

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:

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

Manuscript type: Article.

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

Abstract

Trait-based selection is the force behind differences in fitness; the most extreme example of selection being extinction. Modern experiments and observations have shown that average fitness and selection can vary in time and space as environments and interactions change. This begs the question “as average fitness increases, does the strength of selection increase or decrease?”

The fossil record illustrates that extinction rates have varied through time, with periods of rapid species turnover as well as slow species turnover. With Paleozoic brachiopods as a study system, I developed a model to estimate the average duration of a species (i.e. fitness), how duration varies based on species’ biological traits (i.e. selection), and how these evolutionary forces covary. I analyzed how the effects of genus geographic range, preference for epicontinental seas versus open ocean environments, and body size on duration while allowing these effects to vary over time. Environmental preference is presented as a continuum between epicontinental and open-ocean specialists, with degrees of environmental generality in between, so a possibly nonlinear (i.e. quadratic) relationship between environmental preference and duration was modeled. I find evidence that as extinction intensity increases that the magnitude of the effect of geographic range on extinction risk increases, which then increases the difference in extinction risk between small and large ranged taxa during periods of high extinction risk. Similarly, I find strong evidence for a non-linear relationship between environmental preference for epicontinental versus open-ocean environments and expected taxon duration, where taxa with intermediate or no environmental preference are expected to have greater durations than taxa which appear exclusively in either environmental end-member. Finally, I find that taxa which appear more frequently in epicontinental environments will have a greater expected duration than those taxa which prefer open-ocean environments. My analysis supports the conclusions that as extinction intensity increases and average fitness decreases, as in a mass extinction, the trait-associated differences in fitness (selection) would increase and be greater than average. In contrast, during periods of low extinction intensity when fitness is greater than average, my model predicts that geographic

²⁷ range and environmental preference associated differences in fitness (selection) would decrease and be less than average.

Introduction

30 Trait-based selection is the force behind differences in fitness; the most extreme example of se-
lection being extinction. Modern experiments and paleontological analyses have demonstrated
that selection strength and fitness can vary over time and space. That these evolutionary curren-
33 cies vary over time begs the question “do they covary?” Specifically, how does the strength of
selection change as average fitness also changes? The fossil record demonstrates that extinction
risk has varied continuously over time, from periods of low average extinction rate to very high
36 extinction rates (e.g. mass extinctions) (Foote, 2000*a,b*, 2001). Paleontological analyses have also
demonstrated trait-based differences in extinction risk among taxa (Jablonski, 2008). Conceptually,
extinction is the ultimate manifestation of selection as we would expect that a taxon with a
39 beneficial trait to persist for longer than a similar taxon without that trait due to selection (Jablon-
ski, 2008; Rabosky and McCune, 2010; Raup, 1994; Stanley, 1975). Thus, the expected duration of
a species can be conceived of as a measure of a species’ fitness (Cooper, 1984); this then means
42 that trait-associated differences in species fitness are (species) selection (Rabosky and McCune,
2010).

In order to test for an association between extinction intensity and extinction selectivity, ex-
45 tinction 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 esti-
48 mates (Payne et al., 2016). I find this approach problematic for a few reasons. Modeling each
each time point or cohort independently does not use all of the information present in the data
and their estimates are only based the data from that time point. A hierarchical/mixed-effect
51 modelling approach leverages all data across time points or cohorts by partially pooling infor-
mation across each of the time-points or cohorts. The resulting parameter estimates have better
behaved posteriors (e.g. smaller credible intervals) and limit spurious parameter estimates by
54 weighing estimates by the amount of data associated with each time point or cohort (Gelman

et al., 2013). Importantly, the partial pooling in hierarchical/mixed-effect also mitigates the effect of complete separation which prevent parameter estimates for some time points or cohorts (Gelman et al., 2013; Payne et al., 2016). Finally, treating each time point or cohort as independent any and all post-hoc analyses are at risk of false positive results due to multiple comparisons (Gelman et al., 2013; Gelman and Hill, 2007); hierarchical/mixed-effect models ameliorate this problem as possible comparisons are simultaneously modeled.

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 regimes" 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 are unimodal with continuous variation. This conflict arises because differences between these papers is that Jablonski (1986) is considering qualitative differences while Wang (2003) is considering quantitative differences. This 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) (Payne et al., 2016). 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 develop a statistical model describing the relationship between brachiopod taxon durations and multiple functional taxon traits in order to understand the relationship between extinction intensity and selectivity over time. Trait-dependent differences in extinction risk should be associated with differences in taxon duration (Cooper, 1984; Rabosky and McCune, 2010). Brachiopods are an ideal group for this study as they have an exceptionally complete fossil record (Foote, 2000b; 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 (ap-

proximately 252 My) as this represents the time of greatest global brachiopod diversity (Alroy, 2010) which results in a large sample size.

84 The analysis of taxon durations, or time from a taxon's origination till its 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 (Cram-
87 ton 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;
90 Crampton et al., 2016; Ezard et al., 2012; Foote, 1988; Raup, 1975, 1978; Simpson, 2006; Van Valen,
1973, 1979).

While estimation of trait-dependent speciation and extinction rates from phylogenies of ex-
tant taxa has grown dramatically (Fitzjohn, 2010; Goldberg et al., 2011, 2005; Maddison et al.,
93 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
96 record. These two directions, phylogenetic comparative and paleobiological, are complemen-
tary 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 dis-
99 tinct 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; Ra-
bosky, 2010). The approach used here is thus complementary to the analysis of trait-dependent
102 extinction based on a phylogeny.

Factors affecting brachiopod survival

Conceptually, taxon survival can be considered an aspect of "taxon fitness" (Cooper, 1984; Palmer
105 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 an-
alyzed here are all examples of emergent and aggregate traits (Jablonski, 2008; Rabosky and

¹⁰⁸ 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
¹¹¹ average of a shared trait across all members of a lower level.

¹¹⁴ 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 a broad range is less likely to go extinct at random than a taxon with a
¹¹⁷ restricted range.

¹²⁰ 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
¹²³ 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
¹²⁶ 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).

¹³² 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

135 that open-ocean taxa may have a greater extinction rate because these environments would be
more strongly affected by poisoning of the environment from impact fallout or volcanic events
because water circulates at a higher rate and in greater volume in open-ocean environments
138 compared to the relatively more sluggish circulation in epicontinental environments. These two
environment types represent the primary identifiable environmental dichotomy observed in an-
cient marine systems (Miller and Foote, 2009; Sheehan, 2001). Given these findings, I would
141 hypothesize that as extinction risk increases, the extinction risk associated with open-ocean en-
vironments should generally increase.

Because environmental preference is defined here as the continuum between occurring ex-
144clusively 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
147 (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 func-
tion where a concave down relationship between preference and expected duration indicates that
150 generalists are favored over specialists end-members. Importantly, this approach does not “force”
a non-linear relationship, and only allows one if the second-order term is non-zero.

Body size, measured as shell length, is also considered as a trait that may potentially influence
153 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 a phylogenetic and non-phylogenetic context;
156 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
159 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

¹⁶² 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
¹⁶⁵ 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
¹⁶⁸ 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
¹⁷¹ 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

¹⁷⁴ The brachiopod dataset analyzed here was sourced from the Paleobiology Database (<http://www.paleodb.org>)
and then filtered to a limited selection of higher taxonomic groups (Rhynchonelliformea: Rhyn-
chonellata, Chileata, Obolellida, Kutorginida, Strophomenida, Spiriferida). Additionally, samples
¹⁷⁷ were limited to those taxa present during the post-Cambrian Paleozoic. Temporal, stratigraphic,
and other relevant occurrence information used in this analysis was also downloaded. Analyzed
occurrences were restricted to those with paleolatitude and paleolongitude coordinates, have
¹⁸⁰ 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 assign-
ments for each fossil occurrence are based on those from previous analyses Foote and Miller
¹⁸³ (2013); Miller and Foote (2009); Ritterbush and Foote (2017). 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 genera were included in this
¹⁸⁶ analysis.

Fossil occurrences were analyzed at the genus level, a common practice for paleobiological, macroevolutionary and macroecological studies and this is especially the case for marine invertebrates (Alroy, 2010; Eronen et al., 2011; 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; Ritterbush and Foote, 2017; 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. Additionally, when species and genera can be compared, they often yield similar results (Foote et al., 2007; Jernvall and Fortelius, 2002; Roy D. & Valentine, I. W., 1996). Importantly, it may also be possible genera may represent coherent biological units as there is evidence for congruence between morphologically and genetically defined genera of molluscs and mammals (Jablonski and Finarelli, 2009).

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. Censoring in this context indicates that the genus was observed up to a certain age but its ultimate time of extinction is unknown (Klein and Moeschberger, 2003).

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

A genus's geographic range was calculated as the number of occupied grid cells from a gridded map of all contemporaneous occurrences. 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
 216 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
 219 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 Bi-
 222 nomial 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
 225 sampled to yield relative occupancy, my proxy for geographic range.

The environmental preference for a taxon is defined by their relative occurrences in epicontinental or open-ocean environments, presenting a continuum from exclusive occurrence at
 228 the ends and equal occurrences in the middle. Mathematically, 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
 231 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
 234 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*
 237 (Williams et al., 2007).

The sampling probability for individual taxa, called s was calculated using the standard gap

statistic (Foote, 2000a; Foote and Raup, 1996). The gap statistic is calculated as the number of
240 stages in which the taxon was sampled except for its first stage and last stage. Because taxa that
were right censored only include a first appearance, one was subtracted from the numerator and
denominator instead of two. The inclusion of genus-specific sampling probability as a covariate
243 in the model implicitly deals with my using stratigraphic range as a proxy for genus duration.
The implications of this choice are discussed further later in the Discussion.

The minimum duration for which a gap statistic can be calculated is three stages, so I chose
246 to impute the gap statistic for all observations with a duration less than 3. Imputation is the
“filling in” of missing observations based on the observed values (Gelman and Hill, 2007; Rubin,
1996).

249 Prior to analysis, geographic range was logit transformed and the number of samples was
natural-log transformed; 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
252 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
255 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.

258

Details of model

Hierarchical modelling is a statistical approach which explicitly takes into account the structure
of the observed data in order to model the within and the between group variances (Gelman
261 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,
264 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
267 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).

270 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
273 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
276 with age. Note that the Weibull distribution is equivalent to the exponential distribution when $\alpha = 1$.

Censoring and truncation reflects a number of factors that limit our ability to fully observe
279 a taxon's duration: limited resolution, which leads to left-censoring or truncation; end of study interval, which leads to right censoring; and incomplete sampling, which can left-censor (short-lived taxa are less likely to be preserved at all) or right-censor (durations are truncated). You
282 are talking only about the first two sources, which is fine. But it might make sense to state this explicitly and to say that you are dealing with sampling-related censoring by using empirically estimated sampling rates as a predictor in your model. In the case of the right- and left-censored
285 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
288 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
291 over treating these taxa as being fully observed; additionally, a model with censoring yields

better WAIC and LOOIC values than a model without censoring, which means that the model with censoring is expected to have greater out-of-sample predictive accuracy than the model
294 without censoring (not shown).

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

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

297 where i indexes across all observations where $i = 1, \dots, n$ where n is the total number of obser-
vations, $j[i]$ 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 ones for the inter-
300 cept 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-
303 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.

306 The elements of μ were given independent normally distributed priors. The effects of geo-
graphic 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 ef-
309 fect. The correlation matrix Ω was given an LKJ distributed prior (Lewandowski et al., 2009) that

slightly favors an identity matrix as recommended by Team (2017). 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 $\log(\alpha) \sim \mathcal{N}(0, 1)$
₃₁₂ centered at $\alpha = 1$, which corresponds to the Law of Constant Extinction (Van Valen, 1973).

Imputation of sampling probability

The vector sampling s has two parts: the observed part s^o , and the unobserved part s^u . Of the
₃₁₅ 1130 total observations, 539 have a duration of 3 or more and have an observed gap statistic. The
gap statistic for the remaining 591 observations was imputed. As stated above, the unobserved
part is the imputed, or filled in, based on the observed part s^o . Because sampling varies between
₃₁₈ 0 and 1, I chose to model it as a Beta regression with matrix W being a $N \times (D - 3)$ matrix of
covariates (i.e. geographic range size, environmental preference, body size; no interactions) as
predictors of sampling; this assumes that the sampling value for all taxa come from the same
₃₂₁ distribution. Importantly, I make no assumptions of causal structure.

Predicting sampling probability using the other covariate that are then included in the model
of duration is acceptable and appropriate in the case of imputation where the sample goal is
₃₂₄ accurate prediction (Gelman and Hill, 2007; Rubin, 1996). Not including these covariates can
lead to biased estimates of the imputed variable; if the covariates themselves are related, not

including them will bias this correlation towards zero which then leads to improper imputation
³²⁷ and inference (Rubin, 1996).

The Beta regression is defined

$$s^o \sim \text{Beta}(\phi = \text{logit}^{-1}(X^o\gamma), \lambda), \quad (4)$$

where γ is a length D vector of regression coefficients, and X defined as above. The Beta distribution used in the regression is reparameterized in terms of a mean parameter
³³⁰

$$\phi = \frac{\alpha}{\alpha + \beta} \quad (5)$$

and total count parameter

$$\lambda = \alpha + \beta \quad (6)$$

where α and β are the characteristic parameters of the Beta distribution (Gelman et al., 2013).

³³³ The next step is to then estimate $s^u | s^o, X^o, X^u, \gamma$, the posterior distribution of which are folded back into s and used as a covariate of duration (Eq. 2). All the elements of γ , δ (Eq. 2), and λ (Eq. 4) were given weakly informative priors as recommended by Team (2017):

$$\begin{aligned} \gamma &\sim \mathcal{N}(0, 1) \\ \delta &\sim \mathcal{N}(0, 1) \\ \lambda &\sim \text{Pareto}(0.1, 1.5). \end{aligned} \quad (7)$$

³³⁶ The imputed values are estimated simultaneously at the same time and in the same manner as other parameters are estimated; this ensures that all uncertainty surrounding these unobservable covariate values is propagated through to all estimates.

³³⁹ *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
³⁴² 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
345 half warm-up and half sampling and thinned to every 20th sample for a total of 4000 posterior
samples. Starting conditions for sampling were left at defaults for CmdStan interface except for
the following changes: adapt delta was set 0.95 to ensure no divergent samples, and initial value
was set to 0 which allows for stable initial samples. Posterior convergence was assessed using
348 standard MCMC diagnostics such as the scale reduction factor \hat{R} (target < 1.1) and effective
sample size or ESS (target eff/steps < 0.0001), and HMC specific criteria such as energy (target
 > 0.2), presence and number of divergent samples and number of samples that saturated the
351 maximum trajectory length for avoiding infinite loops (target value 0). For futher explanation of
these diagnostic criteria, see the Stan Manual (Team, 2017).

After the model was fit to the data, 100 datasets were simulated from the posterior predictive
354 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 extinc-
357 tion 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
360 its final observed duration and the simulated duration. For left-censored individuals, their simu-
lated duration was pegged at a minimum of one stage with simulated values less than one stage
set to one.

363 Model adequacy was evaluated using a series of posterior predictive checks. Posterior predic-
tive 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
366 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
369 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-parametric 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 $P(T \geq t)$ 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

I first present the results of the multiple posterior predictive checks for the whole dataset as well as each of the origination cohorts. I next present the parameter posterior estimates and their interpretations. Importantly, the adequacy of model fit associated with each cohort-specific interpretation is presented to demonstrate which results for which I have the greatest confidence.

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 at durations of 2-3 stages. The
396 model slightly under-estimates the number of observations with duration of 2 or 3 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 bulk of our data
399 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 adequacy (Fig. 2). The survival curves of
402 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 underestimating
percentage of taxa with duration 1 stage, and an over-estimate of probability of species surviving
405 at least 10-13 stages. Importantly, the major divergence between the observed and estimated
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.

408 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
411 heterogeneous fit of the model to the entire distribution of the data, the fitted model's ability
to predict the mean and median of the observed data appears adequate (Fig. 3a, 3b). Because
the principle goal of this model is to obtain adequate prediction of how a taxon's expected
414 duration for a given set of ecological covariates, the seemingly adequate fit of our model to mean
taxon duration is reassuring (Fig. 3a). Additionally, given the skewness of the observed taxon
durations (Fig. 1, the ability for the model to closely recapitulates the median observed taxon
417 duration points to the overall good fit of the model to the data.

When considered together, all of the above posterior predictive checks indicate adequate
model fit for key questions such as expected taxon duration (Fig. 3a). However, there is obviously
420 heterogeneity in model fit because, while the model can recapitulate some aspects of the observed
data (Fig. 3a, 3b), there are obvious discrepancies between the model and the data (Fig. 1, 2). By

423 performing the same posterior predictive tests for each of the origination cohorts, it may be
possible to get a better picture of the sources of model misfit.

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. 4, 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. Tournaisian, Visean, Bashirian, Moscovian, Stephanian, Asselian, etc.). The poor model fit to some origination cohorts may indicate that these cohorts are undergoing a different extinction process whose aspects are unmodeled in this analysis. However, for those cohorts where the model recapitulates the empirical survival function it is likely that the 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. 5). In comparison, the model has a much more heterogeneous fit to each origination cohort's median taxon duration (Fig. 6). 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, Artinskian, and Roadian. The remaining cohorts, however, have adequate fit. These results indicate that our model is very good at recapitulate mean taxon duration (Fig. 3a, 5), and it is capable of estimating overall median duration and median duration of most origination cohorts (Fig. 3b, 6).

444 A larger than average geographic range is expected to have a positive effect on taxon survival (Table 1). 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 an approximately 3.7% ($\pm 4.3\%$ SD) probability that this relationships would be reversed ($\text{Pr}N(\mu^r, \tau^r) > 0$)).

Body size measured as valve length is estimated to, on average, have no effect on duration
450 (Table 1).

The group-level estimate of the effect of environmental preference is estimated from μ^v and
 μ^{v^2} . The estimate of μ^v indicates that taxa which slightly prefer epicontinental environments
453 to open-ocean environments 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 approx-
imately 18.1% ($\pm 7.5\%$ SD) probability that, for any given cohort, taxa which favor open-ocean
456 environments would have a greater expected duration than taxa which favor epicontinental envi-
ronments ($\Pr(\mathcal{N}(\mu^v, \tau^v) > 0)$). The estimate of μ^{v^2} indicates that the overall relationship between
environmental preference and $\log(\sigma)$ is concave down (Fig. 7), with only a 2.5% ($\pm 2.9\%$ SD)
459 probability that any given cohort is convex up ($\Pr(\mathcal{N}(\mu^{v^2}, \tau^{v^2}) < 0)$).

The cohort-specific estimates of all the regression coefficients demonstrate a lot of between co-
hort variance, with no obvious long-term trends (Fig. 8). While most cohort-specific estimates are
462 very similar to the overall cohort-level estimate, there are a few notable cohorts for which two or
more individual-level parameter estimates diverge greatly from the group-level averages (Fig. 8).
What's even more interesting is that for many cohort's for which two or more covariate effect
465 estimates are different from the overall group-level mean are well fit by the model (Fig. 4, 5, 6).

The cohort-specific relationships between environmental preference and $\log(\sigma)$ were calcu-
lated from the estimates of β^0 , β^v , and β^{v^2} (Fig. 9) and reflect how these three parameters act in
468 concert and not just individually (Fig. 8). Because of the relationship between β^v and β^{v^2} , it is
important to consider than together when drawing conclusions from the model. In many cases,
the cohort-specific estimated relationship between environmental preference and duration is ap-
471 proximately equal to the group-level average, but 14 of the 33 analyzed origination cohorts have
at least one of these three parameters being noticeably different from the group-level average.

There is an approximately 90.4% probability that cohort estimates of β^0 and β^r are negatively
474 correlated, with median estimate of correlation being -0.35. This result means that for any cohort,
we would expect that if extinction intensity increases (β^0 increases), the effect of geographic range

on duration increases (β^r decreases). This result is strong evidence for a relationship between
477 intensity and selectivity with respect to geographic range size.

There is an approximate 97.9% probability that the cohort-specific estimates of β^0 and β^v
are negatively correlated with median correlation -0.49. This result means that as extinction
480 intensity increases it is expected that epicontinental taxa become more favored over open-ocean
environments (i.e. as β^0 increases, β^v decreases). This result is strong evidence for a relationship
between intensity and selectivity with respect to the linear aspect of environmental preference.

483 Correlations between the non-intercept estimates reflect potential similarities in selective pres-
sures between cohorts, however there is only weak evidence of any potential cross-correlations
in cohort-specific covariate effects. There is an approximate 31.2% probability that β^r and β^v are
486 positively correlated. This lack of cross-correlation may be due in part to the 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
489 not vary enough relative to the much noisier environmental preference.

Conversely, There is a 74.6% probability that estimates of the effect of geographic range (β^r)
and the quadratic aspect of environmental preference (β^v^2 are positively correlated; this is weak
492 evidence of a relationship between the effects of these covariates. Thus, as the effect of geographic
range increases, we might expect with weak evidence that the peakedness of relationship between
environmental preference and duration to increase. However, because there is only a 74.6% prob-
495 ability of a positive correlation, this result cannot interpreted with authority. Instead, this result
is an opportunity for future research to understand a potential relationship between geographic
range, environmental preferece, and species duration.

498 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
501 lived taxa have more opportunities to not be sampled than shorter lived taxa. These two factors
will lead to this result.

The Weibull shape parameter α was found to be approximately 1.41 (± 0.05 SD) with a
504 100% probability of being greater than 1. This result is not consistent with the Law of Con-
stant Extinction (Van Valen, 1973) and is instead consistent with accelerating extinction risk with
507 taxon age. This result is consistent with recent empirical results and may be caused by newly
originating species have a fundamentally lower risk of extinction compared to species which
have already originated (Quental and Marshall, 2013; Smits, 2015; Wagner and Estabrook, 2014).
This result is also consistent with a recently proposed nearly-neutral evolution where compe-
510 tition/selection/evolution drives whole communities to increase in average fitness over time
while still maintaining constant relative fitness across the community, thus older species are
more likely to go extinct because of having a fundamentally lower average fitness than newly
513 originating species (Rosindell et al., 2015). This results, however, is not consistent with other
empirical results from the marine fossil record (Crampton et al., 2016; Finnegan et al., 2008) and
could potentially be caused by the minimum resolution of the fossil record (Sepkoski, 1975). It
516 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

519 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 associated with difference in duration had no effect
522 (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?

525 I find that as intensity increases (β^0 increases), the magnitude of the effect of geographic range
increases (β^r decreases). I also find that as intensity increases, the difference in survival for taxa
favoring epicontinental environments over open-ocean environments is expected to decrease;

528 this is consistent with the results of Miller and Foote (2009). Finally, there is no evidence for
a correlation between the effects of geographic range and environmental preference on taxon
duration.

531 I find consistent support for the “survival of the unspecialized,” with respect to epicontinen-
tal versus open-ocean environmental preference, as a time-invariant generalization of brachiopod
survival (Simpson, 1944). Taxa with intermediate environmental preferences are expected to have
534 lower extinction risk than taxa specializing in either epicontinental or open-ocean environments
(Fig. 7), though the curvature of the relationship varies from rather shallow to very peaked
(Fig. 9). However, this relationship is not symmetric about 0, as taxa favoring epicontinental en-
537 vironments are expected with approximately 75% probability to have a greater duration than taxa
favoring open-ocean environments. This description of environment preference only describes
one major aspect of a taxon’s environmental context, with factors such as bathymetry and tem-
540 perature 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 analy-
ses would potentially improve our understanding the extent and complexity of the “survival of
543 the unspecialized” hypothesis as it applies to all dimensions of an adaptive zone.

544 Hopkins et al. (2014), in their analysis of niche conservatism and substrate preference in
marine invertebrates, found that brachiopods were among the least “conservative” groups; taxa
546 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
549 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
552 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

555 lineage duration in relation to environmental availability to estimate if the generalists–specialists continuum is actually ternary relationship.

The analysis presented in this paper is an example of how to approach the interplay between selection and intensity using a continuous-survival framework. As with any analytical approach, there are many ways to approach this problem. An alternative framework would be a discrete-time survival analysis (Tutz and Schmid, 2016) where species survival is tracked at discrete intervals. An example of a discrete-time survival approach that has become increasingly popular in paleontological analysis the Cormack-Jolly-Seber (CJS) model (Liow et al., 2008; Liow and Nichols, 2010; Royle and Dorazio, 2008; Tomiya, 2013). Discrete-survival analysis has some advantages over continuous-time approaches, specifically the ease of including time-varying covariates and well known extensions for allowing incomplete sampling (e.g. CJS model).

Something that has not been included in these discrete-time analysis is the inclusion of an age-based varying-intercept or covariate as recommended by Tutz and Schmid (2016); these inclusions are extremely important for estimating the effect of taxon age on survival. Those varying-intercept estimates would then be equivalent to the hazard function (probability of going extinct at t given being alive at $t - 1$) when all covariates are equal to 0 (Tutz and Schmid, 2016). A good avenue for future applied research would be a CJS-type model with survival modeled as a multi-level regression as in this study, combined with an age-based varying-intercept as recommended by (Tutz and Schmid, 2016). A potential hurdle to this analysis would be understanding model goodness-of-fit; the continuous time analysis presented here allows for numerous posterior predictive analysis because of the fundamental nature of the data (Fig. 1, 2, 4, 5, 6).

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

582 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 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 difficulty with this particular model expansion is in 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

609 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 macroevolu-
612 tion (Fritz et al., 2013; Harnik et al., 2014; Raia et al., 2012, 2013; Simpson et al., 2011; Slater, 2013,
2015; Slater et al., 2012; Smits, 2015; Tomiya, 2013).

Conclusion

615 My analysis demonstrates that for post-Cambrian Paleozoic brachiopods, as extinction intensity
increases and average fitness decreases, like in a mass extinction, the trait-associated differences
in fitness (selection) would increase and be greater than average. In contrast, during periods of
618 low extinction intensity when fitness is greater than average, my model predicts that geographic
range and environmental preference associated differences in fitness (selection) would decrease
and be less than average. Taken together, these results point to a potential macroevolutionary
621 mechanism behind differences in trait-based survival during and away from mass extinctions
due to a correlation between intensity and selectivity. Additionally, I find continued support
for greater survival in environmental generalists over specialists; this is further evidence that the
624 the long standing “survival of the unspecialized” hypothesis (Baumiller, 1993; Liow, 2004, 2007;
Nürnberg and Aberhan, 2013, 2015; Simpson, 1944, 1953; Smits, 2015) should be considered the
default hypothesis. Overall, this analysis further refines our knowledge of brachiopod extinction
627 dynamics while also revealing a potential macroevolutionary mechanism behind the difference
between so-called mass and background extinction regimes.

Acknowledgments

630 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

633 epicontinental versus open-ocean assignments. Finally, thank you to P. Harnik, L. L. Hsiang, G.
Simpson, and the other anonymous reviewers for their comments at various points during this
manuscripts life that improved this manuscript. This entire study would have not been
636 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-
639 Gottshall, M. Patzkowsky, M. Uhen, L. Villier, and P. Wager. Additionally, I would like to thank
J. Payne and N. Heim for making their body size measurements publically available. This work
was supported by a NASA Exobiology grant (NNX10AQ446) to A. Miller and M. Foote. I declare
642 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

- 645 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*
648 329:1191–1194.
- . 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
651 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
654 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.
- 657 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.
- 660 Eronen, J. T., A. R. Evans, M. Fortelius, and J. Jernvall. 2011. Genera Are Often Better Than Species
for Detecting Evolutionary Change in the Fossil Record: a Reply To Salesa Et Al. *Evolution*
65:1514–1516.
- 663 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* 8:139–142.
- Felsenstein, J. 1985. Phylogenies and the comparative method. *The American Naturalist* 125:1–15.

- 666 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.
- 669 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.
- 672 Foote, M. 1988. Survivorship analysis of Cambrian and Ordovician Trilobites. *Paleobiology*
14:258–271.
- . 2000a. Origination and extinction components of taxonomic diversity: general problems.
675 *Paleobiology* 26:74–102.
- . 2000b. Origination and extinction components of taxonomic diversity: Paleozoic and
post-Paleozoic dynamics. *Paleobiology* 26:578–605.
- 678 ———. 2001. Inferring temporal patterns of preservation, origination, and extinction from taxo-
nomic survivorship analysis. *Paleobiology* 27:602–630.
- 681 ———. 2014. Environmental controls on geographic range size in marine animal genera. *Paleo-
biology* 40:440–458.
- Foote, M., J. S. Crampton, A. G. Beu, B. a. Marshall, R. a. Cooper, P. a. Maxwell, and I. Matcham.
2007. Rise and fall of species occupancy in Cenozoic fossil mollusks. *Science* (New York, N.Y.)
318:1131–1134.
- Foote, M., and A. I. Miller. 2013. Determinants of early survival in marine animal genera. *Paleo-
biology* 39:171–192.
- 687 Foote, M., and D. M. Raup. 1996. Fossil preservation and the stratigraphic ranges of taxa. *Paleo-
biology* 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* 28:509–516.
- 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.
- 693 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* 60:451–465.
- 696 Goldberg, E. E., K. Roy, R. Lande, and D. Jablonski. 2005. Diversity, endemism, and age distributions in macroevolutionary sources and sinks. *The American Naturalist* 165:623–633.
- 699 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.
- 702 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.
- 705 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.
- 708 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.

- 711 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
714 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.
- 717 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 extinctions: the alternation of macroevolutionary
720 regimes. *Science* 231:129–133.
- . 1987. Heritability at the species level: analysis of geographic ranges of cretaceous mol-
lusks. *Science* 238:360–363.
- 723 ———. 2008. Species Selection: Theory and Data. *Annual Review of Ecology, Evolution, and
Systematics* 39:501–524.
- Jablonski, D., and J. A. Finarelli. 2009. Congruence of morphologically-defined genera with
726 molecular phylogenies. *Proceedings of the National Academy of Sciences* 106:8262–8266.
- 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.
- 729 Jernvall, J., and M. Fortelius. 2002. Common mammals drive the evolutionary increase of hyp-
sodonty in the Neogene. *Nature* 417:538–540.
- Johnson, J. G. 1974. Extinction of Perched Faunas. *Geology* 2:479–482.
- 732 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
735 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.
- 738 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
741 fossil crinoids. *The American Naturalist* 164:431–443.
- . 2007. Does versatility as measured by geographic range, bathymetric range and morpho-
logical variability contribute to taxon longevity? *Global Ecology and Biogeography* 16:117–128.
- 744 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.
- 747 Liow, L. H., and J. D. Nichols. 2010. Estimating rates and probabilities of origination and ex-
tinction 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
750 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* 59:646–659.
- 753 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
756 mass extinction and origination. *Science* 326:1106–1109.

- Nürnberg, S., and M. Aberhan. 2013. Habitat breadth and geographic range predict diversity dynamics in marine Mesozoic bivalves. *Paleobiology* 39:360–372.
- 759 ———. 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.
762 *PLoS one* 7:e38025.
- Payne, J. L., A. M. Bush, E. T. Chang, N. A. Heim, M. L. Knope, and S. B. Pruss. 2016. Extinction intensity, selectivity and their combined macroevolutionary influence in the fossil record.
765 *Biology Letters* 12:20160202.
- 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–
768 10511.
- 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
771 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*
774 454:626–629.
- Quental, T. B., and C. R. Marshall. 2009. Extinction during evolutionary radiations: reconciling
the fossil record with molecular phylogenies. *Evolution* 63:3158–3167.
- 777 ———. 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. *Evo-
lution* 64:1816–1824.

- 780 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.
783 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
786 in fossil mammals explains Cope's rule. *The American Naturalist* 179:328–337.
- 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
789 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.
792 ———. 1978. Cohort Analysis of generic survivorship. *Paleobiology* 4:1–15.
———. 1994. The role of extinction in evolution. *Proceedings of the National Academy of Sciences* 91:6758–6763.
- 795 Ritterbush, K. A., and M. Foote. 2017. Association between geographic range and initial survival of Mesozoic marine animal genera: Circumventing the confounding effects of temporal and taxonomic heterogeneity. *Paleobiology* 43:209–223.
- 798 Rosindell, J., L. J. Harmon, and R. S. Etienne. 2015. Unifying ecology and macroevolution with individual-based theory. *Ecology Letters* 18:472–482.
- Roy D. & Valentine, I. W., K. J. 1996. Higher taxa in biodiversity studies: Patterns from eastern
801 pacific marine molluscs. *Phil. Trans. R. Soc., Londo B* 351:1605–1613.
- 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.

- 804 Rubin, D. B. 1996. Multiple imputation after 18+ years. *Journal of the American Statistical
Association* 91:473–489.
- 807 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.
- 810 Sheehan, P. M. 2001. The late Ordovician mass extinction. *Annual Review of Earth and Planetary
Sciences* 29:331–364.
- 813 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.
- 816 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.
- 819 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.
- 822 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.
- 825 Slater, G. J., L. J. Harmon, and M. E. Alfaro. 2012. Integrating fossils with molecular phylogenies
improves inference of trait evolution. *Evolution* 66:3931–3944.

- Smithson, M., and J. Verkuilen. 2006. A better lemon squeezer? Maximum-likelihood regression
828 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.
- 831 Solow, A. R., 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
834 the National Academy of Sciences* 108:6187–6192.
- . 2013. Recovering speciation and extinction dynamics based on phylogenies. *Journal of
Evolutionary Biology* 26:1203–1219.
- 837 Stadler, T., and F. Bokma. 2013. Estimating speciation and extinction rates for phylogenies of
higher taxa. *Systematic biology* 62:220–230.
- Stan Development Team. 2014. Stan: A c++ library for probability and sampling, version 2.5.0.
- 840 Stanley, S. M. 1975. A theory of evolution above the species level. *Proceedings of the National
Academy of Sciences* 72:646–650.
- Team, S. D. 2017. Stan Modeling Language Users Guide and Reference Manual.
- 843 Tomiya, S. 2013. Body Size and Extinction Risk in Terrestrial Mammals Above the Species Level.
The American Naturalist 182:196–214.
- Tutz, G., and M. Schmid. 2016. Modeling discrete time-to-event data. Springer International
846 Publishing, Switzerland.
- Van Valen, L. 1973. A new evolutionary law. *Evolutionary Theory* 1:1–30.
- . 1979. Taxonomic survivorship curves. *Evolutionary Theory* 4:129–142.

- 849 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.
- 852 Wagner, P. J., and G. F. Estabrook. 2014. Trait-based diversification shifts reflect differential ex-
tinction 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,
855 space and taxa: assessment and implications for macroevolutionary studies. *Methods in Ecol-
ogy and Evolution* 4:703–713.
- Wang, S. C. 2003. On the continuity of background and mass extinction. *Paleobiology* 29:455–467.
- 858 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
861 H, Brachiopoda. Geological Society of America, Boulder, Colorado.

Tables

Table 1: Estimates of group-level and invariant parameter values for the fitted model analyzed here.

Category	Parameter	Effect of...	Mean	SD	10%	50%	90%
	μ^i	intercept	-3.04	0.19	-3.29	-3.04	-2.80
	μ^r	geographic	-0.98	0.16	-1.17	-0.98	-0.78
Mean		range					
	μ^v	environmental	-0.76	0.18	-0.99	-0.76	-0.53
		preference					
	μ^{v^2}	environmental	3.15	0.35	2.71	3.15	3.59
		preference ²					
	μ^m	body size	-0.02	0.12	-0.17	-0.02	0.14
	τ^i	intercept	0.50	0.11	0.38	0.50	0.65
	τ^r	geographic	0.49	0.16	0.29	0.49	0.70
Standard deviation		range					
	τ^v	environmental	0.83	0.16	0.63	0.82	1.05
		preference					
	τ^{v^2}	environmental	1.49	0.35	1.08	1.46	1.94
		preference ²					
	τ^m	body size	0.47	0.12	0.32	0.46	0.63
Other	δ	sampling	0.90	0.15	0.71	0.89	1.08
	α	ageing	1.36	0.04	1.30	1.36	1.42

Note: These parameters are the 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 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.

Figure legends

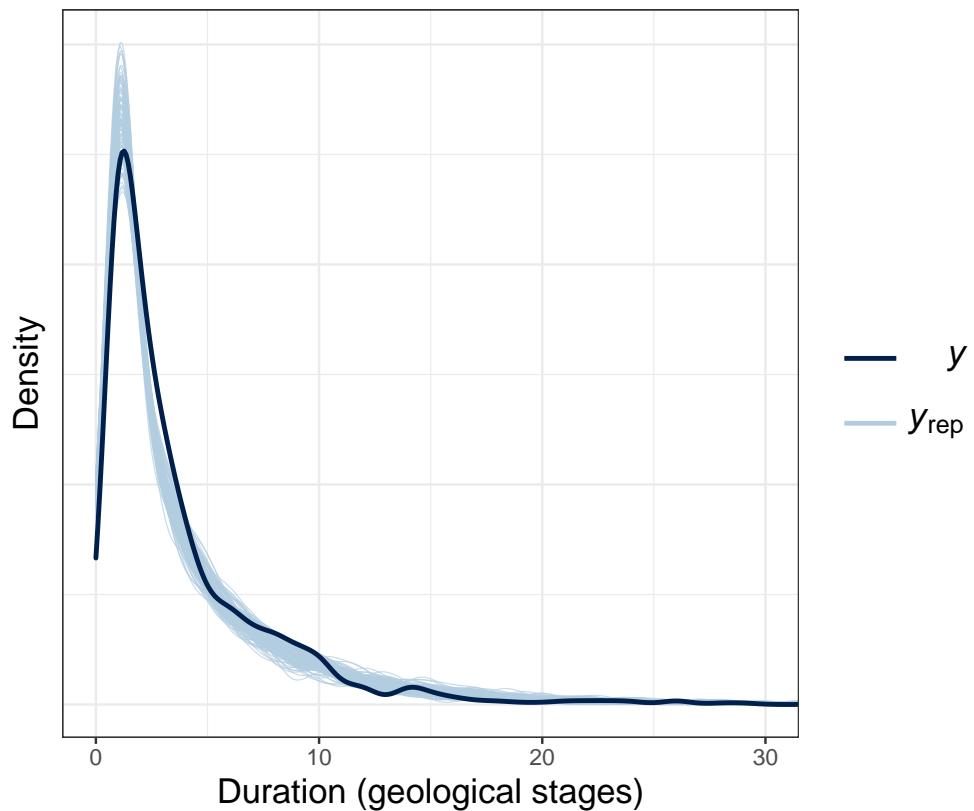


Figure 1: Comparison of the distribution of the observed data (black) to 100 simulated distributions (blue). This is a close-up view of the bulk of the distribution which shows the more subtle aspects of (mis)fit between the data and the model.

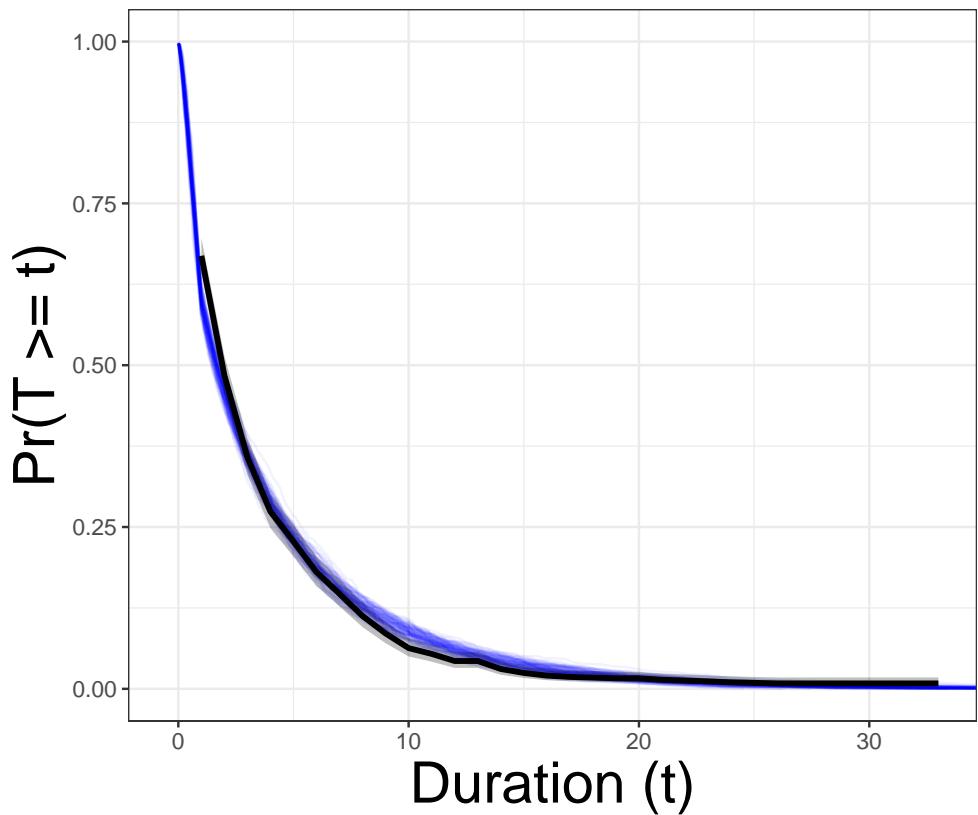


Figure 2: Comparison of the empirical estimate of $S(t)$ (blue) versus estimates from 100 posterior predictive data sets (black). $S(t)$ corresponds to the probability that the age of a genus t is less than the genus' ultimate duration T .

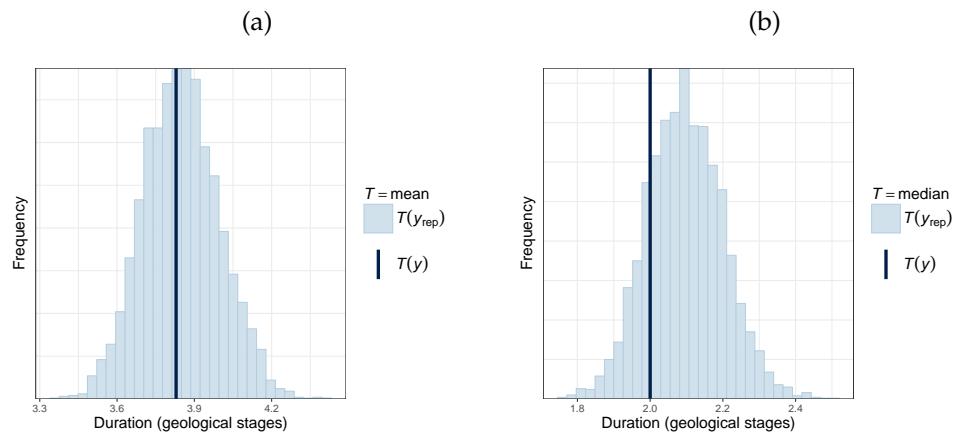


Figure 3: Comparison of the (A) observed mean genus duration (black vertical line) to a distribution of means estimated from 100 simulated datasets (blue), and (B) comparison of the observed median genus duration (black vertical line) to a distribution of medians estimated from 100 simulated datasets (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.

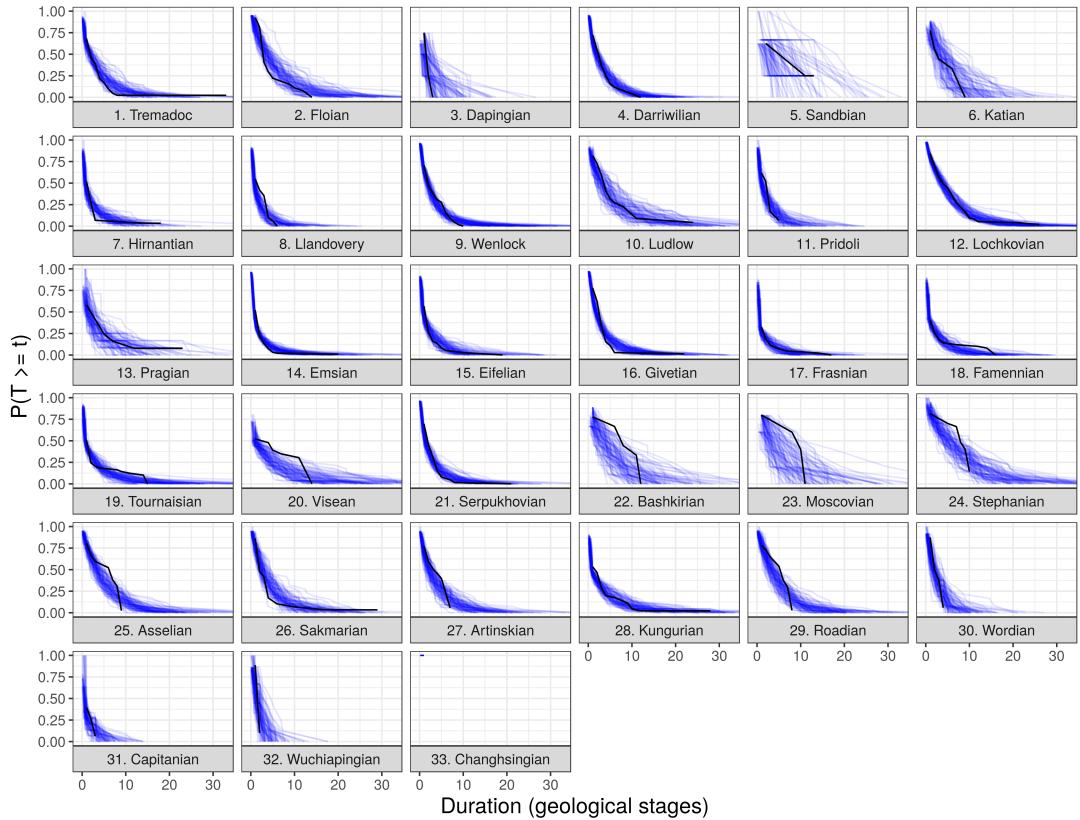


Figure 4: Comparison of the empirical estimate of $S(t)$ (blue) versus estimates from 100 posterior predictive data sets (black) for each of the origination cohorts. $S(t)$ corresponds to the probability that the age of a genus t is less than the genus' ultimate duration T . By comparing the fit of the model to the individual cohorts, when and where the model (mis)fits is more observable.

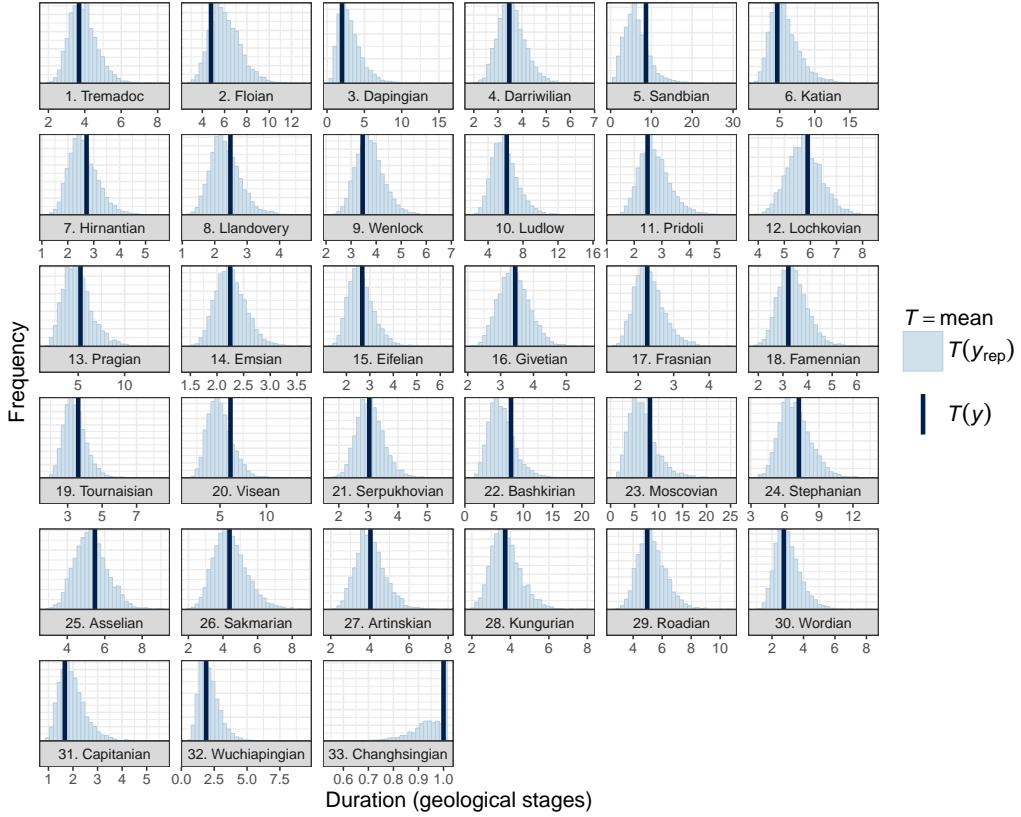


Figure 5: Comparison of the observed mean genus duration (black vertical line) to a distribution of means estimated from 100 simulated datasets (blue) for each of the origination cohorts. Model fit is evaluated by the similarity between the observed and the estimated, where good fit is demonstrated by the vertical line being “within” the simulated distribution.

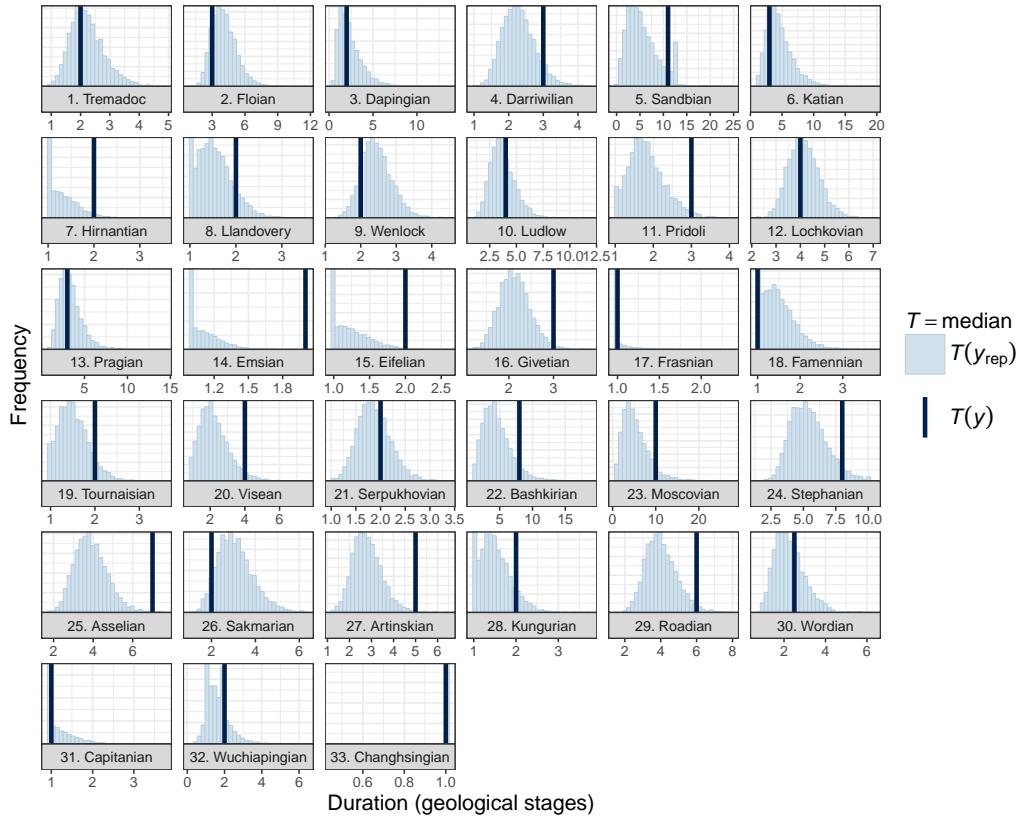


Figure 6: Comparison of the observed median genus duration (black vertical line) to a distribution of medians estimated from 100 simulated datasets (blue) for each of the origination cohorts. Model fit is evaluated by the similarity between the observed and the estimated, where good fit is demonstrated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

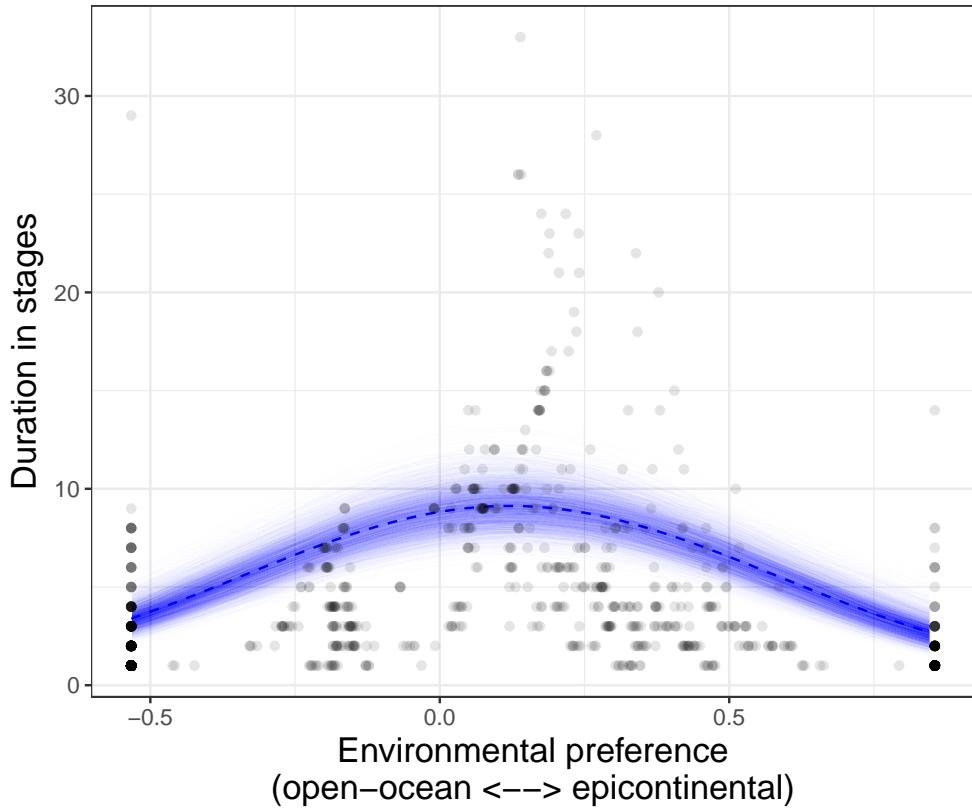


Figure 7: The overall expected relationship between environmental affinity v_i and a $\log(\sigma)$ when $r = 0$ and $m = 0$. The 1000 semi-transparent lines corresponds to a single draw from the posterior predictive distribution, while the highlighted line corresponds to the median of the posterior predictive distribution. The overall relationship demonstrates a greater durations among environmental generalists than specialists. Additionally, because the apex of is rightward from 0, taxa favoring epicontinental environments are expected to have a slightly 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.

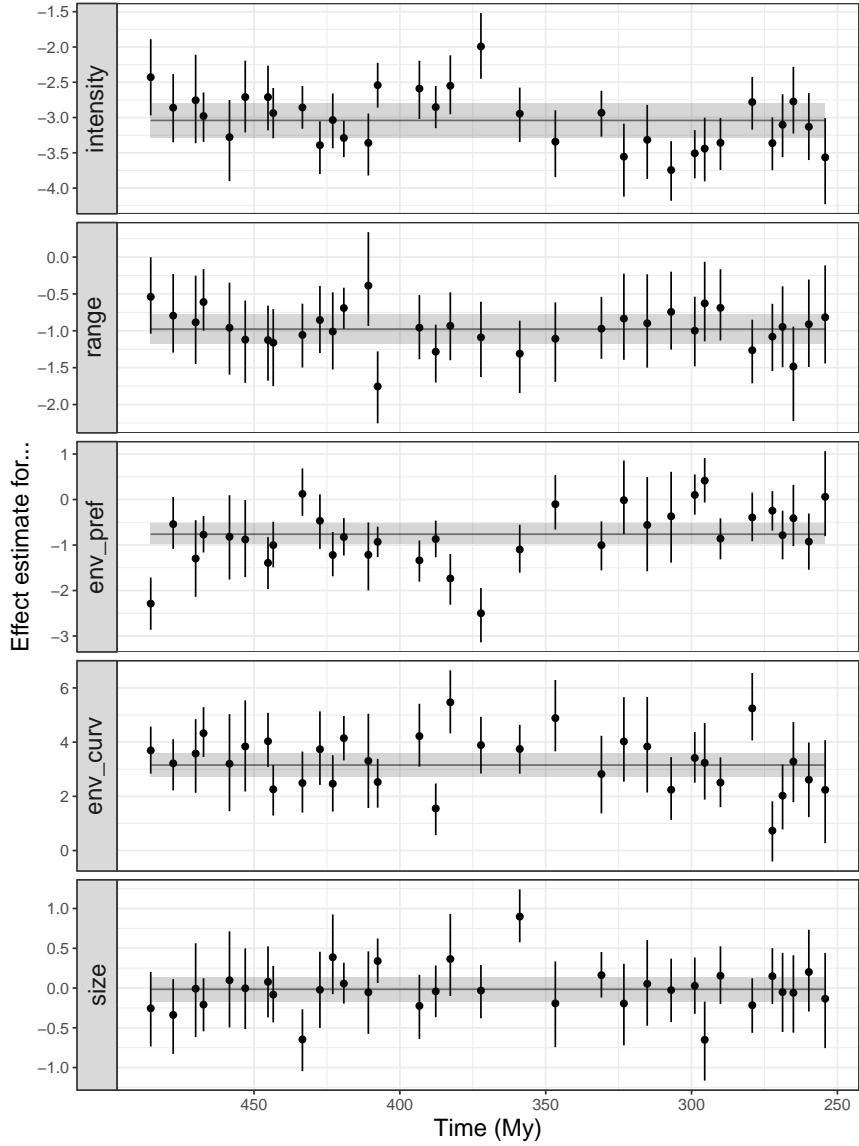


Figure 8: Comparison of cohort-specific estimates of β^0 , the effect of geographic range on extinction risk β^r , the effect of environmental preference β^v and β^{v^2} , and body size β^m . Points correspond to the median of the cohort-specific estimate, along with 80% credible intervals. Points are 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).

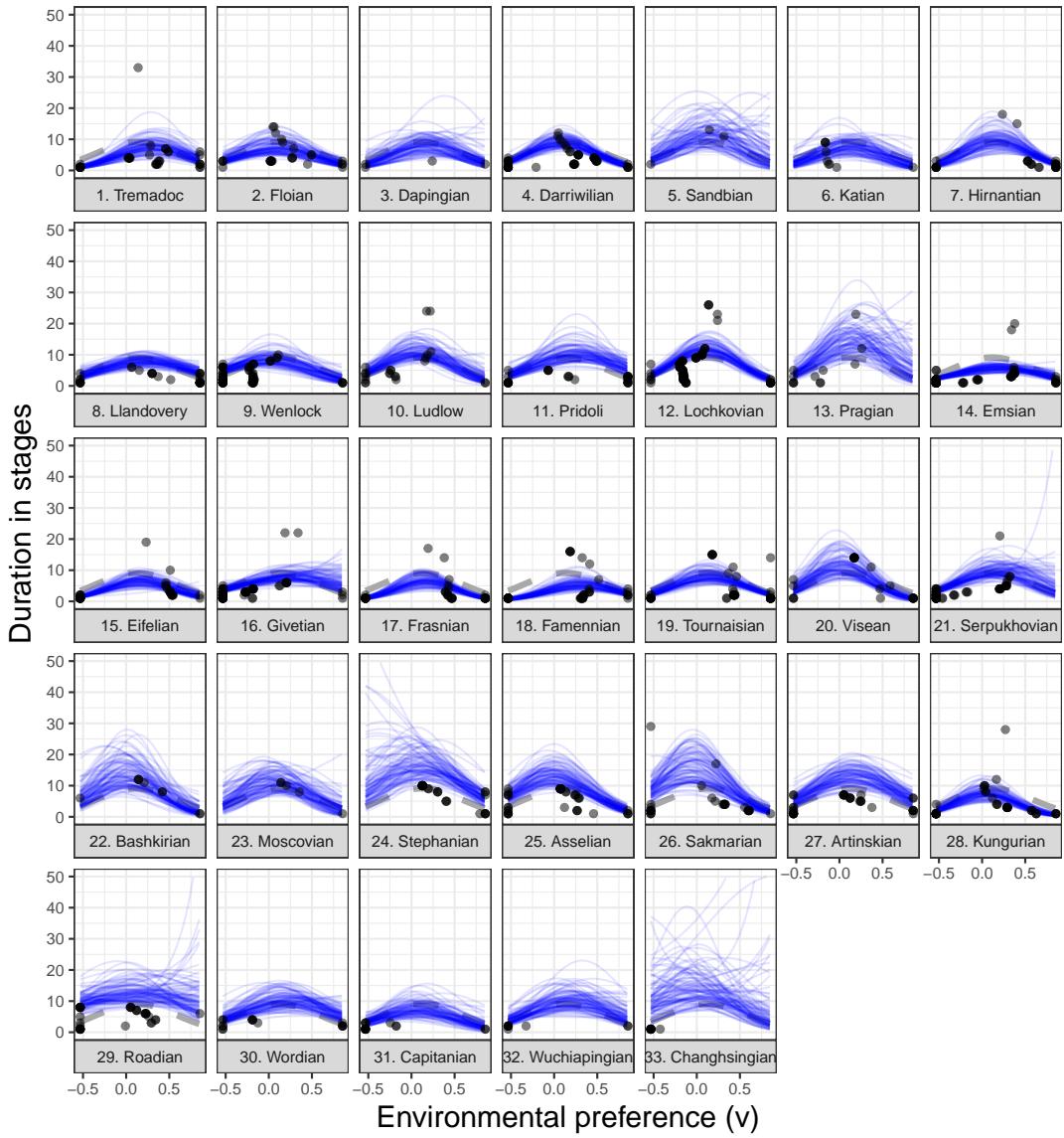


Figure 9: Comparison of origination cohort-specific (posterior predictive) estimates of the effect of environmental preference on $\log(\sigma)$ to the mean overall estimate of the effect of environmental preference. Cohort-specific estimates are from 100 posterior predictive simulations across the range of (transformed and rescaled) observed values of environmental preference. The oldest cohort is in the top-left 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.

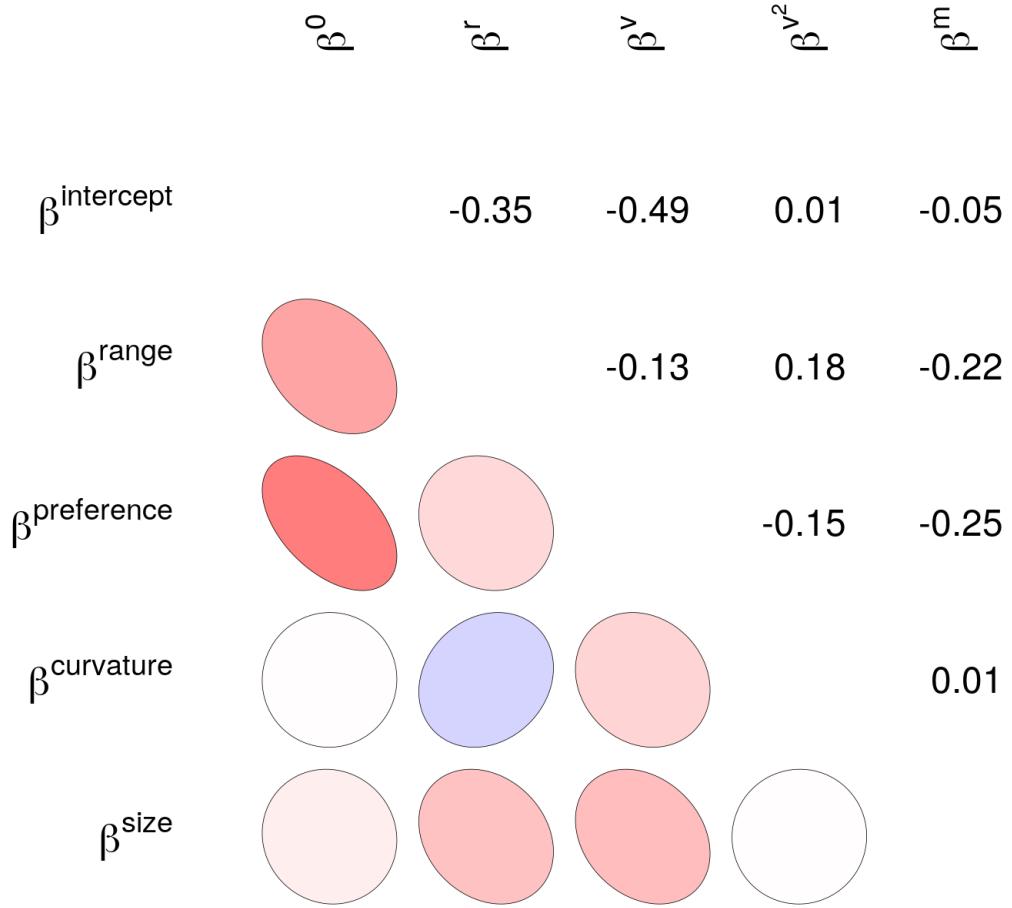


Figure 10: Mixed graphical and numerical representation of the correlation matrix Ω of variation in cohort-specific covariate estimates. These correlations are between the estimates of the cohort-level effects of covariates, along with intercept/baseline extinction risk. The median estimates of the correlations 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.