

**Title:** The interplay between extinction intensity and selectivity:  
2 correlation in trait effects on taxonomic survival

**Running title:** Variation in trait effects on taxonomic survival

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

6 **Keywords:** macroevolution, extinction, macroecology, Bayesian, brachiopods

**Word count:**

8 **Table count:**

**Figure count:**

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

## Abstract

While the effect of geographic range on extinction risk is well documented, how other traits may increase or decrease extinction risk is less well known. I analyze patterns of Paleozoic brachiopod genus durations and their relationship to geographic range, affinity for epicontinental seas versus open ocean environments, and body size. Additionally, I allow for environmental affinity to have a nonlinear effect on duration. Using a hierarchical Bayesian approach, I also model the interaction between the effects of biological traits and a taxon's time of origination. My analysis framework eschews the traditional distinction between background and mass extinction, instead the entire time period is analyzed where these are part of the same continuum. I find evidence that as baseline extinction risk increases, the effect of geographic range increases but the effect of environmental preference tends to decrease. Additionally, I find strong evidence for correlation between the effects of geographic range and the non-linear aspect of environmental preference which may help explain this pattern. For parts of the Paleozoic I find support for a "survival of the generalists" scenario, though there are times where this relationship is absent or even reversed. These results support the hypothesis that as extinction intensity increases, overall extinction selectivity decreases.

## 1 Introduction

How do biological traits affect extinction risk? Biological traits are defined here as descriptors of a taxon's adaptive zone, which is the set of all biotic–biotic and biotic–abiotic interactions that a taxon can experience (Simpson, 1944). In effect, these are descriptors of a taxon's broad-sense ecology. Jablonski (1986) observed that during a mass extinction event, the effects of biological traits on

taxonomic survival decreased in size. However, this pattern was not the case for  
40 the effect of geographic range on survival (Jablonski, 1986).

Jablonski (1986) phrased his conclusions in terms of background versus mass  
42 extinction, but this scenario is readily transferable to a continuous variation  
framework as there is no obvious distinction in terms of extinction rate between  
44 these two states (Wang, 2003). Additionally, the Jablonski (1986) scenario has  
strong model structure requirements in order to test its proposed  
46 macroevolutionary mechanism; not only do the taxon trait effects need to be  
modeled, but the correlation between trait effects need to be modeled as well.

48 There are two end-member macroevolutionary mechanisms which may underlie  
the pattern observed by Jablonski (1986): the effect of geographic range on  
50 predictive survival remains constant and those of other biological traits decrease,  
or the effect of geographic range in predicting survival increases and those of  
52 other biological traits stay constant. Reality, of course, may fall somewhere  
along this continuum.

54 Conceptually, taxon survival can be considered an aspect of “taxon fitness”  
along with expected lineage specific branching/origination rate (Cooper, 1984,  
56 Palmer and Feldman, 2012). A taxon with a beneficial trait should persist for  
longer, on average, than a taxon without that beneficial trait. Here I model  
58 brachiopod taxon durations because trait based differences in extinction risk  
should manifest as differences in taxon durations. Brachiopods are an ideal  
60 group for this study as they are well known for having an exceptionally  
complete fossil record (Foote, 2000). I focus on the brachiopod record from most  
62 of the Paleozoic, from the start of the Ordovician (approximately 485 Mya)  
through the end Permian (approximately 252 Mya) as this represents the time  
64 of greatest global brachiopod diversity (Alroy, 2010).

The analysis of taxon durations, or time from origination to extinction, falls  
66 under the purview of survival analysis, a field of applied statistics commonly  
used in health care (Klein and Moeschberger, 2003) but has a long history in  
68 paleontology (Simpson, 1944, 1953, Van Valen, 1973, 1979). I adopt a  
hierarchical Bayesian survival modeling approach, which represents both a  
70 conceptual and statistical unification of the paleontological dynamic and cohort  
survival analytic approaches (Baumiller, 1993, Foote, 1988, Raup, 1975, 1978,  
72 Simpson, 2006, Van Valen, 1973, 1979). By using a Bayesian framework I am  
able to quantify the uncertainty inherent in the estimates of the effects of  
74 biological traits on survival, especially in cases where the covariates of interest  
(i.e. biological traits) are themselves known with error.

### 76 **1.1 Factors affecting brachiopod survival**

Geographic range is widely considered the most important taxon trait for  
78 estimating differences in extinction risk at nearly all times, with large  
geographic range associated with low extinction risk (Jablonski, 1986, 1987,  
80 Jablonski and Roy, 2003, Payne and Finnegan, 2007), though Foote and Miller  
(2013) find that this generalization does not hold in the Mesozoic. For the  
82 Paleozoic, however, I expect this to hold true for the entire period analyzed.

Epicontinental seas are a shallow-marine environment where the ocean has  
84 spread over the surface of a continental shelf with a depth typically less than  
100m. In contrast, open-ocean coastline environments have much greater  
86 variance in depth, do not cover the continental shelf, and can persist during  
periods of low sea level. Because of this, it is strongly expected that taxa which  
88 favor epicontinental seas would be at great risk during periods of low sea levels,  
such as during glacial periods, when epicontinental seas are drained. During the

90 Paleozoic (approximately 541–252 My), epicontinental seas were widely spread  
91 globally but declined over the Mesozoic (approximately 252–66 My) and have  
92 nearly disappeared during the Cenozoic (approximately 66–0 My) as open-ocean  
93 coastlines became the dominant shallow-marine setting (Johnson, 1974, Miller  
94 and Foote, 2009, Peters, 2008).

Miller and Foote (2009) demonstrated that during several mass extinctions taxa  
96 associated with open-ocean environments tend to have a greater extinction risk  
97 than those taxa associated with epicontinental seas. During periods of  
98 background extinction, however, they found no consistent difference between  
99 taxa favoring either environment. These two environment types represent the  
100 primary environmental dichotomy observed in ancient marine systems (Miller  
101 and Foote, 2009, Peters, 2008, Sheehan, 2001).

102 Given these findings, I predict that as extinction risk increases, the extinction  
103 risk associated with favoring open-ocean environments should generally increase.  
104 Additionally, there is a possible nonlinear relationship between environmental  
105 preference and taxon duration. A long standing hypothesis is that generalists or  
106 unspecialized taxa will have greater survival than specialists (Baumiller, 1993,  
107 Liow, 2004, 2007, Nürnberg and Aberhan, 2013, 2015, Simpson, 1944). In this  
108 analysis I allowed for environmental preference to have a parabolic effect on  
109 taxon duration

110 Body size, measured as shell length, was also considered as a potential trait that  
111 influences extinction risk (Payne et al., 2014). Body size is a proxy for metabolic  
112 activity and other correlated life history traits (Payne et al., 2014). Given no  
113 strong evidence that body size effects extinction risk in brachiopods, the effect  
114 is most likely small and could be either positive, negative, or even absent.

## 2 Materials and Methods

### <sup>116</sup> 2.1 Fossil occurrence information

The dataset analyzed here was sourced from the Paleobiology Database  
<sup>118</sup> (<http://www.paleodb.org>) which was then filtered based on taxonomic,  
temporal, stratigraphic, and other occurrence information used in this analysis.  
<sup>120</sup> 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  
<sup>122</sup> Payne et al. (2014). Epicontinental versus open-ocean assignments for each fossil  
occurrence are partially based on those from Miller and Foote (2009), with  
<sup>124</sup> additional occurrences assigned similarly (Miller and Foote, personal  
communication). Analyzed occurrences were restricted to those with  
<sup>126</sup> paleolatitude and paleolongitude coordinates, assignment to either  
epicontinental or open-ocean environment, and belonging to a genus present in  
<sup>128</sup> the body size dataset (Payne et al., 2014).

Fossil occurrences were analyzed at the genus level which is common for  
<sup>130</sup> paleobiological, macroevolution, or macroecological studies of marine  
invertebrates (Alroy, 2010, Foote and Miller, 2013, Harnik et al., 2013, Kiessling  
<sup>132</sup> and Aberhan, 2007, Miller and Foote, 2009, Nürnberg and Aberhan, 2013, 2015,  
Payne and Finnegan, 2007, Simpson and Harnik, 2009, Vilhena et al., 2013).  
<sup>134</sup> While species diversity dynamics is frequently of much greater interest than  
those of higher taxa (though see Foote 2014, Harnik et al. 2014, Hoehn et al.  
<sup>136</sup> 2015), the nature of the fossil record makes accurate and precise taxonomic  
assignments at the species level for all occurrences. In particular, the simplicity  
<sup>138</sup> of brachiopod external morphology and the quality of preservation makes it very  
difficult for assignments below the genus level. As such, the choice to analyze  
<sup>140</sup> genera as opposed to species was in order to assure a minimum level of

confidence and accuracy in the data analyzed here.

<sup>142</sup> Genus duration was calculated as the number of geologic stages from first  
<sup>144</sup> appearance to last appearance, inclusive. Durations were based on geologic  
<sup>146</sup> stages as opposed to millions of years because of the inherently discrete nature  
of the fossil record; dates are not assigned to fossils themselves but instead  
<sup>148</sup> fossils are known from a geological interval which represents some temporal  
range. Stages are effectively irreducible temporal intervals in which taxa may  
<sup>150</sup> occur. Genera with a last occurrence in or after Changhsingian stage were right  
censored at the Changhsingian; genera with a duration of only one stage were  
<sup>152</sup> left censored (Klein and Moeschberger, 2003). The covariates used to model  
genus duration were geographic range size ( $r$ ), environmental preference ( $v, v^2$ ),  
body size ( $m$ ), and sampling ( $s$ ).

Geographic range was calculated using an occupancy approach. First, all  
<sup>154</sup> occurrences were projected onto an equal-area cylindrical map projection. Each  
occurrence was then assigned to one of the cells from a  $70 \times 34$  regular raster  
<sup>156</sup> grid placed on the map. Each grid cell represents approximately  $250,000 \text{ km}^2$ .  
The map projection and regular lattice were made using shape files from  
<sup>158</sup> <http://www.naturalearthdata.com/> and the **raster** package for R (Hijmans,  
2015). For each stage, the total number of occupied grid cells was calculated.  
<sup>160</sup> Then, for each genus, the number of grid cells occupied by that genus was  
calculated. Dividing the genus occupancy by the total occupancy gives the  
<sup>162</sup> relative occupancy of that genus. Mean relative genus occupancy was then  
calculated as the mean of the per stage relative occupancies of that genus.  
<sup>164</sup> Sampling  $s$  was calcualted as the average “gap ratio” of all stages in which a  
taxon existed. The basis of the gap ratio are range-through taxa, which are  
<sup>166</sup> taxa, for any stage  $t$ , that are present in stages  $t - 1$  and  $t + 1$  CITATION  
FOOTE ALROY. A gap is when a range-through taxon is not present in stage  $t$ .

<sup>168</sup> The gap ratio for a stage is the number of gaps divided by the total number of  
range-through taxa, and the gap ratio for an individual taxon is the average of  
<sup>170</sup> all gap ratios for each stage the taxon is present (inclusive). Additional  
occurrence information for the late Cambrian and earliest Trassic were used to  
<sup>172</sup> calculate the gap ratios of the first and last stages included in this analysis.

<sup>174</sup> Environmental preference  $v$  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  
<sup>176</sup> at the same time ( $e'_i/E'_i$ ). This measure of environmental preference is  
expressed.

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

<sup>178</sup> Body size, measured as shell length, was sourced directly from Payne et al.  
(2014).

<sup>180</sup> Prior to analysis, geographic range  $r$  was logit transformed and body size  $m$  was  
natural-log transformed. All covariates were then standardized by subtracting  
<sup>182</sup> the mean from all values and dividing by twice its standard deviation, which  
follows Gelman and Hill (2007). This standardization means that all regression  
<sup>184</sup> coefficients are comparable as the expected change per 1-unit change in the  
rescaled covariates.

## <sup>186</sup> 2.2 Analytical approach

Hierarchical modelling is a statistical approach which explicitly takes into  
<sup>188</sup> account the structure of the observed data in order to model both the within  
and between group variance (Gelman et al., 2013, Gelman and Hill, 2007). The

<sup>190</sup> units of study (e.g. genera) each belong to a single grouping (e.g. origination cohort). These groups are considered draws from a shared probability distribution (e.g. all cohorts, observed and unobserved). The group-level parameters are then estimated simultaneously as the other parameters of interest (e.g. covariate effects) (Gelman et al., 2013). The subsequent estimates are partially pooled together, where parameters from groups with large samples or effects remain large while those of groups with small samples or effects are pulled towards the overall group mean.

<sup>198</sup> This partial pooling is one of the greatest advantages of hierarchical modeling. By letting the groups “support” each other, parameter estimates then better reflect our statistical uncertainty. Additionally, this partial pooling helps control for multiple comparisons and possibly spurious results as effects with little support are drawn towards the overall group mean (Gelman et al., 2013, Gelman and Hill, 2007).

<sup>204</sup> All covariate effects (regression coefficients), as well as the intercept term (baseline extinction risk), were allowed to vary by group (origination cohort). The covariance/correlation between covariate effects was also modeled. This hierarchical structure allows inference for how covariates effects may change with respect to each other while simultaneously estimating the effects themselves, propagating our uncertainty through all estimates.

<sup>210</sup> Genus durations were assumed to follow a Weibull distribution which allows for age dependent extinction (Klein and Moeschberger, 2003). The Weibull distribution has two parameters: a scale  $\sigma$ , and a shape  $\alpha$ . When  $\alpha = 1$ ,  $\sigma$  is equal to the expected duration of any taxon.  $\alpha$  acts as a time dilation effect where values greater than 1 indicate that extinction risk increases with age, and values less than 1 indicate that extinction risk decreases with age. Note that the Weibull distribution is equivalent to the exponential distribution when  $\alpha = 1$ .

The scale parameter  $\sigma$  was modeled as a regression with both varying intercept  
218 and varying slopes. The following variables are defined:  $y_i$  is the duration of  
genus  $i$  in geologic stages,  $X$  is the matrix of covariates including a column of  
220 ones for the intercept/constant term,  $B_j$  is the vector of regression coefficients  
for origination cohort  $j$ ,  $\mu$  is the vector of means of each regression coefficient,  $\Sigma$   
222 is the covariance matrix of the regression coefficients,  $\tau$  is the vector of the  
standard deviations of the between-cohort variation of the regression coefficient  
estimates, and  $\Omega$  is the correlation matrix of the regression coefficients. The  
224 elements of the vector  $\mu$  were given independent normally distributed priors.  
The effects of geographic range size and the breadth of environmental preference  
226 were given informative priors. The correlation matrix  $\Omega$  was given an almost flat  
228 LKJ distributed prior following CITATION STAN manual.

The shape parameter  $\alpha$  was allowed to vary by cohort. The log of  $\alpha$  was  
230 modeled as a regression with intercept  $\alpha'$  and standard deviation  $\sigma^\alpha$  where the  
effect of origination cohort  $a_j$  is modeled as draws from a shared normal  
232 distribution with mean 0 and standard deviation  $\sigma^a$ .

Except where noted, regression coefficients were given a weakly informative  
234 normally distributed prior, scale (e.g. standard deviation) parameters were  
given a weakly informative half-Cauchy prior following the CITATION Gelman  
236 textbook, STAN manual.

The full sampling statement, along with all necessary transformations and

<sup>238</sup> priors, is expressed as

$$\begin{aligned}
y_i &\sim \text{Weibull}(\alpha_i, \sigma_i) \\
\sigma_i &= \exp\left(\frac{-(\mathbf{X}_i B_{j[i]})}{\alpha_i}\right) \\
B_j &\sim \text{MVN}(\mu, \Sigma) \\
\mu_0 &\sim \mathcal{N}(0, 5) \\
\mu_r &\sim \mathcal{N}(-1, 1) \\
\mu_v &\sim \mathcal{N}(0, 1) \\
\mu_{v^2} &\sim \mathcal{N}(1, 1) \\
\mu_m &\sim \mathcal{N}(0, 1) \\
\mu_s &\sim \mathcal{N}(0, 1) \\
\Sigma &= \text{Diag}(\tau)\Omega\text{Diag}(\tau) \\
\tau &\sim \text{C}^+(1) \\
\Omega &\sim \text{LKJ}(2) \\
\alpha_i &= \exp(\mathcal{N}(\alpha' + a_{j[i]}, \sigma^\alpha)) \\
\alpha' &\sim \mathcal{N}(0, 1) \\
a_j &\sim \mathcal{N}(0, \sigma^a) \\
\sigma^a &\sim \text{C}^+(1) \\
\sigma^\alpha &\sim \text{C}^+(1).
\end{aligned} \tag{2}$$

The joint posterior was approximated using a Markov-chain Monte Carlo  
<sup>240</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>242</sup> programming language Stan (Stan Development Team, 2014). The posterior  
 distribution was approximated from four parallel chains run for 10,000 steps

<sup>244</sup> each, split half warm-up and half sampling and thinned to every 10th sample for  
<sup>245</sup> a total of 4000 posterior samples. Chain convergence was assessed via the scale  
<sup>246</sup> reduction factor  $\hat{R}$  where values close to 1 ( $\hat{R} < 1.1$ ) indicate approximate  
<sup>247</sup> convergence. Convergence means that the chains are approximately stationary  
<sup>248</sup> and the samples are well mixed (Gelman et al., 2013).

<sup>250</sup> The fit of the above model (Eq. 2; the “full” model) was compared to the fits of  
<sup>251</sup> three other sub-models: constant  $\alpha$  across cohorts, no sampling as a covariate,  
<sup>252</sup> or both constant  $\alpha$  and no sampling covariate. These models were compared for  
<sup>253</sup> predicted out-of-sample predictive accuracy using both the widely-applicable  
<sup>254</sup> information criterion (WAIC) and leave-one-out cross-validation estimated via  
<sup>255</sup> Pareto-smoothed importance sampling CITATIONS. Both of these are  
<sup>256</sup> estimates of the out-of-sample predictive accuracy or the expected quality of fit  
of the model to new data.

<sup>258</sup> WAIC is a more fully Bayesian alternative to AIC or DIC (Gelman et al., 2013,  
Watanabe, 2010); comparisons of WAIC values are useful for better  
understanding the effect of model complexity on out-of-sample predictive  
<sup>260</sup> accuracy. The calculation of WAIC used here corresponds to the “WAIC 2”  
formulation recommended by Gelman et al. (2013). Lower values of WAIC  
<sup>262</sup> indicate greater expected out-of-sample predictive accuracy than higher values.

PSIS-LOO explanation here.

<sup>264</sup> Model adequacy was evaluated using a couple of posterior predictive checks.  
The posterior predictive checks are estimates of model adequacy in that  
<sup>266</sup> replicated data sets using the fitted model should be similar to the original data  
where systematic differences between the simulations and observations indicate  
<sup>268</sup> weaknesses of the model fit (Gelman et al., 2013). 1000 posterior predictive  
datasets were generated from 1000 unique draws from the posterior distribution

270 of each parameter. The two posterior predictive checks used in this analysis are  
 271 a comparison of a non-parameteric estimate of the survival function  $S(t)$  from  
 272 the empirical dataset to the non-parameteric estimates of  $S(t)$  from the 1000  
 273 posterior predictive datasets, and comparison of the observed genus durations to  
 274 the average posterior predictive estimate of  $\log(\sigma_i)$  (Eq. 2). The former is to see  
 275 if simulated data has a similar survival pattern to the observed, and the latter is  
 276 to see if the model systematically over or under estimates taxon survival.

### 3 Results

278 The model with greatest estimated out-of-sample predictive accuracy by a large  
 279 margin, as estimted via both LOOIC and WAIC, has both constant  $\alpha$  and no  
 280 sampling covariate (Table 1). Because of this, only the results from this model  
 281 will be presented here.

Table 1: Widely applicable information criterion (WAIC) and leave-one-out cross-validation information criterion (LOOIC) values for the four models compared in this analysis.

model	WAIC	LOOIC
constant $\alpha$	9080.39	4823.13
constant $\alpha$ , no sampling	4676.84	4621.95
no sampling	8250.22	5563.97
full model	759172.46	109417.64

282 Comparison of the distribution of posterior predictive estimates of  $S(t)$  to the  
 283 empirical estimate reveal few obvious biases except for estimates of the far right  
 284 tail of durations (Fig. 1). This interpretation is supported by the additional  
 285 posterior predictive comparison where most estimates have no systematic biases  
 286 except for the consistent under-estimate of  $\log(\sigma)$  for older taxa (Fig. 2). The  
 287 results of both posterior predictive checks indicate that, for the majority of  
 288 observations, model fit is not overly biased towards either over- or

under-estimates of duration.

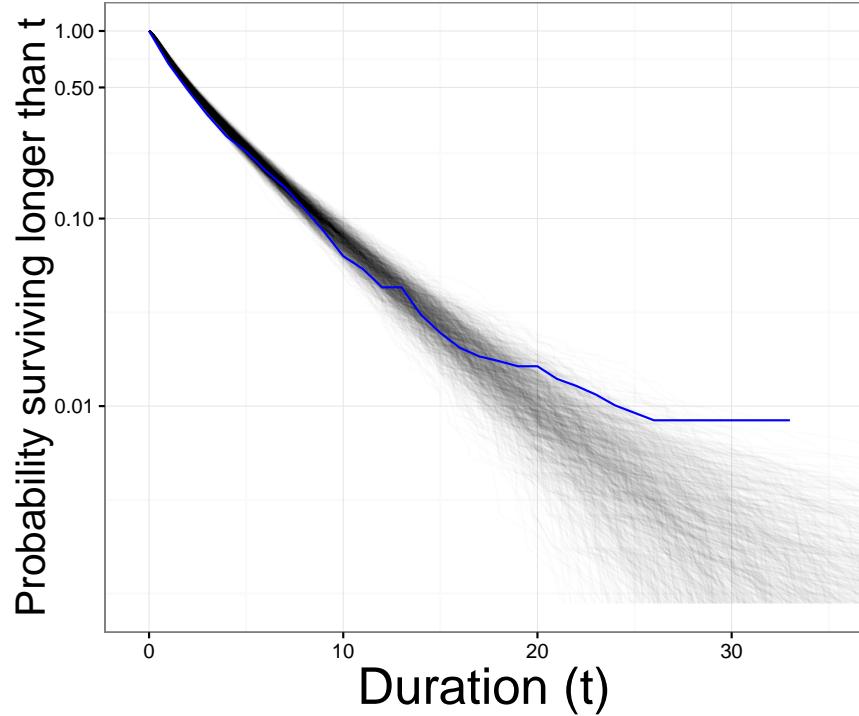


Figure 1: Comparison of the empirical estimate of  $S(t)$  (highlighted) versus estimates from 1000 posterior predictive data sets (black).  $S(t)$  corresponds to the probability that the age of a genus  $t$  is less than the genus' ultimate duration  $T$ .

## 290 4 Discussion

By not knowing the function form of  $f$ , measurement error in duration probably  
292 leads to additional noise to all estimates of model parameters; this is the case  
with linear regression when the dependent variable is known with error.

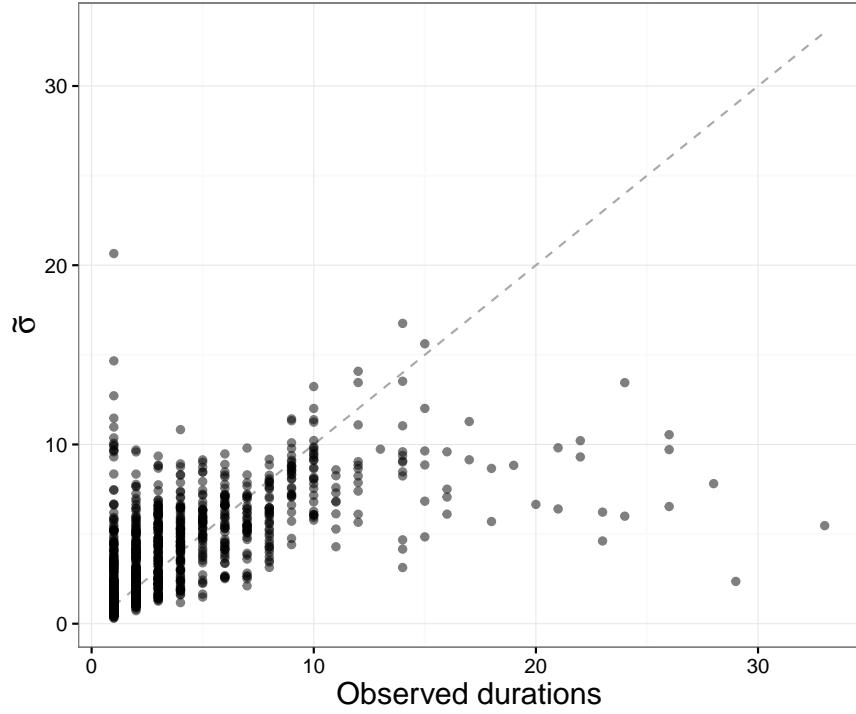


Figure 2: Comparison of all observed genus durations in number of geological stages to the average posterior predictive estimates of  $\log(\sigma)$ . The dashed, diagonal line corresponds to  $x = y$ .

<sup>294</sup> Specifically,

$$\begin{aligned}
 y &= y^* + \nu \\
 y^* &= X\beta + \epsilon \\
 y &= X\beta + \epsilon + \nu \\
 y &= X\beta + \epsilon^*
 \end{aligned} \tag{3}$$

My approach is just checking if duration is associated with a level of sampling ability. If so, it might be misleading results (interaction effects worth thinking about).

<sup>296</sup> My results demonstrate that both the effects of geographic range and the

parameter	mean	standard deviation	10%	50%	90%
$\mu_i$	-2.32	0.14	-2.50	-2.32	-2.15
$\mu_r$	-0.76	0.11	-0.91	-0.76	-0.62
$\mu_v$	-0.66	0.17	-0.88	-0.66	-0.43
$\mu_{v^2}$	2.88	0.31	2.48	2.88	3.27
$\mu_m$	0.04	0.12	-0.12	0.04	0.19
$\tau_i$	0.50	0.10	0.37	0.49	0.63
$\tau_r$	0.27	0.13	0.11	0.26	0.45
$\tau_v$	0.76	0.16	0.56	0.74	0.97
$\tau_{v^2}$	1.24	0.33	0.84	1.21	1.67
$\tau_m$	0.47	0.12	0.33	0.47	0.63

Table 2: 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, 10th, 50th, and 90th quantiles of the posterior are presented.

peakedness/concavity of environmental preference are both negatively  
 300 correlated with baseline extinction risk, meaning that as baseline extinction risk  
 increases the effect size of geographic range increases but the curvature of the  
 302 effect of environmental preference decreases (Fig. 4B).

However, the correlation between geographic range and origination cohort,  
 304 although weak, indicates that neither of the two proposed end-member  
 macroevolutionary mechanisms is strictly right. The evidence for a correlation  
 306 between changes in the effect of geographic range and the curvature of the effect  
 of environmental preference may help explain this result. I hypothesize that  
 308 because taxa with large geographic ranges encompass more possible  
 environments, when extinction intensity is high there is little coherent difference  
 310 in the environmental preference among the surviving taxa. The intensity  
 decreases the selectivity such that the effect of environmental preference is  
 312 effectively washed out by the strength of the effect of geographic range; this is  
 consistent with similar hypotheses presented in Raup (1991).

314 For the approximately 233 My period analyzed there is an approximate 74%

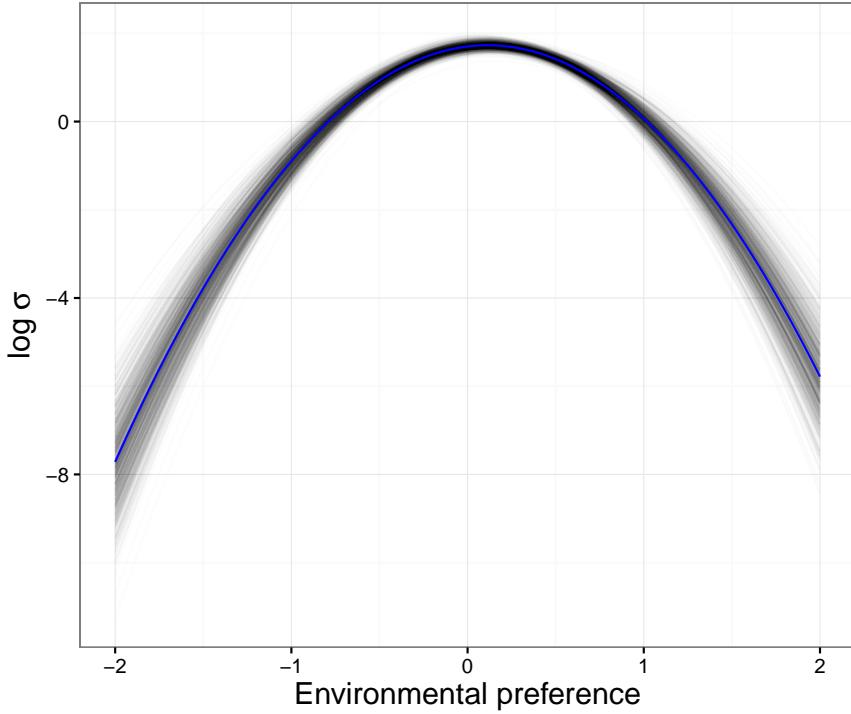


Figure 3: The overall expected relationship between environmental affinity  $v_i$  and  $\log(\sigma)$  when  $r = 0$  and  $m = 0$ . Each grey line corresponds to a single draw from the posterior predictive distribution, while the highlighted line corresponds to the median of the posterior predictive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

posterior probability that brachiopod genera with intermediate environmental  
 316 preferences are expected to have a lower extinction risk than either end  
 members. However, the over all curvature of  $f(v_i)$  is not very peaked meaning  
 318 that when averaged over the entire Phanerozoic this relationship may not lead  
 to large differences in extinction risk (Fig. 3). Note that the duration of the  
 320 period analyzed is approximately four times then length of the Cenozoic (e.g.  
 time since the extinction of the non-avian dinosaurs). This result gives very  
 322 weak support for the universality of the hypothesis that environmental

generalists have greater survival than environmental specialists (Liow, 2004,  
324 2007, Nürnberg and Aberhan, 2013, 2015, Simpson, 1944).

The posterior variance in the estimate of overall  $f(v_i)$  reflects the large between  
326 cohort variance in cohort specific estimates of  $f(v_i)$  (Fig. 6). Given that there is  
only a 74% posterior probability that the expected overall estimate of  $f(v_i)$  is  
328 concave down, it is not surprising that there are some stages where the  
estimated relationship is in fact the reverse of the prior expectation.

330 Additionally, some of those same stages where  $f(v_i)$  does not resemble the prior  
expectation of a concave down nonlinear relation are instead highly skewed  
332 and effectively linear (Fig. 6). These results demonstrate that, while the  
group-level estimate may only weakly support one hypothesis, the cohort-level  
334 estimates may exhibit very different characteristics. These results are also  
consistent with aspects of Miller and Foote (2009) who found that the effect of  
336 environmental preference on extinction risk was quite variable and without  
obvious patterning during times of background extinction.

338 There are two mass extinction events that are captured within the time frame  
considered here: the Ordovician-Silurian and the Frasnian-Famennian. The  
340 cohorts bracketing these events are worth considering in more detail.

The proposed mechanism for the end Ordovician mass extinction is a decrease  
342 in sea level and the draining of epicontinental seas due to protracted glaciation  
(Johnson, 1974, Sheehan, 2001). My results are broadly consistent with this  
344 scenario with both epicontinental and open-ocean specialists having a much  
lower expected duration than intermediate taxa (Fig. 6). All of the stages  
346 between the Darriwillian and the Llandovery, except the Hirnantian, have a  
high probability (90+%) that  $f(v)$  is concave down. The pattern for the  
348 Darriwillian, which proceeds the supposed start of Ordovician glacial activity,  
demonstrates that taxa tend to favor open-ocean environments are expected to

<sup>350</sup> have a greater duration than either intermediate or epicontinental specialists, in decreasing order.

<sup>352</sup> For nearly the entire Devonian estimates of  $f(v)$  indicate that one of the environmental end members is favored over the other end member of intermediate preference (Fig. 6). For almost the entirety the Givetian through the end of the Devonian and into the Viséan, I find that epicontinental favoring taxa are expected to have a greater duration than either intermediate or open-ocean specialists. Additionally, for nearly the entire Devonian and through <sup>354</sup> to the Visean, the cohort-specific estimates of  $f(v)$  are concave-up. This result, however, seems to reflect the intensity of the seemingly nearly-linear difference <sup>356</sup> in expected duration across the range of  $v$  as opposed to an inversion of the weakly expected curvilinear pattern.

<sup>362</sup> The use of genera as the unit of the study and how to exactly interpret the effects of the biological traits is a remaining concern. For example, if any of the <sup>364</sup> traits analyzed here are associated with increases in speciation rates, this might increase the duration of genera through self-renewal (Raup, 1991, 1994), which <sup>366</sup> would be an example of the difference in biological pattern between species and genera (Jablonski, 1987, 2007, 2008). This could lead to a trait appearing to <sup>368</sup> decrease generic level extinction risk by increasing species level origination rate instead of decreasing species level extinction risk. However, given the nature of <sup>370</sup> the fossil record and maintaining a minimum level of data consistency/quality, there is no simple solution to decreasing this uncertainty in the interpretations <sup>372</sup> of how the biological traits studied at the genus-level may translate to the species-level.

<sup>374</sup> The model used here could be improved through either increasing the number of analyzed taxon traits, expanding the hierarchical structure of the model to <sup>376</sup> include other major taxonomic groups of interest, and the inclusion of explicit

phylogenetic relationships between the taxa in the model as an additional  
378 hierarchical effect. An example taxon trait that may be of particular interest is  
the affixing strategy or method of interaction with the substrate of the taxon,  
380 which has been found to be related to brachiopod survival where, for  
cosmopolitan taxa, taxa that are attached to the substrate are expected to have  
382 a greater duration than those that are not (Alexander, 1977).

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

390 Taxon traits like environmental preference or geographic range (Hunt et al.,  
2005, Jablonski, 1987) are most likely heritable, at least phylogenetically  
392 (Housworth et al., 2004, Lynch, 1991). Without phylogenetic context, this  
analysis assumes that differences in extinction risk between taxa are  
394 independent of the shared evolutionary history of those taxa (Felsenstein, 1985).

In contrast, the origination cohorts only capture shared temporal context. For  
396 example, if taxon duration is phylogenetically heritable, then closely related  
taxa may have more similar durations as well as more similar biological traits.

398 Without taking into account phylogenetic similarity the effects of these  
biological traits would be inflated solely due to inheritance. The inclusion of  
400 phylogenetic context as an additional individual-level hierarchical effect  
independent of origination cohort would allow for determining how much of the  
402 observed variability is due to shared evolutionary history versus shared temporal  
context versus actual differences associated with biological traits (Smits, 2015).

404 In summary, patterns of Paleozoic brachiopod survival were analyzed using a  
405 fully Bayesian hierarchical survival modelling approach while also eschewing the  
406 traditional separation between background and mass extinction. I find that as  
407 baseline extinction risk increases, the form of the selectivity of extinction  
408 changes such that during periods of low extinction risk the effect environmental  
409 preference is expected to change from nonlinear to potentially linear or even  
410 absent while the effect of geographic range increases. In particular, the  
411 correlation between the effect of geographic range and the curvature of the  
412 effect of environmental preference on taxon survival supports the hypothesis  
413 that during periods of high extinction intensity of the effect of geographic range  
414 effectively washes out the effects of other biological traits (Jablonski, 1987,  
415 Raup, 1991). Finally, I find weak support for “survival of the unspecialized”  
416 (Liow, 2004, 2007, Nürnberg and Aberhan, 2013, 2015, Simpson, 1944) as a  
417 general characterization of the effect of environmental preference on extinction  
418 risk (Fig. 3), most origination cohorts conforming to this hypothesis (Fig. 6).

## Acknowledgements

420 I would like to thank K. Angielczyk, M. Foote, P. D. Polly, and R. Ree for  
421 helpful discussion during the conception of this study. I'd also like to thank M.  
422 Villarosa Garcia for draft comments. Additionally, thank you A. Miller for the  
423 epicontinental versus open-ocean assignments. This entire study would  
424 not have been possible without the Herculean effort of the many contributors to  
425 the Paleobiology Database. In particular, I would like to thank J. Alroy, M.  
426 Aberhan, D. Bottjer, M. Clapham, F. Fürsich, N. Heim, A. Hendy, S. Holland,  
427 L. Ivany, W. Kiessling, B. Kröger, A. McGowan, T. Olszewski, P.  
428 Novack-Gottshall, M. Patzkowsky, M. Uhen, L. Villier, and P. Wager. This work

was supported by a NASA Exobiology grant (NNX10AQ446) to A. Miller and  
430 M. Foote. I declare no conflicts of interest. This is Paleobiology Database  
publication XXX.

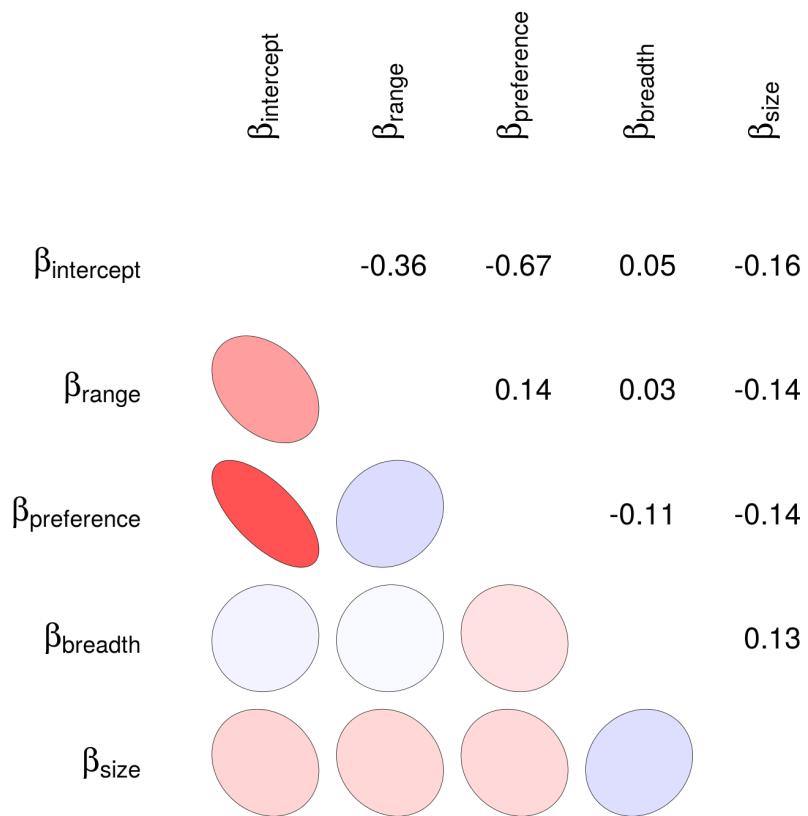


Figure 4: Mixed graphical and numerical representation of the correlation matrix  $\Omega$  of variation in cohort-specific covariate estimates. These correlations are between the estimates of the cohort-level effects of covariates, along with intercept/baseline extinction risk. The median estimates of the correlations are presented numerically (upper-triangle) and as idealized ellipses representing that much correlation (lower-triangle).

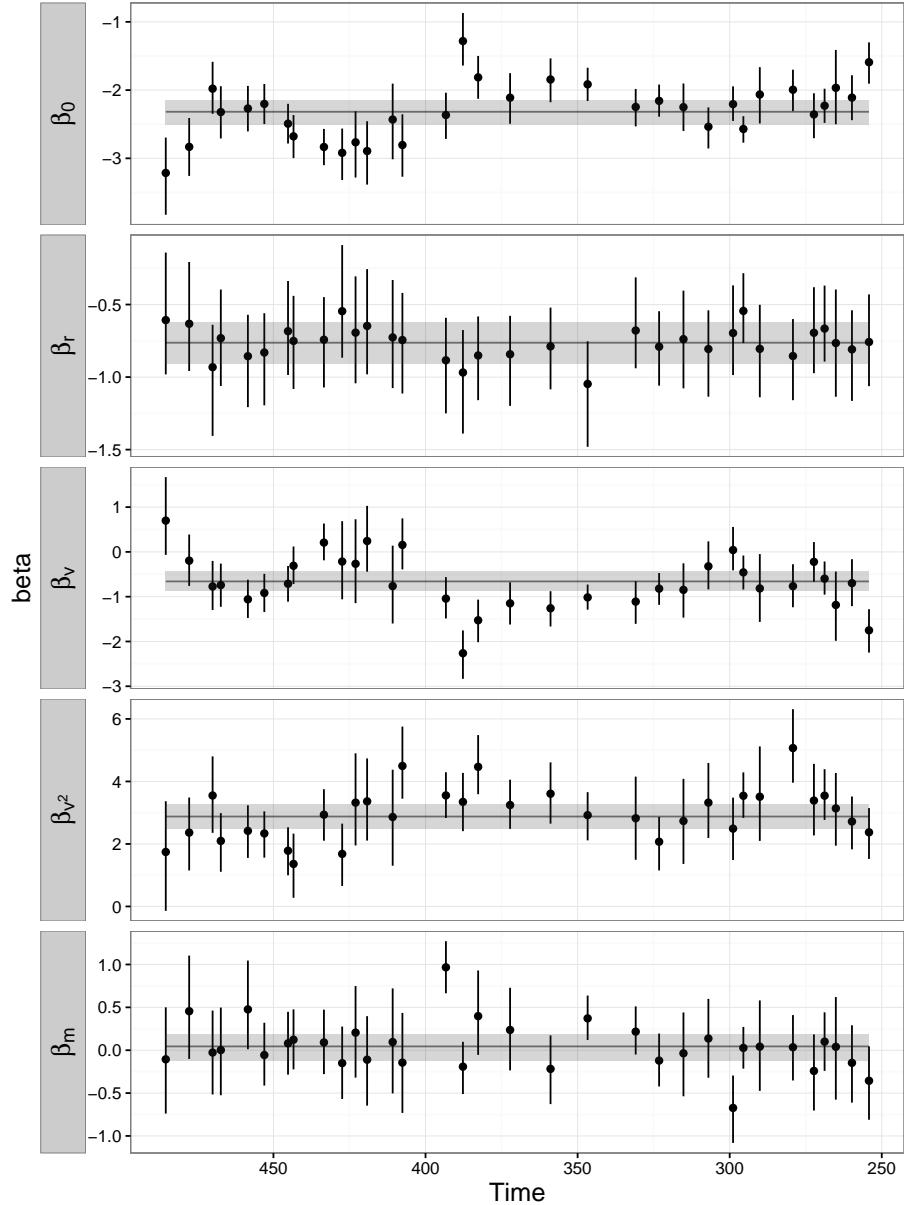


Figure 5: Comparison of cohort-specific estimates of  $\beta_0$  (first row), the effect of geographic range on extinction risk  $\beta_r$  (second row), the effect of environmental preference  $\beta_v$  and  $\beta_{v^2}$  (third, fourth rows), and body size  $\beta_m$  (fifth row). Points correspond to the median of the cohort-specific estimate, along with 80% credible intervals. The horizontal axis is time in millions of years; points are plotted at the midpoint of the origination cohort. Black, horizontal lines are the overall estimates of covariate effects along with 80% credible intervals (shaded).

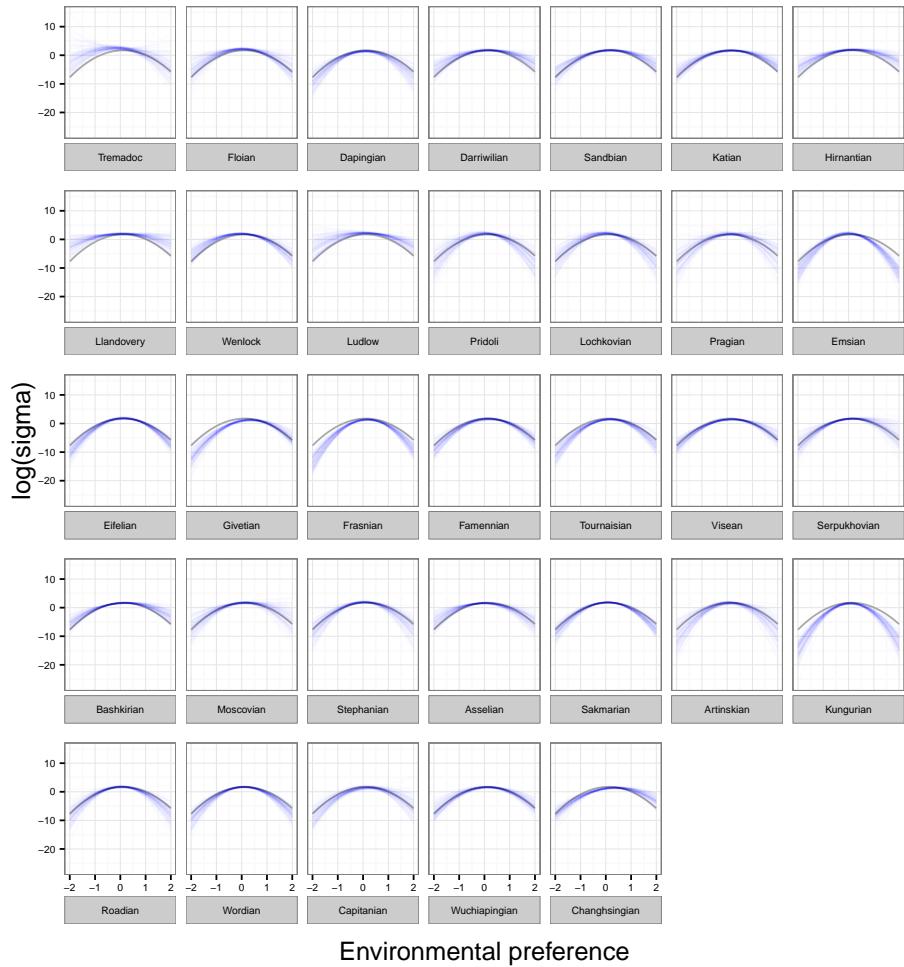


Figure 6

<sup>432</sup> **References**

- Alexander, R. R., 1977. Generic longevity of articulate brachiopods in relation  
<sup>434</sup> to the mode of stabilization on the substrate. *Palaeogeography, Palaeoclimatology, Palaeoecology* 21:209–226.
- <sup>436</sup> Alroy, J., 2010. The Shifting Balance of Diversity Among Major Marine Animal Groups. *Science* 329:1191–1194.
- <sup>438</sup> Baumiller, T. K., 1993. Survivorship analysis of Paleozoic Crinoidea: effect of filter morphology on evolutionary rates. *Paleobiology* 19:304–321.
- <sup>440</sup> Cooper, W. S., 1984. Expected time to extinction and the concept of fundamental fitness. *Journal of Theoretical Biology* 107:603–629.
- <sup>442</sup> Felsenstein, J., 1985. Phylogenies and the comparative method. *American Naturalist* 125:1–15.
- <sup>444</sup> Foote, M., 1988. Survivorship analysis of Cambrian and Ordovician Trilobites. *Paleobiology* 14:258–271.
- <sup>446</sup> ———, 2000. Origination and extinction components of taxonomic diversity: Paleozoic and post-Paleozoic dynamics. *Paleobiology* 26:578–605.
- <sup>448</sup> ———, 2014. Environmental controls on geographic range size in marine animal genera. *Paleobiology* 40:440–458.
- <sup>450</sup> Foote, M. and A. I. Miller, 2013. Determinants of early survival in marine animal genera. *Paleobiology* 39:171–192.
- <sup>452</sup> 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,  
<sup>454</sup> FL.

- Gelman, A. and J. Hill, 2007. Data Analysis using Regression and  
456 Multilevel/Hierarchical Models. Cambridge University Press, New York, NY.
- Harnik, P. G., P. C. Fitzgerald, J. L. Payne, and S. J. Carlson, 2014.  
458 Phylogenetic signal in extinction selectivity in Devonian terebratulide  
brachiopods. *Paleobiology* 40:675–692.
- Harnik, P. G., C. Simpson, and J. L. Payne, 2013. Long-term differences in  
460 extinction risk among the seven forms of rarity. *Proceedings of the Royal  
Society B: Biological Sciences* 282.
- Hijmans, R. J., 2015. raster: Geographic data analysis and modeling. URL  
464 <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  
466 natural selection on traits above the species level. *Methods in Ecology and  
Evolution* .
- Hoffman, M. D. and A. Gelman, 2014. The No-U-Turn Sampler: Adaptively  
468 Setting Path Lengths in Hamiltonian Monte Carlo. *Journal of Machine  
Learning Research* 15:1351–1381.
- Housworth, E. A., P. Martins, and M. Lynch, 2004. The Phylogenetic Mixed  
472 Model. *The American Naturalist* 163:84–96.
- Hunt, G., K. Roy, and D. Jablonski, 2005. Species-level heritability reaffirmed: a  
474 comment on "On the heritability of geographic range sizes". *American  
Naturalist* 166:129–135.
- Jablonski, D., 1986. Background and mass extinctions: the alternation of  
476 macroevolutionary regimes. *Science* 231:129–133.
- , 1987. Heritability at the species level: analysis of geographic ranges of  
478 cretaceous mollusks. *Science* 238:360–363.

- 480 ———, 2007. Scale and hierarchy in macroevolution. *Palaeontology* 50:87–109.
- 482 ———, 2008. Species Selection: Theory and Data. *Annual Review of Ecology, Evolution, and Systematics* 39:501–524.
- 484 Jablonski, D. and K. Roy, 2003. Geographical range and speciation in fossil and living molluscs. *Proceedings. Biological sciences / The Royal Society* 270:401–6.
- 486 Johnson, J. G., 1974. Extinction of Perched Faunas. *Geology* 2:479–482.
- 488 Kiessling, W. and M. Aberhan, 2007. Environmental determinants of marine benthic biodiversity dynamics through Triassic–Jurassic time. *Paleobiology* 33:414–434.
- 490 Klein, J. P. and M. L. Moeschberger, 2003. *Survival Analysis: Techniques for Censored and Truncated Data*. 2nd ed. Springer, New York.
- 492 Liow, L. H., 2004. A test of Simpson's "rule of the survival of the relatively unspecialized" using fossil crinoids. *The American Naturalist* 164:431–43.
- 494 ———, 2007. Does versatility as measured by geographic range, bathymetric range and morphological variability contribute to taxon longevity? *Global Ecology and Biogeography* 16:117–128.
- 498 Lynch, M., 1991. Methods for the analysis of comparative data in evolutionary biology. *Evolution* 45:1065–1080.
- 500 Miller, A. I. and M. Foote, 2009. Epicontinental seas versus open-ocean settings: the kinetics of mass extinction and origination. *Science* 326:1106–9.
- 502 Nürnberg, S. and M. Aberhan, 2013. Habitat breadth and geographic range predict diversity dynamics in marine Mesozoic bivalves. *Paleobiology* 39:360–372.

- 504 ———, 2015. Interdependence of specialization and biodiversity in Phanerozoic  
marine invertebrates. *Nature communications* 6:6602.
- 506 Palmer, M. E. and M. W. Feldman, 2012. Survivability is more fundamental  
than evolvability. *PloS one* 7:e38025.
- 508 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.
- 510 Payne, J. L., N. A. Heim, M. L. Knope, and C. R. McClain, 2014. Metabolic  
512 dominance of bivalves predates brachiopod diversity decline by more than 150  
million years. *Proceedings of the Royal Society B* 281:20133122.
- 514 Peters, S. E., 2008. Environmental determinants of extinction selectivity in the  
fossil record. *Nature* 454:626–9.
- 516 Raup, D. M., 1975. Taxonomic survivorship curves and Van Valen's Law.  
*Paleobiology* 1:82–96.
- 518 ———, 1978. Cohort Analysis of generic survivorship. *Paleobiology* 4:1–15.
- , 1991. Extinction: Bad Genes or Bad Luck? Norton, New York.
- 520 ———, 1994. The role of extinction in evolution. *Proceedings of the National  
Academy of Sciences* 91:6758–6763.
- 522 Sepkoski Jr., J. J., 1981. A factor analytic description of the Phanerozoic  
marine fossil record. *Paleobiology* 7:36–53.
- 524 Sheehan, P., 2001. The late Ordovician mass extinction. *Annual Review of  
Earth and Planetary Sciences* 29:331–364.
- 526 Simpson, C., 2006. Levels of selection and large-scale morphological trends.  
Ph.D. thesis, University of Chicago.

- 528 Simpson, C. and P. G. Harnik, 2009. Assessing the role of abundance in marine  
bivalve extinction over the post-Paleozoic. *Paleobiology* 35:631–647.
- 530 Simpson, G. G., 1944. *Tempo and Mode in Evolution*. Columbia University  
Press, New York.
- 532 ———, 1953. *The Major Features of Evolution*. Columbia University Press,  
New York.
- 534 Smits, P. D., 2015. Expected time-invariant effects of biological traits on  
mammal species duration. *Proceedings of the National Academy of Sciences* .
- 536 Stan Development Team, 2014. Stan: A c++ library for probability and  
sampling, version 2.5.0. URL <http://mc-stan.org/>.
- 538 Van Valen, L., 1973. A new evolutionary law. *Evolutionary Theory* 1:1–30.
- , 1979. Taxonomic survivorship curves. *Evolutionary Theory* 4:129–142.
- 540 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  
542 latitudinal selectivity gradient at the end-Cretaceous mass extinction.  
*Scientific Reports* 3:1790.
- 544 Wang, S. C., 2003. On the continuity of background and mass extinction.  
*Paleobiology* 29:455–467.
- 546 Watanabe, S., 2010. Asymptotic Equivalence of Bayes Cross Validation and  
Widely Applicable Information Criterion in Singular Learning Theory.
- 548 Journal of Machine Learning Research 11:3571–3594.