensity. Additionally, I find support for greater survival of environmental generalists versus specialists in all origination cohorts. These results support the conclusion that for Paleozoic brachiopods, as extinction intensity increases overall extinction selectivity increases.

Introduction

Extinction is one half of the diversification process (Raup, 1994; Stanley, 1975, 1979), second only to speciation or origination in shaping changes to diversity; it can also be the ultimate manifestation of selection as a taxon with a beneficial trait should persist for longer on average than a taxon without that beneficial trait (Jablonski, 2008; Rabosky and McCune, 2010; Raup, 1994; Stanley, 1975). Species duration is a measure of species fitness CITATION, and trait-associated differences in fitness is the hallmark of (species) selection CITATION.

Jablonski (1986) observed that for bivalves at the end Cretaceous mass extinction event, previous trait-associated differences in survival no longer mattered except for the case of geographic range. Based on this evidence, Jablonski (1986) proposed the idea of "macroevolutionary modes" and that mass extinction and background extinction are fundamentally different processes. However, based on estimates of extinction rates over time, there is no evidence of there being two or more "types" of extinction (Wang, 2003). Instead, extinction rates for marine invertebrates form a unimodal distribution where estimates of extinction rate/intensity show continuous variation.

The apparent disconnect between the theory of macroevolutionary modes and the observation of continuous variation in extinction rates implies the possibility of a relationship between the strength of selection (extinction **intensity**) and the association between traits and differences in fitness (extinction **selectivity**) CITATION PAYNE. As extinction intensity increases, what happens to extinction selectivity? How do trait-associated differences in fitness change as average extinction rate changes over time?

Here I model brachiopod taxon durations as a function of multiple functional taxon traits because trait-dependent differences in extinction risk should be associated with differences in taxon duration CITATION. Brachiopods are an ideal group for this study as they have an exceptionally complete fossil record (Foote, 2000; Foote and Raup, 1996). I focus on the brachiopod record from the post-Cambrian Paleozoic, from the start of the Ordovician (approximately 485 My) through the end Permian (approximately 252 My) as this represents the time of greatest

global brachiopod diversity (Alroy, 2010) which results in a large sample size.

The analysis of taxon durations, or time from origination to extinction, falls under the purview of survival analysis, a field of applied statistics commonly used in health care and engineering (Klein and Moeschberger, 2003) but has a long history in paleontology (Crampton et al., 2016; Simpson, 1944, 1953; Smits, 2015; Van Valen, 1973, 1979). I adopt a hierarchical Bayesian modeling approach (Gelman et al., 2013; Gelman and Hill, 2007) in order to unify the previously distinct dynamic and cohort paleontological survival approaches (Baumiller, 1993; Crampton et al., 2016; Ezard et al., 2012; Foote, 1988; Raup, 1975, 1978; Simpson, 2006; Van Valen, 1973, 1979).

In order to test for an association between extinction intensity and extinction selectivity, both extinction rate and trait-based differences in extinction rate need to be estimated. Previous work has approached this problem by estimating the extinction intensity and selectivity at different time points or for different origination cohorts independently and then comparing those estimates CITATION. I find this approach problematic for a few reasons. First, by treating each time point or cohort as independent the estimates from each model are not relative to each other and they do not represent the greatest compromise across all available data CITATIONS. Second, by treating each time point or cohort as independent any and all post-hoc analyses are at risk of false positive results because of multiple comparisons CITATIONS. Third, post-hoc analysis of correlations between maximum likelihood estimates is much less satisfying or coherent than including those correlations as estimates of the initial model, something that is not possible when each time point or cohort is considered independent CITATION.

While estimation of both trait-dependent speciation and extinction rates from phylogenies of extant taxa has grown dramatically (Fitzjohn, 2010; Goldberg et al., 2011, 2005; Maddison et al., 2007; Rabosky et al., 2013; Stadler, 2011, 2013; Stadler and Bokma, 2013), there are two major ways to estimate trait-dependent extinction: analysis of phylogenies, and analysis of the fossil record. These two directions, phylogenetic comparative and paleobiological, are complementary and intertwined in the field of macroevolution (Hunt and Rabosky, 2014; Jablonski, 2008; Rabosky and McCune, 2010). In the case of extinction, analysis of the fossil record has the distinct advantage

over phylogenies of only extant taxa because extinction is observable; this means that extinction rate is possible to estimate (Liow et al., 2010; Quental and Marshall, 2009; Rabosky, 2010). The approach used here is thus complementary to the analysis of trait-dependent extinction based on a phylogeny.

Factors affecting brachiopod survival

Conceptually, taxon survival can be considered an aspect of “taxon fitness” (Cooper, 1984; Palmer and Feldman, 2012). Traits associated with taxon survival are thus examples of species (or higher-level) selection, as differences in survival are analogous to differences in fitness. The traits analyzed here are all examples of emergent and aggregate traits (Jablonski, 2008; Rabosky and McCune, 2010); specifically I analyze genus-level traits. Emergent traits are those which are not measurable at a lower level (e.g. species versus individual organism) such as geographic range, or even fossil sampling rate. Aggregate traits, like body size or environmental preference, are the average of a shared trait across all members of a lower level.

Geographic range is widely considered the most important biological trait for estimating differences in extinction risk at nearly all times, with large geographic range associated with low extinction risk (Finnegan et al., 2012; Harnik et al., 2012; Jablonski, 1986, 1987, 2008; Jablonski and Roy, 2003; Payne and Finnegan, 2007). This stands to reason even if extinction is completely at random; a taxon with an unrestricted range is less likely to go extinct at random than a taxon with a restricted range.

Epicontinental seas are a shallow-marine environment where the ocean has spread over the continental interior or craton with a depth typically less than 100m. In contrast, open-ocean coastline environments have much greater variance in depth, do not cover the continental craton, and can persist during periods of low sea level (Miller and Foote, 2009). Because of this, a simple hypothesis that taxa which favor epicontinental seas would be at great risk during periods of low sea levels, such as during glacial periods, when epicontinental seas are drained. During the Paleozoic (approximately 541–252 My), epicontinental seas were widely spread globally but

declined over the Mesozoic (approximately 252–66 My) and have nearly disappeared during the Cenozoic (approximately 66–0 My) as open-ocean coastlines became the dominant shallow-marine setting (Johnson, 1974; Miller and Foote, 2009; Peters, 2008; Sheehan, 2001). Taxa in epicontinental environments could also have a greater extinction susceptibility than taxa in open-ocean environments due to anoxic events due to enhanced water stratification or poor water circulation (Peters, 2007).

Miller and Foote (2009) demonstrated that during several mass extinctions taxa associated with open-ocean environments tend to have a greater extinction risk than those taxa associated with epicontinental seas. During periods of background extinction, however, they found no consistent difference between taxa favoring either environment. Miller and Foote (2009) hypothesize that open-ocean taxa may have a greater extinction rate because these environments would be more strongly affected by waterborne hazards such as fallout from impacts or volcanic events which would propagate more quickly than in epicontinental environments with sluggish circulation. These two environment types represent the primary identifiable environmental dichotomy observed in ancient marine systems (Miller and Foote, 2009; Sheehan, 2001). Given these findings, I would hypothesize that as extinction risk increases, the extinction risk associated with open-ocean environments should generally increase.

Because environmental preference is defined here as the continuum between occurring exclusively in open-ocean environments versus epicontinental environments, intermediate values are considered “generalists” in the sense that they favor neither end member. A long-standing hypothesis is that generalists or unspecialized taxa will have greater survival than specialists (Baumiller, 1993; Liow, 2004, 2007; Nürnberg and Aberhan, 2013, 2015; Simpson, 1944; Smits, 2015). Because of this, the effect of environmental preference was modeled as a quadratic function where a concave down relationship between preference and expected duration indicates that generalists are favored over specialists end-members.

Body size, measured as shell length, is also considered as a trait that may potentially influence extinction risk (Harnik, 2011; Payne et al., 2014). Body size is a proxy for metabolic activity and

other correlated life history traits (Payne et al., 2014). Harnik et al. (2014) analyzed the effect of body size selectivity in Devonian brachiopods in both a phylogenetic and non-phylogenetic context; finding that body size was not found to be associated with differences in taxonomic duration. It has also been found that, at least in the case of some bivalve subclades, body size can be as important a factor as geographic range size in determining extinction risk (Harnik, 2011). Given these results, I expect that if body size has any effect on brachiopod taxonomic survival it is very small.

It is well known that, given the incompleteness of the fossil record, the observed duration of a taxon is an underestimate of that taxon's true duration (Alroy, 2014; Foote and Raup, 1996; Liow and Nichols, 2010; Solow and Smith, 1997; Wagner and Marcot, 2013; Wang and Marshall, 2004). Because of this, the concern is that a taxon's observed duration may reflect its relative chance of being sampled and not any of the effects of the covariates of interest. In this case, for sampling to be a confounding factor there must be consistent relationship between the quality of sampling of a taxon and its apparent duration (e.g. greater sampling, longer duration). If there is no relationship between sampling and duration then interpretation can be made clearly; while observed durations are obviously truncated true durations, a lack of a relationship would indicate that the amount and form of this truncation is not a major determinant of the taxon's apparent duration. By including sampling as a covariate in the model, this effect can be quantified and can be taken into account when interpreting the estimates of the effects of the other covariates.

Methods

The brachiopod dataset analyzed here was sourced from the Paleobiology Database (<http://www.paleodb.org>) which was then filtered based on taxonomic (Rhynchonelliformea: Rhynchonellata, Chileata, Obolellida, Kutorginida, Strophomenida, Spiriferida), temporal (post-Cambrian Paleozoic), stratigraphic, and other occurrence information used in this analysis. Analyzed occurrences were restricted to those with paleolatitude and paleolongitude coordinates, have been assigned to either

epicontinental or open-ocean environment, and belonging to a genus present in the body size dataset (Payne et al., 2014). Epicontinental versus open-ocean assignments for each fossil occurrence are partially based on those from Miller and Foote (2009), with additional occurrences assigned similarly (Miller and Foote, personal communication). These filtering criteria are very similar to those from Foote and Miller (2013) with an additional constraint of being present in the body size data set from Payne et al. (2014). In total, there 1130 were genera included in the dataset.

Fossil occurrences were analyzed at the genus level which is common for paleobiological, macroevolutionary and macroecological studies of marine invertebrates (Alroy, 2010; Foote and Miller, 2013; Harnik et al., 2012; Kiessling and Aberhan, 2007; Miller and Foote, 2009; Nürnberg and Aberhan, 2013, 2015; Payne and Finnegan, 2007; Simpson and Harnik, 2009; Vilhena et al., 2013). While species diversity dynamics are frequently of much greater interest than those of higher taxa (though see Foote 2014; Hoehn et al. 2015), the nature of the fossil record makes accurate, precise, and consistent taxonomic assignments at the species level difficult for all occurrences. As such, the choice to analyze genera as opposed to species was in order to assure a minimum level of confidence and accuracy in the data analyzed here.

Genus duration was calculated as the number of geologic stages from first appearance to last appearance, inclusive. Durations were based on geologic stages as opposed to millions of years because of the inherently discrete nature of the fossil record; dates are not assigned to individual fossils themselves but instead fossils are assigned to a geological interval which represents some temporal range. In this analysis, stages are effectively irreducible temporal intervals in which taxa may occur. Genera with a last occurrence in or after Changhsingian stage (e.g. the final stage of the study interval) were right censored at the Changhsingian. How the likelihood of censored observations is calculated is detailed in section .

The covariates of duration included in this analysis are geographic range size (r), environmental preference (v, v^2), shell length (m) and sampling (s).

Geographic range was calculated as relative occupancy corrected for incomplete sampling.

First, the paleolatitude-paleolongitude coordinates for all occurrences were projected onto an equal-area cylindrical map projection. Each occurrence was then assigned to one of the cells from a 70×34 regular raster grid placed on the map. Each grid cell represents approximately 250,000 km². The map projection and regular lattice were made using shape files from <http://www.naturalearthdata.com/> and the raster package for R (Hijmans, 2015). For each stage, the total number of occupied grid cells was calculated. Then, for each temporal bin, the relative occurrence probability of the observed taxa was calculated using the JADE method developed by Chao et al. (2015). This method accounts for the fact that taxa with an occupancy of 0 cannot be observed which means that occupancy follows a truncated Binomial distribution. This correction is critical when comparing occupancies from different times with different geographic sampling. Finally, for each genus, the mean relative occurrence probability was calculated as the average of that genus' occurrence probabilities for all stages it was sampled to yield relative occupancy, my proxy for geographic range.

Environmental preference was defined as probability of observing the ratio of epicontinental occurrences to total occurrences ($\theta_i = e_i/E_i$) or greater given the background occurrence probability θ'_i as estimated from all other taxa occurring at the same time (e'_i/E'_i). This measure of environmental preference is expressed.

$$\begin{aligned} p(\theta'_i | e'_i, E'_i) &\propto \text{Beta}(e'_i, E'_i - e'_i)\text{Beta}(1, 1) \\ &= \text{Beta}(e'_i + 1, E'_i - e'_i + 1), \end{aligned} \tag{1}$$

where v is the percent of the distribution defined in equation 1 less than or equal to θ_i . The Beta distribution is used here because it is a continuous distribution bounded at 0 and 1, which is ideal for modeling percentages.

Body size, measured as shell length, was sourced directly from Payne et al. (2014). These measurements were made from brachiopod taxa figured in the *Treatise on Invertebrate Paleontology* (Williams et al., 2007).

To control for the effect of incomplete sampling on species duration, the number of samples of each genus over their duration was used as a covariate. This measure is a simple proxy

of differences in fossilization between genera that has been used before CITATION. the basic assumption of using this measure is that more abundant taxa may have more precise estimate of their duration while species with low abundance may have unduly truncated durations. By including this covariate as a predictor in this model, the hope is to help mitigate the effect of the incompleteness of the fossil record when estimating ecological differences in duration. The effect of sampling was assumed constant through time.

Prior to analysis, geographic range was logit transformed and the number of samples was natural-log transformed; both of these transformations make these variables defined for the entire real line. Sampling probability was transformed as $(s(n - 1) + 0.5)/n$ where n is the sample size as recommended by Smithson and Verkuilen (2006); this serves to slightly shrink the range of the data so that there are no values of 0 or 1. All covariates except for sampling were standardized by subtracting the mean from all values and dividing by twice its standard deviation, which follows Gelman and Hill (2007). This standardization means that the associated regression coefficients are comparable as the expected change per 1-unit change in the rescaled covariates. Finally, D is defined as the total number of covariates, excluding sampling, plus one for the intercept term.

Details of model

Hierarchical modelling is a statistical approach which explicitly takes into account the structure of the observed data in order to model both the within and between group variance (Gelman et al., 2013; Gelman and Hill, 2007). The units of study (e.g. genera) each belong to a single group (e.g. origination cohort). Each group is considered a draw from a shared probability distribution (e.g. prior) of all cohorts, observed and unobserved. The group-level parameters, or the hyperparameters of this shared prior, are themselves given (hyper)prior distributions and are also estimated like the other parameters of interest (e.g. covariate effects) (Gelman et al., 2013). The subsequent estimates are partially pooled together, where parameters from groups with large samples or effects remain large while those of groups with small samples or effects are pulled towards the overall group mean. All covariate effects (regression coefficients), as well as

the intercept term (baseline extinction risk), were allowed to vary by group (origination cohort). The covariance between covariate effects was also modeled.

Genus durations were assumed to follow a Weibull distribution which allows for age-dependent extinction (Klein and Moeschberger, 2003): $y \sim \text{Weibull}(\alpha, \sigma)$. The Weibull distribution has two parameters: scale σ , and shape α . When $\alpha = 1$, σ is equal to the expected duration of any taxon. α is a measure of the effect of age on extinction risk where values greater than 1 indicate that extinction risk increases with age, and values less than 1 indicate that extinction risk decreases with age. Note that the Weibull distribution is equivalent to the exponential distribution when $\alpha = 1$.

In the case of the right- and left-censored observations mentioned above, the probability of those observations has a different calculation (Klein and Moeschberger, 2003). For right-censored observations, the likelihood is calculated $p(y|\theta) = 1 - F(y) = S(y)$ where $F(y)$ is the cumulative distribution function. Taxa that existed for only a single stage were left-censored, which implies that that taxon went extinct at any point between 0 and 1 stages. In contrast to right-censored data, the likelihood of a left-censored observation is calculated from $p(y|\theta) = F(y)$. This censoring strategy improves model fit greatly over treating these taxa as being fully observed (see Appendix).

The scale parameter σ was modeled as a regression following Kleinbaum and Klein (2005) with both varying intercept and varying slopes and the effect of sampling; this is expressed

$$\sigma_i = \exp\left(\frac{-\mathbf{X}_i B_{j[i]} + \delta s_i}{\alpha}\right) \quad (2)$$

where i indexes across all observations where $i = 1, \dots, n$ where n is the total number of observations, $j[i]$ is the cohort membership of the i th observation where $j = 1, \dots, J$ where J is the total number of cohorts, X is a $N \times D$ matrix of covariates along with a column of 1's for the intercept term, B is a $J \times D$ matrix of cohort-specific regression coefficients, and δ is the regression coefficient for the effect of sampling s . δ does not vary by cohort.

Each of the rows of matrix B are modeled as realizations from a multivariate normal distri-

bution with length D location vector μ and $J \times J$ covariance matrix Σ : $B_j \sim \text{MVN}(\mu, \Sigma)$. The covariance matrix was then decomposed into a length J vector of scales τ and a $J \times J$ correlation matrix Ω , defined $\Sigma = \text{diag}(\tau)\Omega\text{diag}(\tau)$ where “diag” indicates a diagonal matrix.

The elements of μ were given independent normally distributed priors. The effects of geographic range size and the breadth of environmental preference were given informative priors reflecting the previous findings while the others were given weakly informative favoring no effect. The correlation matrix Ω was given an LKJ distributed prior (Lewandowski et al., 2009) that slightly favors an identity matrix as recommended by ?. These priors are defined

$$\begin{aligned}\mu^0 &\sim \mathcal{N}(0, 5) \\ \mu^r &\sim \mathcal{N}(-1, 1) \\ \mu^v &\sim \mathcal{N}(0, 1) \\ \mu^{v^2} &\sim \mathcal{N}(1, 1) \\ \mu^m &\sim \mathcal{N}(0, 0.5) \\ \delta &\sim \mathcal{N}(1) \\ \tau &\sim C^+(1) \\ \Omega &\sim \text{LKJ}(2).\end{aligned}\tag{3}$$

The log of the shape parameter α was given a weakly informative prior $\log(\alpha) \sim \mathcal{N}(0, 1)$ centered at $\alpha = 1$, which corresponds to the Law of Constant Extinction (Van Valen, 1973).

Posterior inference and posterior predictive checks

The joint posterior was approximated using a Markov-chain Monte Carlo routine that is a variant of Hamiltonian Monte Carlo called the No-U-Turn Sampler (Hoffman and Gelman, 2014) as implemented in the probabilistic programming language Stan (Stan Development Team, 2014). The posterior distribution was approximated from four parallel chains run for 40,000 steps, split half warm-up and half sampling and thinned to every 20th sample for a total of 4000 posterior samples. Chain convergence was assessed via the scale reduction factor \hat{R} where values close to

$\hat{R} < 1.1$) indicate approximate convergence. Convergence means that the chains are approximately stationary and the samples are well mixed (Gelman et al., 2013). After the model was fit to the data, 100 datasets were simulated from the posterior predictive distribution of the model. These simulations were used to test for adequacy of model fit as described below.

Survival analysis is complicated by censored observations where the ultimate time of extinction for some taxa could not be fully observed during the study window. Importantly, posterior predictive simulations for these observations must be similarly censored. To accomplish this, posterior predictive simulated durations for right-censored observations were the minimum of its final observed duration and the simulated duration. For left-censored individuals, their simulated duration was pegged at a minimum of one stage with simulated values less than one stage set to one.

Model adequacy was evaluated using a series of posterior predictive checks. Posterior predictive checks are a means for understanding model fit or adequacy. The concept of model adequacy is that if our model is an adequate descriptor of the observed data, then data simulated from the posterior predictive distribution should be similar to the observed given the same covariates, etc. (Gelman et al., 2013). Posterior predictive checks can take many forms but the basic idea is to compare some property of the empirical data to that property estimated from each of the simulated datasets. Additionally, for structured datasets like the one analyzed here, the fit of the model to different parts of the data can be assessed, which in turn can reveal a great deal if the model has good fit to some aspects of data but not others; it is in these scenarios when knowledge about the biology, geology, and paleoenvironment becomes important in order to explain what unmodeled processes might lead to these discrepancies between our data and the model (Gelman et al., 2013).

The types of posterior predictive tests used in this analysis fall into two categories: comparison of observed mean and median genus duration to a distribution of mean and median genus duration estimates from the posterior simulations, comparison of a non-parameteric estimate of the survival function from the observed data to estimates of that same survival function from the

simulations.

The survival function describes the probability of a taxon persisting given that it has survived up to time t ; this is expressed $P(T > t)$ because T is the true extinction time of the species and t is some arbitrary time of observation and we are estimating that probability that t is less than T . It is important to note, however, that the survival function does not reflect density of observations unlike e.g. histograms. Instead, this posterior predictive check reflects the model's ability to predict genus survival.

These posterior predictive tests were done for the entire data and for each of the origination cohorts.

Results

Model adequacy and posterior predictive checks

Comparisons between the observed distribution of durations to the distributions of 100 simulated datasets reveals the relatively good but heterogeneous fit of the model to the data (Fig. 1). The two major aspects of possible misfit that are observable are at durations of 1 and 8-9 stages. In both of these cases, the model slightly over-estimates the number of observations with duration of 1 stage, and slightly under-estimates the number of individuals with a duration of 8-9 stages. The goal of this model is estimating the expected duration of a genus given its covariate information. While the model estimates are not exact, it is possible that our model fits the “middle” of our data well but does poorly towards the tails.

Comparisons between the survival functions estimated from the empirical data and from 100 simulated datasets further expands the picture of model (mis)fit (Fig. 2). The survival curves of the 100 simulated datasets are very similar to the survival function estimated from the empirical data. The major points of potential misfit between the model and the data are underestimated percentage of taxa with duration 1 stage, and an over-estimate of probability of species surviving at least 9-16 stages. Importantly, the major divergence between the observed and estimated

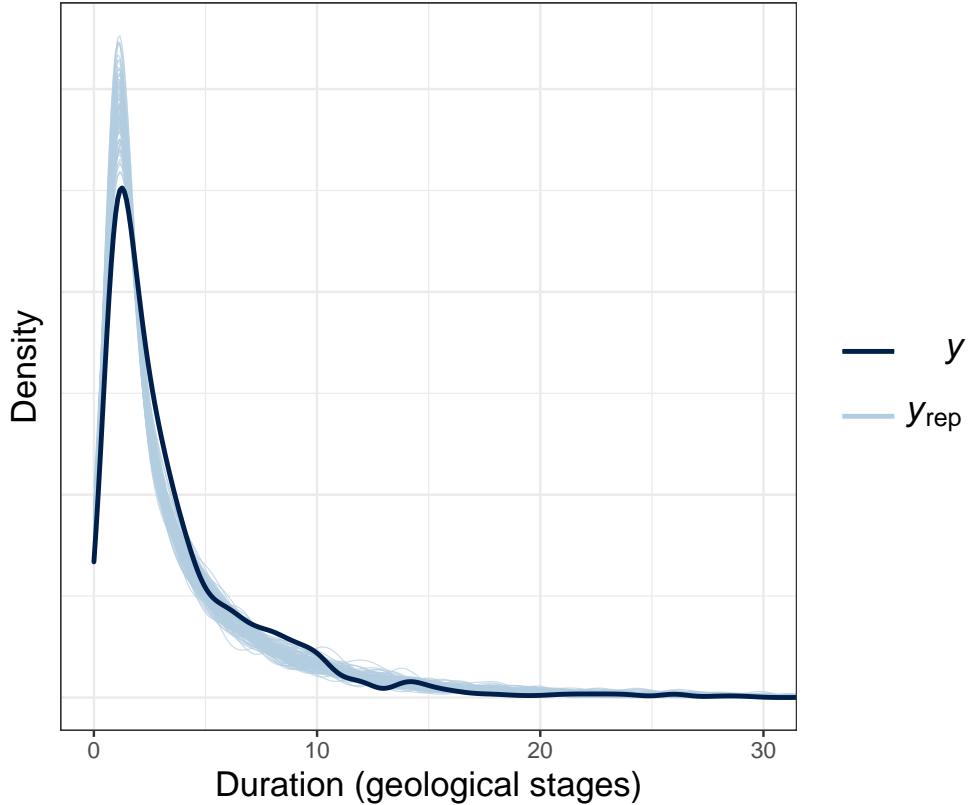


Figure 1: Comparison of the distribution of the observed data (black) to 100 simulated distributions (blue). This is a close-up view of the bulk of the distribution which shows the more subtle aspects of (mis)fit between the data and the model.

applies to taxa with a less than 15% probability of continuing to surviving. Keep in mind, also, that the survival curve as presented does not depict density of observations as in Fig. 1.

In addition to distributional comparisons, model adequacy at the total data level was assessed through comparison of the mean and median of the observed data to those from simulated data sets. While the previous posterior predictive checks have focused on the relatively good but heterogeneous fit of the model to the entire distribution of the data, assessment of the models ability to predict the mean and median of the observed data appears adequate (Fig. ??). The principle goal of this model is to obtain adequate prediction of how a taxon's expected duration for a given set of ecological covariates, as such the seemingly adequate fit of our model to average

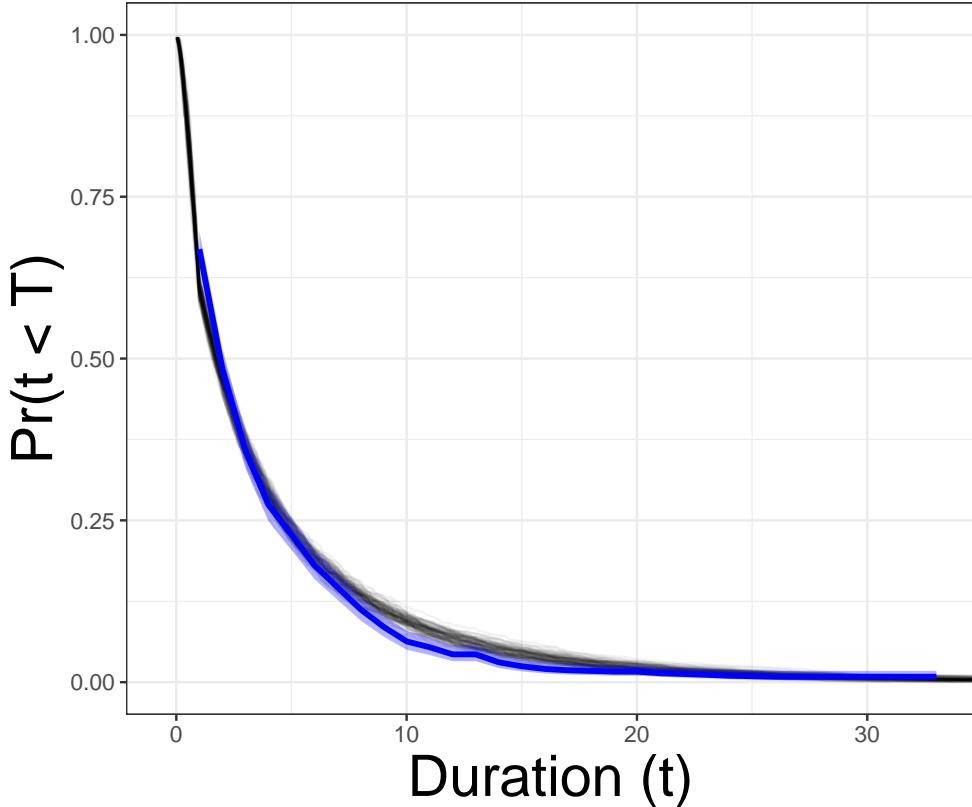


Figure 2: Comparison of the empirical estimate of $S(t)$ (highlighted) versus estimates from 100 posterior predictive data sets (black). $S(t)$ corresponds to the probability that the age of a genus t is less than the genus' ultimate duration T .

taxon duration is reassuring (Fig. 3). Additionally, given the skewness of the observed taxon durations (Fig. 1, the ability for the model to recapitulate the median observed taxon duration is important.

When considered together, all of the above posterior predictive checks indicate weak, but approximately adequate model fit for key questions such as expected taxon duration (Fig. 3). However, there is obviously heterogeneity in model fit because, while the model can recapitulate some aspects of the observed data (Fig. ??), there are obvious discrepancies between the model and the data (Fig. ??). By performing the same posterior predictive tests for each of the origination cohorts, it may be possible to get a better picture of the source of model misfit.

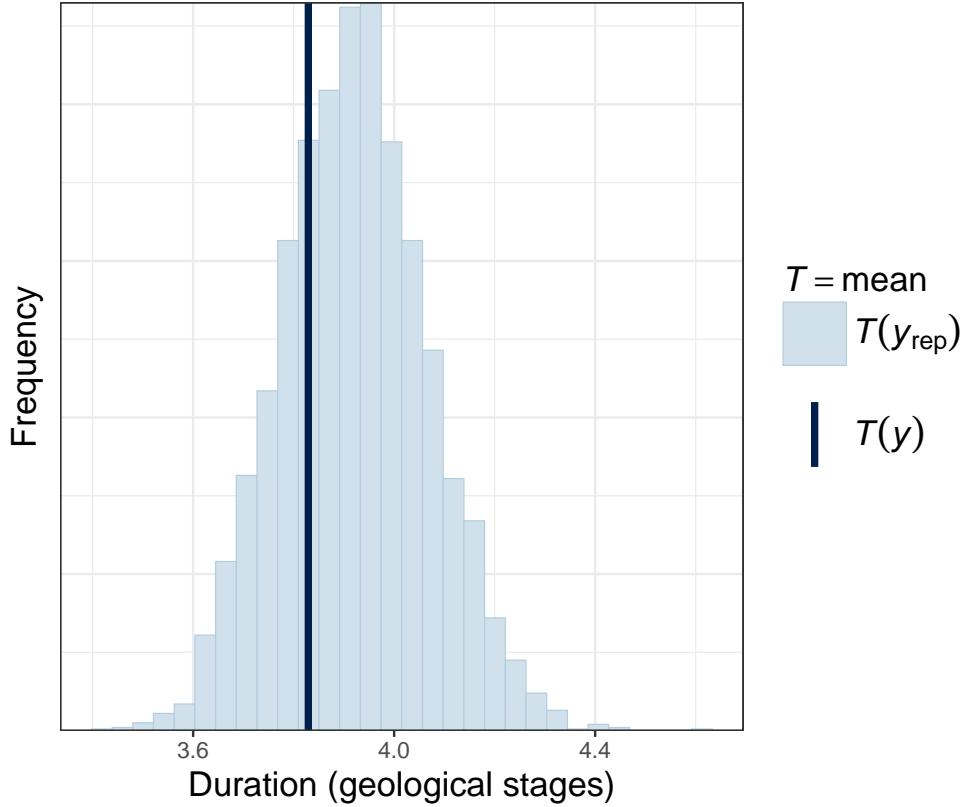


Figure 3: Comparison of the observed mean genus duration (black vertical line) to a distribution of means estimated from 100 simulated datasets (blue). Model fit is evaluated by the similarity between the observed and the estimated, where good fit is demonstrated by the vertical line being “within” the simulated distribution.

When the posterior predictive tests are visualized for each of the origination cohorts, a complex picture of model fit emerges. Comparison between the empirical survival functions estimated for each cohort to those estimated from the simulated datasets reveals the degree of heterogeneity in model fit (Fig. 5, as some origination cohorts appear to be very well fit by the model (e.g. Tremadoc, Darriwilian Wenlock, Ludlow, Lochkovian, etc.) while others are more poorly fit by the model (e.g. Visean, Bashirian, Moscovian, Stephanian, Roadian, etc.). The poverty of model fit to these cohorts may indicate that these origination cohorts are undergoing a different extinction process whose aspects are unmodeled in this analysis. However, for those cohorts

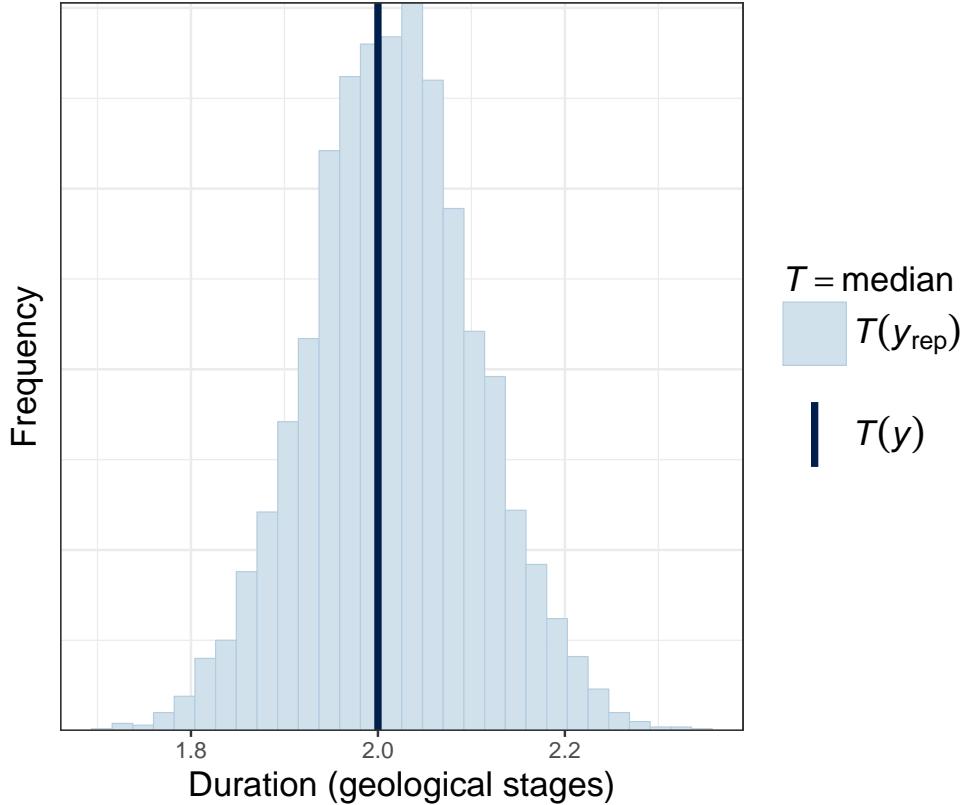


Figure 4: Comparison of the observed median genus duration (black vertical line) to a distribution of medians estimated from 100 simulated datasets (blue). Model fit is evaluated by the similarity between the observed and the estimated, where good fit is demonstrated by the vertical line being “within” the simulated distribution.

where our model recapitulates the empirical survival function indicates approximately adequate fit and that our model may be capturing some of the processes underlying taxon extinction.

For nearly every origination cohort, the model is able to approximately recapitulate the observed mean duration (Fig. 6). The only exceptions to this observed mean is within the tails of the posterior predictive distributions (e.g. Moscovian, Sakmarian, Wordian, etc.). In general, however, posterior predictive distribution of the model is able to adequately estimate the mean duration of nearly all origination cohorts.

In comparison, the model has a much more heterogeneous fit to each origination cohort’s

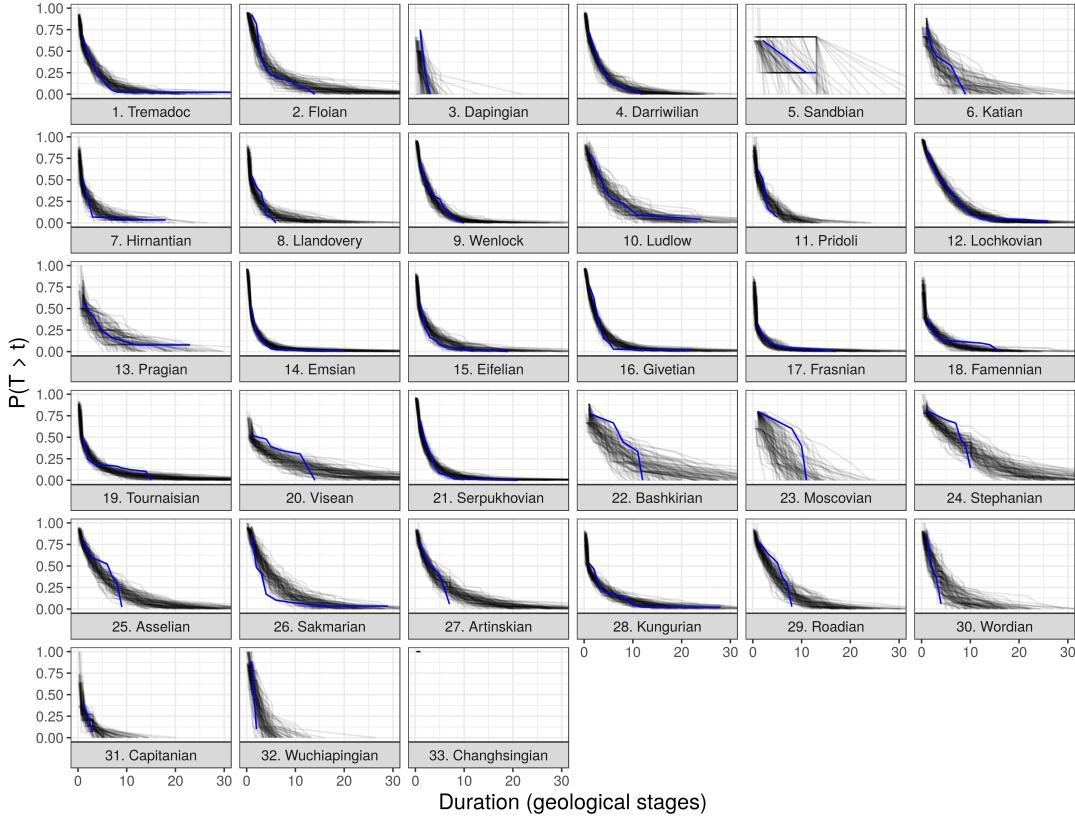


Figure 5: Comparison of the empirical estimate of $S(t)$ (highlighted) versus estimates from 100 posterior predictive data sets (black) for each of the origination cohorts. $S(t)$ corresponds to the probability that the age of a genus t is less than the genus' ultimate duration T . By comparing the fit of the model to the individual cohorts, when and where the model (mis)fits is more observable.

median taxon duration (Fig. 7). The skewness of the distribution underlying (Fig. 1) means that for some origination cohorts, median duration might be pegged at 1 stage; this means that the posterior predictive distributions for some cohorts can be extremely skewed. The cohorts with notably inadequate poor predictions are the Hirnantian, Priodoli, Emsian, Eifelian, Givetian, Asselian, Sakmarian, Artinskian, Kungurian, and Roadian. The remaining 13 cohorts, however, have decently adequate fit. These results indicate that our model is very good at recapitulate mean taxon duration (Fig. 3, 6), and it is capable of esimtaing overall median duration and median duration of most origination cohorts (Fig. 4, 7).

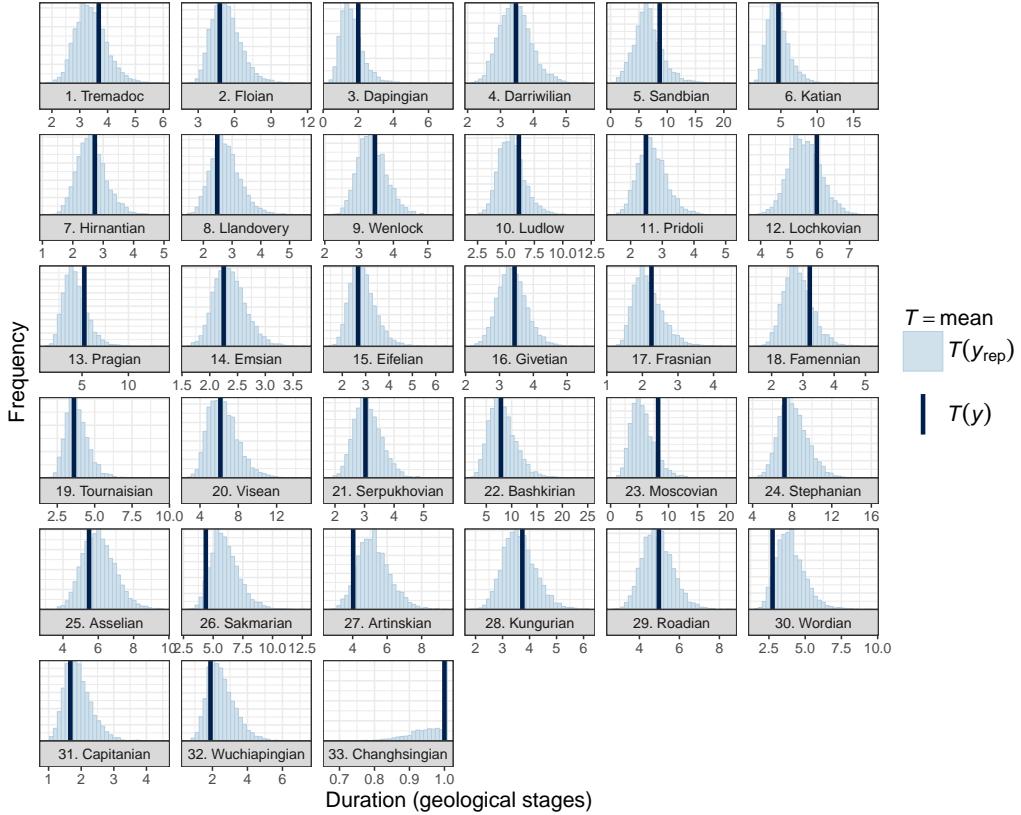


Figure 6: Comparison of the observed mean genus duration (black vertical line) to a distribution of means estimated from 100 simulated datasets (blue) for each of the origination cohorts. Model fit is evaluated by the similarity between the observed and the estimated, where good fit is demonstrated by the vertical line being “within” the simulated distribution.

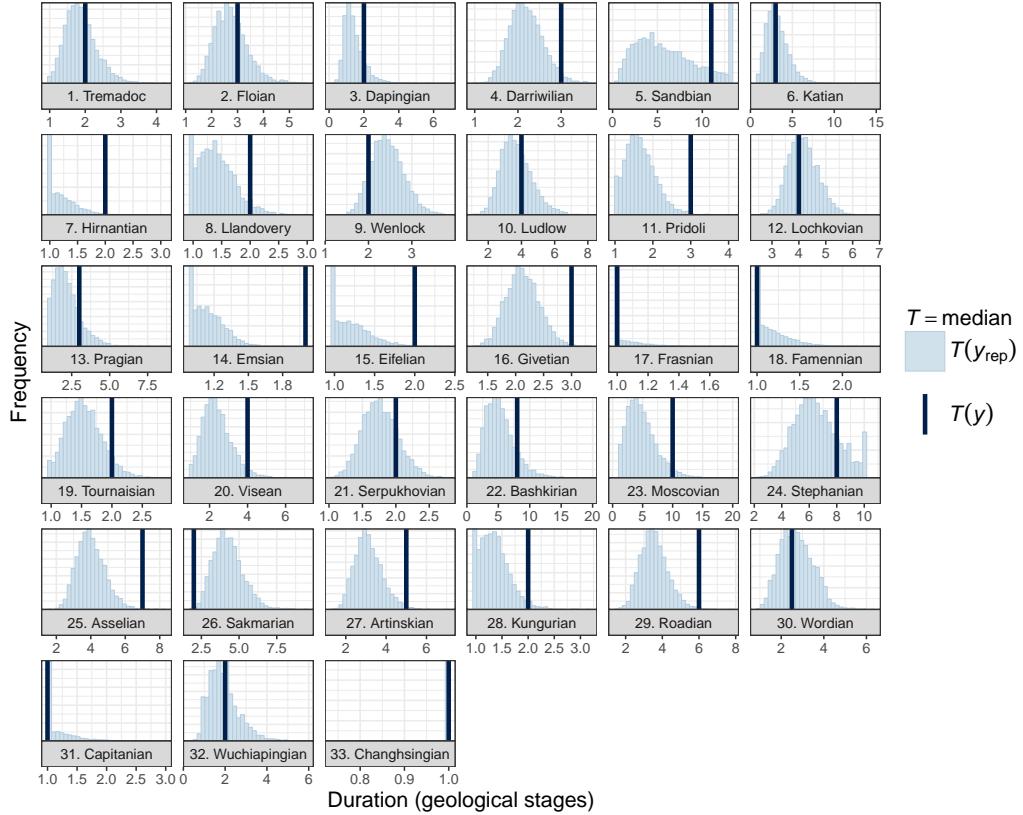


Figure 7: Comparison of the observed median genus duration (black vertical line) to a distribution of medians estimated from 100 simulated datasets (blue) for each of the origination cohorts. Model fit is evaluated by the similarity between the observed and the estimated, where good fit is demonstrated by the vertical line being “within” the simulated distribution.

Table 1: Estimates of various parameters in the model used here. These include group-level estimates of the effects of biological traits on brachiopod generic survival, the standard deviation of the between-cohort effects, as well as the estimates of both the effect of sampling δ and the Weibull shape parameter α . The mean, standard deviation (SD), 10th, 50th, and 90th quantiles of the marginal posteriors are presented.

type	parameter	effect of	mean	SD	10%	50%	90%
Mean	μ^i	intercept	-2.34	0.14	-2.52	-2.34	-2.16
	μ^r	geographic	1.19	0.21	0.93	1.19	1.45
		range					
	μ^v	environmental	-0.34	0.20	-0.59	-0.34	-0.07
		preference					
Standard deviation	μ^{v^2}	environmental	2.15	0.30	1.76	2.15	2.53
		preference ²					
	μ^m	body size	0.05	0.11	-0.09	0.05	0.19
	τ^i	intercept	0.51	0.11	0.38	0.50	0.65
	τ^r	geographic	0.77	0.17	0.55	0.75	1.00
Other	τ^v	range					
		environmental	0.97	0.18	0.76	0.95	1.20
	τ^{v^2}	preference					
		environmental	1.08	0.34	0.66	1.07	1.52
		preference ²					
	τ^m	body size	0.41	0.11	0.27	0.41	0.56
Other	δ	sampling	-2.77	0.15	-2.97	-2.77	-2.58
	α	"time"	1.61	0.05	1.54	1.61	1.67

The cohort-level estimate of the effect of geographic range size indicates that as a taxon's geographic range increases, that taxon's duration is expected to increase (Table 1). Given the estimates of μ^r and τ^r , there is a less than 3.7% ($\pm 0.04\%$ SD) probability that this relationships would be reversed ($\text{Pr}(\mathcal{N}(\mu^r, \tau^r) > 0)$). The between-cohort variance τ^r is the lowest of all the regression coefficients (Table 1).

Body size is estimated to have no effect on taxon duration, with the estimate being nearly 0 (Table 1). The variance between the cohort-specific estimates of the effect of body size τ^m is estimated to be greater than the variance of between-cohort estimates of the effect of geographic range size τ^r .

The group-level estimate of the effect of environmental preference is estimated from both μ^v and μ^{v^2} .

The estimate of μ^v indicates that epicontinental favoring taxa are expected to have a greater duration than open-ocean favoring taxa (Table 1). Additionally, given the estimate of between-cohort variance τ^v , there is approximately 18% ($\pm 7\%$ SD) probability that, for any given cohort, taxa favoring open-ocean environments would have a greater expected duration than taxa favoring epicontinental environments ($\text{Pr}(\mathcal{N}(\mu^v, \tau^v) > 0)$).

The estimate of μ^{v^2} indicates that the overall relationship between environmental preference and $\log(\sigma)$ is concave down (Fig. 8), with only a 2.7% ($\pm 3\%$ SD) probability that any given cohort is convex up ($\text{Pr}(\mathcal{N}(\mu^{v^2}, \tau^{v^2}) < 0)$).

The cohort-specific estimates of all the regression coefficients demonstrate a lot of between cohort variance, with no obvious trends. As indicated in Table 1 and detectable visually (Fig. 9), the between-cohort estimates for β^0 , β^r , and β^m all have much lower variance than the between-cohort estimates of both β^v and β^{v^2} .

While most cohort-specific estimates are very similar to the overall cohort-level estimate, there are a few notable excursions away from the overall mean (Fig. 9). There are simultaneous excursions in both β^0 and β^v for cohorts originating in the Givetian (387-382 My) and Frasnian (382-372 My) stages; both of which directly precede the late Devonian mass extinction event at

the Frasnian/Famennian boundary. These cohorts are marked by both a high extinction intensity and an increase in expected duration for taxa favoring epicontinental environments over open-ocean ones; this is consistent with the results of Miller and Foote (2009).

Cohorts originating from the Silurian through the Early Devonian have a slightly lower extinction intensity than the overall mean; these cohorts are those originating in the Llandovery (443-443 My) through the Emsian (407-393 My). This is also a time period is also when there is the lowest overall probability that epicontinental favoring taxa are expected to have greater duration than open-ocean favoring taxa. Both the Silurian and Devonian periods are notable for having been periods with a mostly “hothouse” climate, with no polar icecaps and a high sea-level (Edwards and Fanning, 1985; Joachimski et al., 2009; Munnecke et al., 2010).

The cohort-specific relationships between environmental preference and $\log(\sigma)$ were calculated from the estimates of β^0 , β^v , and β^{v^2} (Fig. 10) and reflect how these three parameters act in concert and not just individually (Fig. 9). Beyond results already discussed above in the context of the parameters individually, it is notable that the cohort originating in the Kungurian (279-272 My) is least like the overall expected relationship and has the most sharply curved appearance due to a high estimate β^{v^2} (Fig. 9). This cohort has the biggest difference in extinction risk between environmental generalists and specialists. The cohorts originating during the Emsian (407-393 My) and Frasnian (382 - 372 My) are tied for second in sharpness of curvature. The least sharply curved cohorts include those originating during Tremadocian (484-477 My), Hirnantian (445-443 My), Llandovery (443-433 My), and Ludlow (427-423 My). Except for the Tremadocian cohort, most of these cohorts originate during the Silurian through the Early Devonian range identified earlier as having lower expected extinction intensity than what is expected from the group-level estimate.

The correlations of the cohort-specific estimates of the regression coefficients are estimated as the off-diagonal elements of the correlation matrix Ω . Only two of the elements of Ω are distinguishable from 0: the correlation of β^0 (extinction intensity) with both β^r and β^v (Fig. 11).

There is an approximate 90% probability that the cohort-specific estimates of baseline ex-

tinction intensity β^0 and the effect of geographic range β^r are negatively correlated; this means that for cohorts experiencing a lower extinction intensity (β^0 decreases), the magnitude of the effect of geographic range is expected to decrease as well, and *vice versa*; this is in contrast to the observation made by Jablonski (1986) with regards to late Cretaceous bivalves.

Similarly, there is an approximate 97.4% probability that the cohort-specific estimates of β^0 and β^v are negatively correlated; this means that as extinction intensity increases it is expected that epicontinental taxa become more favored over open-ocean environments (i.e. as β^0 increases, β^v decreases).

There is only an approximate 30% probability that β^r and β^v are positively correlated. This lack of cross-correlation may be due in part to the much higher between-cohort variance of the effect of environmental preference τ^v than the very small between-cohort variance in the effect of geographic range τ^r (Table 1); the effect of geographic range might simply not vary enough relative to the much noisier environmental preference.

Sampling was found to have a negative effect (positive δ) on duration: greater sampling, shorter duration (Table 1). While potentially counter intuitive, this result is most likely due to some long lived taxa only be sampled in the stages of the first and last appearance. Also, longer lived taxa have more opportunities to not be sampled than shorter lived taxa. These two factors will lead to this result.

While the effect of sampling appears large compared to the other regression coefficients, this is only because sampling was not standardized like the other covariates. To make the coefficients comparable, δ is multiplied by twice the posterior mean of the standard deviation of sampling probability; the transformed value of δ is then 0.642 (± 0.1 SD). This effect is relatively small compared to the other covariate effects (Table 1). There is then a 98.6% probability that the effect of geographic range μ^r has a greater magnitude than δ . Similarly, μ^v has a 71.8% probability of having a greater magnitude of effect than δ . Finally, μ^{v^2} has a 100% probability of having a greater magnitude of effect than δ .

The Weibull shape parameter α was found to be approximately 1.36 (± 0.05 SD) with a 100%

probability of being greater than 1. This result is not consistent with the Law of Constant Extinction (Van Valen, 1973) and is instead consistent with accelerating extinction risk with taxon age. This may indicate that older taxa are out-competed by younger taxa, a result consistent with some empirical results (Quental and Marshall, 2013; Smits, 2015; Wagner and Estabrook, 2014) and (ironically) with a recently proposed Red Queen-like model of evolution (Rosindell et al., 2015). This results, however, is not consistent with other empirical results (Crampton et al., 2016; Finnegan et al., 2008) and could potentially be caused by the minimum resolution of the fossil record (Sepkoski, 1975). It is thus unclear if a strong biological inference can be made from this result, which means that further work is necessary on the effect of taxon age on extinction risk.

Discussion

The generating observation behind this study was that for bivalves at the end Cretaceous mass extinction event, the only biological trait that was found the affect extinction risk was geographic range while traits that had previously been beneficial had no effect (Jablonski, 1986). This observation raises two linked questions: how does the effect of geographic range change with changing extinction intensity, and how does the effect of other biological traits change with changing extinction intensity?

I find that as intensity increases (β^0 decreases), the magnitude of the effect of geographic range increases. I also find that as intensity increases, the effect of favoring epicontinental environments of open-ocean environments is expected to be increase; this is consistent with the results of Miller and Foote (2009). There is no evidence for a correlation between the effect of geographic range and environmental preference. Additionally, the between-cohort variance in effect of geographic range is much less then the between-cohort variance of the effect of environmental preference which may underlie the lack of correlation between these two effects.

Additionally, the lower between-cohort variance in the effect of geographic range versus that higher between-cohort variance implies that for cohorts with a greater than average extinction

intensity, the difference in the effect geographic range and the group-level effect of geographic range is expected to be smaller than the difference between the effect of environmental preference and the group-level effect of environmental preference.

I find consistent support for the “survival of the unspecialized,” with respect to epicontinental versus open-ocean environmental preference, as a time-invariant generalization of brachiopod survival; taxa with intermediate environmental preferences are expected to have lower extinction risk than taxa specializing in either epicontinental or open-ocean environments (Fig. 8), though the curvature of the relationship varies from rather shallow to very peaked (Fig. 10). However, this relationship is not symmetric about 0, as taxa favoring epicontinental environments are expected to have a greater duration than taxa favoring open-ocean environments. This description of environment only describes one major aspect of a taxon’s environmental context, with factors such as bathymetry and temperature being further descriptors of a taxon’s adaptive zone (Harnik, 2011; Harnik et al., 2012; Heim and Peters, 2011; Nürnberg and Aberhan, 2013); inclusion of these factors in future analyses would potentially improve our understanding of the “survival of the unspecialized” hypothesis (Simpson, 1944).

Hopkins et al. (2014), in their analysis of niche conservatism and substrate lithological preference in marine invertebrates, found that brachiopods were among the least “conservative” groups; taxa were found to easily change substrate preference on short time scales. While substrate preference is not the same as environmental preference (as defined here), a question does arise: are there three classes of environmental preference instead of two? These classes would be taxa with broad tolerance (“true” generalists), inflexible specialists (“true” specialists), and flexible but with a narrow tolerance. A flexible taxon is one with a narrow habitat preference at one time, but with preference that changes over time with changing environmental availability. My analysis assumes that traits are constant over the duration of the taxon meaning that this scenario is not detectable; taxa with broad tolerances and flexible taxa with narrow per-stage preference end up being treated the same way. Future work should explore how environmental preference changes over lineage duration in relation to environmental availability to estimate if

the generalists–specialists continuum is actually ternary relationship.

An alternative approach for specifically modeling survival that can take into account imperfect observation than the method used here is the Cormack-Jolly-Seber (CJS) model (Liow et al., 2008; Liow and Nichols, 2010; Royle and Dorazio, 2008; Tomiya, 2013). This model is a type of hidden Markov model with an absorbing state (i.e. extinction). In this model, survival is defined as the probability of surviving from time t to time $t + 1$. Additionally, the effect of preservation and sighting is estimated as probability of observing a taxon that is present; this can extend the duration of a taxon beyond its last occurrence. This approach is a fundamentally different from the method used in my analysis: I am estimating the biasing effect of sampling probability on taxon duration to then compare with effects of other covariates, while the CJS model estimates the pre-sampling fossil record and then estimates per-time unit survival probability.

The use of genera as the unit of the study and how to exactly interpret the effects of the biological traits is an important question. For example, if any of the traits analyzed here are associated with increases in speciation rates, this might increase the duration of genera through self-renewal (Raup, 1991, 1994), which would be an example of the difference in biological pattern between species and genera (Jablonski, 1987, 2007, 2008). This could lead to a trait appearing to decrease generic level extinction risk by that trait increasing species level origination rate instead of decreasing species level extinction risk.

The model used here could be improved through either increasing the number of analyzed traits, expanding the hierarchical structure of the model to include other major taxonomic groups of interest, and the inclusion of explicit phylogenetic relationships between the taxa in the model as an additional hierarchical effect. An example trait that may be of particular interest is the affixing strategy or method of interaction with the substrate of the taxon, which has been found to be related to brachiopod survival where, for cosmopolitan taxa, taxa that are attached to the substrate are expected to have a greater duration than those that are not (Alexander, 1977).

It is theoretically possible to expand this model to allow for comparisons both within and between major taxonomic groups which would better constrain the brachiopod estimates while

also allowing for estimation of similarities and differences in cross-taxonomic patterns. The major issue surrounding this particular expansion involves finding a similarly well sampled taxonomic group that is present during the Paleozoic. Potential groups include Crinoidea, Ostracoda, and other members of the “Paleozoic fauna” (Sepkoski, 1981).

With significant updates, it would also be possible to compare the brachiopod record with modern groups such as bivalves or gastropods (Sepkoski, 1981), though remembering that the groups may not necessarily share all cohorts with the brachiopods. This particular model expansion would act as a test of any universal cross-taxonomic patterns in the effects of emergent traits on extinction such as has been proposed for geographic range (Payne and Finnegan, 2007). Additionally, this expanded model would also act as a test of the distinctness of the Sepkoski (1981) three-fauna hypothesis in terms of trait-dependent extinction.

Traits like environmental preference or geographic range (Hunt et al., 2005; Jablonski, 1987) are most likely heritable. Without phylogenetic context, this analysis assumes that differences in extinction risk between taxa are independent of the shared evolutionary history of those taxa (Felsenstein, 1985). In contrast, the origination cohorts only capture shared temporal context. For example, if taxon duration is phylogenetically heritable, then closely related taxa may have more similar durations as well as more similar biological traits. Without taking into account phylogenetic similarity the effects of these biological traits would be inflated solely due to inheritance. The inclusion of phylogenetic context as an additional individual-level hierarchical effect, independent of origination cohort, would allow for determining how much of the observed variability is due to shared evolutionary history versus shared temporal context versus actual differences associated with biological traits (Smits, 2015).

The combination and integration of the phylogenetic comparative and paleontological approaches requires both sources of data, something which is not possible for this analysis because there is no phylogenetic hypothesis for all Paleozoic taxa, something that is frequently the case for marine invertebrates with a good fossil record. When both data sources are available has been possible, however, the analysis can more fully address the questions of interest in macroevolution

(Fritz et al., 2013; Harnik et al., 2014; Raia et al., 2012, 2013; Simpson et al., 2011; Slater et al., 2012; Slater, 2013, 2015; Smits, 2015; Tomiya, 2013).

Conclusion

In summary, patterns of Paleozoic brachiopod survival were analyzed using a fully Bayesian hierarchical survival modelling approach while also eschewing the traditional separation between background and mass extinction. I find that cohort extinction intensity is negatively correlated with both the cohort-specific effects of geographic range and environmental preference. These results imply that as extinction intensity increases (β^0) increases, it is expected that both effects will increase in magnitude. However, the change in effect of environmental preference is expected to be greater than the change in the effect of geographic range. Additionally, I find support for greater survival in environmental generalists over specialists in all origination cohorts analyzed; this is consistent with the long standing “survival of the unspecialized” hypothesis (Baumiller, 1993; Liow, 2004, 2007; Nürnberg and Aberhan, 2013, 2015; Simpson, 1944, 1953; Smits, 2015). The results of this analysis support the conclusion that for Paleozoic brachiopods, as extinction intensity increases overall extinction selectivity is expected to increase as well.

Acknowledgments

I would like to thank K. Angielczyk, M. Foote, P. D. Polly, R. Ree, and G. Slater for helpful discussion during the conception of this study. I'd also like to thank D. Bapst, N. Pierrehumbert and M. Villarosa Garcia for additional comments. An additional thank you to A. Miller for the epicontinental versus open-ocean assignments. Finally, thank you to the reviewers for their helpful comments that improved this manuscript. This entire study would not have been possible without the Herculean effort of the many contributors to the Paleobiology Database. In particular, I would like to thank J. Alroy, M. Aberhan, D. Bottjer, M. Clapham, F. Fürsich, N. Heim, A. Hendy, S. Holland, L. Ivany, W. Kiessling, B. Kröger, A. McGowan, T. Olszewski, P.

Novack-Gottshall, M. Patzkowsky, M. Uhen, L. Villier, and P. Wager. This work was supported by a NASA Exobiology grant (NNX10AQ446) to A. Miller and M. Foote. I declare no conflicts of interest. This is Paleobiology Database publication XXX. The code necessary for reproducing these results is available as a Zenodo archive DOI 10.5281/zenodo.46928.

Literature Cited

- Alexander, R. R. 1977. Generic longevity of articulate brachiopods in relation to the mode of stabilization on the substrate. *Palaeogeography, Palaeoclimatology, Palaeoecology* 21:209–226.
- Alroy, J. 2010. The Shifting Balance of Diversity Among Major Marine Animal Groups. *Science* 329:1191–1194.
- . 2014. A simple Bayesian method of inferring extinction. *Paleobiology* 40:584–607.
- Baumiller, T. K. 1993. Survivorship analysis of Paleozoic Crinoidea: effect of filter morphology on evolutionary rates. *Paleobiology* 19:304–321.
- Chao, A., T. C. Hsieh, R. L. Chazdon, R. K. Colwell, N. J. Gotelli, and B. D. Inouye. 2015. Unveiling the species-rank abundance distribution by generalizing the Good-Turing sample coverage theory. *Ecology* 96:1189–1201.
- Cooper, W. S. 1984. Expected time to extinction and the concept of fundamental fitness. *Journal of Theoretical Biology* 107:603–629.
- Crampton, J. S., R. A. Cooper, P. M. Sadler, and M. Foote. 2016. Greenhouse–icehouse transition in the Late Ordovician marks a step change in extinction regime in the marine plankton. *Proceedings of the National Academy of Sciences* 113:1498–1503.
- Edwards, D., and U. Fanning. 1985. Evolution and environment in the late Silurian–early Devonian: the rise of pteridophytes. *Philosophical Transactions of the Royal Society B: Biological Sciences* 309:147–165.
- Ezard, T. H. G., P. N. Pearson, T. Aze, and A. Purvis. 2012. The meaning of birth and death (in macroevolutionary birth-death models). *Biology Letters* 8:139–42.
- Felsenstein, J. 1985. Phylogenies and the comparative method. *The American Naturalist* 125:1–15.

- Finnegan, S., N. A. Heim, S. E. Peters, and W. W. Fischer. 2012. Climate change and the selective signature of the Late Ordovician mass extinction. *Proceedings of the National Academy of Sciences* 109:6829–6834.
- Finnegan, S., J. L. Payne, and S. C. Wang. 2008. The Red Queen revisited: reevaluating the age selectivity of Phanerozoic marine genus extinctions. *Paleobiology* 34:318–341.
- Fitzjohn, R. G. 2010. Quantitative Traits and Diversification. *Systematic Biology* 59:619–633.
- Foote, M. 1988. Survivorship analysis of Cambrian and Ordovician Trilobites. *Paleobiology* 14:258–271.
- . 2000. Origination and extinction components of taxonomic diversity: Paleozoic and post-Paleozoic dynamics. *Paleobiology* 26:578–605.
- . 2014. Environmental controls on geographic range size in marine animal genera. *Paleobiology* 40:440–458.
- Foote, M., and A. I. Miller. 2013. Determinants of early survival in marine animal genera. *Paleobiology* 39:171–192.
- Foote, M., and D. Raup. 1996. Fossil preservation and the stratigraphic ranges of taxa. *Paleobiology* 22:121–140.
- Fritz, S. A., J. Schnitzler, J. T. Eronen, C. Hof, K. Böhning-Gaese, and C. H. Graham. 2013. Diversity in time and space: wanted dead and alive. *Trends in Ecology & Evolution* 28:509–16.
- Gelman, A., J. B. Carlin, H. S. Stern, D. B. Dunson, A. Vehtari, and D. B. Rubin. 2013. Bayesian data analysis. 3rd ed. Chapman and Hall, Boca Raton, FL.
- Gelman, A., and J. Hill. 2007. Data Analysis using Regression and Multilevel/Hierarchical Models. Cambridge University Press, New York, NY.

- Goldberg, E. E., L. T. Lancaster, and R. H. Ree. 2011. Phylogenetic inference of reciprocal effects between geographic range evolution and diversification. *Systematic Biology* 60:451–65.
- Goldberg, E. E., K. Roy, R. Lande, and D. Jablonski. 2005. Diversity, endemism, and age distributions in macroevolutionary sources and sinks. *The American Naturalist* 165:623–33.
- Harnik, P. G. 2011. Direct and indirect effects of biological factors on extinction risk in fossil bivalves. *Proceedings of the National Academy of Sciences* 108:13594–13599.
- Harnik, P. G., P. C. Fitzgerald, J. L. Payne, and S. J. Carlson. 2014. Phylogenetic signal in extinction selectivity in Devonian terebratulide brachiopods. *Paleobiology* 40:675–692.
- Harnik, P. G., C. Simpson, and J. L. Payne. 2012. Long-term differences in extinction risk among the seven forms of rarity. *Proceedings of the Royal Society B: Biological Sciences* 279:4969–4976.
- Heim, N. A., and S. E. Peters. 2011. Regional environmental breadth predicts geographic range and longevity in fossil marine genera. *PloS one* 6:e18946.
- Hijmans, R. J. 2015. raster: Geographic data analysis and modeling. R package version 2.3-24.
- Hoehn, K. B., P. G. Harnik, and V. L. Roth. 2015. A framework for detecting natural selection on traits above the species level. *Methods in Ecology and Evolution* pages doi: 10.1111/2041-210X.12461.
- Hoffman, M. D., and A. Gelman. 2014. The No-U-Turn Sampler: Adaptively Setting Path Lengths in Hamiltonian Monte Carlo. *Journal of Machine Learning Research* 15:1351–1381.
- Hopkins, M. J., C. Simpson, and W. Kiessling. 2014. Differential niche dynamics among major marine invertebrate clades. *Ecology Letters* 17:314–323.
- Hunt, G., and D. L. Rabosky. 2014. Phenotypic Evolution in Fossil Species: Pattern and Process. *Annual Review of Earth and Planetary Sciences* 42:421–441.

- Hunt, G., K. Roy, and D. Jablonski. 2005. Species-level heritability reaffirmed: a comment on "On the heritability of geographic range sizes". *The American Naturalist* 166:129–135.
- Jablonski, D. 1986. Background and mass extincitons: the alternation of macroevolutionary regimes. *Science* 231:129–133.
- . 1987. Heritability at the species level: analysis of geographic ranges of cretaceous mollusks. *Science* 238:360–363.
- . 2007. Scale and hierarchy in macroevolution. *Palaeontology* 50:87–109.
- . 2008. Species Selection: Theory and Data. *Annual Review of Ecology, Evolution, and Systematics* 39:501–524.
- Jablonski, D., and K. Roy. 2003. Geographical range and speciation in fossil and living molluscs. *Proceedings of the Royal Society B: Biological Sciences* 270:401–406.
- Joachimski, M. M., S. Breisig, W. Buggisch, J. A. Talent, R. Mawson, M. Gereke, J. R. Morrow, J. Day, and K. Weddige. 2009. Devonian climate and reef evolution: Insights from oxygen isotopes in apatite. *Earth and Planetary Science Letters* 284:599–609.
- Johnson, J. G. 1974. Extinction of Perched Faunas. *Geology* 2:479–482.
- Kiessling, W., and M. Aberhan. 2007. Environmental determinants of marine benthic biodiversity dynamics through Triassic – Jurassic time. *Paleobiology* 33:414–434.
- Klein, J. P., and M. L. Moeschberger. 2003. *Survival Analysis: Techniques for Censored and Truncated Data*. 2nd ed. Springer, New York.
- Kleinbaum, D. G., and M. Klein. 2005. *Survival analysis: a self-learning text*. 2nd ed. Springer, New York, NY.
- Lewandowski, D., D. Kurowicka, and H. Joe. 2009. Generating random correlation matrices based on vines and extended onion method. *Journal of Multivariate Analysis* 100:1989–2001.

- Liow, L. H. 2004. A test of Simpson's "rule of the survival of the relatively unspecialized" using fossil crinoids. *The American Naturalist* 164:431–43.
- . 2007. Does versatility as measured by geographic range, bathymetric range and morphological variability contribute to taxon longevity? *Global Ecology and Biogeography* 16:117–128.
- Liow, L. H., M. Fortelius, E. Bingham, K. Lintulaakso, H. Mannila, L. Flynn, and N. C. Stenseth. 2008. Higher origination and extinction rates in larger mammals. *Proceedings of the National Academy of Sciences* 105:6097–6102.
- Liow, L. H., and J. D. Nichols. 2010. Estimating rates and probabilities of origination and extinction using taxonomic occurrence data: Capture-mark-recapture (CMR) approaches. Pages 81–94 in J. Alroy and G. Hunt, eds. *Quantitative Methods in Paleobiology*. The Paleontological Society.
- Liow, L. H., T. B. Quental, and C. R. Marshall. 2010. When can decreasing diversification rates be detected with molecular phylogenies and the fossil record? *Systematic biology* 59:646–59.
- Maddison, W. P., P. E. Midford, and S. P. Otto. 2007. Estimating a binary character's effect on speciation and extinction. *Systematic Biology* 56:701.
- Miller, A. I., and M. Foote. 2009. Epicontinental seas versus open-ocean settings: the kinetics of mass extinction and origination. *Science* 326:1106–9.
- Munnecke, A., M. Calner, D. A. T. Harper, and T. Servais. 2010. Ordovician and Silurian seawater chemistry, sea level, and climate: A synopsis. *Palaeogeography, Palaeoclimatology, Palaeoecology* 296:389–413.
- Nürnberg, S., and M. Aberhan. 2013. Habitat breadth and geographic range predict diversity dynamics in marine Mesozoic bivalves. *Paleobiology* 39:360–372.
- . 2015. Interdependence of specialization and biodiversity in Phanerozoic marine invertebrates. *Nature Communications* 6:6602.

Palmer, M. E., and M. W. Feldman. 2012. Survivability is more fundamental than evolvability. *PloS one* 7:e38025.

Payne, J. L., and S. Finnegan. 2007. The effect of geographic range on extinction risk during background and mass extinction. *Proceedings of the National Academy of Sciences* 104:10506–11.

Payne, J. L., N. A. Heim, M. L. Knope, and C. R. McClain. 2014. Metabolic dominance of bivalves predates brachiopod diversity decline by more than 150 million years. *Proceedings of the Royal Society B: Biological Sciences* 281:20133122.

Peters, S. E. 2007. The problem with the Paleozoic. *Paleobiology* 33:165–181.

———. 2008. Environmental determinants of extinction selectivity in the fossil record. *Nature* 454:626–629.

Quental, T. B., and C. R. Marshall. 2009. Extinction during evolutionary radiations: reconciling the fossil record with molecular phylogenies. *Evolution* 63:3158–3167.

———. 2013. How the Red Queen Drives Terrestrial Mammals to Extinction. *Science* 341:290–292.

Rabosky, D. L. 2010. Extinction rates should not be estimated from molecular phylogenies. *Evolution* 64:1816–1824.

Rabosky, D. L., and A. R. McCune. 2010. Reinventing species selection with molecular phylogenies. *Trends in Ecology & Evolution* 25:68–74.

Rabosky, D. L., F. Santini, J. Eastman, S. a. Smith, B. Sidlauskas, J. Chang, and M. E. Alfaro. 2013. Rates of speciation and morphological evolution are correlated across the largest vertebrate radiation. *Nature Communications* 4:1–8.

Raia, P., F. Carotenuto, F. Passaro, D. Fulgione, and M. Fortelius. 2012. Ecological specialization in fossil mammals explains Cope's rule. *The American Naturalist* 179:328–37.

- Raia, P., F. Carotenuto, F. Passaro, P. Piras, D. Fulgione, L. Werdelin, J. Saarinen, and M. Fortelius. 2013. Rapid action in the Palaeogene, the relationship between phenotypic and taxonomic diversification in Coenozoic mammals. *Proceedings of the Royal Society B: Biological Sciences* 280:20122244.
- Raup, D. M. 1975. Taxonomic survivorship curves and Van Valen's Law. *Paleobiology* 1:82–96.
- _____. 1978. Cohort Analysis of generic survivorship. *Paleobiology* 4:1–15.
- _____. 1991. *Extinction: Bad Genes or Bad Luck?* Norton, New York.
- _____. 1994. The role of extinction in evolution. *Proceedings of the National Academy of Sciences* 91:6758–6763.
- Rosindell, J., L. J. Harmon, and R. S. Etienne. 2015. Unifying ecology and macroevolution with individual-based theory. *Ecology Letters* 18:472–482.
- Royle, J. A., and R. M. Dorazio. 2008. *Hierarchical modeling and inference in ecology: the analysis of data from populations, metapopulations and communities.* Elsevier, London.
- Sepkoski, J. J. 1975. Stratigraphic biases in the analysis of taxonomic survivorship. *Paleobiology* 1:343–355.
- _____. 1981. A factor analytic description of the Phanerozoic marine fossil record. *Paleobiology* 7:36–53.
- Sheehan, P. 2001. The late Ordovician mass extinction. *Annual Review of Earth and Planetary Sciences* 29:331–364.
- Simpson, C. 2006. Levels of selection and large-scale morphological trends. Ph.D. thesis. University of Chicago.
- Simpson, C., and P. G. Harnik. 2009. Assessing the role of abundance in marine bivalve extinction over the post-Paleozoic. *Paleobiology* 35:631–647.

- Simpson, C., W. Kiessling, H. Mewis, R. C. Baron-Szabo, and J. Müller. 2011. Evolutionary diversification of reef corals: a comparison of the molecular and fossil records. *Evolution* 65:3274–3284.
- Simpson, G. G. 1944. *Tempo and Mode in Evolution*. Columbia University Press, New York.
- . 1953. *The Major Features of Evolution*. Columbia University Press, New York.
- Slater, G., L. Harmon, and M. Alfaro. 2012. Integrating fossils with molecular phylogenies improves inference of trait evolution. *Evolution* 66:3931–3944.
- Slater, G. J. 2013. Phylogenetic evidence for a shift in the mode of mammalian body size evolution at the Cretaceous-Palaeogene boundary. *Methods in Ecology and Evolution* 4:734–744.
- . 2015. Iterative adaptive radiations of fossil canids show no evidence for diversity-dependent trait evolution. *Proceedings of the National Academy of Sciences* 112:4897–4902.
- Smithson, M., and J. Verkuilen. 2006. A better lemon squeezer? Maximum-likelihood regression with beta-distributed dependent variables. *Psychological Methods* 11:54–71.
- Smits, P. D. 2015. Expected time-invariant effects of biological traits on mammal species duration. *Proceedings of the National Academy of Sciences* 112:13015–13020.
- Solow, A., and W. Smith. 1997. On fossil preservation and the stratigraphic ranges of taxa. *Paleobiology* 23:271–277.
- Stadler, T. 2011. Mammalian phylogeny reveals recent diversification rate shifts. *Proceedings of the National Academy of Sciences* 108:6187–6192.
- . 2013. Recovering speciation and extinction dynamics based on phylogenies. *Journal of Evolutionary Biology* 26:1203–1219.
- Stadler, T., and F. Bokma. 2013. Estimating speciation and extinction rates for phylogenies of higher taxa. *Systematic biology* 62:220–30.

- Stan Development Team. 2014. Stan: A c++ library for probability and sampling, version 2.5.0.
- Stanley, S. M. 1975. A theory of evolution above the species level. *Proceedings of the National Academy of Sciences* 72:646–650.
- . 1979. Macroevolution: pattern and process. W. H. Freeman, San Francisco.
- Tomiya, S. 2013. Body Size and Extinction Risk in Terrestrial Mammals Above the Species Level. *The American Naturalist* 182:196–214.
- Van Valen, L. 1973. A new evolutionary law. *Evolutionary Theory* 1:1–30.
- . 1979. Taxonomic survivorship curves. *Evolutionary Theory* 4:129–142.
- Vilhena, D. A., E. B. Harris, C. T. Bergstrom, M. E. Maliska, P. D. Ward, C. A. Sidor, C. A. E. Strömberg, and G. P. Wilson. 2013. Bivalve network reveals latitudinal selectivity gradient at the end-Cretaceous mass extinction. *Scientific Reports* 3:1790.
- Wagner, P. J., and G. F. Estabrook. 2014. Trait-based diversification shifts reflect differential extinction among fossil taxa. *Proceedings of the National Academy of Sciences* 111:16419–16424.
- Wagner, P. J., and J. D. Marcot. 2013. Modelling distributions of fossil sampling rates over time, space and taxa: assessment and implications for macroevolutionary studies. *Methods in Ecology and Evolution* 4:703–713.
- Wang, S. C. 2003. On the continuity of background and mass extinction. *Paleobiology* 29:455–467.
- Wang, S. C., and C. R. Marshall. 2004. Improved confidence intervals for estimating the position of a mass extinction boundary. *Paleobiology* 30:5–18.
- Williams, C., C. H. C. Brunton, and S. J. Carlson. 2007. Treatise on invertebrate paleontology. Part H, Brachiopoda. Geological Society of America, Boulder, Colorado.

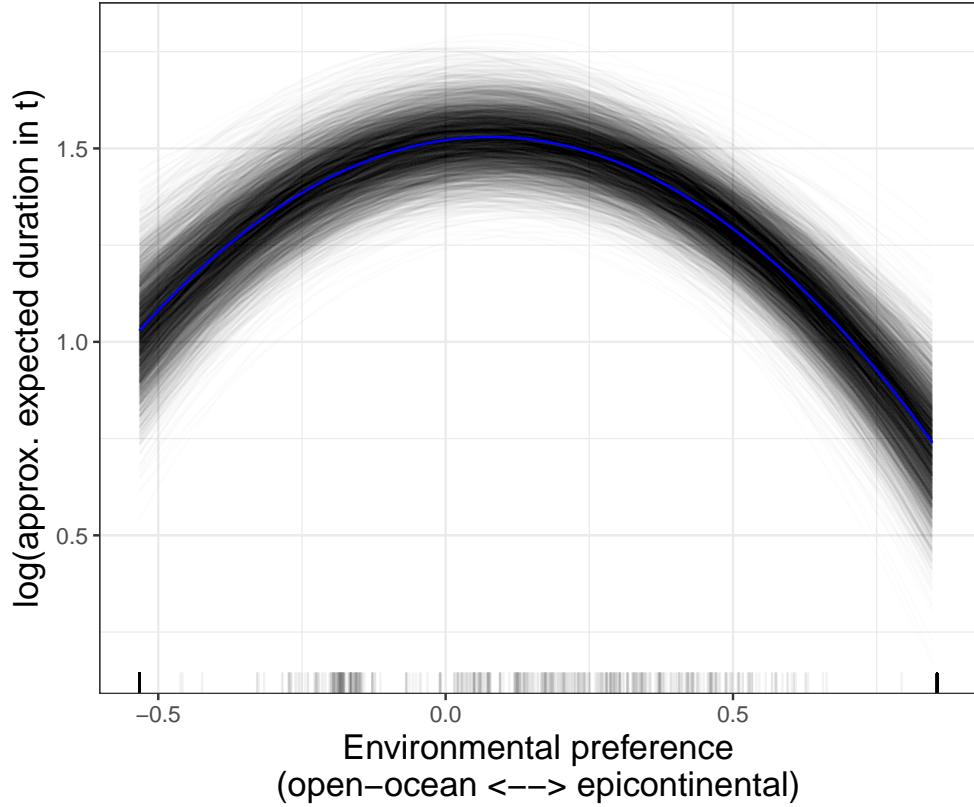


Figure 8: The overall expected relationship between environmental affinity v_i and a $\log(\sigma)$ when $r = 0$ and $m = 0$. The 1000 semi-transparent lines corresponds to a single draw from the posterior predictive distribution, while the highlighted line corresponds to the median of the posterior predictive distribution. The overall relationship is concave down with an optimum greater than 0, which means that taxa favoring epicontinental environments are expected to have longer durations than those favoring open-ocean environments. The tick marks along the bottom of the plot correspond to the (rescaled) observed values of environmental preference.

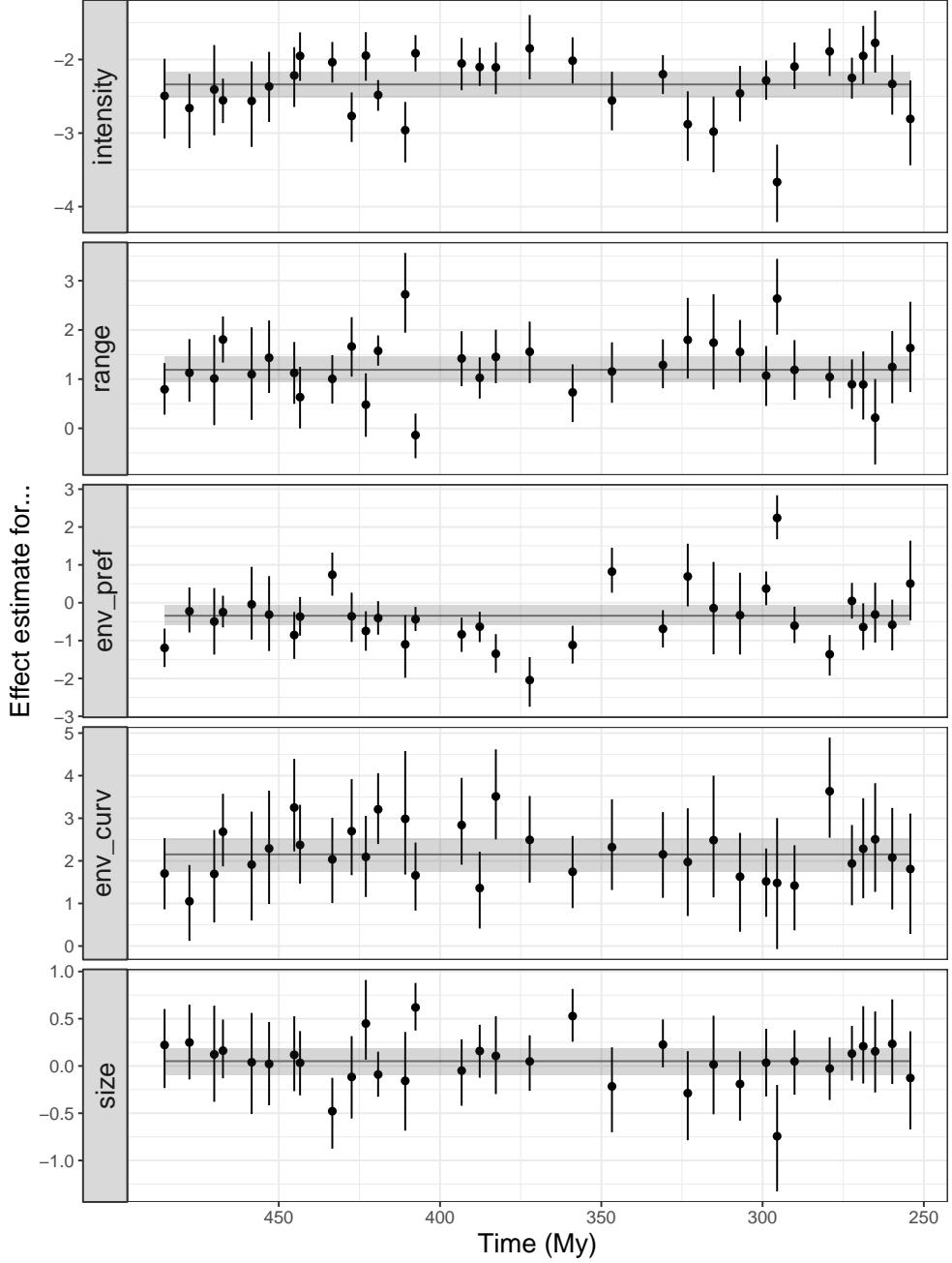


Figure 9: Comparison of cohort-specific estimates of β^0 , the effect of geographic range on extinction risk β^r , the effect of environmental preference β^v and β^{v^2} , and body size β^m . Points correspond to the median of the cohort-specific estimate, along with 80% credible intervals. Points are plotted at the midpoint of the cohorts stage of origination in millions of years before present (My). Black, horizontal lines are the overall estimates of covariate effects along with 80% credible intervals (shaded).

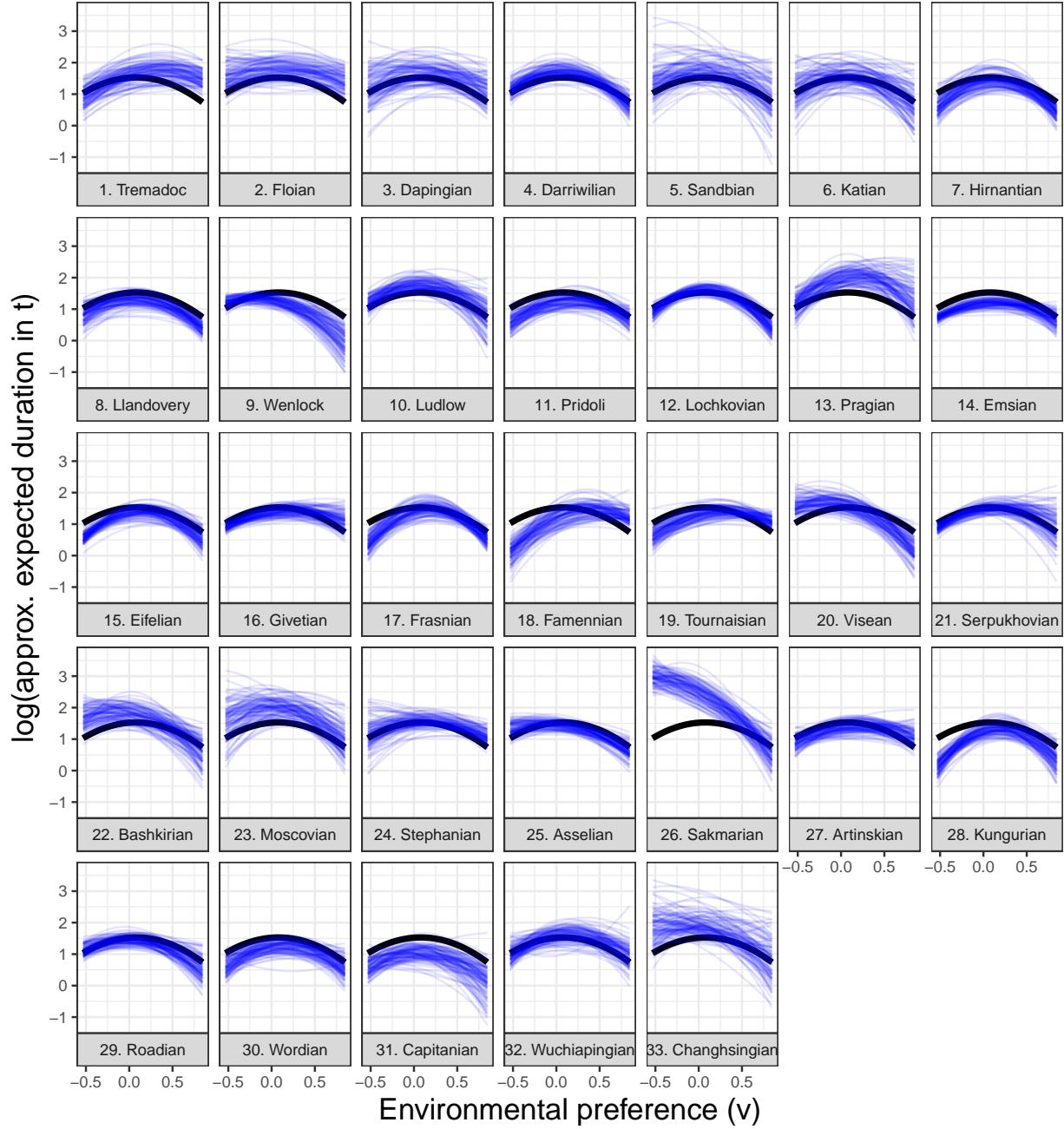


Figure 10: Comparison of origination cohort-specific (posterior predictive) estimates of the effect of environmental preference on $\log(\sigma)$ to the mean overall estimate of the effect of environmental preference. Cohort-specific estimates are from 100 posterior predictive simulations across the range of (transformed and rescaled) observed values of environmental preference. The oldest cohort is in the top-right and younger cohorts proceed left to right, with the youngest cohort being the right-most facet of the last row. Panel names correspond to the name of the stage in which that cohort originated.

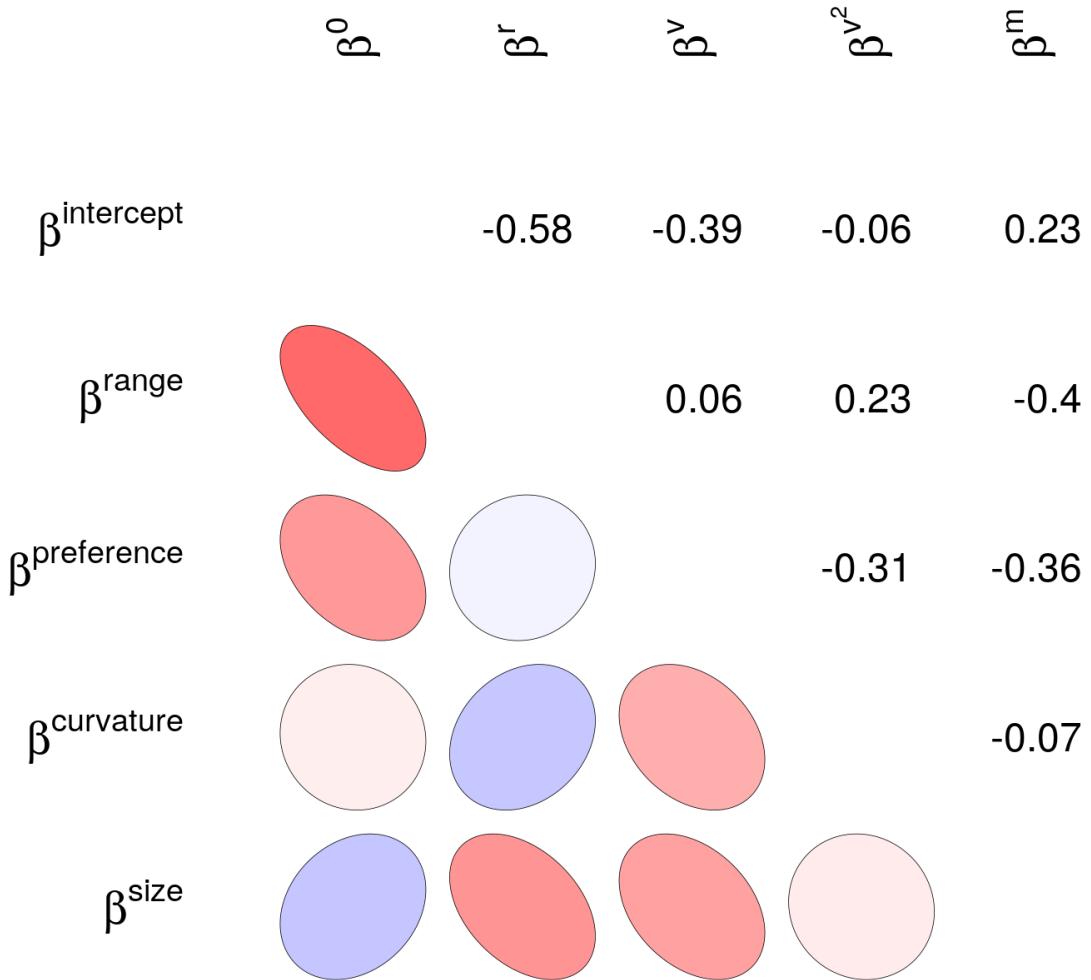


Figure 11: Mixed graphical and numerical representation of the correlation matrix Ω of variation in cohort-specific covariate estimates. These correlations are between the estimates of the cohort-level effects of covariates, along with intercept/baseline extinction risk. The median estimates of the correlations are presented numerically (upper-triangle) and as idealized ellipses representing that much correlation (lower-triangle). The darkness of the ellipse corresponds to the magnitude of the correlation.