

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

4 **Running title:** Trait-dependent extinction in brachiopods

5 **Author:** Peter D Smits, psmits@uchicago.edu, Committee on Evolutionary
Biology, University of Chicago, IL, USA.

Keywords: extinction, brachiopods, Paleozoic, imputation, Bayesian

8 **Word count:**

Table count: 1

10 **Figure count:** 6

11 **Data archival location:** If accepted, all data and code necessary to duplicate
this analysis will be made available on DRYAD.

Abstract

As extinction intensity increases, how do the effects of traits on taxonomic survival change? Does the extinction rate associated with certain traits increase while that of others decreases? 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. I also test for the potential confounding effect of fossil preservation on my analysis. 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. I also infer that as extinction intensity increases, though both effects will increase, the variance in magnitude of the effect of environmental preference will be greater than the variance in the effect of geographic range. Additionally, I find support for greater survival of environmental generalists versus specialists in all origination cohorts, though taxa favoring epicontinental environments have a slightly greater duration than open-ocean favoring taxa. These results support the conclusion that for Paleozoic brachiopods, as extinction intensity increases overall extinction selectivity increases.

⁴⁰ 1 Introduction

How do biological traits affect extinction risk? Extinction is one half of the
⁴² diversification process (Raup, 1994, Stanley, 1975, 1979), second only to
speciation or origination; it can also be the ultimate manifestation of selection
⁴⁴ 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,
⁴⁶ 2010, Raup, 1994, Stanley, 1975).

While estimation of both trait-dependent speciation and extinction rates from
⁴⁸ phylogenies of extant taxa has grown dramatically (Fitzjohn, 2010, Goldberg
et al., 2011, 2005, Maddison et al., 2007, Rabosky et al., 2013, Stadler, 2011,
⁵⁰ 2013, Stadler and Bokma, 2013), there are two major ways to estimate
trait-dependent extinction: analysis of phylogenies, and analysis of the fossil
⁵² record. These two directions, phylogenetic comparative and paleobiological, are
complementary and intertwined in the field of macroevolution (Hunt and
⁵⁴ Rabosky, 2014, Jablonski, 2008, Rabosky and McCune, 2010). In the case of
extinction, analysis of the fossil record has the distinct advantage over
⁵⁶ phylogenies of only extant taxa because extinction is observable; this means
that extinction rate is possible to estimate (Liow et al., 2010, Quental and
⁵⁸ Marshall, 2009, Rabosky, 2010). The approach used here is thus complementary
to the analysis of trait-dependent extinction based on a phylogeny.

⁶⁰ Jablonski (1986) observed that for bivalves at the end Cretaceous mass
extinction event, the effects of some biological traits on taxonomic survival
⁶² decreased. However, this pattern was not the case for the effect of geographic
range on survival (Jablonski, 1986, Payne and Finnegan, 2007). There are
⁶⁴ multiple possible macroevolutionary mechanisms which may underlie this
pattern: the effect of geographic range on survival remains constant and those of

66 other biological traits decrease, the effect of geographic range on survival
increases and those of other biological traits stay constant, or the effects of all
68 traits decrease potentially by different degrees.

70 While Jablonski (1986) phrased his conclusions in terms of background versus
mass extinction, these states are not distinguishable in terms of extinction rate
alone; my analysis treats the time period analyzed as part of the same
72 continuum (Payne and Finnegan, 2007, Simpson and Harnik, 2009, Wang, 2003).
Additionally, in order to test the proposed macroevolutionary mechanism behind
74 the Jablonski (1986) scenario; not only do the taxon trait effects need to be
modeled, but the correlation between trait effects need to be modeled as well.

76 Here I model brachiopod taxon durations because trait based differences in
extinction risk should manifest as differences in taxon durations. Brachiopods
78 are an ideal group for this study as they are well known for having an
exceptionally complete fossil record (Foote, 2000b, Foote and Raup, 1996). I
80 focus on the brachiopod record from the post-Cambrian Paleozoic, from the
start of the Ordovician (approximately 485 My) through the end Permian
82 (approximately 252 My) as this represents the time of greatest global
brachiopod diversity (Alroy, 2010).

84 The analysis of taxon durations, or time from origination to extinction, falls
under the purview of survival analysis, a field of applied statistics commonly
86 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,
88 2015, Van Valen, 1973, 1979). I adopt a hierarchical modeling approach
(Gelman et al., 2013, Gelman and Hill, 2007, Klein and Moeschberger, 2003),
90 which represents both a conceptual and statistical unification of the
paleontological dynamic and cohort survival analytic approaches (Baumiller,
92 1993, Crampton et al., 2016, Ezard et al., 2012, Foote, 1988, Raup, 1975, 1978,

Simpson, 2006, Smits, 2015, Van Valen, 1973, 1979).

94 1.1 Factors affecting brachiopod survival

Conceptually, taxon survival can be considered an aspect of “taxon fitness”
96 along with expected lineage specific branching/origination rate (Cooper, 1984,
Palmer and Feldman, 2012). Traits associated with taxon survival are thus
98 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
100 emergent and aggregate traits (Jablonski, 2008, Rabosky and McCune, 2010);
specifically they are genus-level traits. Emergent traits are those which are not
102 measurable at a lower level (e.g. individual organism) such as geographic range,
environmental affinity, or even fossil sampling rate. Aggregate traits, like body
104 size, 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
106 estimating differences in extinction risk at nearly all times, with large
geographic range associated with low extinction risk (Finnegan et al., 2012,
108 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
110 random; a taxon with an unrestricted range is less likely to go extinct randomly
than a taxon with a restricted range.

112 Epicontinental seas are a shallow-marine environment where the ocean has
spread over the continental interior or craton of a continental shelf with a depth
114 typically less than 100m. In contrast, open-ocean coastline environments have
much greater variance in depth, do not cover the continental craton, and can
116 persist during periods of low sea level (Miller and Foote, 2009). Because of this,
it is strongly expected that taxa which favor epicontinental seas would be at

¹¹⁸ great risk during periods of low sea levels, such as during glacial periods, when
¹²⁰ epicontinental seas are drained. During the Paleozoic (approximately 541–252
My), epicontinental seas were widely spread globally but declined over the
¹²² Mesozoic (approximately 252–66 My) and have nearly disappeared during the
Cenozoic (approximately 66–0 My) as open-ocean coastlines became the
¹²⁴ dominant shallow-marine setting (Johnson, 1974, Miller and Foote, 2009, Peters,
2008).

¹²⁶ 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
¹²⁸ background extinction, however, they found no consistent difference between
taxa favoring either environment. These two environment types represent the
¹³⁰ primary identifiable environmental dichotomy observed in ancient marine
systems (Miller and Foote, 2009, Peters, 2008, Sheehan, 2001). Given these
¹³² findings, I predict that as extinction risk increases, the extinction risk associated
with open-ocean environments should generally increase.

¹³⁴ Because environmental preference is defined here as the continuum between
occurring exclusively in open-ocean environments versus epicontinental
¹³⁶ environments, intermediate values are considered “generalists” in the sense that
they favor neither end member. A long-standing hypothesis is that generalists or
¹³⁸ unspecialized taxa will have greater survival than specialists (Baumiller, 1993,
Liow, 2004, 2007, Nürnberg and Aberhan, 2013, 2015, Simpson, 1944). Because
¹⁴⁰ of this, the effect of environmental preference was modeled as a quadratic
function where a concave down relationship between preference and expected
¹⁴² duration indicates that generalists are favored over specialists end-members.

¹⁴⁴ 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
146 (Payne et al., 2014). Harnik et al. (2014) analyzed the effect of body size
selectivity in Devonian brachiopods in both a phylogenetic and
148 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
150 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
152 results, I expect that if body size has any effect on brachiopod taxonomic
survival it is very small.

154 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,
156 Foote and Raup, 1996, Liow and Nichols, 2010, Solow and Smith, 1997,
Wagner and Marcot, 2013, Wang and Marshall, 2004). Because of this, the
158 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
160 case, for sampling to be a confounding factor there must be consistent
relationship between the quality of sampling of a taxon and its apparent
162 duration (e.g. greater sampling, longer duration). By including sampling as a
covariate in the model, this effect can be quantified and can be taken into
164 account when estimating the effects of the other covariates. If there is no
relationship between sampling and duration then interpretation can be made
166 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
168 not a major determinant of the taxon's apparent duration.

2 Materials and Methods

170 2.1 Fossil occurrence information

The brachiopod dataset analyzed here was sourced from the Paleobiology
172 Database (<http://www.paleodb.org>) which was then filtered based on
taxonomic (Rhynchonelliformea: Rhynchonellata, Chileata, Obolellida,
174 Kutorginida, Strophomenida, Spiriferida)), temporal (post-Cambrian Paleozoic),
stratigraphic, and other occurrence information used in this analysis. Analyzed
176 occurrences were restricted to those with paleolatitude and paleolongitude
coordinates, assignment to either epicontinental or open-ocean environment,
178 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
180 partially based on those from Miller and Foote (2009), with additional
occurrences assigned similarly (Miller and Foote, personal communication).
182 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
184 Payne et al. (2014). In total, there were 1130 genera included in the dataset.

Fossil occurrences were analyzed at the genus level which is common for
186 paleobiological, macroevolutionary and macroecological studies of marine
invertebrates (Alroy, 2010, Foote and Miller, 2013, Harnik et al., 2012, Kiessling
188 and Aberhan, 2007, Miller and Foote, 2009, Nürnberg and Aberhan, 2013, 2015,
Payne and Finnegan, 2007, Simpson and Harnik, 2009, Vilhena et al., 2013).
190 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
192 the fossil record makes accurate and precise taxonomic assignments at the
species level difficult for all occurrences. As such, the choice to analyze genera
194 as opposed to species was in order to assure a minimum level of confidence and

accuracy in the data analyzed here.

196 Genus duration was calculated as the number of geologic stages from first
appearance to last appearance, inclusive. Durations were based on geologic
198 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
200 instead fossils are assigned to a geological interval which represents some
temporal range. In this analysis, stages are effectively irreducible temporal
202 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
204 censored at the Changhsingian; genera with a duration of only one stage were
left censored (Klein and Moeschberger, 2003). How the likelihood of censored
206 observations is calculated is detailed below.

The covariates detailed below are geographic range size (r), environmental
208 preference (v, v^2), body size (m), and sampling (s).

Geographic range was calculated using an occupancy approach corrected for
210 incomplete sampling. First, all occurrences were projected onto an equal-area
cylindrical map projection. Each occurrence was then assigned to one of the
212 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
214 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
216 occupied grid cells was calculated. Then, for each temporal bin, the relative
occurrence probability of the observed taxa was calculated using the JADE
218 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 is
220 estimated from a truncated Binomial distribution. This correction is critical
when comparing occupancies from different times with different geographic

²²² sampling. Finally, for each genus, the mean relative occurrence probability was
calculated as the average of that genus' occurrence probabilities for all stages it
²²⁴ was sampled.

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.

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

²²⁶ where v is the percent of the distribution defined in equation 2.1 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 idea for modeling percentages.

²²⁸ Body size, measured as shell length, was sourced directly from Payne et al.
(2014). These measurements were made from brachiopod taxa figured in the
²³⁰ *Treatise on Invertebrate Paleontology* (Williams et al., 2007).

The sampling record for individual taxa was calculated using the standard gap
²³² statistic (Foote, 2000a, Foote and Raup, 1996). The gap statistic is calculated as
the number of stages in which the taxon was sampled minus two divided by the
²³⁴ duration of the taxon minus two. Subtracting two from both the numerator and
denominator is because the first and last appearance stages are by definition
²³⁶ sampled. Because the minimum duration for which a gap statistic can be
calculated is three stages, I chose to impute the gap statistic for all
²³⁸ observations with a duration of 2 or less. Imputation is the “filling in” of
missing observations based on the observed values (Gelman and Hill, 2007).

²⁴⁰ This is fairly straight forward in a Bayesian framework because both covariates

and parameters are considered random variables, meaning that the missing
242 values of sampling can be modeled as coming from some probability distribution.
The model for sampling is described below.

244 Prior to analysis, geographic range was logit transformed and body size was
natural-log transformed; both of these transformations make these variables
246 defined for the entire real line. Sampling was transformed where

$$s* = (s(n - 1) + 0.5)/n \text{ where } n \text{ is the sample size as recommended by}$$

248 CITATION; this serves to shrink the range of the data slightly so that there are
no values of 0 or 1. All covariates except for sampling were standardized by
250 subtracting the mean from all values and dividing by twice its standard
deviation, which follows Gelman and Hill (2007). This standardization means
252 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
254 total number of covariates, excluding sampling, plus one for the intercept term.

2.2 Analytical approach

256 Hierarchical modelling is a statistical approach which explicitly takes into
account the structure of the observed data in order to model both the within
258 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
260 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,
262 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.
264 covariate effects) (Gelman et al., 2013). The subsequent estimates are partially
pooled together, where parameters from groups with large samples or effects

266 remain large while those of groups with small samples or effects are pulled
 towards the overall group mean. All covariate effects (regression coefficients), as
 268 well as the intercept term (baseline extinction risk), were allowed to vary by
 group (origination cohort). The covariance between covariate effects was also
 270 modeled.

Genus durations were assumed to follow a Weibull distribution which allows for
 272 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
 274 $\alpha = 1$, 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
 276 extinction risk increases with age, and values less than 1 indicate that
 extinction risk decreases with age. Note that the Weibull distribution is
 278 equivalent to the exponential distribution when $\alpha = 1$.

In the case of the right- and left-censored observations mentioned above, the
 280 probability of those observations has a different calculation (Klein and
 Moeschberger, 2003). For right-censored observations, the likelihood is
 282 calculated $p(y|\theta) = 1 - F(y) = S(y)$ where $F(y)$ is the cumulative distribution
 function. In contrast, the likelihood of a left-censored observation is calculated
 284 $p(y|\theta) = F(y)$.

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

$$\sigma_i = \exp\left(\frac{-\mathbf{X}_i B_{j[i]} + \delta s}{\alpha}\right) \quad (1)$$

288 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

290 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 $J \times D$
 292 matrix of cohort-specific regression coefficients, and δ is the regression
 coefficient for the effect of sampling s . δ does not vary by cohort.

294 Each of the rows of matrix B are modeled as realizations from a multivariate
 normal distribution with length D location vector μ and $J \times J$ covariance
 296 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
 298 $\Sigma = \text{diag}(\tau)\Omega\text{diag}(\tau)$ 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 almost uniform LKJ distributed prior following (Team, 2016). These
 priors are defined

$$\begin{aligned}
 \mu_0 &\sim \mathcal{N}(0, 5) \\
 \mu_r &\sim \mathcal{N}(-1, 1) \\
 \mu_v &\sim \mathcal{N}(0, 1) \\
 \mu_{v^2} &\sim \mathcal{N}(1, 1) \\
 \mu_m &\sim \mathcal{N}(0, 1) \\
 \tau &\sim \text{C}^+(1) \\
 \Omega &\sim \text{LKJ}(2).
 \end{aligned}$$

The vector sampling s has two parts: the observed part s^o , and the unobserved
 part s^u . Of the 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 0 and 1, I chose to model it has a beta regression with the other covariates X (e.g. geographic range size) as predictors of sampling CITATION; this assumes that the sampling value for all taxa come from the same distribution. This beta regression is defined

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

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

$$\phi = \frac{\alpha}{\alpha + \beta}$$

and total count parameter

$$\lambda = \alpha + \beta \quad (2)$$

300 where α and β are the characteristic parameters of the Beta distribution.

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. 1). All the elements of γ , δ , and λ were given weakly informative priors where

$$\gamma \sim \mathcal{N}(0, 1)$$

$$\delta \sim \mathcal{N}(0, 1)$$

$$\lambda \sim \text{Pareto}(0.1, 1.5).$$

The log of the shape parameter α was given a weakly informative prior

302 $\log(\alpha) \sim \mathcal{N}(0, 1)$ centered at $\alpha = 1$, which corresponds to the Law of Constant Extinction (Van Valen, 1973).

304 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

³⁰⁶ Sampler (Hoffman and Gelman, 2014) as implemented in the probabilistic
programming language Stan (Stan Development Team, 2014). The posterior
³⁰⁸ distribution was approximated from four parallel chains run for 10,000 steps
each, split half warm-up and half sampling and thinned to every 10th sample for
³¹⁰ a total of 4000 posterior samples. Chain convergence was assessed via the scale
reduction factor \hat{R} where values close to 1 ($\hat{R} < 1.1$) indicate approximate
³¹² convergence. Convergence means that the chains are approximately stationary
and the samples are well mixed (Gelman et al., 2013).

³¹⁴ Model adequacy was evaluated using a couple of posterior predictive checks.
Posterior predictive checks are a means for understanding model fit or adequacy
³¹⁶ where the basic idea is that replicated data sets simulated from the fitted model
should be similar to the original data and systematic differences between the
³¹⁸ simulations and observations indicate weaknesses of the model fit (Gelman
et al., 2013). For both approaches used here, 1000 posterior predictive datasets
³²⁰ were generated from 1000 unique draws from the posterior distribution of each
parameter. The two posterior predictive checks used in this analysis are a
³²² comparison of a non-parameteric estimate of the survival function $S(t)$ from the
empirical dataset to the non-parameteric estimates of $S(t)$ from the 1000
³²⁴ posterior predictive datasets, and comparison of the observed genus durations to
the average posterior predictive estimate of $\log(\sigma)$ (Eq. 1). The former is to see
³²⁶ if simulated data has a similar survival pattern to the observed, while the latter
is to see if the model systematically over- or under- estimates taxon survival.

³²⁸ 3 Results

Comparison of the posterior predictive estimates of $S(t)$ to the empirical
³³⁰ estimate reveal few obvious biases except for the case of estimates from the far

right tail of observed durations (Fig. 1). This result is reinforced by the
 332 additional posterior predictive comparison where most estimates are not
 systematically biased except for a consistent under-estimate of $\log(\sigma)$ for older
 334 taxa (Fig. 2). The results of both posterior predictive checks indicate that, for
 the majority of observations, model fit is generally not biased.

 336 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
 338 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
 340 ($\Pr(\mathcal{N}(\mu_r, \tau_r) > 0)$). The between-cohort variance τ_r is the lowest of all the
 regression coefficients (Table 1).

 342 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
 344 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 .

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

 348 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,
 350 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
 352 environments would have a greater expected duration than taxa favoring
 epicontinental environments ($\Pr(\mathcal{N}(\mu_v, \tau_v) > 0)$).

 354 The estimate of μ_{v^2} indicates that the overall relationship between
 environmental preference and $\log(\sigma)$ is concave down (Fig. 3), with only a 2.7%
 356 ($\pm 3\%$ SD) 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
358 of between cohort variance, with no obvious trends. As indicated in Table 1 and
detectable visually (Fig. 4), the between-cohort estimates for β_0 , β_r , and β_m all
360 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
362 estimate, there are a few notable excursions away from the overall mean (Fig.
4). There are simultaneous excursions in both β_0 and β_v for cohorts originating
364 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
366 Frasnian/Famennian boundary. These cohorts are marked by both a high
extinction intensity and an increase in expected duration for taxa favoring
368 epicontinental environments over open-ocean ones; this is consistent with the
results of Miller and Foote (2009).

370 Cohorts originating from the Silurian through the Early Devonian have a
slightly lower extinction intensity than the overall mean; these cohorts are those
372 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
374 that epicontinental favoring taxa are expected to have greater duration than
open-ocean favoring taxa. Both the Silurian and Devonian periods are notable
376 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,
378 Munnecke et al., 2010).

The cohort-specific relationships between environmental preference and $\log(\sigma)$
380 were calculated from the estimates of β_0 , β_v , and β_{v^2} (Fig. 5) and reflect how
these three parameters act in concert and not just individually (Fig. 4). Beyond
382 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. 4). 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. 6).

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

⁴¹⁰ 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
⁴¹² 1); the effect of geographic range might simply not vary enough relative to the
much noisier environmental preference.

⁴¹⁴ Sampling was found to have a negative effect on duration: greater sampling,
shorter duration (Table 1). While this affect appears large compared to the
⁴¹⁶ other regression coefficients, this is only because sampling was not standardized
like the other covariates. To make the coefficients comparable, μ_r is multiplied
⁴¹⁸ by twice the standard deviation of geographic range which yields mean -1.51
(± 0.24 SD), which has a 99.25% probability of being greater than the estimate
⁴²⁰ of δ .

⁴²² The effect of sampling, however, has a 96.8% probability of being greater in
magnitude than the effect of environmental preference μ_v though there are a
few cases when the effect of environmental preference is greater in magnitude
⁴²⁴ than the effect of sampling such as the Givetian (387-382 My; 99.8% $|\beta_v| > |\delta|$)
and Frasnian (382-372 My; 85.3% $|\beta_v| > |\delta|$). Finally, there is a 0% probability
⁴²⁶ that δ is greater in magnitude than μ_v^2 .

⁴²⁸ 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
⁴³⁰ consistent with accelerating extinction risk with taxon age. This may indicate
that older taxa are out-competed by younger taxa, a result consistent
⁴³² (ironically) with a recently proposed Red Queen-like scenario (Rosindell et al.,
2015). This results, however, may be caused by the minimum resolution of the
⁴³⁴ 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
⁴³⁶ effect of taxon age on extinction risk.

4 Discussion

438 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
440 affect extinction risk was geographic range while traits that had previously been
beneficial had no effect (Jablonski, 1986). This observation raises two linked
442 questions: how does the effect of geographic range change with changing
extinction intensity, and how does the effect of other biological traits change
444 with changing extinction intensity?

I find that as intensity increases (β_0 decreases), the magnitude of the effect of
446 geographic range increase. I also find that as intensity increases, the effect of
favoring epicontinental environments of open-ocean environments is expected to
448 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
450 environmental preference. However, the between-cohort variance in effect of
geographic range is much less than the between-cohort variance of the effect of
452 environmental preference. As stated earlier, this may underlie the lack of
correlation between these two effects.

454 Additionally, the lower between-cohort variance in the effect of geographic range
versus that higher between-cohort variance implies that for cohorts with a
456 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
458 than the difference between the effect of environmental preference and the
group-level effect of environmental preference.

460 I find consistent support for the “survival of the unspecialized,” with respect to
epicontinental versus open-ocean environmental preference, as a time-invariant
462 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. 3), though the curvature
of the relationship varies from rather shallow to very peaked (Fig. 5). 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: do
“generalists” have a greater duration because they are “true” environmental
generalists from origination (e.g. can inhabit both environments) or have taxa
with longer durations evolved their preference over their duration in the face of
changing environments (i.e. niche evolution and adaptation)? The nature of the
analysis here assumes that traits are constant over the duration of the taxon;
future work should explore how environmental preference changes over lineage
duration in relation to environmental availability.

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,

⁴⁹⁰ Royle and Dorazio, 2008, Tomiya, 2013). This model is a type of hidden Markov
⁴⁹¹ model with an absorbing state (i.e. extinction). In this model, survival is defined
⁴⁹² as the probability of surviving from time t to time $t + 1$. Additionally, the effect
⁴⁹³ of preservation and sighting is estimated as probability of observing a taxon
⁴⁹⁴ that is present; this can extend the duration of a taxon beyond its last
⁴⁹⁵ occurrence. This approach is a fundamentally different from the method used in
⁴⁹⁶ the above analysis; for example, instead of modeling persistence from time t to
⁴⁹⁷ time $t + 1$, I model the distribution of taxon durations (lifetime) which then
⁴⁹⁸ decomposes into a function describing survival ($S(t)$) and a function describing
⁴⁹⁹ instantaneous extinction rate with taxon age (Klein and Moeschberger, 2003,
⁵⁰⁰ Royle and Dorazio, 2008).

The use of genera as the unit of the study and how to exactly interpret the
⁵⁰¹ effects of the biological traits is an important question. For example, if any of
⁵⁰² the traits analyzed here are associated with increases in speciation rates, this
⁵⁰³ might increase the duration of genera through self-renewal (Raup, 1991, 1994),
⁵⁰⁴ which would be an example of the difference in biological pattern between
⁵⁰⁵ species and genera (Jablonski, 1987, 2007, 2008). This could lead to a trait
⁵⁰⁶ appearing to decrease generic level extinction risk by that trait increasing
⁵⁰⁷ species level origination rate instead of decreasing species level extinction risk.

The model used here could be improved through either increasing the number of
⁵⁰⁸ analyzed traits, expanding the hierarchical structure of the model to include
⁵⁰⁹ other major taxonomic groups of interest, and the inclusion of explicit
⁵¹⁰ phylogenetic relationships between the taxa in the model as an additional
⁵¹¹ hierarchical effect. An example trait that may be of particular interest is the
⁵¹² affixing strategy or method of interaction with the substrate of the taxon, which
⁵¹³ has been found to be related to brachiopod survival where, for cosmopolitan
⁵¹⁴ taxa, taxa that are attached to the substrate are expected to have a greater

duration than those that are not (Alexander, 1977).

518 It is theoretically possible to expand this model to allow for comparisons both
within and between major taxonomic groups which would better constrain the
520 brachiopod estimates while also allowing for estimation of similarities and
differences in cross-taxonomic patterns. The major issue surrounding this
522 particular expansion involves finding a similarly well sampled taxonomic group
that is present during the Paleozoic. Example groups include Crinoidea,
524 Ostracoda, and other members of the “Paleozoic fauna” (Sepkoski, 1981).

Traits like environmental preference or geographic range (Hunt et al., 2005,
526 Jablonski, 1987) are most likely heritable. Without phylogenetic context, this
analysis assumes that differences in extinction risk between taxa are independent
528 of the shared evolutionary history of those taxa (Felsenstein, 1985). In contrast,
the origination cohorts only capture shared temporal context. For example, if
530 taxon duration is phylogenetically heritable, then closely related taxa may have
more similar durations as well as more similar biological traits. Without taking
532 into account phylogenetic similarity the effects of these biological traits would
be inflated solely due to inheritance. The inclusion of phylogenetic context as an
534 additional individual-level hierarchical effect independent of origination cohort
would allow for determining how much of the observed variability is due to
536 shared evolutionary history versus shared temporal context versus actual
differences associated with biological traits (Harnik et al., 2014, Smits, 2015).

538 In summary, patterns of Paleozoic brachiopod survival were analyzed using a
fully Bayesian hierarchical survival modelling approach while also eschewing the
540 traditional separation between background and mass extinction. I find that
cohort extinction intensity is negatively correlated with both the cohort-specific
542 effects of geographic range and environmental preference. These results imply
that as extinction intensity increases (β_0) increases, it is expected that both

544 effects will increase in magnitude. However, the change in effect of
environmental preference is expected to be greater than the change in the effect
546 of geographic range. Additionally, I find support for greater survival in
environmental generalists over specialists in all origination cohorts analyzed;
548 this is consistent with the long standing “survival of the unspecialized”
hypothesis (Baumiller, 1993, Liow, 2004, 2007, Nürnberg and Aberhan, 2013,
550 2015, Simpson, 1944, 1953, Smits, 2015). I also find that taxa favoring
epicontinental environments are expected to have a slightly greater duration
552 than open-ocean favoring taxa. The results of this analysis support the
conclusion that for Paleozoic brachiopods, as extinction intensity increases
554 overall extinction selectivity is expected to increase as well.

Acknowledgements

556 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
558 D. Bapst, N. Pierrehumbert and M. Villarosa Garcia for additional comments.
An additional thank you to A. Miller for the epicontinental versus open-ocean
560 assignments. Finally, thank you to the reviewers for their helpful comments that
improved this manuscript. This entire study would not have been possible
562 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,
564 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.
566 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
568 conflicts of interest. This is Paleobiology Database publication XXX.

References

- 570 Alexander, R. R., 1977. Generic longevity of articulate brachiopods in relation
to the mode of stabilization on the substrate. *Palaeogeography,*
572 *Palaeoclimatology, Palaeoecology* 21:209–226.
- Alroy, J., 2010. The Shifting Balance of Diversity Among Major Marine Animal
574 Groups. *Science* 329:1191–1194.
- , 2014. A simple Bayesian method of inferring extinction. *Paleobiology*
576 40:584–607.
- Baumiller, T. K., 1993. Survivorship analysis of Paleozoic Crinoidea: effect of
578 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.
580 Inouye, 2015. Unveiling the species-rank abundance distribution by
generalizing the Good-Turing sample coverage theory. *Ecology* 96:1189–1201.
- 582 Cooper, W. S., 1984. Expected time to extinction and the concept of
fundamental fitness. *Journal of Theoretical Biology* 107:603–629.
- 584 Crampton, J. S., R. A. Cooper, P. M. Sadler, and M. Foote, 2016.
Greenhouseicehouse transition in the Late Ordovician marks a step change in
586 extinction regime in the marine plankton. *Proceedings of the National
Academy of Sciences P.* 201519092.
- 588 Edwards, D. and U. Fanning, 1985. Evolution and environment in the late
Silurian–early Devonian: the rise of pteridophytes. *Philosophical Transactions
590 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
592 birth and death (in macroevolutionary birth-death models). *Biology letters*
8:139–42.

- 594 Felsenstein, J., 1985. Phylogenies and the comparative method. *American
Naturalist* 125:1–15.
- 596 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.
598 *Proceedings of the National Academy of Sciences* 109:6829–6834.
- 600 Fitzjohn, R. G., 2010. Quantitative Traits and Diversification. *Systematic
Biology* Pp. 1–15.
- 602 Foote, M., 1988. Survivorship analysis of Cambrian and Ordovician Trilobites.
Paleobiology 14:258–271.
- 604 ———, 2000a. Origination and extinction components of taxonomic diversity:
general problems. *Paleobiology* 26:74–102.
- 606 ———, 2000b. Origination and extinction components of taxonomic diversity:
Paleozoic and post-Paleozoic dynamics. *Paleobiology* 26:578–605.
- 608 ———, 2014. Environmental controls on geographic range size in marine animal
genera. *Paleobiology* 40:440–458.
- 610 Foote, M. and A. I. Miller, 2013. Determinants of early survival in marine
animal genera. *Paleobiology* 39:171–192.
- 612 Foote, M. and D. Raup, 1996. Fossil preservation and the stratigraphic ranges
of taxa. *Paleobiology* 22:121–140.
- 614 Gelman, A., J. B. Carlin, H. S. Stern, D. B. Dunson, A. Vehtari, and D. B.
Rubin, 2013. *Bayesian data analysis*. 3 ed. Chapman and Hall, Boca Raton,
FL.
- 616 Gelman, A. and J. Hill, 2007. *Data Analysis using Regression and
Multilevel/Hierarchical Models*. Cambridge University Press, New York, NY.

- 618 Goldberg, E. E., L. T. Lancaster, and R. H. Ree, 2011. Phylogenetic inference
of reciprocal effects between geographic range evolution and diversification.
- 620 Systematic biology 60:451–65.
- Goldberg, E. E., K. Roy, R. Lande, and D. Jablonski, 2005. Diversity,
622 endemism, and age distributions in macroevolutionary sources and sinks. The
American naturalist 165:623–33.
- 624 Harnik, P. G., 2011. Direct and indirect effects of biological factors on
extinction risk in fossil bivalves. Proceedings of the National Academy of
Sciences of the United States of America 108:13594–9.
- Harnik, P. G., P. C. Fitzgerald, J. L. Payne, and S. J. Carlson, 2014.
628 Phylogenetic signal in extinction selectivity in Devonian terebratulide
brachiopods. Paleobiology 40:675–692.
- 630 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
632 Society B: Biological Sciences 279:4969–4976.
- Heim, N. A. and S. E. Peters, 2011. Regional environmental breadth predicts
634 geographic range and longevity in fossil marine genera. PloS one 6:e18946.
- Hijmans, R. J., 2015. raster: Geographic data analysis and modeling. URL
636 <http://CRAN.R-project.org/package=raster>. R package version 2.3-24.
- Hoehn, K. B., P. G. Harnik, and V. L. Roth, 2015. A framework for detecting
638 natural selection on traits above the species level. Methods in Ecology and
Evolution .
- 640 Hoffman, M. D. and A. Gelman, 2014. The No-U-Turn Sampler: Adaptively
Setting Path Lengths in Hamiltonian Monte Carlo. Journal of Machine
642 Learning Research 15:1351–1381.

- Hopkins, M. J., C. Simpson, and W. Kiessling, 2014. Differential niche dynamics
644 among major marine invertebrate clades. *Ecology Letters* 17:314–323.
- Hunt, G. and D. L. Rabosky, 2014. Phenotypic Evolution in Fossil Species:
646 Pattern and Process. *Annual Review of Earth and Planetary Sciences*
42:421–441.
- 648 Hunt, G., K. Roy, and D. Jablonski, 2005. Species-level heritability reaffirmed: a
comment on "On the heritability of geographic range sizes". *American
650 Naturalist* 166:129–135.
- Jablonski, D., 1986. Background and mass extinctions: the alternation of
652 macroevolutionary regimes. *Science* 231:129–133.
- , 1987. Heritability at the species level: analysis of geographic ranges of
654 cretaceous mollusks. *Science* 238:360–363.
- , 2007. Scale and hierarchy in macroevolution. *Palaeontology* 50:87–109.
- 656 ———, 2008. Species Selection: Theory and Data. *Annual Review of Ecology,
Evolution, and Systematics* 39:501–524.
- 658 Jablonski, D. and K. Roy, 2003. Geographical range and speciation in fossil and
living molluscs. *Proceedings. Biological sciences / The Royal Society*
660 270:401–6.
- Joachimski, M. M., S. Breisig, W. Buggisch, J. A. Talent, R. Mawson,
662 M. Gereke, J. R. Morrow, J. Day, and K. Weddige, 2009. Devonian climate
and reef evolution: Insights from oxygen isotopes in apatite. *Earth and
664 Planetary Science Letters* 284:599–609.
- Johnson, J. G., 1974. Extinction of Perched Faunas. *Geology* 2:479–482.
- 666 Kiessling, W. and M. Aberhan, 2007. Environmental determinants of marine

- benthic biodiversity dynamics through Triassic–Jurassic time. *Paleobiology*
668 33:414–434.
- Klein, J. P. and M. L. Moeschberger, 2003. Survival Analysis: Techniques for
670 Censored and Truncated Data. 2nd ed. Springer, New York.
- Kleinbaum, D. G. and M. Klein, 2005. Survival analysis: a self-learning text. 2
672 ed. Springer, New York, NY.
- Liow, L. H., 2004. A test of Simpson's "rule of the survival of the relatively
674 unspecialized" using fossil crinoids. *The American Naturalist* 164:431–43.
- , 2007. Does versatility as measured by geographic range, bathymetric
676 range and morphological variability contribute to taxon longevity? *Global
Ecology and Biogeography* 16:117–128.
- 678 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
680 mammals. *Proceedings of the National Academy of Sciences* 105:6097.
- Liow, L. H. and J. D. Nichols, 2010. Estimating rates and probabilities of
682 origination and extinction using taxonomic occurrence data:
Capture-mark-recapture (CMR) approaches. *Quantitative Methods in
684 Paleobiology* Pp. 81–94.
- Liow, L. H., T. B. Quental, and C. R. Marshall, 2010. When can decreasing
686 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
688 character's effect on speciation and extinction. *Systematic Biology* 56:701.
- Miller, A. I. and M. Foote, 2009. Epicontinental seas versus open-ocean settings:
690 the kinetics of mass extinction and origination. *Science* 326:1106–9.

- 692 Munnecke, A., M. Calner, D. A. T. Harper, and T. Servais, 2010. Ordovician
and Silurian sea-water chemistry, sea level, and climate: A synopsis.
- 694 Palaeogeography, Palaeoclimatology, Palaeoecology 296:389–413.
- Nürnberg, S. and M. Aberhan, 2013. Habitat breadth and geographic range
696 predict diversity dynamics in marine Mesozoic bivalves. Paleobiology
39:360–372.
- 698 ———, 2015. Interdependence of specialization and biodiversity in Phanerozoic
marine invertebrates. Nature communications 6:6602.
- 700 Palmer, M. E. and M. W. Feldman, 2012. Survivability is more fundamental
than evolvability. PloS one 7:e38025.
- 702 Payne, J. L. and S. Finnegan, 2007. The effect of geographic range on
extinction risk during background and mass extinction. Proceedings of the
704 National Academy of Sciences 104:10506–11.
- Payne, J. L., N. A. Heim, M. L. Knope, and C. R. McClain, 2014. Metabolic
706 dominance of bivalves predates brachiopod diversity decline by more than 150
million years. Proceedings of the Royal Society B 281:20133122.
- 708 Peters, S. E., 2008. Environmental determinants of extinction selectivity in the
fossil record. Nature 454:626–9.
- 710 Quental, T. B. and C. R. Marshall, 2009. Extinction during evolutionary
radiations: reconciling the fossil record with molecular phylogenies. Evolution
712 63:3158–67.
- Rabosky, D. L., 2010. Extinction rates should not be estimated from molecular
714 phylogenies. Evolution 64:1816–24.
- Rabosky, D. L. and A. R. McCune, 2010. Reinventing species selection with
716 molecular phylogenies. Trends in ecology & evolution 25:68–74.

- Rabosky, D. L., F. Santini, J. Eastman, S. a. Smith, B. Sidlauskas, J. Chang,
718 and M. E. Alfaro, 2013. Rates of speciation and morphological evolution are
correlated across the largest vertebrate radiation. *Nature Communications*
720 4:1–8.
- Raup, D. M., 1975. Taxonomic survivorship curves and Van Valen's Law.
722 *Paleobiology* 1:82–96.
- , 1978. Cohort Analysis of generic survivorship. *Paleobiology* 4:1–15.
- 724 ———, 1991. *Extinction: Bad Genes or Bad Luck?* Norton, New York.
- , 1994. The role of extinction in evolution. *Proceedings of the National
726 Academy of Sciences* 91:6758–6763.
- Rosindell, J., L. J. Harmon, and R. S. Etienne, 2015. Unifying ecology and
728 macroevolution with individual-based theory. *Ecology letters* 18:472–82.
- Royle, J. A. and R. M. Dorazio, 2008. *Hierarchical modeling and inference in
730 ecology: the analysis of data from populations, metapopulations and
communities.* Elsevier, London.
- 732 Sepkoski, J. J., 1975. Stratigraphic biases in the analysis of taxonomic
survivorship. *Paleobiology* 1:343–355.
- 734 ———, 1981. A factor analytic description of the Phanerozoic marine fossil
record. *Paleobiology* 7:36–53.
- 736 Sheehan, P., 2001. The late Ordovician mass extinction. *Annual Review of
Earth and Planetary Sciences* 29:331–364.
- 738 Simpson, C., 2006. Levels of selection and large-scale morphological trends.
Ph.D. thesis, University of Chicago.

- 740 Simpson, C. and P. G. Harnik, 2009. Assessing the role of abundance in marine
bivalve extinction over the post-Paleozoic. *Paleobiology* 35:631–647.
- 742 Simpson, G. G., 1944. *Tempo and Mode in Evolution*. Columbia University
Press, New York.
- 744 ———, 1953. *The Major Features of Evolution*. Columbia University Press,
New York.
- 746 Smits, P. D., 2015. Expected time-invariant effects of biological traits on
mammal species duration. *Proceedings of the National Academy of Sciences*
748 112:13015–13020.
- Solow, A. and W. Smith, 1997. On fossil preservation and the stratigraphic
750 ranges of taxa. *Paleobiology* 23:271–277.
- 752 Stadler, T., 2011. Mammalian phylogeny reveals recent diversification rate
shifts. *Proceedings of the National Academy of Sciences* Pp. 1–6.
- 754 ———, 2013. Recovering speciation and extinction dynamics based on
phylogenies. *Journal of evolutionary biology* 26:1203–1219.
- 756 Stadler, T. and F. Bokma, 2013. Estimating speciation and extinction rates for
phylogenies of higher taxa. *Systematic biology* 62:220–30.
- 758 Stan Development Team, 2014. Stan: A c++ library for probability and
sampling, version 2.5.0. URL <http://mc-stan.org/>.
- 760 Stanley, S. M., 1975. A theory of evolution above the species level. *Proc. Natl.
Acad. Sci. USA* 72:646.
- 762 ———, 1979. *Macroevolution: pattern and process*. W. H. Freeman, San
Francisco.
- Team, S. D., 2016. *Stan Modeling Language Users Guide and Reference Manual*.

- 764 Tomiya, S., 2013. Body Size and Extinction Risk in Terrestrial Mammals Above
the Species Level. *The American Naturalist* 182:196–214.
- 766 Van Valen, L., 1973. A new evolutionary law. *Evolutionary Theory* 1:1–30.
———, 1979. Taxonomic survivorship curves. *Evolutionary Theory* 4:129–142.
- 768 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
770 latitudinal selectivity gradient at the end-Cretaceous mass extinction.
Scientific Reports 3:1790.
- 772 Wagner, P. J. and J. D. Marcot, 2013. Modelling distributions of fossil sampling
rates over time, space and taxa: assessment and implications for
774 macroevolutionary studies. *Methods in Ecology and Evolution* 4:703–713.
- Wang, S. C., 2003. On the continuity of background and mass extinction.
776 *Paleobiology* 29:455–467.
- Wang, S. C. and C. R. Marshall, 2004. Improved confidence intervals for
778 estimating the position of a mass extinction boundary. *Paleobiology* 30:5.
- Williams, C., C. H. C. Brunton, and S. J. Carlson, 2007. Treatise on
780 invertebrate paleontology. Part H, Brachiopoda. Geological Society of
America, Boulder, Colorado.

Table 1: Group-level estimates of the effects of biological traits on brachiopod generic survival, along with estimates of the effect of sampling δ and the estimate of the Weibull shape parameter α . μ values are the location parameters of the effects, while τ values are the scale terms describing the variation between cohorts. The mean, standard deviation (SD), 10th, 50th, and 90th quantiles of the posterior are presented.

parameter	interpretation	mean	SD	10%	50%	90%
μ_i	intercept	-3.05	0.20	-3.30	-3.05	-2.80
μ_r	geographic range	-0.98	0.16	-1.18	-0.98	-0.79
μ_v	environmental preference	-0.76	0.19	-0.99	-0.76	-0.52
μ_{v^2}	environmental preference ²	3.15	0.36	2.69	3.15	3.62
μ_m	body size	-0.01	0.13	-0.17	-0.01	0.15
τ_i	SD(intercept)	0.51	0.11	0.38	0.50	0.65
τ_r	SD(geographic range)	0.50	0.16	0.30	0.49	0.71
τ_v	SD(environmental preference)	0.84	0.17	0.63	0.82	1.05
τ_{v^2}	SD(environmental preference ²)	1.51	0.36	1.08	1.48	1.97
τ_m	SD(body size)	0.47	0.13	0.32	0.46	0.64
δ	sampling	0.90	0.15	0.71	0.90	1.09
α	“time”	1.36	0.04	1.30	1.36	1.42

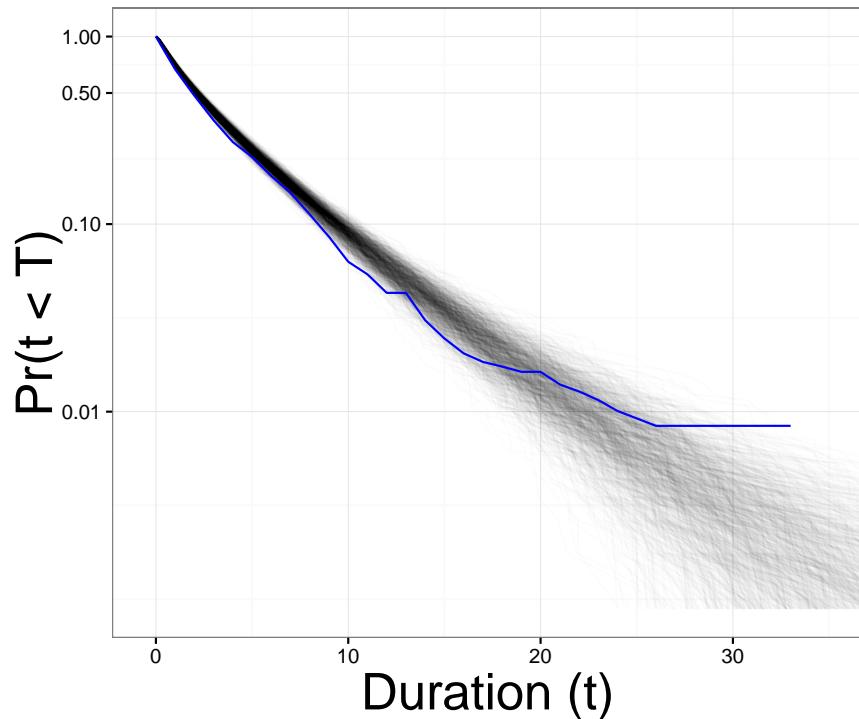


Figure 1: Comparison of the empirical estimate of $S(t)$ (highlighted) versus estimates from 1000 posterior predictive data sets (black). $S(t)$ corresponds to the probability that the age of a genus t is less than the genus' ultimate duration T . The vertical axis is log10 transformed.

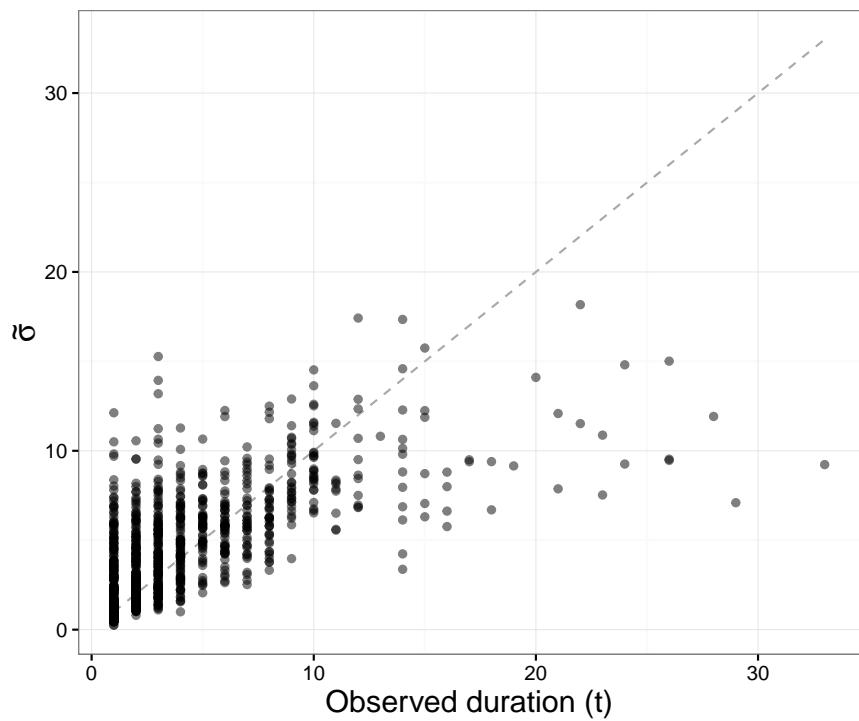


Figure 2: Comparison of all observed genus durations in number of geological stages to the average posterior predictive estimates of $\log(\sigma)$. The dashed, diagonal line corresponds to $x = y$.

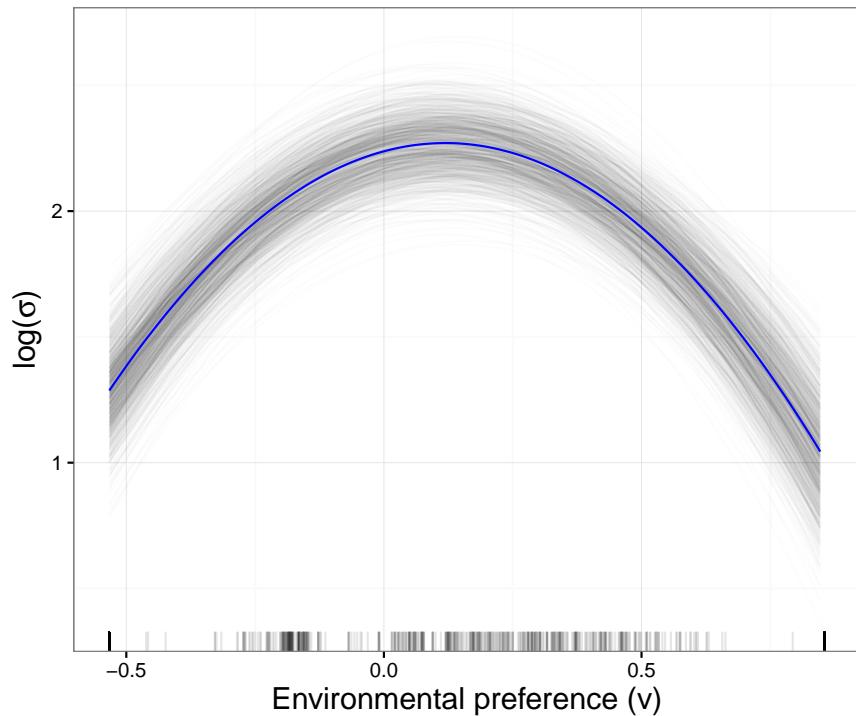


Figure 3: The overall expected relationship between environmental affinity v_i and a $\log(\sigma)$ when $r = 0$ and $m = 0$. Each grey line corresponds to a single draw from the posterior predictive distribution, while the highlighted line corresponds to the median of the posterior predictive distribution. 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.

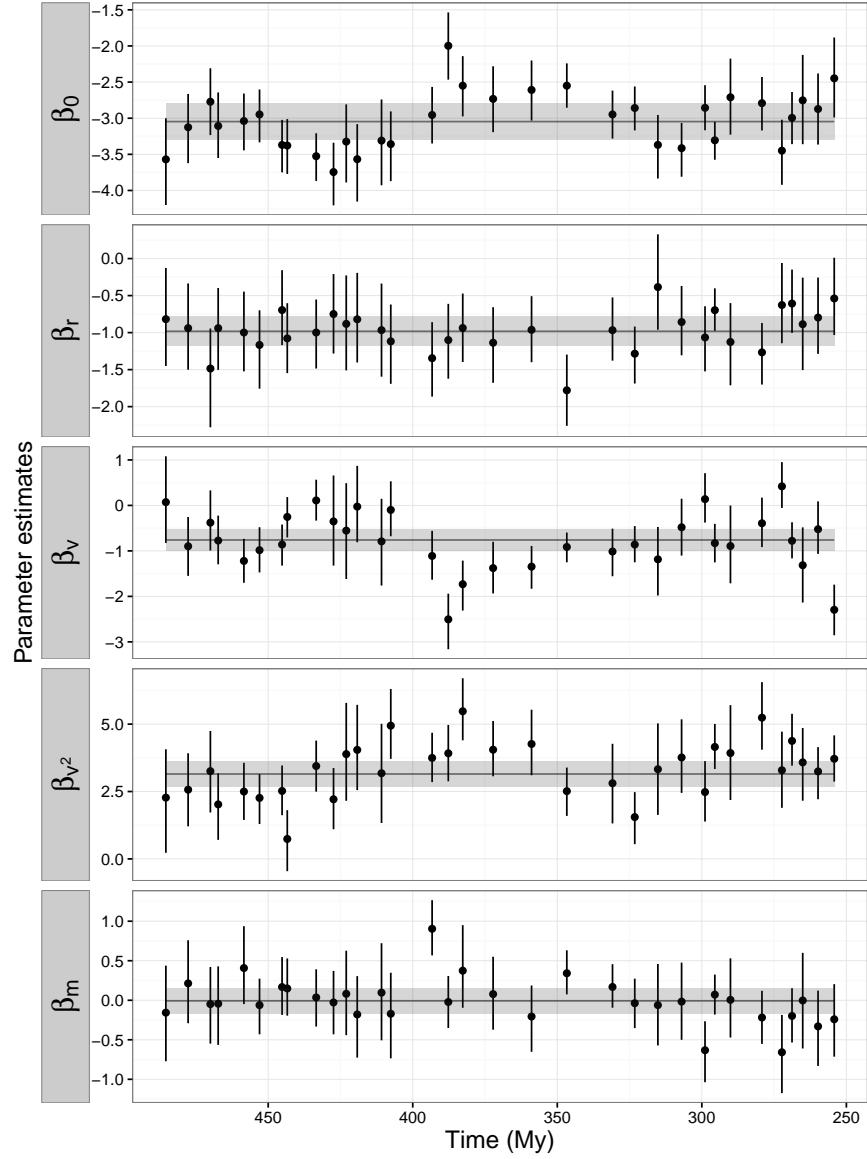


Figure 4: Comparison of cohort-specific estimates of β_0 , the effect of geographic range on extinction risk β_r , the effect of environmental preference β_v and β_{v^2} , and body size β_m . Points correspond to the median of the cohort-specific estimate, along with 80% credible intervals. Points are 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).

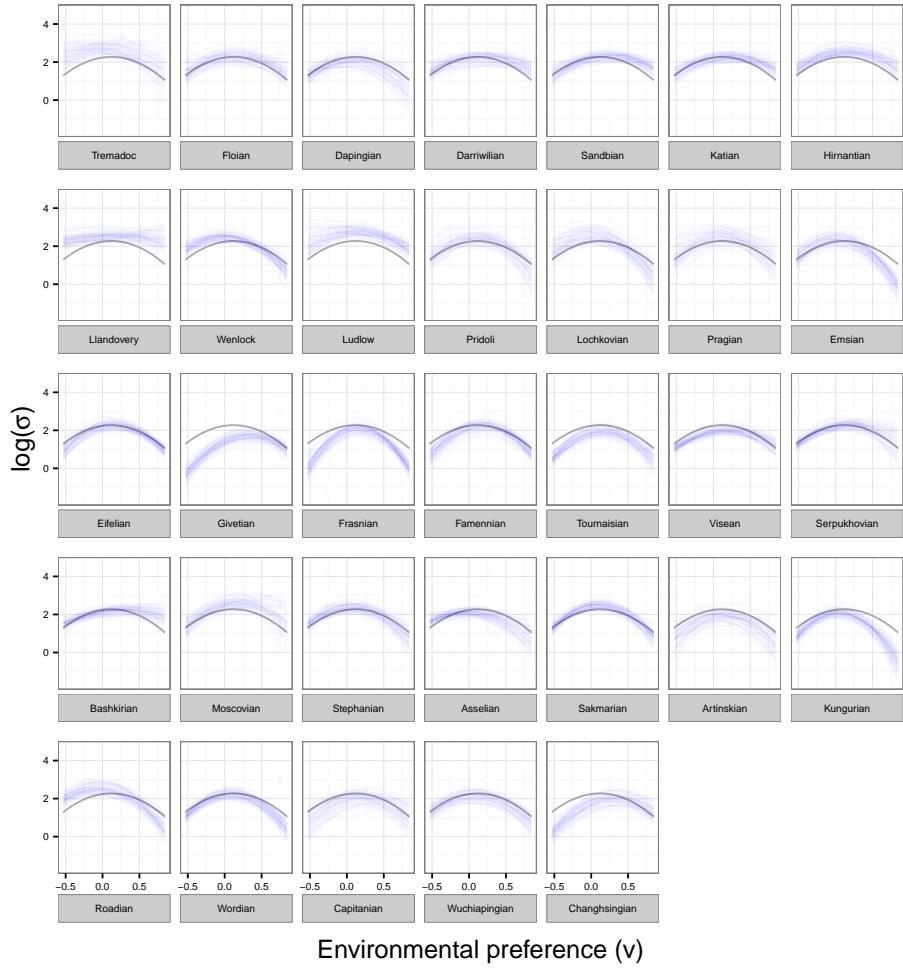


Figure 5: Comparison of origination cohort-specific (posterior predictive) estimates of the effect of environmental preference on $\log(\sigma)$ to the mean overall estimate of the effect of environmental preference. Cohort-specific estimates are from 100 posterior predictive simulations across the range 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. Facet names correspond to the name of the stage in which that cohort originated.

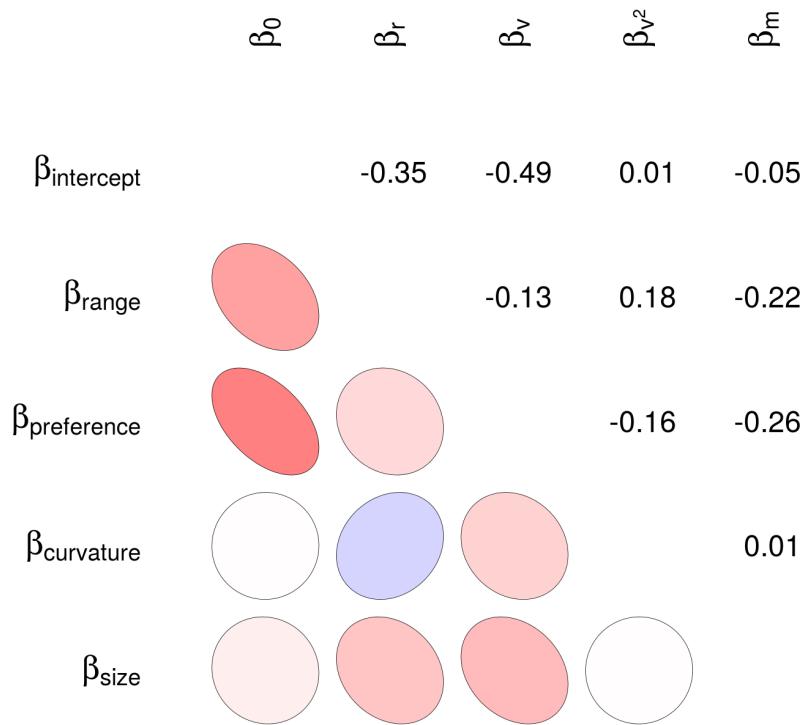


Figure 6: Mixed graphical and numerical representation of the correlation matrix Ω of variation in cohort-specific covariate estimates. These correlations are between the estimates of the cohort-level effects of covariates, along with intercept/baseline extinction risk. The median estimates of the correlations are presented numerically (upper-triangle) and as idealized ellipses representing that much correlation (lower-triangle).