Title: How do biological traits affect brachiopod taxonomic survival? A hierarchical Bayesian approach.

Running title: How do biological traits affect taxonomic survival?

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

Keywords: extinction, macroevolution, macroecology, Paleozoic, selection

Word count: ?

Table count: 0.

Figure count: 5 main text, 3 supplement.

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, the effects of other traits are less well known. Using a hierarchical Bayesian modeling approach, I also model the possible interaction between the effects of the biological traits and a taxon's time of origination. 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. My analysis framework eschews the traditional distinction between background and mass extinction, instead the entire time period is analyzed where these "states" are part of a continuum. I find evidence that as extinction risk increases, the expected strength of the selection gradient on biological traits (except body size) increases. This manifests as greater expected differences in extinction risk for each unit change in geographic range and environmental preference during periods of high extinction risk, as opposed to a much flatter expected selection gradient during periods of low extinction risk. I find weak evidence for a universally nonlinear relationship between environmental preference and extinction risk such that "generalists" have a lower expected extinction risk than either "specialists". While for the many parts of the Paleozoic this hypothesis is supported, there are many times where this hypothesized relationship is absent or even reversed. Importantly, I find that as extinction risk increases, the steepness of this relationship is expected to increases as well.

1   Introduction

How do biological traits affect extinction risk? Jablonski (1986) observed that during periods of high extinction risk, the effects of biological traits on survival

1

decreased in size. However, this pattern was weakest/absent in the effect of geographic range on survival (Jablonski, 1986). Biological traits are defined here as descriptors of a taxon's adaptive zone, which is the set of 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) phrased his conclusions in terms of background versus mass extinction, but this scenario is readily transferable to a continuous variation framework as there is no obvious distinction in terms of extinction rate between these two states (Wang, 2003). Additionally, the Jablonski (1986) scenario has strong model structure requirements in order to test its proposed macroevolutionary mechanism; not only do the taxon trait effects need to be modeled, but the correlation between the trait effects need to be modeled as well.

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

I model brachiopod taxon durations because trait based differences in extinction risk should manifest as differences in taxon durations. Namely, a taxon with a beneficial trait should survive longer, on average, than a taxon without that beneficial trait. Conceptually, taxon survival can be considered an aspect of "taxon fitness" along with expected lineage specific branching/origination rate (Cooper, 1984, Palmer and Feldman, 2012). Brachiopods are an ideal group for this study as they are are well known for having an exceptionally complete fossil records (Foote, 2000). Specifically, I focus on the brachiopod record from most

2

of the Paleozoic, specifically from the start of the Ordovician (approximately 485 Mya) through the end Permian (approximately 252 Mya) as this represents the time of greatest global brachiopod diversity (Alroy, 2010).

he analysis of taxon durations, or time from origination to extinction, falls under the purview of survival analysis, a field of applied statistics commonly used in health care (Klein and Moeschberger, 2003) but has a long history in paleontology (Simpson, 1944, 1953, Van Valen, 1973, 1979).

Geographic range is widely considered the most important taxon trait for estimating differences in extinction risk at nearly all times with large geographic range associated with low extinction risk (Jablonski, 1986, 1987, Jablonski and Roy, 2003, Payne and Finnegan, 2007). I expect this to hold true nearly always.

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

Epicontinental seas are a shallow-marine environment where the ocean has spread over the surface of a continental shelf with a depth typically less than 100m. In contrast, open-ocean coastline environments have much greater 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 favor epicontinental seas would be at great risk during periods of low sea levels, such as during glacial periods, where these seas are drained. During the Paleozoic (approximately 541–252 My), epicontinental seas were widely spread

globally but declined over the Mesozoic (approximately 252–66 My) and eventually diminished disappearing during the Cenozoic (approximately 66–0 My) as open-ocean coastlines became the dominant shallow-marine setting (Johnson, 1974, Miller and Foote, 2009, Peters, 2008).

Given the above information, I predict that as extinction risk increases, taxa associated with open-ocean environments should generally increase in extinction risk versus those that favor epicontinental seas. Additionally, there is a possible nonlinear relationship between environmental preference and taxon duration. A long standing hypothesis is that generalists or unspecialized taxa will have greater survival than specialists (Baumiller, 1993, Liow, 2004, 2007, Nürnberg and Aberhan, 2013, 2015, Simpson, 1944). In this analysis I allowed for environmental preference to possibly have a parabolic effect on taxon duration

Body size, measured as shell length (Payne et al., 2014), was also considered as a potentially informative covariate. Body size is a proxy for metabolic activity and other correlated life history traits (Payne et al., 2014). There is no strong hypothesis of how body size effects extinction risk in brachiopods, meaning a positive, negative, or zero effect are all plausible.

I adopt a hierarchical Bayesian survival modeling approach, which represents a conceptual and statistical unification of the paleontological dynamic and cohort survival analytic approaches (Baumiller, 1993, Foote, 1988, Raup, 1975, 1978, 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 biological traits on survival, especially in cases where the covariates of interest (i.e. biological traits) are themselves known with error.

4

2 Materials and Methods

2.1 Fossil occurrence information

118 The dataset analyzed here was sourced from the Paleobiology Database
(http://www.paleodb.org) which was then filtered based on taxonomic,
120 temporal, stratigraphic, and other occurrence information that was necessary
for this analysis. These filtering criteria are very similar to those from Foote and
122 Miller (2013) with an additional constraint of being present in the body size
data set from Payne et al. (2014). Epicontinental versus open-ocean assignments
124 for each fossil occurrence are partially based on those from Miller and Foote
(2009), with additional occurrences assigned similarly (Miller and Foote,
126 personal communication). Sampled occurrences were restricted to those with
paleolatitude and paleolongitude coordinates, assignment to either
128 epicontinental or open-ocean environment, and belonging to a genus present in
the body size dataset (Payne et al., 2014).

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

Genus duration was calculated as the number of geologic stages from first appearance to last appearance, inclusive. Durations were based on geologic stages as opposed to millions of years because of the inherently discrete nature of the fossil record; dates are not assigned to fossils themselves but instead fossils are known from a geological interval which represents some temporal range. Stages are effectively irreducible temporal intervals in which taxa may 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 left censored (Appendix **??**). The covariates used to model genus duration were geographic range size (r), environmental preference (v, v^2), and body size (m).

Geographic range was calculated using an occupancy approach. First, all occurrences were projected onto an equal-area cylindrical map projection. Each occurrence was then assigned to one of the cells from a 70×34 regular raster grid placed on the map. Each grid cell represents approximately 250,000 km^2. The map projection and regular lattice were made using shape files from http://www.naturalearthdata.com/ and the `raster` package for R (Hijmans, 2015). For each stage, the total number of occupied grid cells was calculated. Then, for each genus, the number of grid cells occupied by that genus was calculated. Dividing the genus occupancy by the total occupancy gives the relative occupancy of that genus. Mean relative genus occupancy was then calculated as the mean of the per stage relative occupancies of that genus.

Sampling was calculated as the average relative number of occurrences per temporal unit. That is, for each geographic unit between the first and last appearance of a taxon, the total number of occurrences of that taxon is divided by the total number of occurrences during that temporal unit. The average of all these ratios is that taxon's average sampling s. Sampling was included as a covariate along with its interactions with geographic range and environmental

preference. No other interaction terms were considered.

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

$$
\begin{aligned}
p\left(\theta'_i \middle| \frac{e'_i}{E'_i}\right) &\propto \mathrm{Beta}(e'_i, E'_i - e'_i)\mathrm{Beta}(1,1) \\
&= \mathrm{Beta}(e'_i + 1, E'_i - e'_i + 1) \\
v &= p(\theta_i > \theta'_i)
\end{aligned}
\tag{1}
$$

Body size data was sourced directly from Payne et al. (2014).

Prior to analysis, geographic range r and sampling s were both logit transformed and body size m was natural-log transformed prior to analysis. All covariates were standardized by subtracting the mean from all values and dividing by twice its standard deviation, which follows Gelman and Hill (2007). This standardization means that all regression coefficients are comparable as the expected change per 1-unit change in any of the covariates.

2.2 Analytical approach

Hierarchical modelling is a statistical approach which explicitly takes into account the structure of the observed data in order to model both the within and between group variance (Gelman et al., 2013, Gelman and Hill, 2007). The units of study (e.g. genera) each belong to a single 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

7

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.

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).

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.

Genus durations were modeled as time-till-event data (Klein and Moeschberger, 2003), with covariate information used in estimates of extinction risk as a hierarchical regression model. Genus durations were assumed to follow a Weibull distribution. While the exponential distribution assumes that extinction risk is independent of duration, the Weibull distribution allows for age dependent extinction (Klein and Moeschberger, 2003). The Weibull distribution has two parameters: a scale σ, and a shape α. When $\alpha = 1$, σ is equal to the expected duration of any taxon. α 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 $\alpha = 1$.

8

The scale parameter σ was modeled as a regression with both varying intercept 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 ones for the intercept/constant term, B_j is the vector of regression coefficients for origination cohort j, μ is the vector of means of each regression coefficient, Σ is the covariance matrix of the regression coefficients, τ is the vector of the standard deviations of the between-cohort variation of the regression coefficient estimates, and Ω is the correlation matrix of the regression coefficients. The elements of the vector μ were given independent normally distributed priors. The effects of geographic range size and the breadth of environmental preference were given informative priors. The correlation matrix Ω was given an almost flat LKJ distributed prior following CITATION STAN manual.

The shape parameter α was also modeled as a regression with intercept α' and standard deviation σ^α. The effect of origination cohort a_j is modeled as draws from a shared normal distribution with mean 0 and standard deviation σ^a. γ is the regression coefficient for the effect of the rescaled logarithm of the number of samples s of taxon i.

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

The full sampling statement, along with all necessary transformations and

9

priors, is expressed as

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

The joint posterior was approximated using a Markov-chain Monte Carlo routine that is a variant of Hamiltonian Monte Carlo called the No-U-Turn Sampler (Hoffman and Gelman, 2014) as implemented in the probabilistic programming language Stan (Stan Development Team, 2014). The posterior distribution was approximated from four parallel chains run for 10,000 draws

each, split half warm-up and half sampling and thinned to every 10th sample for
a total of 5000 posterior samples. Chain convergence was assessed via the scale
reduction factor $\hat{R}$ where values close to 1 ($\hat{R} < 1.1$) indicate approximate
convergence. Convergence means that the chains are approximately stationary
and the samples are well mixed (Gelman et al., 2013).

The fit of the above model (Eq. 2; the "full" model) was compared to the fits of
three other sub-models: constant α across cohorts, no sampling or sampling
interaction terms as covariates, or both constant α and no sampling covariates.
These models were compared for predicted out-of-sample predictive accuracy
using both the widely-applicable information criterion (WAIC) and
leave-one-out cross-validation estimated via Pareto-smoothed importance
sampling CITATIONS.

Model adequacy was evaluated using a couple of posterior predictive checks.
The posterior predictive checks are estimates of model adequacy in that
replicated data sets using the fitted model should be similar to the original data
where systematic differences between the simulations and observations indicate
weaknesses of the model fit (Gelman et al., 2013). 1000 posterior predictive
datasets were generated from 1000 unique draws from the posterior distribution
of each parameter. The two posterior predictive checks used in this analysis are
a 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
posterior predictive datasets, and comparison of the observed genus durations to
estimates of $\log(\sigma)$ of each observation (Eq. 2). The former is to see if simulated
data has a similar survival pattern to the observed, and the latter is to see if the
model systematically over or under estimates taxon survival.

11

3 Results

As stated above, posterior approximations for both the exponential and Weibull models achieved approximate stationarity after 10,000 steps, as all parameter estimates have an $\hat{R} < 1.1$.

Comparisons of the survival functions estimated from 1000 posterior predictive data sets to the estimated survival function of the observed genera demonstrates that both the exponential and Weibull models approximately capture the observed pattern of extinction (Fig. ??). The major difference in fit between the two models is that the Weibull model has a slightly better fit for longer lived taxa than the exponential model.

Additionally, the Weibull model is expected to have slightly better out-of-sample predictive accuracy when compared to the exponential model (WAIC 4576 versus 4604, respectively). ??). Because the difference in WAIC between these two models is large, while results from both the exponential and Weibull models will be presented, only those from the Weibull model will be discussed.

Estimates of the overall mean covariate effects μ can be considered time-invariant generalizations for brachiopod survival during the Paleozoic (Fig. ??). Consistent with prior expectations, geographic range size has a negative effect on extinction risk, where genera with large ranges having greater durations than genera with small ranges.

I find that while the mean estimate of the effect of body size on extinction risk is negative, implying that increasing body size decreases extinction risk, this estimate is within 2 standard deviations of 0 (mean $\mu_m = -0.09$, standard deviation 0.09; Fig. ??). Because of this, I infer that body size has no distinguishable effect on brachiopod taxonomic survival.

12

Interpretation of the effect of environmental preference v on duration is slightly

more involved. Because a quadratic term is the equivalent of an interaction

term, both μ_v and μ_{v^2} have to be interpreted together because it is illogical to

change values of v without also changing values v^2. To determine the nature of

the effect of v on duration I calculated the multiplicative effect of environmental

preference on extinction risk.

Given mean estimated extinction risk $\tilde{\sigma}$, we can define the extinction risk

multiplier of an observation with environmental preference v_i as

$$\frac{\tilde{\sigma}_i}{\tilde{\sigma}} = f(v_i) = \exp\left(\frac{-(\mu_v v_i + \mu_{v^2} v^2)}{\alpha}\right). \tag{3}$$

This function $f(v_i)$ has a y-intercept of $\exp(0)$ or 1 because it does not have a

non-zero intercept term. Equation 3 can be either concave up or down. A

concave down $f(v_i)$ may indicate that genera of intermediate environmental

preference have greater durations than either extreme, and *vice versa* for

concave up function.

The expected effect of environmental preference as a multiplier of expected

extinction risk can then be visualized (Fig. **??**). This figure depicts 1000

posterior predictive estimates of Eq. 3 across all possible values of v. The

number indicates the posterior probability that the function is concave down,

with generalists having lower extinction risk/greater duration than either type

of specialist. Note that the inflection point/optimum of Fig. **??** is approximately

$x = 0$, something that is expected given the estimate of μ_v (Fig. **??**).

The matrix Σ describing the covariance between the different coefficients

describes how these coefficients might vary together across the origination

cohorts. Similar to how this was modeled (Eq. **??**, 2), for interpretation

purposes Σ can be decomposed into a vector of standard deviations $\vec{\tau}$ and a

13

correlation matrix Ω.

The estimates of the standard deviation of between-cohort coefficient estimates τ inidicate that some effects can vary greatly between-cohorts (Fig. **??**). Coefficients with greater values of τ have greater between-cohort variation. The covariate effects with the greatest between origination cohort variation are β_r, β_v, and β_{v^2}. Estimates of β_m have negligible between cohort variation, as there is less between cohort variation than the between cohort variation in baseline extinction risk β_0. However the amount of between cohort variation in estimates of β_{v^2} means that it is possible for the function describing the effect of environmental affinity to be upward facing for some cohorts (Eq. 3), which corresponds to environmental generalists being shorter lived than specialists in that cohort.

The correlation terms of Ω (Fig. **??**) describe the relationship between the coefficients and how their estimates may vary together across cohorts. The correlations between the intercept term β_0 and the effects of the taxon traits are of particular interest for evaluating the Jablonski (1986) scenario (Fig. **??** first column/last row). Keep in mind that when β_0 is low, extinction risk is low; and conversely, when β_0 is high, then extinction risk is high.

Marginal posterior probabilities of the correlations between the level of baseline extinction risk β_0 and the effects of the taxon traits indicate that the correlation between expected extinction risk and both geographic range β_r and β_{v^2} are of particular note (Fig. **??**).

There is approximately a 98% probability that β_0 and β_r are negatively correlated (Fig. **??**), meaning that as extinction risk increases, the effect/importance of geographic range on genus duration increases. This means that increases in baseline extinction rate are correlated with an increased

14

importance of geographic range size. There is a 93% probability that β_0 and β_{v^2} are negatively correlated (Fig. **??**), meaning that as extinction risk increases, the peakedness of $f(v_i)$ increases and the relationship tends towards concave down. Additionally, there is a 97% probability that values of β_r and β_{v^2} are positively correlated (Mean correlation 0.51, standard deviation 0.23).

While the overall group level estimates are of particular importance when defining time-invariant differences in extinction risk, it is also important and useful to analyze the individual level parameter estimates in order to better understand how parameters actually vary across cohorts.

In comparison to the overall mean extinction risk μ_0, cohort level estimates β_0 show some amount of variation through time as expected by estimates of τ_0 (Fig. **??**). A similar, if slightly greater, amount of variation is also observable in cohort estimates of the effect of geographic range β_r (Fig. **??**). Again, smaller values of β_0 correspond to lower expected extinction risk. Similarly, smaller values of β_r correspond to greater decrease in extinction risk with increasing geographic range

How the effect of environmental affinity varies between cohorts can be observed by using the cohort specific coefficients estimates. Following the same procedure used earlier (Fig. **??**), but substituting cohort specific estimates of β_v and β_{v^2} for μ_v and μ_{v^2}, the cohort specific effect of environmental preference as a multiplier of mean extinction risk can be calculated. This was done only for the Weibull model, though the observed pattern should be similar for the exponential model.

As expected based on the estimates of τ_v and τ_{v^2}, there is greater variation in the peakedness of $f(v_i)$ than there is variation between concave up and down functions (Fig. **??**). 12 of the 33 cohorts have less than 50% posterior

15

probability that generalists are potentially expected to be shorter lived than specialists, though two of those cases have approximately a 50% probability of being either concave up or down. This is congruent with the 0.72 posterior probability that μ_{v^2} is positive/$f(v_i)$ is concave down.

Additionally, for some cohorts there is a quite striking pattern where the effect of environmental preference v has a nearly-linear relationship (Fig. **??**). These are primarily scenarios where one of the end member preferences is expected to have a greater duration than either intermediate or the opposite end member preference. Whatever curvature is present in these nearly-linear cases is due to the defintion of $f(v)$ as it is not defined for non-negative values of σ (Eq. 3). For all stages between the Emsian through the Viséan, inclusive, intermediate preferences are of intermediate extinction risk when compared with epicontinental specialists (lowest risk) or open-ocean specialists (highest risk). This time period represents most of the Devonian through the early Carboniferous.

4   Discussion

My results demonstrate that both the effects of geographic range and the peakedness/concavity of environmental preference are both negatively correlated with baseline extinction risk, meaning that as baseline extinction risk increases the effect sizes of both these traits are expected to increase (Fig. **??**). This result supports neither of the two proposed macroevolutionary mechanisms for how biological traits should correlate with extinction risk. The observed correlation between the two effects as well as between the effects and baseline extinction risk instead implies that as baseline extinction risk increases, the strength of the total selection gradient on biological traits (except body size)

16

increases. This manifests as greater differences in extinction risk for each unit difference in the biological covariates during periods of high extinction risk, while a relatively flatter selection gradient during periods of low extinction risk.

For the approximately 233 My period analyzed there is an approximate 75% posterior probability that brachiopod genera with intermediate environmental 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 that when averaged over the entire Phanerozoic this relationship may not lead to large differences in extinction risk (Fig. **??**). Note that the duration of the 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 weak support for the universality of the hypothesis that environmental generalists have greater survival than environmental specialists (Liow, 2004, 2007, Nürnberg and Aberhan, 2013, 2015, Simpson, 1944).

The posterior variance in the estimate of overall $f(v_i)$ reflects the large between cohort variance in cohort specific estimates of $f(v_i)$ (Fig. **??**). Given that there is only a 75% posterior probability that the expected overall estimate of $f(v_i)$ is concave down, it is not surprising that there are some stages where the estimated relationship is in fact the reverse of the prior expectation. Additionally, some of those same stages where $f(v_i)$ does not resemble the prior expectation of a concave down nonlinear relation are instead is highly skewed and effectively linear (Fig. **??**). These results demonstrate that, while the group-level estimate may only weakly support one hypothesis, the cohort-level estimates may exhibit very different characteristics.These results are also consistent with aspects of Miller and Foote (2009) who found that the effect of environmental preference on extinction risk was quite variable and without obvious patterning during times of background extinction.

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

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

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. ??). This is consistent with the predictions of Miller and Foote (2009). For almost the entirely the Givetian though 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 to the Visean, the cohort-specific estimates of $f(v)$ are concave-up. This is the opposite pattern than what is expected (Fig. ??). This result, however, seems to reflect the intensity of the seemingly nearly-linear difference in expected duration across the range of v) as opposed to an inversion of the weakly expected curvilinear pattern.

Of concern is the use of genera as the unit of the study and how to exactly

18

interpret the effects of the biological traits. For example, if any of the traits analyzed here are associated with increases in speciation rates, this might "artificially" increase the duration of genera through self-renewal (Raup, 1991, 1994). This could lead to a trait appearing to decrease generic level extinction risk by increasing species level origination rate instead of decreasing species level extinction risk. However, given the nature of the brachiopod fossil record and the difficulty of identifying individual specimens to the species level, there is no simple solution to decreasing this uncertainty in the interpretations of how the biological traits studied here actually affect extinction risk.

This model could be improved through either increasing the number of analyzed taxon traits, expanding the hierarchical structure of the model to include other major taxonomic groups of interest, and the inclusion of explicit phylogenetic relationships between the taxa in the model as an additional hierarchical effect.

An example taxon trait that may be of particular interest is the affixing strategy or method of interaction with the substrate of the taxon. This trait has been found to be related to brachiopod survival (Alexander, 1977) so its inclusion may be of particular interest.

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

Taxon traits like environmental preference or geographic range (Hunt et al., 2005, Jablonski, 1987) are most likely heritable, at least phylogenetically

19

₄₇₆ (Housworth et al., 2004, Lynch, 1991). Without phylogenetic context, this analysis assumes that differences in extinction risk between taxa are

₄₇₈ independent of those taxa's shared evolutionary history (Felsenstein, 1985). In contrast, the origination cohorts only capture shared temporal context. The

₄₈₀ inclusion of phylogenetic context as an addition individual level hierarchical structure independent of origination cohort would allow for determining how

₄₈₂ much of the observed variability is due to shared evolutionary history versus actual differences associated with these taxonomic traits.

₄₈₄ In summary, patterns of Paleozoic brachiopod survival were analyzed using a fully Bayesian hierarchical survival modelling approach while also eschewing the

₄₈₆ traditional separation between background and mass extinction. I modeled both the overall mean effect of biological covariates on extinction risk while also

₄₈₈ modeling the correlation between cohort-specific estimates of covariate effects. I find that as baseline extinction risk increases, the strength of the selection

₄₉₀ gradient on biological traits (except body size) increases. This manifests as greater differences in extinction risk for each unit difference in the biological

₄₉₂ covariates during periods of high extinction risk, while a much flatter total selection gradient during periods of low extinction risk. I also find some support

₄₉₄ for "survival of the unspecialized" (Liow, 2004, 2007, Nürnberg and Aberhan, 2013, 2015, Simpson, 1944) as a general characterization of the effect of

₄₉₆ environmental preference on extinction risk (Fig. ??), though there is heterogeneity between origination cohorts with most periods of time conforming

₄₉₈ to this hypothesis (Fig. ??).

20

Acknowledgements

References

Alexander, R. R., 1977. Generic longevity of articulate brachiopods in relation to the mode of stabilization on the substrate. Palaeogeography, Palaeoclimatology, Palaeoecology 21:209–226.

Alroy, J., 2010. The Shifting Balance of Diversity Among Major Marine Animal Groups. Science 329:1191–1194.

Baumiller, T. K., 1993. Survivorship analysis of Paleozoic Crinoidea: effect of filter morphology on evolutionary rates. Paleobiology 19:304–321.

Cooper, W. S., 1984. Expected time to extinction and the concept of fundamental fitness. Journal of Theoretical Biology 107:603–629.

Felsenstein, J., 1985. Phylogenies and the comparative method. American Naturalist 125:1–15.

Foote, M., 1988. Survivorship analysis of Cambrian and Ordovician Trilobites. Paleobiology 14:258–271.

———, 2000. Origination and extinction components of taxonomic diversity: Paleozoic and post-Paleozoic dynamics. Paleobiology 26:578–605.

Foote, M. and A. I. Miller, 2013. Determinants of early survival in marine animal genera. Paleobiology 39:171–192.

Gelman, A., J. B. Carlin, H. S. Stern, D. B. Dunson, A. Vehtari, and D. B. Rubin, 2013. Bayesian data analysis. 3 ed. Chapman and Hall, Boca Raton, FL.

Gelman, A. and J. Hill, 2007. Data Analysis using Regression and Multilevel/Hierarchical Models. Cambridge University Press, New York, NY.

Harnik, P. G., C. Simpson, and J. L. Payne, 2013. Long-term differences in

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 `http://CRAN.R-project.org/package=raster`. R package version 2.3-24.

Hoffman, M. D. and A. Gelman, 2014. The No-U-Turn Sampler: Adaptively Setting Path Lengths in Hamiltonian Monte Carlo. Journal of Machine Learning Research 15:1351–1381.

Housworth, E. A., P. Martins, and M. Lynch, 2004. The Phylogenetic Mixed Model. The American Naturalist 163:84–96.

Hunt, G., K. Roy, and D. Jablonski, 2005. Species-level heritability reaffirmed: a comment on "On the heritability of geographic range sizes". American Naturalist 166:129–135.

Jablonski, D., 1986. Background and mass extincitons: the alternation of macroevolutionary regimes. Science 231:129–133.

———, 1987. Heritability at the species level: analysis of geographic ranges of cretaceous mollusks. Science 238:360–363.

Jablonski, D. and K. Roy, 2003. Geographical range and speciation in fossil and living molluscs. Proceedings. Biological sciences / The Royal Society 270:401–6.

Johnson, J. G., 1974. Extinction of Perched Faunas. Geology 2:479–482.

Kiessling, W. and M. Aberhan, 2007. Environmental determinants of marine benthic biodiversity dynamics through Triassic Jurassic time. Paleobiology 33:414–434.

Klein, J. P. and M. L. Moeschberger, 2003. Survival Analysis: Techniques for Censored and Truncated Data. 2nd ed. Springer, New York.

Liow, L. H., 2004. A test of Simpson's "rule of the survival of the relatively unspecialized" using fossil crinoids. The American naturalist 164:431–43.

———, 2007. Does versatility as measured by geographic range, bathymetric range and morphological variability contribute to taxon longevity? Global Ecology and Biogeography 16:117–128.

Lynch, M., 1991. Methods for the analysis of comparative data in evolutionary biology. Evolution 45:1065–1080.

Miller, A. I. and M. Foote, 2009. Epicontinental seas versus open-ocean settings: the kinetics of mass extinction and origination. Science 326:1106–9.

Nürnberg, S. and M. Aberhan, 2013. Habitat breadth and geographic range predict diversity dynamics in marine Mesozoic bivalves. Paleobiology 39:360–372.

———, 2015. Interdependence of specialization and biodiversity in Phanerozoic marine invertebrates. Nature communications 6:6602.

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

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

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

24

Peters, S. E., 2008. Environmental determinants of extinction selectivity in the fossil record. Nature 454:626–9.

Raup, D. M., 1975. Taxonomic survivorship curves and Van Valen's Law. Paleobiology 1:82–96.

———, 1978. Cohort Analysis of generic survivorship. Paleobiology 4:1–15.

———, 1991. Extinction: Bad Genes or Bad Luck? Norton, New York.

———, 1994. The role of extinction in evolution. Proceedings of the National Academy of Sciences 91:6758–6763.

Sepkoski Jr., J. J., 1981. A factor analytic description of the Phanerozoic marine fossil record. Paleobiology 7:36–53.

Sheehan, P., 2001. The late Ordovician mass extinction. Annual Review of Earth and Planetary Sciences 29:331–364.

Simpson, C., 2006. Levels of selection and large-scale morphological trends. Ph.D. thesis, University of Chicago.

Simpson, C. and P. G. Harnik, 2009. Assessing the role of abundance in marine bivalve extinction over the post-Paleozoic. Paleobiology 35:631–647.

Simpson, G. G., 1944. Tempo and Mode in Evolution. Columbia University Press, New York.

———, 1953. The Major Features of Evolution. Columbia University Press, New York.

Stan Development Team, 2014. Stan: A c++ library for probability and sampling, version 2.5.0. URL `http://mc-stan.org/`.

Van Valen, L., 1973. A new evolutionary law. Evolutionary Theory 1:1–30.

———, 1979. Taxonomic survivorship curves. Evolutionary Theory 4:129–142.

25

Vilhena, D. A., E. B. Harris, C. T. Bergstrom, M. E. Maliska, P. D. Ward, C. A.

606   Sidor, C. A. E. Strömberg, and G. P. Wilson, 2013. Bivalve network reveals

latitudinal selectivity gradient at the end-Cretaceous mass extinction.

608   Scientific Reports 3:1790.

Wang, S. C., 2003. On the continuity of background and mass extinction.

610   Paleobiology 29:455–467.