

**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

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

**Keywords:** extinction, brachiopods, Paleozoic, epicontinental seas, Bayesian

8 **Word count:**

**Table count:** 2

10 **Figure count:** 6

**Data archival location:** If accepted, all data and code necessary to duplicate  
12 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. 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.

## <sup>38</sup> 1 Introduction

How do biological traits affect extinction risk? Extinction is one half of the  
<sup>40</sup> diversification process (Raup, 1994, Stanley, 1975, 1979), second only to  
speciation or origination; it can also be the ultimate manifestation of selection  
<sup>42</sup> 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,  
<sup>44</sup> 2010, Raup, 1994, Stanley, 1975).

While estimation of both trait-dependent speciation and extinction rates from  
<sup>46</sup> phylogenies of extant taxa has grown dramatically (Fitzjohn, 2010, Goldberg  
et al., 2011, 2005, Maddison et al., 2007, Rabosky et al., 2013, Stadler, 2011,  
<sup>48</sup> 2013, Stadler and Bokma, 2013), there are two major ways to estimate  
trait-dependent extinction: analysis of phylogenies, and analysis of the fossil  
<sup>50</sup> record. These two directions, phylogenetic comparative and paleobiological, are  
complementary and intertwined in the field of macroevolution (Hunt and  
Rabosky, 2014, Jablonski, 2008, Rabosky and McCune, 2010). In the case of  
extinction, analysis of the fossil record has the distinct advantage over  
<sup>54</sup> phylogenies of only extant taxa because extinction is observable; this means  
that extinction rate is possible to estimate (Liow et al., 2010, Quental and  
Marshall, 2009, Rabosky, 2010). The approach used here is thus complementary  
to the analysis of trait-dependent extinction based on a phylogeny.

<sup>58</sup> Jablonski (1986) observed that for bivalves at the end Cretaceous mass  
extinction event, the effects of some biological traits on taxonomic survival  
<sup>60</sup> decreased. However, this pattern was not the case for the effect of geographic  
range on survival (Jablonski, 1986, Payne and Finnegan, 2007). There are  
<sup>62</sup> multiple possible macroevolutionary mechanisms which may underlie this  
pattern: the effect of geographic range on survival remains constant and those of

64 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  
66 traits decrease potentially by different degrees.

68 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  
70 continuum (Payne and Finnegan, 2007, Simpson and Harnik, 2009, Wang, 2003).  
Additionally, in order to test the proposed macroevolutionary mechanism behind  
72 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.

74 Here I model brachiopod taxon durations because trait based differences in  
extinction risk should manifest as differences in taxon durations. Brachiopods  
76 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  
78 focus on the brachiopod record from the post-Cambrian Paleozoic, from the  
start of the Ordovician (approximately 485 My) through the end Permian  
80 (approximately 252 My) as this represents the time of greatest global  
brachiopod diversity (Alroy, 2010).

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

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

## **92    1.1 Factors affecting brachiopod survival**

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

110 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  
112 typically less than 100m. In contrast, open-ocean coastline environments have  
much greater variance in depth, do not cover the continental craton, and can  
114 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

<sup>116</sup> great risk during periods of low sea levels, such as during glacial periods, when  
<sup>118</sup> epicontinental seas are drained. During the Paleozoic (approximately 541–252  
My), epicontinental seas were widely spread globally but declined over the  
<sup>120</sup> Mesozoic (approximately 252–66 My) and have nearly disappeared during the  
Cenozoic (approximately 66–0 My) as open-ocean coastlines became the  
dominant shallow-marine setting (Johnson, 1974, Miller and Foote, 2009, Peters,  
<sup>122</sup> 2008).

<sup>124</sup> 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  
<sup>126</sup> background extinction, however, they found no consistent difference between  
taxa favoring either environment. These two environment types represent the  
<sup>128</sup> primary identifiable environmental dichotomy observed in ancient marine  
systems (Miller and Foote, 2009, Peters, 2008, Sheehan, 2001). Given these  
<sup>130</sup> findings, I predict that as extinction risk increases, the extinction risk associated  
with open-ocean environments should generally increase.

<sup>132</sup> Because environmental preference is defined here as the continuum between  
occurring exclusively in open-ocean environments versus epicontinental  
<sup>134</sup> environments, intermediate values are considered “generalists” in the sense that  
they favor neither end member. A long-standing hypothesis is that generalists or  
<sup>136</sup> unspecialized taxa will have greater survival than specialists (Baumiller, 1993,  
Liow, 2004, 2007, Nürnberg and Aberhan, 2013, 2015, Simpson, 1944). Because  
<sup>138</sup> of this, the effect of environmental preference was modeled as a quadratic  
function where a concave down relationship between preference and expected  
<sup>140</sup> duration indicates that generalists are favored over specialists end-members.

<sup>142</sup> 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  
144 (Payne et al., 2014). Harnik et al. (2014) analyzed the effect of body size  
selectivity in Devonian brachiopods in both a phylogenetic and  
146 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  
148 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  
150 results, I expect that if body size has any effect on brachiopod taxonomic  
survival it is very small.

152 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,  
154 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  
156 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  
158 case, for sampling to be a confounding factor there must be consistent  
relationship between the quality of sampling of a taxon and its apparent  
160 duration (e.g. greater sampling, longer duration). If there is no relationship  
between sampling and duration then interpretation can be made clearly; while  
162 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  
164 major determinant of the taxon's apparent duration.

## 2 Materials and Methods

### **166 2.1 Fossil occurrence information**

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

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

accuracy in the data analyzed here.

192 Genus duration was calculated as the number of geologic stages from first  
appearance to last appearance, inclusive. Durations were based on geologic  
194 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  
196 instead fossils are assigned to a geological interval which represents some  
temporal range. In this analysis, stages are effectively irreducible temporal  
198 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  
200 censored at the Changhsingian; genera with a duration of only one stage were  
left censored (Klein and Moeschberger, 2003). How the likelihood of censored  
202 observations is calculated is detailed below.

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

Geographic range was calculated using an occupancy approach corrected for  
206 incomplete sampling. First, all occurrences were projected onto an equal-area  
cylindrical map projection. Each occurrence was then assigned to one of the  
208 cells from a  $70 \times 34$  regular raster grid placed on the map. Each grid cell  
represents approximately 250,000 km<sup>2</sup>. The map projection and regular lattice  
210 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  
212 occupied grid cells was calculated. Then, for each temporal bin, the relative  
occurrence probability of the observed taxa was calculated using the JADE  
214 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  
216 estimated from a truncated Binomial distribution. This correction is critical  
when comparing occupancies from different times with different geographic

<sup>218</sup> sampling. Finally, for each genus, the mean relative occurrence probability was  
<sup>219</sup> calculated as the average of that genus' occurrence probabilities for all stages it  
<sup>220</sup> 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  $\theta'_i$  as estimated from all other taxa occurring  
at the same time ( $e'_i/E'_i$ ). This measure of environmental preference is  
expressed.

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

<sup>221</sup> where  $v$  is the percent of the distribution defined in equation 2.1 less than or  
<sup>222</sup> equal to  $\theta_i$ . The Beta distribution is used here because it is a continuous  
distribution bounded at 0 and 1, which is idea for modeling percentages.

<sup>224</sup> Body size, measured as shell length, was sourced directly from Payne et al.  
(2014). These measurements were made from brachiopod taxa figured in the  
<sup>226</sup> *Treatise on Invertebrate Paleontology* (Williams et al., 2007).

The sampling record for individual taxa was calculated using the standard gap  
<sup>228</sup> 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  
<sup>230</sup> 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  
<sup>232</sup> sampled. The minimum duration for which a gap statistic can be calculated is  
three stages. Sampling was only included as a covariate in the truncated model,  
<sup>234</sup> described later.

Prior to analysis, geographic range was logit transformed and body size was  
<sup>236</sup> natural-log transformed; both of these transformations make these variables

defined for the entire real line. All covariates were then standardized by  
238 subtracting the mean from all values and dividing by twice its standard  
deviation, which follows Gelman and Hill (2007). This standardization means  
240 that all regression coefficients are comparable as the expected change per 1-unit  
change in the rescaled covariates.  $D$  is defined as the total number of covariate  
242 effects included in the model plus one for the intercept term.

## 2.2 Analytical approach

244 Hierarchical modelling is a statistical approach which explicitly takes into  
account the structure of the observed data in order to model both the within  
246 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  
248 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,  
250 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.  
252 covariate effects) (Gelman et al., 2013). The subsequent estimates are partially  
pooled together, where parameters from groups with large samples or effects  
254 remain large while those of groups with small samples or effects are pulled  
towards the overall group mean. All covariate effects (regression coefficients), as  
256 well as the intercept term (baseline extinction risk), were allowed to vary by  
group (origination cohort). The covariance between covariate effects was also  
258 modeled.

Genus durations were assumed to follow a Weibull distribution which allows for  
260 age-dependent extinction (Klein and Moeschberger, 2003):  $y \sim \text{Weibull}(\alpha, \sigma)$ .  
The Weibull distribution has two parameters: scale  $\sigma$ , and shape  $\alpha$ . When

<sup>262</sup>  $\alpha = 1$ ,  $\sigma$  is equal to the expected duration of any taxon.  $\alpha$  is a measure of the  
<sup>263</sup> effect of age on extinction risk where values greater than 1 indicate that  
<sup>264</sup> extinction risk increases with age, and values less than 1 indicate that  
<sup>265</sup> extinction risk decreases with age. Note that the Weibull distribution is  
<sup>266</sup> equivalent to the exponential distribution when  $\alpha = 1$ .

In the case of the right- and left-censored observations mentioned above, the  
<sup>268</sup> probability of those observations has a different calculation (Klein and  
<sup>269</sup> Moeschberger, 2003). For right-censored observations, the likelihood is  
<sup>270</sup> calculated  $p(y|\theta) = 1 - F(y) = S(y)$  where  $F(y)$  is the cumulative distribution  
<sup>271</sup> function. In contrast, the likelihood of a left-censored observation is calculated  
<sup>272</sup>  $p(y|\theta) = F(y)$ .

The scale parameter  $\sigma$  was modeled as a regression following Kleinbaum and  
<sup>274</sup> Klein (2005) with both varying intercept and varying slopes; this is expressed

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

where  $i$  indexes across all observations,  $j[i]$  is the cohort membership of the  $i$ th  
<sup>276</sup> observation,  $X$  is a  $N \times D$  matrix of covariates along with a column of 1's for  
<sup>277</sup> the intercept term, and  $B$  is a  $J \times D$  matrix of cohort-specific regression  
<sup>278</sup> coefficients.

Each of the rows of matrix  $B$  are modeled as realizations from a multivariate  
<sup>280</sup> normal distribution with length  $D$  location vector  $\mu$  and  $J \times J$  covariance  
<sup>281</sup> matrix  $\Sigma$ :  $B_j \sim \text{MVN}(\mu, \Sigma)$ . The covariance matrix was then decomposed into  
<sup>282</sup> a length  $J$  vector of scales  $\tau$  and a  $J \times J$  correlation matrix  $\Omega$ , defined  
 $\Sigma = \text{diag}(\tau)\Omega\text{diag}(\tau)$  where “diag” indicates a diagonal matrix.

The elements of  $\mu$  were given independent normally distributed priors. The  
<sup>283</sup> 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  $\Omega$  was given an almost uniform LKJ distributed prior following (Team, 2016). These priors are defined

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

<sup>284</sup> The log of the shape parameter  $\alpha$  was given a weakly informative prior  
 $\log(\alpha) \sim \mathcal{N}(0, 1)$  centered at  $\alpha = 1$ , which corresponds to the Law of Constant  
<sup>286</sup> Extinction (Van Valen, 1973).

The above model is for all taxa and does not include sampling as a covariate. In  
<sup>288</sup> order to determine if sampling is acting as a confounding factor in this analysis,  
an additional model was developed because sampling was only estimated for  
<sup>290</sup> taxa with a duration of three or more which creates a left-truncated distribution  
of durations (Klein and Moeschberger, 2003). The sampling statement and  
<sup>292</sup> log-probability for a left-truncated Weibull distribution, truncated at time  $Y$   
(e.g. three), is

$$\begin{aligned}p(y|\theta) &= \frac{\text{Weibull}(y, \alpha, \sigma)}{1 - \text{Weibull}_{cdf}(Y, \alpha, \sigma)} \\ &= \frac{\text{Weibull}(y, \alpha, \sigma)}{\text{Weibull}_{ccdf}(Y, \alpha, \sigma)}\end{aligned}\tag{2}$$

$$\log(p(y|\theta)) = \log(\text{Weibull}(y, \alpha, \sigma)) - \log(\text{Weibull}_{ccdf}(Y, \alpha, \sigma)).$$

<sup>294</sup> Note that cdf stands for cumulative distribution function and ccdf is the  
complementary cumulative distribution function.

<sup>296</sup> The definition of  $\sigma$  (Eq. 1) is then updated so that  $X$ , the matrix of covariates,  
and  $B$ , the matrix of regression coefficients, now include an additional column  
<sup>298</sup> for the sampling estimates and the cohort-specific effects of sampling. This  
addition then modifies the dimensions of  $\mu$  and  $\Sigma$ ; the new group-level effect of  
<sup>300</sup>  $\mu_s$  is given a weakly informative prior:  $\mu_s \sim \mathcal{N}(0, 1)$ .

For this left-truncated model, I've excluded one observation that is  
<sup>302</sup> right-censored with a duration equal to the truncation time; the second line of  
equation 2 becomes  $p(y|\theta) = \text{Weibull}_{ccdf}(y, \alpha, \sigma)/\text{Weibull}_{ccdf}(Y, \alpha, \sigma)$  which  
<sup>304</sup> yields a log-probability of 0.

The joint posterior was approximated using a Markov-chain Monte Carlo  
<sup>306</sup> routine that is a variant of Hamiltonian Monte Carlo called the No-U-Turn  
Sampler (Hoffman and Gelman, 2014) as implemented in the probabilistic  
<sup>308</sup> programming language Stan (Stan Development Team, 2014). The posterior  
distribution was approximated from four parallel chains run for 10,000 steps  
<sup>310</sup> each, split half warm-up and half sampling and thinned to every 10th sample for  
a total of 4000 posterior samples. Chain convergence was assessed via the scale  
<sup>312</sup> reduction factor  $\hat{R}$  where values close to 1 ( $\hat{R} < 1.1$ ) indicate approximate  
convergence. Convergence means that the chains are approximately stationary  
<sup>314</sup> and the samples are well mixed (Gelman et al., 2013).

Model adequacy was evaluated using a couple of posterior predictive checks.  
<sup>316</sup> Posterior predictive checks are a means for understanding model fit or adequacy  
where the basic idea is that replicated data sets simulated from the fitted model  
<sup>318</sup> should be similar to the original data and systematic differences between the  
simulations and observations indicate weaknesses of the model fit (Gelman

<sup>320</sup> et al., 2013). For both approaches used here, 1000 posterior predictive datasets  
<sup>322</sup> were generated from 1000 unique draws from the posterior distribution of each  
parameter. The two posterior predictive checks used in this analysis are a  
<sup>324</sup> 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  
<sup>326</sup> 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  
<sup>328</sup> 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.

### 3 Results

<sup>330</sup> Comparison of the posterior predictive estimates of  $S(t)$  to the empirical  
estimate reveal few obvious biases except for the case of estimates from the far  
<sup>332</sup> right tail of observed durations (Fig. 1). This result is reinforced by the  
additional posterior predictive comparison where most estimates are not  
<sup>334</sup> systematically biased except for a consistent under-estimate of  $\log(\sigma)$  for older  
taxa (Fig. 2). The results of both posterior predictive checks indicate that, for  
<sup>336</sup> the majority of observations, model fit is generally not biased.

The cohort-level estimate of the effect of geographic range size indicates that as  
<sup>338</sup> a taxon's geographic range increases, that taxon's duration is expected to  
increase (Table 1). Given the estimates of  $\mu_r$  and  $\tau_r$ , there is a less than 0.008%  
<sup>340</sup> ( $\pm 0.05$  SD) probability that this relationships would be reversed  
( $\Pr(\mathcal{N}(\mu_r, \tau_r) > 0)$ ). The between-cohort variance  $\tau_r$  is the lowest of all the  
<sup>342</sup> regression coefficients (Table 1).

Body size is estimated to have no effect on taxon duration, with the estimate  
<sup>344</sup> being nearly 0 (Table 1). The variance between the cohort-specific estimates of

the effect of body size  $\tau_m$  is estimated to be greater than the variance of  
346 between-cohort estimates of the effect of geographic range size  $\tau_r$ .

The group-level estimate of the effect of environmental preference is estimated  
348 from both  $\mu_v$  and  $\mu_{v^2}$ .

The estimate of  $\mu_v$  indicates that epicontinental favoring taxa are expected to  
350 have a greater duration than open-ocean favoring taxa (Table 1). Additionally,  
given the estimate of between-cohort variance  $\tau_v$ , there is approximately 19%  
352 ( $\pm 8$  SD) probability that, for any given cohort, taxa favoring open-ocean  
environments would have a greater expected duration than taxa favoring  
354 epicontinental environments ( $\Pr(\mathcal{N}(\mu_v, \tau_v) > 0)$ ).

The estimate of  $\mu_{v^2}$  indicates that the overall relationship between  
356 environmental preference and  $\log(\sigma)$  is concave down (Fig. 3), with only a 1.9%  
( $\pm 2.4$  SD) probability that any given cohort is convex up ( $\Pr(\mathcal{N}(\mu_{v^2}, \tau_{v^2}) < 0)$ ).

358 The cohort-specific estimates of all the regression coefficients demonstrate a lot  
of between cohort variance, with no obvious trends. As indicated in Table 1 and  
360 detectable visually (Fig. 4), the between-cohort estimates for  $\beta_0$ ,  $\beta_r$ , and  $\beta_m$  all  
have much lower variance than the between-cohort estimates of both  $\beta_v$  and  $\beta_{v^2}$ .

362 While most cohort-specific estimates are very similar to the overall cohort-level  
estimate, there are a few notable excursions away from the overall mean (Fig.

364 4). There are simultaneous excursions in both  $\beta_0$  and  $\beta_v$  for cohorts originating  
in the Givetian (387-382 My) and Frasnian (382-372 My) stages; both of which  
366 directly precede the late Devonian mass extinction event at the

Frasnian/Famennian boundary. These cohorts are marked by both a high

368 extinction intensity and an increase in expected duration for taxa favoring  
epicontinental environments over open-ocean ones; this is consistent with the  
370 results of Miller and Foote (2009).

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

380 The cohort-specific relationships between environmental preference and  $\log(\sigma)$   
were calculated from the estimates of  $\beta_0$ ,  $\beta_v$ , and  $\beta_{v^2}$  (Fig. 5) and reflect how  
382 these three parameters act in concert and not just individually (Fig. 4). Beyond  
results already discussed above in the context of the parameters individually, it  
384 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  
386 appearance due to a high estimate  $\beta_{v^2}$  (Fig. 4). This cohort has the biggest  
difference in extinction risk between environmental generalists and specialists.  
388 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  
390 cohorts include those originating during Tremadocian (484-477 My), Hirnantian  
(445-443 My), Llandovery (443-433 My), and Ludlow (427-423 My). Except for  
392 the Tremadocian cohort, most of these cohorts originate during the Silurian  
through the Early Devonian range identified earlier as having lower expected  
394 extinction intensity than what is expected from the group-level estimate.

The correlations of the cohort-specific estimates of the regression coefficients are  
396 estimated as the off-diagonal elements of the correlation matrix  $\Omega$ . Only two of  
the elements of  $\Omega$  are distinguishable from 0: the correlation of  $\beta_0$  (extinction

<sup>398</sup> intensity) with both  $\beta_r$  and  $\beta_v$  (Fig. 6).

There is an approximate 86% probability that the cohort-specific estimates of  
<sup>400</sup> baseline extinction intensity  $\beta_0$  and the effect of geographic range  $\beta_r$  are  
negatively correlated; this means that for cohorts experiencing a lower  
<sup>402</sup> extinction intensity ( $\beta_0$  decreases), the magnitude of the effect of geographic  
range is expected to decrease as well, and *vice versa*; this is in contrast to the  
<sup>404</sup> observation made by Jablonski (1986) with regards to late Cretaceous bivalves.

Similarly, there is an approximate 99.9% probability that the cohort-specific  
<sup>406</sup> estimates of  $\beta_0$  and  $\beta_v$  are negatively correlated; this means that as extinction  
intensity increases it is expected that epicontinental taxa become more favored  
<sup>408</sup> over open-ocean environments (i.e. as  $\beta_0$  increases,  $\beta_v$  decreases). Additionally,  
there is only a 19% ( $\pm 8.1$  SD) probability that, for any given cohort,  
<sup>410</sup> open-ocean environments will be favored ( $\Pr(|\mu_v| - 2\tau_v > 0)$ ).

There is only an approximate 68% probability that  $\beta_r$  and  $\beta_v$  are positively  
<sup>412</sup> correlated. This lack of cross-correlation may be due in part to the much higher  
between-cohort variance of the effect of environmental preference  $\tau_v$  than the  
<sup>414</sup> very small between-cohort variance in the effect of geographic range  $\tau_r$  (Table  
1); the effect of geographic range might simply not vary enough relative to the  
<sup>416</sup> much noisier environmental preference.

## 4 Discussion

<sup>418</sup> 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  
<sup>420</sup> affect extinction risk was geographic range while traits that had previously been  
beneficial had no effect (Jablonski, 1986). This observation raises two linked

<sup>422</sup> questions: how does the effect of geographic range change with changing  
<sup>423</sup> extinction intensity, and how does the effect of other biological traits change  
<sup>424</sup> with changing extinction intensity?

<sup>426</sup> I find that as intensity increases ( $\beta_0$  decreases), the magnitude of the effect of  
geographic range increase. I also find that as intensity increases, the effect of  
favoring epicontinental environments of open-ocean environments is expected to  
<sup>428</sup> 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  
<sup>430</sup> environmental preference. However, the between-cohort variance in effect of  
geographic range is much less then the between-cohort variance of the effect of  
<sup>432</sup> environmental preference. As stated earlier, this may underlie the lack of  
correlation between these two effects.

<sup>434</sup> Additionally, the lower between-cohort variance in the effect of geographic range  
versus that higher between-cohort variance implies that for cohorts with a  
<sup>436</sup> 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  
<sup>438</sup> than the difference between the effect of environmental preference and the  
group-level effect of environmental preference.

<sup>440</sup> I find consistent support for the “survival of the unspecialized,” with respect to  
epicontinental versus open-ocean environmental preference, as a time-invariant  
<sup>442</sup> generalization of brachiopod survival; taxa with intermediate environmental  
preferences are expected to have lower extinction risk than taxa specializing in  
<sup>444</sup> either epicontinental or open-ocean environments (Fig. 3), though the curvature  
of the relationship varies from rather shallow to very peaked (Fig. 5). However,  
<sup>446</sup> this relationship is not symmetric about 0, as taxa favoring epicontinental  
environments are expected to have a greater duration than taxa favoring  
<sup>448</sup> open-ocean environments. This description of environment only describes one

major aspect of a taxon's environmental context, with factors such as  
450 bathymetry and temperature being further descriptors of a taxon's adaptive  
zone (Harnik, 2011, Harnik et al., 2012, Heim and Peters, 2011, Nürnberg and  
452 Aberhan, 2013); inclusion of these factors in future analyses would potentially  
improve our understanding of the "survival of the unspecialized" hypothesis  
454 (Simpson, 1944).

Hopkins et al. (2014), in their analysis of niche conservatism and substrate  
456 lithological preference in marine invertebrates, found that brachiopods were  
among the least "conservative" groups; taxa were found to easily change  
458 substrate preference on short time scales. While substrate preference is not the  
same as environmental preference (as defined here), a question does arise: do  
460 "generalists" have a greater duration because they are "true" environmental  
generalists from origination (e.g. can inhabit both environments) or have taxa  
462 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  
464 analysis here assumes that traits are constant over the duration of the taxon;  
future work should explore how environmental preference changes over lineage  
466 duration in relation to environmental availability.

An alternative approach for specifically modeling survival that can take into  
468 account imperfect observation than the method used here is the  
Cormack-Jolly-Seber (CJS) model (Liow et al., 2008, Liow and Nichols, 2010,  
470 Royle and Dorazio, 2008, Tomiya, 2013). This model is a type of hidden Markov  
model with an absorbing state (i.e. extinction). In this model, survival is defined  
472 as the probability of surviving from time  $t$  to time  $t + 1$ . Additionally, the effect  
of preservation and sighting is estimated as probability of observing a taxon  
474 that is present; this can extend the duration of a taxon beyond its last  
occurrence. This approach is a fundamentally different from the method used in

<sup>476</sup> the above analysis; for example, instead of modeling persistence from time  $t$  to  
<sup>477</sup> time  $t + 1$ , I model the distribution of taxon durations (lifetime) which then  
<sup>478</sup> decomposes into a function describing survival ( $S(t)$ ) and a function describing  
<sup>479</sup> instantaneous extinction rate with taxon age (Klein and Moeschberger, 2003,  
<sup>480</sup> Royle and Dorazio, 2008).

<sup>482</sup> The use of genera as the unit of the study and how to exactly interpret the  
<sup>483</sup> effects of the biological traits is an important question. For example, if any of  
<sup>484</sup> the traits analyzed here are associated with increases in speciation rates, this  
<sup>485</sup> might increase the duration of genera through self-renewal (Raup, 1991, 1994),  
<sup>486</sup> which would be an example of the difference in biological pattern between  
<sup>487</sup> species and genera (Jablonski, 1987, 2007, 2008). This could lead to a trait  
<sup>488</sup> appearing to decrease generic level extinction risk by that trait increasing  
species level origination rate instead of decreasing species level extinction risk.

<sup>490</sup> The model used here could be improved through either increasing the number of  
<sup>491</sup> analyzed traits, expanding the hierarchical structure of the model to include  
<sup>492</sup> other major taxonomic groups of interest, and the inclusion of explicit  
<sup>493</sup> phylogenetic relationships between the taxa in the model as an additional  
<sup>494</sup> hierarchical effect. An example trait that may be of particular interest is the  
<sup>495</sup> affixing strategy or method of interaction with the substrate of the taxon, which  
<sup>496</sup> has been found to be related to brachiopod survival where, for cosmopolitan  
taxa, taxa that are attached to the substrate are expected to have a greater  
duration than those that are not (Alexander, 1977).

<sup>498</sup> It is theoretically possible to expand this model to allow for comparisons both  
<sup>499</sup> within and between major taxonomic groups which would better constrain the  
<sup>500</sup> brachiopod estimates while also allowing for estimation of similarities and  
differences in cross-taxonomic patterns. The major issue surrounding this  
<sup>502</sup> particular expansion involves finding a similarly well sampled taxonomic group

that is present during the Paleozoic. Example groups include Crinoidea,  
504 Ostracoda, and other members of the “Paleozoic fauna” (Sepkoski, 1981). Traits like environmental preference or geographic range (Hunt et al., 2005,  
506 Jablonski, 1987) are most likely heritable. Without phylogenetic context, this analysis assumes that differences in extinction risk between taxa are independent  
508 of the shared evolutionary history of those taxa (Felsenstein, 1985). In contrast, the origination cohorts only capture shared temporal context. For example, if  
510 taxon duration is phylogenetically heritable, then closely related taxa may have more similar durations as well as more similar biological traits. Without taking  
512 into account phylogenetic similarity the effects of these biological traits would be inflated solely due to inheritance. The inclusion of phylogenetic context as an  
514 additional individual-level hierarchical effect independent of origination cohort would allow for determining how much of the observed variability is due to  
516 shared evolutionary history versus shared temporal context versus actual differences associated with biological traits (Harnik et al., 2014, Smits, 2015).  
518 In summary, patterns of Paleozoic brachiopod survival were analyzed using a fully Bayesian hierarchical survival modelling approach while also eschewing the  
520 traditional separation between background and mass extinction. I find that cohort extinction intensity is negatively correlated with both the cohort-specific  
522 effects of geographic range and environmental preference. These results imply that as extinction intensity increases ( $\beta_0$ ) increases, it is expected that both  
524 effects will increase in magnitude. However, the change in effect of environmental preference is expected to be greater than the change in the effect  
526 of geographic range. Additionally, I find support for greater survival in environmental generalists over specialists in all origination cohorts analyzed; this is consistent with the long standing “survival of the unspecialized” hypothesis (Baumiller, 1993, Liow, 2004, 2007, Nürnberg and Aberhan, 2013,

530 2015, Simpson, 1944, 1953, Smits, 2015). I also find that taxa favoring  
epicontinental environments are expected to have a slightly greater duration  
532 than open-ocean favoring taxa. The results of this analysis support the  
conclusion that for Paleozoic brachiopods, as extinction intensity increases  
534 overall extinction selectivity is expected to increase as well.

## Acknowledgements

536 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  
538 D. Bapst, N. Pierrehumbert and M. Villarosa Garcia for additional comments.  
An additional thank you to A. Miller for the epicontinental versus open-ocean  
540 assignments. Finally, thank you to the reviewers for their helpful comments that  
improved this manuscript. This entire study would not have been possible  
542 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,  
544 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.  
546 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  
548 conflicts of interest. This is Paleobiology Database publication XXX.

## References

- 550 Alexander, R. R., 1977. Generic longevity of articulate brachiopods in relation  
to the mode of stabilization on the substrate. *Palaeogeography,*  
552 *Palaeoclimatology, Palaeoecology* 21:209–226.
- Alroy, J., 2010. The Shifting Balance of Diversity Among Major Marine Animal  
554 Groups. *Science* 329:1191–1194.
- , 2014. A simple Bayesian method of inferring extinction. *Paleobiology*  
556 40:584–607.
- Baumiller, T. K., 1993. Survivorship analysis of Paleozoic Crinoidea: effect of  
558 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.  
560 Inouye, 2015. Unveiling the species-rank abundance distribution by  
generalizing the Good-Turing sample coverage theory. *Ecology* 96:1189–1201.
- 562 Cooper, W. S., 1984. Expected time to extinction and the concept of  
fundamental fitness. *Journal of Theoretical Biology* 107:603–629.
- 564 Crampton, J. S., R. A. Cooper, P. M. Sadler, and M. Foote, 2016.  
Greenhouseicehouse transition in the Late Ordovician marks a step change in  
566 extinction regime in the marine plankton. *Proceedings of the National  
Academy of Sciences P.* 201519092.
- 568 Edwards, D. and U. Fanning, 1985. Evolution and environment in the late  
Silurian–early Devonian: the rise of pteridophytes. *Philosophical Transactions  
570 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  
572 birth and death (in macroevolutionary birth-death models). *Biology letters*  
8:139–42.

- 574 Felsenstein, J., 1985. Phylogenies and the comparative method. *American  
Naturalist* 125:1–15.
- 576 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.  
578 *Proceedings of the National Academy of Sciences* 109:6829–6834.
- 580 Fitzjohn, R. G., 2010. Quantitative Traits and Diversification. *Systematic  
Biology* Pp. 1–15.
- Foote, M., 1988. Survivorship analysis of Cambrian and Ordovician Trilobites.  
582 *Paleobiology* 14:258–271.
- , 2000a. Origination and extinction components of taxonomic diversity:  
584 general problems. *Paleobiology* 26:74–102.
- , 2000b. Origination and extinction components of taxonomic diversity:  
586 Paleozoic and post-Paleozoic dynamics. *Paleobiology* 26:578–605.
- , 2014. Environmental controls on geographic range size in marine animal  
588 genera. *Paleobiology* 40:440–458.
- Foote, M. and A. I. Miller, 2013. Determinants of early survival in marine  
590 animal genera. *Paleobiology* 39:171–192.
- Foote, M. and D. Raup, 1996. Fossil preservation and the stratigraphic ranges  
592 of taxa. *Paleobiology* 22:121–140.
- Gelman, A., J. B. Carlin, H. S. Stern, D. B. Dunson, A. Vehtari, and D. B.  
594 Rubin, 2013. *Bayesian data analysis*. 3 ed. Chapman and Hall, Boca Raton,  
FL.
- 596 Gelman, A. and J. Hill, 2007. *Data Analysis using Regression and  
Multilevel/Hierarchical Models*. Cambridge University Press, New York, NY.

- 598 Goldberg, E. E., L. T. Lancaster, and R. H. Ree, 2011. Phylogenetic inference  
of reciprocal effects between geographic range evolution and diversification.
- 600 Systematic biology 60:451–65.
- Goldberg, E. E., K. Roy, R. Lande, and D. Jablonski, 2005. Diversity,  
602 endemism, and age distributions in macroevolutionary sources and sinks. The  
American naturalist 165:623–33.
- 604 Harnik, P. G., 2011. Direct and indirect effects of biological factors on  
extinction risk in fossil bivalves. Proceedings of the National Academy of  
606 Sciences of the United States of America 108:13594–9.
- Harnik, P. G., P. C. Fitzgerald, J. L. Payne, and S. J. Carlson, 2014.  
608 Phylogenetic signal in extinction selectivity in Devonian terebratulide  
brachiopods. Paleobiology 40:675–692.
- 610 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  
612 Society B: Biological Sciences 279:4969–4976.
- Heim, N. A. and S. E. Peters, 2011. Regional environmental breadth predicts  
614 geographic range and longevity in fossil marine genera. PloS one 6:e18946.
- Hijmans, R. J., 2015. raster: Geographic data analysis and modeling. URL  
616 <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  
618 natural selection on traits above the species level. Methods in Ecology and  
Evolution .
- 620 Hoffman, M. D. and A. Gelman, 2014. The No-U-Turn Sampler: Adaptively  
Setting Path Lengths in Hamiltonian Monte Carlo. Journal of Machine  
622 Learning Research 15:1351–1381.

- Hopkins, M. J., C. Simpson, and W. Kiessling, 2014. Differential niche dynamics  
624 among major marine invertebrate clades. *Ecology Letters* 17:314–323.
- Hunt, G. and D. L. Rabosky, 2014. Phenotypic Evolution in Fossil Species:  
626 Pattern and Process. *Annual Review of Earth and Planetary Sciences*  
42:421–441.
- Hunt, G., K. Roy, and D. Jablonski, 2005. Species-level heritability reaffirmed: a  
628 comment on "On the heritability of geographic range sizes". *American  
Naturalist* 166:129–135.
- Jablonski, D., 1986. Background and mass extinctions: the alternation of  
632 macroevolutionary regimes. *Science* 231:129–133.
- , 1987. Heritability at the species level: analysis of geographic ranges of  
634 cretaceous mollusks. *Science* 238:360–363.
- , 2007. Scale and hierarchy in macroevolution. *Palaeontology* 50:87–109.
- 636 ———, 2008. Species Selection: Theory and Data. *Annual Review of Ecology,  
Evolution, and Systematics* 39:501–524.
- Jablonski, D. and K. Roy, 2003. Geographical range and speciation in fossil and  
638 living molluscs. *Proceedings. Biological sciences / The Royal Society*  
640 270:401–6.
- Joachimski, M. M., S. Breisig, W. Buggisch, J. A. Talent, R. Mawson,  
642 M. Gereke, J. R. Morrow, J. Day, and K. Weddige, 2009. Devonian climate  
and reef evolution: Insights from oxygen isotopes in apatite. *Earth and  
644 Planetary Science Letters* 284:599–609.
- Johnson, J. G., 1974. Extinction of Perched Faunas. *Geology* 2:479–482.
- 646 Kiessling, W. and M. Aberhan, 2007. Environmental determinants of marine

- benthic biodiversity dynamics through Triassic–Jurassic time. *Paleobiology*  
648 33:414–434.
- Klein, J. P. and M. L. Moeschberger, 2003. Survival Analysis: Techniques for  
650 Censored and Truncated Data. 2nd ed. Springer, New York.
- Kleinbaum, D. G. and M. Klein, 2005. Survival analysis: a self-learning text. 2  
652 ed. Springer, New York, NY.
- Liow, L. H., 2004. A test of Simpson's "rule of the survival of the relatively  
654 unspecialized" using fossil crinoids. *The American Naturalist* 164:431–43.
- , 2007. Does versatility as measured by geographic range, bathymetric  
656 range and morphological variability contribute to taxon longevity? *Global  
Ecology and Biogeography* 16:117–128.
- 658 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  
660 mammals. *Proceedings of the National Academy of Sciences* 105:6097.
- Liow, L. H. and J. D. Nichols, 2010. Estimating rates and probabilities of  
662 origination and extinction using taxonomic occurrence data:  
Capture-mark-recapture (CMR) approaches. *Quantitative Methods in  
664 Paleobiology* Pp. 81–94.
- Liow, L. H., T. B. Quental, and C. R. Marshall, 2010. When can decreasing  
666 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  
668 character's effect on speciation and extinction. *Systematic Biology* 56:701.
- Miller, A. I. and M. Foote, 2009. Epicontinental seas versus open-ocean settings:  
670 the kinetics of mass extinction and origination. *Science* 326:1106–9.

- 672 Munnecke, A., M. Calner, D. A. T. Harper, and T. Servais, 2010. Ordovician  
and Silurian sea-water chemistry, sea level, and climate: A synopsis.
- 674 676 Palaeogeography, Palaeoclimatology, Palaeoecology 296:389–413.  
Nürnberg, S. and M. Aberhan, 2013. Habitat breadth and geographic range  
predict diversity dynamics in marine Mesozoic bivalves. Paleobiology  
39:360–372.
- 678 ———, 2015. Interdependence of specialization and biodiversity in Phanerozoic  
marine invertebrates. Nature communications 6:6602.
- 680 Palmer, M. E. and M. W. Feldman, 2012. Survivability is more fundamental  
than evolvability. PloS one 7:e38025.
- 682 Payne, J. L. and S. Finnegan, 2007. The effect of geographic range on  
extinction risk during background and mass extinction. Proceedings of the  
National Academy of Sciences 104:10506–11.
- 684 Payne, J. L., N. A. Heim, M. L. Knope, and C. R. McClain, 2014. Metabolic  
dominance of bivalves predates brachiopod diversity decline by more than 150  
million years. Proceedings of the Royal Society B 281:20133122.
- 686 Peters, S. E., 2008. Environmental determinants of extinction selectivity in the  
fossil record. Nature 454:626–9.
- 688 Quental, T. B. and C. R. Marshall, 2009. Extinction during evolutionary  
radiations: reconciling the fossil record with molecular phylogenies. Evolution;  
international journal of organic evolution 63:3158–67.
- 690 Rabosky, D. L., 2010. Extinction rates should not be estimated from molecular  
phylogenies. Evolution 64:1816–24.
- 692 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,  
698 and M. E. Alfaro, 2013. Rates of speciation and morphological evolution are  
correlated across the largest vertebrate radiation. *Nature Communications*  
700 4:1–8.
- Raup, D. M., 1975. Taxonomic survivorship curves and Van Valen's Law.  
702 *Paleobiology* 1:82–96.
- , 1978. Cohort Analysis of generic survivorship. *Paleobiology* 4:1–15.
- 704 ———, 1991. *Extinction: Bad Genes or Bad Luck?* Norton, New York.
- , 1994. The role of extinction in evolution. *Proceedings of the National  
706 Academy of Sciences* 91:6758–6763.
- Royle, J. A. and R. M. Dorazio, 2008. Hierarchical modeling and inference in  
708 ecology: the analysis of data from populations, metapopulations and  
communities. Elsevier, London.
- Sepkoski, J. J., 1981. A factor analytic description of the Phanerozoic marine  
710 fossil record. *Paleobiology* 7:36–53.
- Sheehan, P., 2001. The late Ordovician mass extinction. *Annual Review of  
712 Earth and Planetary Sciences* 29:331–364.
- 714 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  
716 bivalve extinction over the post-Paleozoic. *Paleobiology* 35:631–647.
- 718 Simpson, G. G., 1944. *Tempo and Mode in Evolution*. Columbia University  
Press, New York.

- 720        ———, 1953. *The Major Features of Evolution*. Columbia University Press,  
New York.
- 722        Smits, P. D., 2015. Expected time-invariant effects of biological traits on  
mammal species duration. *Proceedings of the National Academy of Sciences*  
724        112:13015–13020.
- 726        Solow, A. and W. Smith, 1997. On fossil preservation and the stratigraphic  
ranges of taxa. *Paleobiology* 23:271–277.
- 728        Stadler, T., 2011. Mammalian phylogeny reveals recent diversification rate  
shifts. *Proceedings of the National Academy of Sciences* Pp. 1–6.
- 730        ———, 2013. Recovering speciation and extinction dynamics based on  
phylogenies. *Journal of evolutionary biology* 26:1203–1219.
- 732        Stadler, T. and F. Bokma, 2013. Estimating speciation and extinction rates for  
phylogenies of higher taxa. *Systematic biology* 62:220–30.
- 734        Stan Development Team, 2014. Stan: A c++ library for probability and  
sampling, version 2.5.0. URL <http://mc-stan.org/>.
- 736        Stanley, S. M., 1975. A theory of evolution above the species level. *Proc. Natl.  
Acad. Sci. USA* 72:646.
- 738        ———, 1979. *Macroevolution: pattern and process*. W. H. Freeman, San  
Francisco.
- Team, S. D., 2016. Stan Modeling Language Users Guide and Reference Manual.
- 740        Tomiya, S., 2013. Body Size and Extinction Risk in Terrestrial Mammals Above  
the Species Level. *The American Naturalist* 182:196–214.
- 742        Van Valen, L., 1973. A new evolutionary law. *Evolutionary Theory* 1:1–30.
- , 1979. Taxonomic survivorship curves. *Evolutionary Theory* 4:129–142.

- 744 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  
746 latitudinal selectivity gradient at the end-Cretaceous mass extinction.  
Scientific Reports 3:1790.
- 748 Wagner, P. J. and J. D. Marcot, 2013. Modelling distributions of fossil sampling  
rates over time, space and taxa: assessment and implications for  
750 macroevolutionary studies. Methods in Ecology and Evolution 4:703–713.
- Wang, S. C., 2003. On the continuity of background and mass extinction.  
752 Paleobiology 29:455–467.
- Wang, S. C. and C. R. Marshall, 2004. Improved confidence intervals for  
754 estimating the position of a mass extinction boundary. Paleobiology 30:5.
- Williams, C., C. H. C. Brunton, and S. J. Carlson, 2007. Treatise on  
756 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.  $\mu$  values are the location parameters of the effects, while  $\tau$  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%
$\mu_i$	intercept	-2.32	0.14	-2.50	-2.32	-2.15
$\mu_r$	geographic range	-0.76	0.11	-0.91	-0.76	-0.62
$\mu_v$	environmental preference	-0.66	0.17	-0.88	-0.66	-0.43
$\mu_{v^2}$	environmental preference <sup>2</sup>	2.88	0.31	2.48	2.88	3.27
$\mu_m$	body size	0.04	0.12	-0.12	0.04	0.19
$\tau_i$	SD(intercept)	0.50	0.10	0.37	0.49	0.63
$\tau_r$	SD(geographic range)	0.27	0.13	0.11	0.26	0.45
$\tau_v$	SD(environmental preference)	0.76	0.16	0.56	0.74	0.97
$\tau_{v^2}$	SD(environmental preference <sup>2</sup> )	1.24	0.33	0.84	1.21	1.67
$\tau_m$	SD(body size)	0.47	0.12	0.33	0.47	0.63

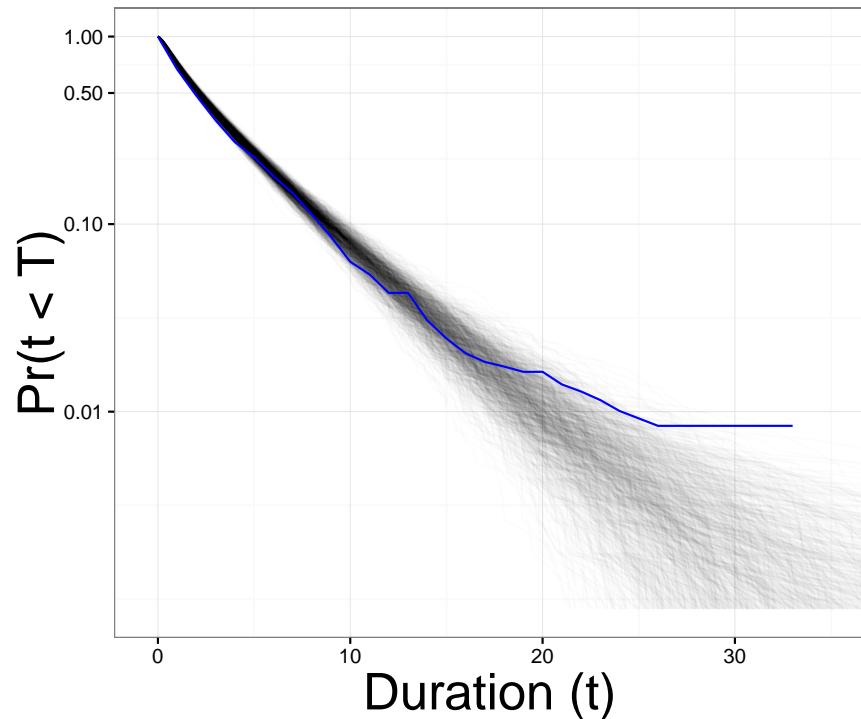


Figure 1: Comparison of the empirical estimate of  $S(t)$  (highlighted) versus estimates from 1000 posterior predictive data sets (black).  $S(t)$  corresponds to the probability that the age of a genus  $t$  is less than the genus' ultimate duration  $T$ . The vertical axis is log10 transformed.

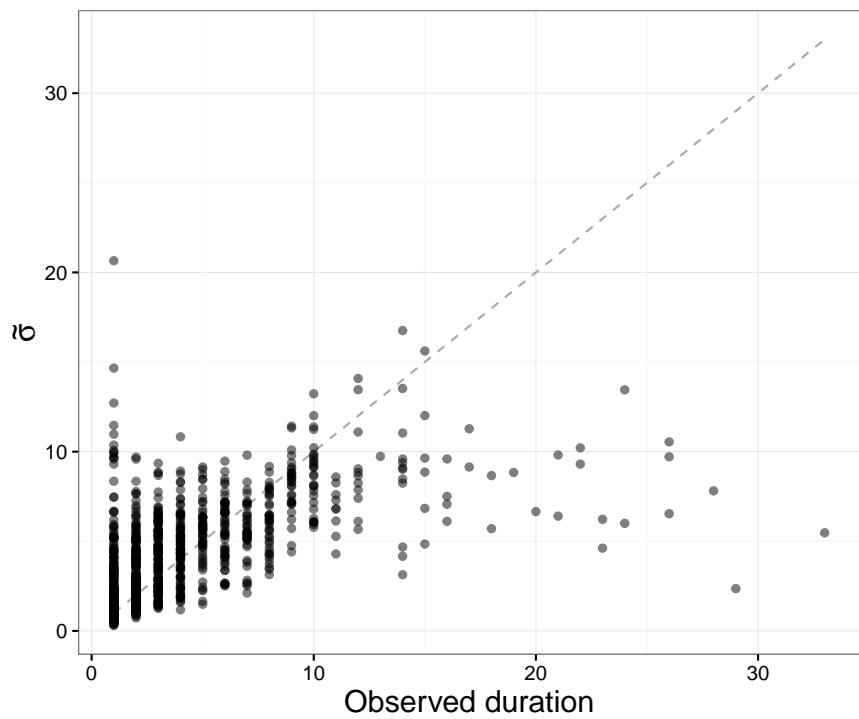


Figure 2: Comparison of all observed genus durations in number of geological stages to the average posterior predictive estimates of  $\log(\sigma)$ . The dashed, diagonal line corresponds to  $x = y$ .

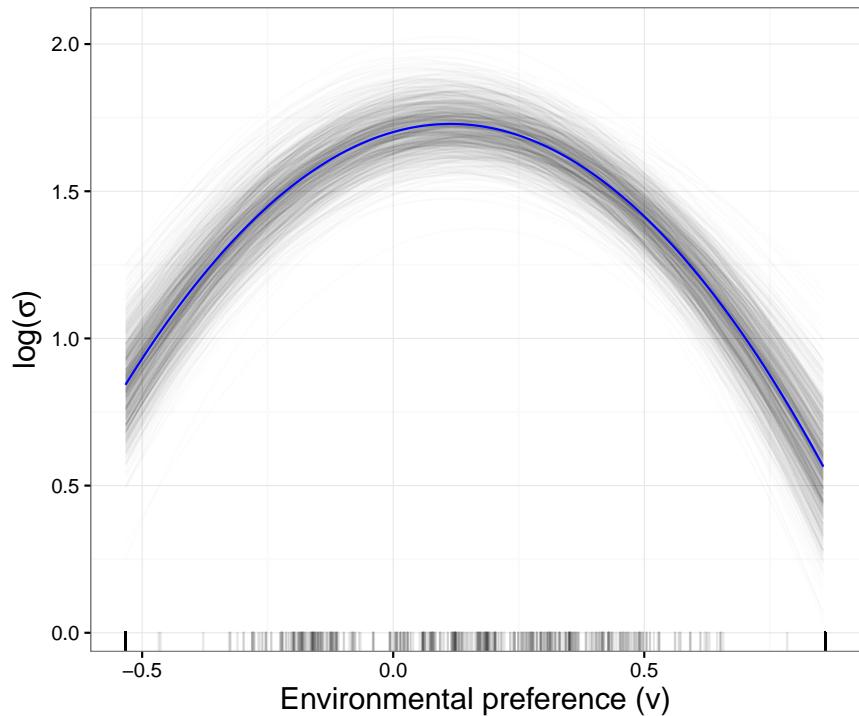


Figure 3: The overall expected relationship between environmental affinity  $v_i$  and a  $\log(\sigma)$  when  $r = 0$  and  $m = 0$ . Each grey line corresponds to a single draw from the posterior predictive distribution, while the highlighted line corresponds to the median of the posterior predictive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

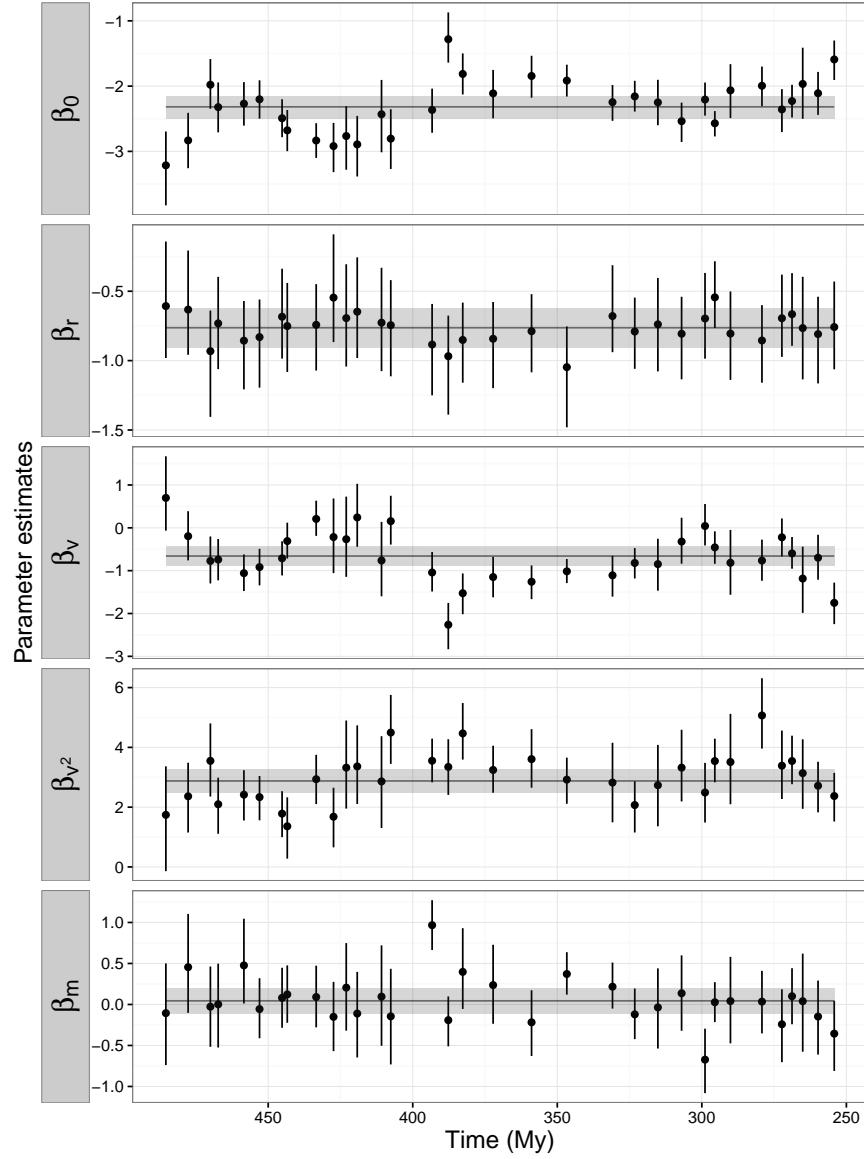


Figure 4: Comparison of cohort-specific estimates of  $\beta_0$ , the effect of geographic range on extinction risk  $\beta_r$ , the effect of environmental preference  $\beta_v$  and  $\beta_{v^2}$ , and body size  $\beta_m$ . Points correspond to the median of the cohort-specific estimate, along with 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).

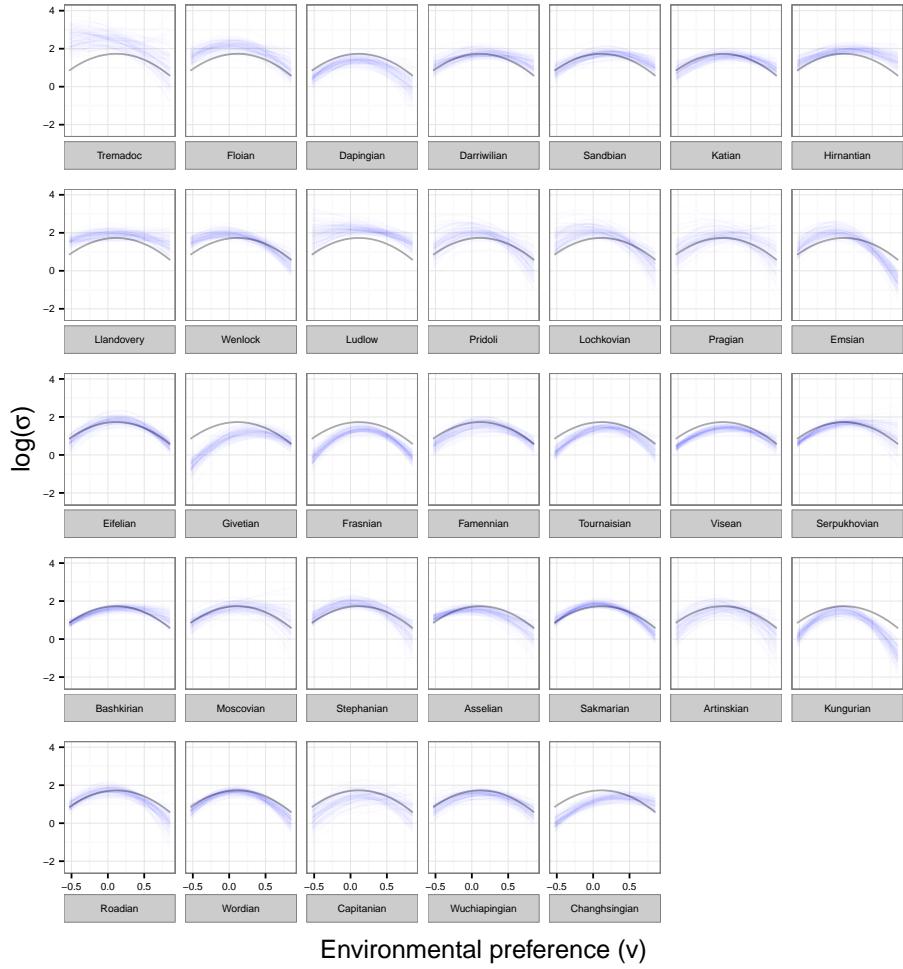


Figure 5: Comparison of origination cohort-specific (posterior predictive) estimates of the effect of environmental preference on  $\log(\sigma)$  to the mean overall estimate of the effect of environmental preference. Cohort-specific estimates are from 100 posterior predictive simulations across the 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.

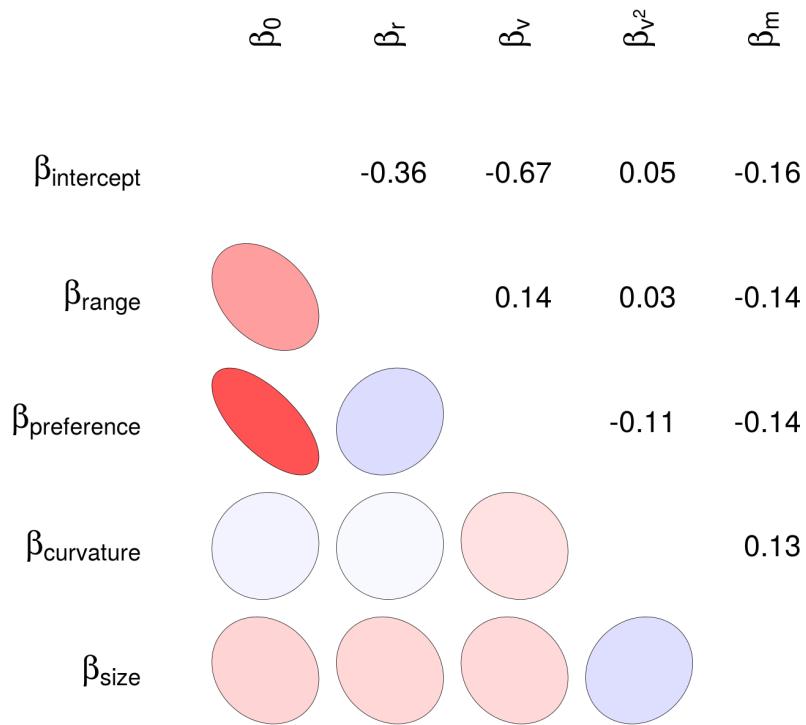


Figure 6: Mixed graphical and numerical representation of the correlation matrix  $\Omega$  of variation in cohort-specific covariate estimates. These correlations are between the estimates of the cohort-level effects of covariates, along with intercept/baseline extinction risk. The median estimates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