

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

Peter D. Smits^{1,*}

1. University of California – Berkeley, Berkeley, California 94720;

* Corresponding author; e-mail: psmits@berkeley.edu

Manuscript elements: Figure 1, figure 2, table 1, online appendices A and B (including figure A1 and figure A2). Figure 2 is to print in color.

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

Manuscript type: Article.

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

Abstract

As average species fitness decreases, do trait-based differences in species' fitness (selection) increase or decrease? Similarly, as taxonomic extinction intensity increases, how do the effects of species traits on taxonomic survival change? Do the selective differences associated with some traits change more than others? 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 strong evidence for a non-linear relationship between environmental preference for epicontinental versus open-ocean environments and expected taxon duration, where taxa with intermediate or no environmental preference are expected to have greater durations than taxa which appear exclusively in either environmental end-member. I find evidence for a correlation between average cohort extinction risk and the effect of geographic range on extinction risk for that cohort where I estimate that as extinction intensity increases that the magnitude of the effect of geographic range on extinction risk increases, thus increasing the difference in extinction risk between small and large ranged taxa. I also find evidence a similar relationship between average cohort extinction risk and the effect of the linear aspect of environmental preference on expected duration, so that cohort's with above average estimates of extinction intensity increases are also expected to have above average estimates of the linear aspect of environmental preference. This result means that for taxa from those cohort's which preferentially appear in epicontinental environments will have a greater expected duration than those taxa which prefer open-ocean environments. These re-

²⁷ sults support the conclusion that, for Paleozoic brachiopods and the effects of geographic range and environmental preference, as extinction intensity increases the magnitude of selection also increases.

Introduction

Extinction is one half of the diversification process (Raup, 1994; Stanley, 1975, 1979), second only to speciation or origination in shaping changes to diversity; it can also be the ultimate
 33 manifestation of selection 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, 2010; Raup,
 1994; Stanley, 1975). Species duration is a measure of species fitness (Cooper, 1984), and trait-
 36 associated differences in fitness is the hallmark of (species) selection (Rabosky and McCune,
 2010).

Jablonski (1986) observed that for bivalves at the end Cretaceous mass extinction event, previous
 39 trait-associated differences in survival no longer mattered except for the case of geographic range. Based on this evidence, Jablonski (1986) proposed the idea of "macroevolutionary modes" and that mass extinction and background extinction are fundamentally different processes. How-
 42 ever, based on estimates of extinction rates over time, there is no evidence of there being two or more "types" of extinction (Wang, 2003). Instead, extinction rates for marine invertebrates are unimodal with continuous variation. The apparent disconnect between the theory of macroevo-
 45 lutionary modes and the observation of continuous variation in extinction rates implies the pos-
 sibility of a relationship between the strength of selection (extinction intensity) and the associa-
 tion between traits and differences in fitness (extinction selectivity) (Payne et al., 2016). As
 48 extinction intensity increases, what happens to extinction selectivity? How do trait-associated differences in fitness change as average extinction rate changes over time?

Here I develop a statistical model describing the relationship between brachiopod taxon du-
 51 rations and multiple functional taxon traits in order to understand the relationship between ex-
 tinction intensity and selectivity over time. Trait-dependent differences in extinction risk should
 be associated with differences in taxon duration (Cooper, 1984; Rabosky and McCune, 2010). Bra-
 54 chiopods are an ideal group for this study as they have an exceptionally complete fossil record (Foote, 2000b; Foote and Raup, 1996). I focus on the brachiopod record from the post-Cambrian

Paleozoic, from the start of the Ordovician (approximately 485 My) through the end Permian (approximately 252 My) as this represents the time of greatest global brachiopod diversity (Alroy, 2010) which results in a large sample size.

The analysis of taxon durations, or time from a taxon's origination till its extinction, falls under the purview of survival analysis, a field of applied statistics commonly used in health care and engineering (Klein and Moeschberger, 2003) but has a long history in paleontology (Crampton et al., 2016; Simpson, 1944, 1953; Smits, 2015; Van Valen, 1973, 1979). I adopt a hierarchical Bayesian modeling approach (Gelman et al., 2013; Gelman and Hill, 2007) in order to unify the previously distinct dynamic and cohort paleontological survival approaches (Baumiller, 1993; Crampton et al., 2016; Ezard et al., 2012; Foote, 1988; Raup, 1975, 1978; Simpson, 2006; Van Valen, 1973, 1979).

In order to test for an association between extinction intensity and extinction selectivity, extinction rate and trait-based differences in extinction rate need to be estimated. Previous work has approached this problem by estimating the extinction intensity and selectivity at different time points or for different origination cohorts independently and then comparing those estimates (Payne et al., 2016). I find this approach problematic for a few reasons. First, by treating each time point or cohort as independent the estimates from each model are not relative to each other and they do not represent the greatest compromise across all available data (Gelman et al., 2013). Second, by treating each time point or cohort as independent any and all post-hoc analyses are at risk of false positive results because of multiple comparisons (Gelman et al., 2013). Third, post-hoc analysis of correlations between maximum likelihood estimates is much less satisfying or coherent than including those correlations as estimates of the initial model, something that is not possible when each time point or cohort is considered independent (Gelman et al., 2013).

While estimation of trait-dependent speciation and extinction rates from phylogenies of extant taxa has grown dramatically (Fitzjohn, 2010; Goldberg et al., 2011, 2005; Maddison et al., 2007; Rabosky et al., 2013; Stadler, 2011, 2013; Stadler and Bokma, 2013), there are two major ways to estimate trait-dependent extinction: analysis of phylogenies, and analysis of the fossil

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 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.

90 *Factors affecting brachiopod survival*

Conceptually, taxon survival can be considered an aspect of “taxon fitness” (Cooper, 1984; Palmer and Feldman, 2012). Traits associated with taxon survival are thus examples of species (or higher-level) selection, as differences in survival are analogous to differences in fitness. The traits analyzed here are all examples of emergent and aggregate traits (Jablonski, 2008; Rabosky and McCune, 2010); specifically I analyze genus-level traits. Emergent traits are those which are not measurable at a lower level (e.g. species versus individual organism) such as geographic range, or even fossil sampling rate. Aggregate traits, like body size or environmental preference, are the average of a shared trait across all members of a lower level.

99 Geographic range is widely considered the most important biological trait for estimating
differences in extinction risk at nearly all times, with large geographic range associated with low
extinction risk (Finnegan et al., 2012; Harnik et al., 2012; Jablonski, 1986, 1987, 2008; Jablonski
102 and Roy, 2003; Payne and Finnegan, 2007). This stands to reason even if extinction is completely
at random; a taxon with an unrestricted range is less likely to go extinct at random than a taxon
with a restricted range.

105 Epicontinental seas are a shallow-marine environment where the ocean has spread over the
continental interior or craton with a depth typically less than 100m. In contrast, open-ocean
coastline environments have much greater variance in depth, do not cover the continental craton,
108 and can persist during periods of low sea level (Miller and Foote, 2009). Because of this, a simple

hypothesis that taxa which favor epicontinental seas would be at great risk during periods of low sea levels, such as during glacial periods, when epicontinental seas are drained. During
111 the Paleozoic (approximately 541–252 My), epicontinental seas were widely spread globally but declined over the Mesozoic (approximately 252–66 My) and have nearly disappeared during the Cenozoic (approximately 66–0 My) as open-ocean coastlines became the dominant shallow-
114 marine setting (Johnson, 1974; Miller and Foote, 2009; Peters, 2008; Sheehan, 2001). Taxa in epicontinental environments could also have a greater extinction susceptibility than taxa in open-ocean environments due to anoxic events due to enhanced water stratification or poor water
117 circulation (Peters, 2007).

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
120 with epicontinental seas. During periods of background extinction, however, they found no consistent difference between taxa favoring either environment. Miller and Foote (2009) hypothesize that open-ocean taxa may have a greater extinction rate because these environments would be
123 more strongly affected by waterborne hazards such as fallout from impacts or volcanic events which would propagate more quickly than in epicontinental environments with sluggish circulation. These two environment types represent the primary identifiable environmental dichotomy
126 observed in ancient marine systems (Miller and Foote, 2009; Sheehan, 2001). Given these findings, I would hypothesize that as extinction risk increases, the extinction risk associated with open-ocean environments should generally increase.

Because environmental preference is defined here as the continuum between occurring exclusively in open-ocean environments versus epicontinental environments, intermediate values are considered “generalists” in the sense that they favor neither end member. A long-standing
129 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; Smits,
132 2015). Because of this, the effect of environmental preference was modeled as a quadratic function where a concave down relationship between preference and expected duration indicates that
135

generalists are favored over specialists end-members.

Body size, measured as shell length, is also considered as a trait that may potentially influence
138 extinction risk (Harnik, 2011; Payne et al., 2014). Body size is a proxy for metabolic activity and other correlated life history traits (Payne et al., 2014). Harnik et al. (2014) analyzed the effect of body size selectivity in Devonian brachiopods in a phylogenetic and non-phylogenetic context;
141 finding that body size was not found to be associated with differences in taxonomic duration. It has also been found that, at least in the case of some bivalve subclades, body size can be as important a factor as geographic range size in determining extinction risk (Harnik, 2011). Given
144 these results, I expect that if body size has any effect on brachiopod taxonomic survival it is very small.

It is well known that, given the incompleteness of the fossil record, the observed duration
147 of a taxon is an underestimate of that taxon's true duration (Alroy, 2014; Foote and Raup, 1996; Liow and Nichols, 2010; Solow and Smith, 1997; Wagner and Marcot, 2013; Wang and Marshall,
2004). Because of this, the concern is that a taxon's observed duration may reflect its relative
150 chance of being sampled and not any of the effects of the covariates of interest. In this case, for sampling to be a confounding factor there must be consistent relationship between the quality of sampling of a taxon and its apparent duration (e.g. greater sampling, longer duration). If there
153 is no relationship between sampling and duration then interpretation can be made clearly; while observed durations are obviously truncated true durations, a lack of a relationship would indicate that the amount and form of this truncation is not a major determinant of the taxon's apparent
156 duration. By including sampling as a covariate in the model, this effect can be quantified and can be taken into account when interpreting the estimates of the effects of the other covariates.

Methods

159 The brachiopod dataset analyzed here was sourced from the Paleobiology Database (<http://www.paleodb.org>) and then filtered to a limited selection of higher taxonomic groups (Rhynchonelliformea: Rhyn-

chonellata, Chileata, Obolellida, Kutorginida, Strophomenida, Spiriferida). Additionally, samples
162 were limited to those taxa present during the post-Cambrian Paleozoic. Temporal, stratigraphic,
and other relevant occurrence information used in this analysis was also downloaded. Analyzed
occurrences were restricted to those with paleolatitude and paleolongitude coordinates, have
165 been assigned to either epicontinental or open-ocean environment, and belonging to a genus
present in the body size dataset (Payne et al., 2014). Epicontinental versus open-ocean assign-
ments for each fossil occurrence are based on those from previous analyses Foote and Miller
168 (2013); Miller and Foote (2009); Ritterbush and Foote (2017). These filtering criteria are very
similar to those from Foote and Miller (2013) with an additional constraint of being present in
the body size data set from Payne et al. (2014). In total, there 1130 genera were included in this
171 analysis.

Fossil occurrences were analyzed at the genus level, a common proactice for paleobiological,
macroevolutionary and macroecological studies of marine invertebrates (Alroy, 2010; Foote and
174 Miller, 2013; Harnik et al., 2012; Kiessling and Aberhan, 2007; Miller and Foote, 2009; Nürnberg
and Aberhan, 2013, 2015; Payne and Finnegan, 2007; Ritterbush and Foote, 2017; Simpson and
Harnik, 2009; Vilhena et al., 2013). While species diversity dynamics are frequently of much
177 greater interest than those of higher taxa (though see Foote 2014; Hoehn et al. 2015), the nature
of the fossil record makes accurate, precise, and consistent taxonomic assignments at the species
level difficult for all occurrences. As such, the choice to analyze genera as opposed to species
180 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
183 because of the inherently discrete nature of the fossil record; dates are not assigned to individual
fossils themselves but instead fossils are assigned to a geological interval which represents some
temporal range. In this analysis, stages are effectively irreducible temporal intervals in which
186 taxa may occur. Genera with a last occurrence in or after Changhsingian stage (e.g. the final
stage of the study interval) were right censored at the Changhsingian.

The covariates of duration included in this analysis are geographic range size (r), environmental preference (v, v^2), shell length (m) and sampling (s).
189

Geographic range was calculated as relative occupancy corrected for incomplete sampling. First, the paleolatitude-paleolongitude coordinates for all occurrences were projected onto an
192 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 km². The map projection and regular lattice were made using shape files from
195 <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 temporal bin, the relative occurrence probability of the observed taxa was calculated using the JADE method developed by Chao et al. (2015). This method accounts for the fact that taxa with an occupancy of 0 cannot be observed which means that occupancy follows a truncated Binomial distribution. This correction is critical when comparing occupancies from different times with different geographic
198 sampling. Finally, for each genus, the mean relative occurrence probability was calculated as the average of that genus' occurrence probabilities for all stages it was sampled to yield relative occupancy, my proxy for geographic range.
201

Environmental preference was defined as probability of observing the ratio of epicontinental
204 occurrences to total occurrences ($\theta_i = e_i/E_i$) or greater given the background occurrence probability θ'_i as estimated from all other taxa occurring at the same time (e'_i/E'_i). This measure of environmental preference is expressed.
207

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

where v is the percent of the distribution defined in equation 1 less than or equal to θ_i . The Beta distribution is used here because it is a continuous distribution bounded at 0 and 1, which is ideal
210 for modeling percentages.

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

213 (Williams et al., 2007).

The sampling probability for individual taxa was calculated using the standard gap 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 duration of the taxon minus two. Subtracting two from the numerator and denominator is because the first and last appearance stages are by definition sampled. Because taxa that were right censored only include a first appearance, one was subtracted from the numerator and denominator instead of two.

216 The minimum duration for which a gap statistic can be calculated is three stages, so I chose to impute the gap statistic for all observations with a duration less than 3. Imputation is the “filling in” of missing observations based on the observed values (Gelman and Hill, 2007; Rubin, 1996). This is fairly straight forward in a Bayesian framework because covariates and parameters are considered random variables, meaning that the missing values of sampling can be modeled as coming from some probability distribution.

219 Prior to analysis, geographic range was logit transformed and the number of samples was natural-log transformed; these transformations make these variables defined for the entire real line. Sampling probability was transformed as $(s(n - 1) + 0.5)/n$ where n is the sample size as recommended by Smithson and Verkuilen (2006); this serves to slightly shrink the range of the data so that there are no values of 0 or 1. All covariates except for sampling were standardized by subtracting the mean from all values and dividing by twice its standard deviation, which follows Gelman and Hill (2007). This standardization means that the associated regression coefficients are comparable as the expected change per 1-unit change in the rescaled covariates. Finally, D is defined as the total number of covariates, excluding sampling, plus one for the intercept term.

Details of model

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

group (e.g. origination cohort). Each group is considered a draw from a shared probability
240 distribution (e.g. prior) of all cohorts, observed and unobserved. The group-level parameters,
or the hyperparameters of this shared prior, are themselves given (hyper)prior distributions and
are also estimated like the other parameters of interest (e.g. covariate effects) (Gelman et al.,
243 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. All covariate effects (regression coefficients), as well as
246 the intercept term (baseline extinction risk), were allowed to vary by group (origination cohort).

The covariance between covariate effects was also modeled.

Genus durations were assumed to follow a Weibull distribution which allows for age-dependent
249 extinction (Klein and Moeschberger, 2003): $y \sim \text{Weibull}(\alpha, \sigma)$. The Weibull distribution has two
parameters: scale σ , and shape α . When $\alpha = 1$, σ is equal to the expected duration of any taxon.
 α is a measure of the effect of age on extinction risk where values greater than 1 indicate that
252 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
 $\alpha = 1$.

255 In the case of the right- and left-censored observations mentioned above, the probability
of those observations has a different calculation (Klein and Moeschberger, 2003). For right-
censored observations, the likelihood is calculated $p(y|\theta) = 1 - F(y) = S(y)$ where $F(y)$ is the
258 cumulative distribution function. Taxa that existed for only a single stage were left-censored,
which implies that that taxon went extinct at any point between 0 and 1 stages. In contrast to
right-censored data, the likelihood of a left-censored observation is calculated from $p(y|\theta) =$
261 $F(y)$. This censoring strategy improves model fit greatly over treating these taxa as being fully
observed (see Appendix).

The scale parameter σ was modeled as a regression following Kleinbaum and Klein (2005)

²⁶⁴ with varying intercept and varying slopes and the effect of sampling; this is expressed

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

where i indexes across all observations where $i = 1, \dots, n$ where n is the total number of observations, $j[i]$ is the cohort membership of the i th observation where $j = 1, \dots, J$ where J is the total number of cohorts, X is a $N \times D$ matrix of covariates along with a column of ones for the intercept term, B is a $J \times D$ matrix of cohort-specific regression coefficients, and δ is the regression coefficient for the effect of sampling s . δ does not vary by cohort.

²⁷⁰ Each of the rows of matrix B are modeled as realizations from a multivariate normal distribution with length D location vector μ and $J \times J$ covariance matrix Σ : $B_j \sim \text{MVN}(\mu, \Sigma)$. The covariance matrix was then decomposed into a length J vector of scales τ and a $J \times J$ correlation matrix Ω , defined $\Sigma = \text{diag}(\tau)\Omega\text{diag}(\tau)$ where “diag” indicates a diagonal matrix.

²⁷³ The elements of μ were given independent normally distributed priors. The effects of geographic range size and the breadth of environmental preference were given informative priors reflecting the previous findings while the others were given weakly informative favoring no effect. The correlation matrix Ω was given an LKJ distributed prior (Lewandowski et al., 2009) that slightly favors an identity matrix as recommended by Team (2017). These priors are defined

$$\begin{aligned} \mu^0 &\sim \mathcal{N}(0, 5) \\ \mu^r &\sim \mathcal{N}(-1, 1) \\ \mu^v &\sim \mathcal{N}(0, 1) \\ \mu^{v^2} &\sim \mathcal{N}(1, 1) \\ \mu^m &\sim \mathcal{N}(0, 0.5) \\ \delta &\sim \mathcal{N}(1) \\ \tau &\sim C^+(1) \\ \Omega &\sim \text{LKJ}(2). \end{aligned} \quad (3)$$

²⁷⁹ The log of the shape parameter α was given a weakly informative prior $\log(\alpha) \sim \mathcal{N}(0, 1)$

centered at $\alpha = 1$, which corresponds to the Law of Constant Extinction (Van Valen, 1973).

Imputation of sampling probability

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

Predicting sampling probability using the other covariate that are then included in the model
291 of duration is acceptable and appropriate in the case of imputation where the sample goal is
accurate prediction (Gelman and Hill, 2007; Rubin, 1996). Not including these covariates can
lead to biased estimates of the imputed variable; if the covariates themselves are related, not
294 including them will bias this correlation towards zero which then leads to improper imputation
and inference (Rubin, 1996).

The Beta regression is defined

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

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

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

and total count parameter

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

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

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

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

Posterior inference and posterior predictive checks

The joint posterior was approximated using a Markov-chain Monte Carlo routine that is a variant of Hamiltonian Monte Carlo called the No-U-Turn Sampler (Hoffman and Gelman, 2014) as implemented in the probabilistic programming language Stan (Stan Development Team, 2014). The posterior distribution was approximated from four parallel chains run for 40,000 steps, split half warm-up and half sampling and thinned to every 20th sample for a total of 4000 posterior samples. Chain convergence was assessed via the scale reduction factor \hat{R} where values close to 1 ($\hat{R} < 1.1$) indicate approximate convergence. Convergence means that the chains are approximately stationary and the samples are well mixed (Gelman et al., 2013). After the model was fit to the data, 100 datasets were simulated from the posterior predictive distribution of the model. These simulations were used to test for adequacy of model fit as described below.

Survival analysis is complicated by censored observations where the ultimate time of extinction for some taxa could not be fully observed during the study window. Importantly, posterior predictive simulations for these observations must be similarly censored. To accomplish this, posterior predictive simulated durations for right-censored observations were the minimum of its final observed duration and the simulated duration. For left-censored individuals, their simulated duration was pegged at a minimum of one stage with simulated values less than one stage set to one.

Model adequacy was evaluated using a series of posterior predictive checks. Posterior predic-

tive checks are a means for understanding model fit or adequacy. The concept of model adequacy
324 is that if our model is an adequate descriptor of the observed data, then data simulated from the posterior predictive distribution should be similar to the observed given the same covariates, etc. (Gelman et al., 2013). Posterior predictive checks can take many forms but the basic idea
327 is to compare some property of the empirical data to that property estimated from each of the simulated datasets. Additionally, for structured datasets like the one analyzed here, the fit of the model to different parts of the data can be assessed, which in turn can reveal a great deal if the
330 model has good fit to some aspects of data but not others; it is in these scenarios when knowledge about the biology, geology, and paleoenvironment becomes important in order to explain what unmodeled processes might lead to these discrepancies between our data and the model
333 (Gelman et al., 2013).

The types of posterior predictive tests used in this analysis fall into two categories: comparison of observed mean and median genus duration to a distribution of mean and median genus
336 duration estimates from the posterior simulations, comparison of a non-parameteric estimate of the survival function from the observed data to estimates of that same survival function from the simulations.

339 The survival function describes the probability of a taxon persisting given that it has survived up to time t ; this is expressed $P(T \geq t)$ because T is the true extinction time of the species and t is some arbitrary time of observation and we are estimating that probability that t is less than T . It
342 is important to note, however, that the survival function does not reflect density of observations unlike e.g. histograms. Instead, this posterior predictive check reflects the model's ability to predict genus survival.

345 These posterior predictive tests were done for the entire data and for each of the origination cohorts.

Results

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

Comparisons between the observed distribution of durations to the distributions of 100 sim-
ulated datasets reveals the relatively good but heterogeneous fit of the model to the data (Fig. 1).

354 The two major aspects of possible misfit that are observable are at at durations of 2-3 stages. The
model slightly under-estimates the number of observations with duration of 2 or 3 stages. The
goal of this model is estimating the expected duration of a genus given its covariate information.
357 While the model estimates are not exact, it is possible that our model fits the bulk of our data
well but does poorly towards the tails.

Comparisons between the survival functions estimated from the empirical data and from 100
360 simulated datasets further expands the picture of model adequacy (Fig. 2). The survival curves of
the 100 simulated datasets are very similar to the survival function estimated from the empirical
data. The major points of potential misfit between the model and the data are underestimated
363 percentage of taxa with duration 1 stage, and an over-estimate of probability of species surviving
at least 10-13 stages. Importantly, the major divergence between the observed and estimated
applies to taxa with a less than 15% probability of continuing to surviving. Keep in mind, also,
366 that the survival curve as presented does not depict density of observations as in Fig. 1.

In addition to distributional comparisons, model adequacy at the total data level was assessed
through comparison of the mean and median of the observed data to those from simulated data
369 sets. While the previous posterior predictive checks have focused on the relatively good but
heterogeneous fit of the model to the entire distribution of the data, the fitted model's ability
to predict the mean and median of the observed data appears adequate (Fig. 3a, 3b). Because
372 the principle goal of this model is to obtain adequate prediction of how a taxon's expected

duration for a given set of ecological covariates, the seemingly adequate fit of our model to mean taxon duration is reassuring (Fig. 3a). Additionally, given the skewness of the observed taxon durations (Fig. 1, the ability for the model to closely recapitulates the median observed taxon duration points to the overall good fit of the model to the data.

When considered together, all of the above posterior predictive checks indicate approximately adequate model fit for key questions such as expected taxon duration (Fig. 3a). However, there is obviously heterogeneity in model fit because, while the model can recapitulate some aspects of the observed data (Fig. 3a, 3b), there are obvious discrepancies between the model and the data (Fig. 1, 2). By performing the same posterior predictive tests for each of the origination cohorts, it may be possible to get a better picture of the sources of model misfit.

When the posterior predictive tests are visualized for each of the origination cohorts, a complex picture of model fit emerges. Comparison between the empirical survival functions estimated for each cohort to those estimated from the simulated datasets reveals the degree of heterogeneity in model fit (Fig. 4, as some origination cohorts appear to be very well fit by the model (e.g. Tremadoc, Darriwilian Wenlock, Ludlow, Lochkovian, etc.) while others are more poorly fit by the model (e.g. Tournaisian, Visean, Bashirian, Moscovian, Stephanian, Asselian, etc.). The poverty of model fit to some origination cohorts may indicate that these cohorts are undergoing a different extinction process whose aspects are unmodeled in this analysis. However, for those cohorts where the model recapitulates the empirical survival function it is likely that the model may be capturing some of the processes underlying taxon extinction.

For nearly every origination cohort, the model is able to approximately recapitulate the observed mean duration (Fig. 5). In comparison, the model has a much more heterogeneous fit to each origination cohort's median taxon duration (Fig. 6). The skewness of the distribution underlying (Fig. 1) means that for some origination cohorts, median duration might be pegged at 1 stage; this means that the posterior predictive distributions for some cohorts can be extremely skewed. The cohorts with notably inadequate poor predictions are the Hirnantian, Priodoli, Emsian, Eifelian, Givetian, Asselian, Artinskian, and Roadian. The remaining cohorts, however, have

approximately adequate fit. These results indicate that our model is very good at recapitulate mean taxon duration (Fig. 3a, 5), and it is capable of estimating overall median duration and
402 median duration of most origination cohorts (Fig. 3b, 6).

Geographic range is estimated to have, on average, a positive effect on survival (Table 1). The cohort-level estimate of the effect of geographic range size indicates that as a taxon's geographic
405 range increases, that taxon's duration is expected to increase (Table 1). Given the estimates of μ^r and τ^r , there is an approximately XX% (\pm XX% SD) probability that this relationships would be reversed ($\text{Pr}(\mathcal{N}(\mu^r, \tau^r) > 0)$).

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

The group-level estimate of the effect of environmental preference is estimated from μ^v and
411 μ^{v^2} . The estimate of μ^v indicates that taxa which slightly prefer epicontinental environments to open-ocean environments are expected to have a greater duration than open-ocean favoring
414 taxa (Table 1). Additionally, given the estimate of between-cohort variance τ^v , there is approximately XX% (\pm XX% SD) probability that, for any given cohort, taxa which favor open-ocean environments would have a greater expected duration than taxa which favor epicontinental environments ($\text{Pr}(\mathcal{N}(\mu^v, \tau^v) > 0)$). The estimate of μ^{v^2} indicates that the overall relationship between
417 environmental preference and $\log(\sigma)$ is concave down (Fig. 7), with only a XX% (\pm XX% SD) probability that any given cohort is convex up ($\text{Pr}(\mathcal{N}(\mu^{v^2}, \tau^{v^2}) < 0)$).

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

In chronological order, the Wenlock (433.4-427.4 Mya) is the first origination cohort for which
426 at least two effect estimates diverge from their group-level averages: the effect of the linear

aspect of environmental preference β^v increases (shifting the parabola left-ward) while the effect of average body size β^m decreases meaning that greater than average size is associated with a
429 longer duration than taxa with less than average body sizes (Fig. 8). Additionally, this cohort is well fit by our model (Fig. 4, 5, 6) which allows us to have greater confidence in these results. This cohort is one of the few for which the estimated effect of body size is different from the
432 group-level average.

The effects of the linear and quadratic aspects of environmental preference, β^v and β^{v^2} , for taxa originating in the Frasnian are divergent from the overall group-level mean. For this cohort,
435 β^v is estimated to be less than expected while β^{v^2} is estimated to be greater than expected. This result means that the curvature of the relationship increased while the apex was right-shifted from average. Additionally, this origination cohort is adequately fit by the model (Fig. 4, 5, 6)
438 which increases confidence in interpreting these results.

Taxa originating during the Famennian (372.2-358.9 Mya) were estimated to have a lower than average expected duration than the group-level average (i.e. greater than average β^0) as
441 well as lower than average estimate of β^v (i.e. right-shifted) (Fig. 8). Finally, this cohort has only moderate quality of fit as the model is able to estimate average cohort duration (Fig. 5) but has difficulty reproducing the survival function and median duration (Fig. 4, 6), meaning that these
444 excursions from the group-level mean may be driven by the difference in processes underlying extinction of taxa in this origination cohort.

Two of the effect estimates for the Sakmarian origination cohort (295.5-290.1 Mya) are divergent from the group-level averages (Fig. 8). Specifically, β^v has an above average estimate indicating a left-shifted relationship between environmental preference and duration, and finally
447 β^m was estimated to be below the group-level average which means that taxa with greater than average body sizes are expected to have greater durations than taxa originating in that cohort
450 with below average body size. The model has a moderate fit to the data, reproducing the mean duration (Fig. 5) while slightly under estimating median durations and not fully recapitulating
453 the survival pattern of the observed data (Fig. 4, 6).

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

The Wenlock, Asseilian, and Sakmarian cohorts all have estimates of the linear aspect of environmental preference, β^v , that are greater than the group-level average. Additionally, this is the only aspect of the relationship between environmental preference and expected taxon duration for these cohorts than is divergent from the group-level means (Fig. 8). These estimates mean that the relationship between environmental preference and duration is right-shifted such that taxa which prefer epicontinental environments over open-ocean are expected to have greater durations than average. Specifically, this means that taxa that can tolerate open-ocean and epicontinental environments but preferentially occur in epicontinental environments are expected to have greater durations than those that occur equally in these two environments (Fig. 9).

The Tremadocian, Frasnian, and Famennian origination cohorts are all estimated to have lower than average estimates of the linear aspect of environmental preference (β^v) (Fig. 8). This change to β^v manifests as a left-shifted relationship between environmental preference and duration such that the apex of the relationship is closer to 0 which means that there is no “advantage” for favoring either an epicontinental and open-ocean environment; this is slightly different from the group-level average relationship which has an approximate 64% probability of the apex being right-shifted so that taxa which slightly favor epicontinental environments are expected to have a greater duration than those that favor open-ocean environments (Fig. 7). Keep in mind, however, there is an approximately XX% probability that any cohort would have a concave down relationship between environmental preference and duration which means that taxa which favor purely open-ocean or epicontinental environments are expected to have shorter durations than

taxa which can exist in either environment (Fig. 9).

Frasnian, Visean, Kungurian origination cohorts are all estimated to have greater than average estimates of the quadratic aspect of the relationship between environmental preference and duration β^v ² (Fig. 8). This relationship means that the relationship between taxon environmental preference is more peaked than expected from the group-level average (Fig. 9), where taxa of intermediate preference are expected to have an even greater duration than those taxa which exclusively prefer a single environment than would be expected on average.

In contrast, the Givetian and Roadian origination cohorts are estimated to have lower than average estimates of the quadratic aspect of the relationship between environmental preference and duration β^v ² (Fig. 8). This result means that the relationship between taxon environmental preference has almost entirely linear as opposed to the peaked relationship expected from the group-level average (Fig. 9).

There is an approximately XX% probability that cohort estimates of β^0 and β^r are negatively correlated, with median estimate of correlation being -0.36. This result means that for any cohort, we would expect that if extinction intensity increases (β^0 increases), the estimated effect of geographic range on duration increases (β^r decreases). This result is strong evidence for a relationship between intensity and selectivity with respect to geographic range size.

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

Correlations between the non-intercept estimates reflect potential similarities in selective pressures between cohorts, however there is only weak evidence of any potential cross-correlations in cohort-specific covariate effects. There is an approximate XX% probability that β^r and β^v are positively correlated. This lack of cross-correlation may be due in part to the higher between-cohort variance of the effect of environmental preference τ^v than the very small between-cohort variance

in the effect of geographic range τ^r (Table 1); the effect of geographic range might simply not vary enough relative to the much noisier environmental preference.

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

Sampling was found to have a negative effect (positive δ) on duration: greater sampling, shorter duration (Table 1). While potentially counter intuitive, this result is most likely due to some long lived taxa only be sampled in the stages of the first and last appearance. Also, longer lived taxa have more opportunities to not be sampled than shorter lived taxa. These two factors will lead to this result.

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

could potentially be caused by the minimum resolution of the fossil record (Sepkoski, 1975). It is thus unclear if a strong biological inference can be made from this result, which means that
537 further work is necessary on the effect of taxon age on extinction risk.

Discussion

The generating observation behind this study was that for bivalves at the end Cretaceous mass
540 extinction event, the only biological trait that was found to affect extinction risk was geographic range while traits that had previously been associated with difference in duration had no effect (Jablonski, 1986). This observation raises two linked questions: how does the effect of geographic
543 range change with changing extinction intensity, and how does the effect of other biological traits change with changing extinction intensity?

I find that as intensity increases (β^0 increases), the magnitude of the effect of geographic range
546 increases (β^r decreases). I also find that as intensity increases, the difference in survival for taxa favoring epicontinental environments over open-ocean environments is expected to decrease; this is consistent with the results of Miller and Foote (2009). Finally, there is no evidence for
549 a correlation between the effects of geographic range and environmental preference on taxon duration.

I find consistent support for the “survival of the unspecialized,” with respect to epicontinen-
552 tal versus open-ocean environmental preference, as a time-invariant generalization of brachiopod survival (Simpson, 1944). Taxa with intermediate environmental preferences are expected to have lower extinction risk than taxa specializing in either epicontinental or open-ocean environments
555 (Fig. 7), though the curvature of the relationship varies from rather shallow to very peaked (Fig. 9). However, this relationship is not symmetric about 0, as taxa favoring epicontinental environments are expected with approximately 75% probability to have a greater duration than taxa
558 favoring open-ocean environments. This description of environment preference only describes one major aspect of a taxon’s environmental context, with factors such as bathymetry and tem-

perature being further descriptors of a taxon's adaptive zone (Harnik, 2011; Harnik et al., 2012; 561 Heim and Peters, 2011; Nürnberg and Aberhan, 2013); inclusion of these factors in future analyses would potentially improve our understanding the extent and complexity of the "survival of the unspecialized" hypothesis as it applies to all dimensions of an adaptive zone.

564 Hopkins et al. (2014), in their analysis of niche conservatism and substrate lithological preference in marine invertebrates, found that brachiopods were among the least "conservative" groups; taxa were found to easily change substrate preference on short time scales. While substrate 567 preference is not the same as environmental preference (as defined here), a question does arise: are there three classes of environmental preference instead of two? These classes would be taxa with broad tolerance ("true" generalists), inflexible specialists ("true" specialists), and 570 flexible but with a narrow tolerance. A flexible taxon is one with a narrow habitat preference at one time, but with preference that changes over time with changing environmental availability. My analysis assumes that traits are constant over the duration of the taxon meaning that this 573 scenario is not detectable; taxa with broad tolerances and flexible taxa with narrow per-stage preference end up being treated the same way. Future work should explore how environmental preference changes over lineage duration in relation to environmental availability to estimate if 576 the generalists–specialists continuum is actually ternary relationship.

An alternative approach for specifically modeling survival that can take into account imperfect observation than the method used here is the Cormack-Jolly-Seber (CJS) model (Liow et al., 579 2008; Liow and Nichols, 2010; 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 as the probability of surviving from time t to time $t + 1$. Additionally, the effect of preservation 582 and sighting is estimated as probability of observing a taxon 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 my analysis: I am estimating the biasing effect of sampling probability on 585 taxon duration to then compare with effects of other covariates, while the CJS model estimates the pre-sampling fossil record and then estimates per-time unit survival probability.

The use of genera as the unit of the study and how to exactly interpret the effects of the biological traits is an important question. For example, if any of the traits analyzed here are associated with increases in speciation rates, this might increase the duration of genera through self-renewal (Raup, 1991, 1994), which 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 decrease generic level extinction risk by that trait increasing species level origination rate instead of decreasing species level extinction risk.

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

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

With significant updates, it would also be possible to compare the brachiopod record with Moden groups such as bivalves or gastropods (Sepkoski, 1981), though remembering that the groups may not necessarily share all cohorts with the brachiopods. This particular model expansion would act as a test of any universal cross-taxonomic patterns in the effects of emergent traits on extinction such as has been proposed for geographic range (Payne and Finnegan, 2007). Additionally, this expanded model would also act as a test of the distinctness of the Sepkoski (1981) three-fauna hypothesis in terms of trait-dependent extinction.

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

The combination and integration of the phylogenetic comparative and paleontological approaches requires both sources of data, something which is not possible for this analysis because there is no phylogenetic hypothesis for all Paleozoic taxa, something that is frequently the case for marine invertebrates with a good fossil record. When both data sources are available has been possible, however, the analysis can more fully address the questions of interest in macroevolution (Fritz et al., 2013; Harnik et al., 2014; Raia et al., 2012, 2013; Simpson et al., 2011; Slater, 2013, 2015; Slater et al., 2012; Smits, 2015; Tomiya, 2013).

Conclusion

In summary, patterns of Paleozoic brachiopod survival were analyzed using a fully Bayesian hierarchical survival modelling approach while also eschewing the traditional separation between background and mass extinction. I find that cohort extinction intensity is negatively correlated with the cohort-specific effects of geographic range and environmental preference. These results imply that as extinction intensity increases (β^0) increases, it is expected that these effects will increase in magnitude. Taken together, these results point to a potential macroevolutionary

639 mechanism behind differences in trait-based survival during and away from mass extinctions
due to a correlations between intensity and selectivity. Additionally, I find support for greater
survival in environmental generalists over specialists in all origination cohorts analyzed; this is
642 consistent with the long standing “survival of the unspecialized” hypothesis (Baumiller, 1993;
Liow, 2004, 2007; Nürnberg and Aberhan, 2013, 2015; Simpson, 1944, 1953; Smits, 2015). The re-
sults of this analysis support the conclusion that for Paleozoic brachiopods, as extinction intensity
645 increases overall extinction selectivity is expected to increase as well.

Acknowledgments

I would like to thank K. Angielczyk, M. Foote, P. D. Polly, R. Ree, and G. Slater for helpful
648 discussion during the conception of this study. I'd also like to thank D. Bapst, N. Pierrehumbert
and M. Villarosa Garcia for additional comments. An additional thank you to A. Miller for the
epicontinental versus open-ocean assignments. Finally, thank you to P. Harnik, L. L. Hsiang, G.
651 Simpson, and the other anonymous reviewers for their comments at various points during this
manuscripts life that improved this manuscript. This entire study would would not have been
possible without the Herculean effort of the many contributors to the Paleobiology Database. In
654 particular, I would like to thank J. Alroy, M. Aberhan, D. Bottjer, M. Clapham, F. Fürsich, N.
Heim, A. Hendy, S. Holland, L. Ivany, W. Kiessling, B. Kröger, A. McGowan, T. Olszewski, P.
Novack-Gottshall, M. Patzkowsky, M. Uhen, L. Villier, and P. Wager. This work was supported
657 by a NASA Exobiology grant (NNX10AQ446) to A. Miller and M. Foote. I declare no conflicts
of interest. This is Paleobiology Database publication XXX. The code necessary for reproducing
these results is available as a Zenodo archive DOI 10.5281/zenodo.46928.

Literature Cited

- 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.
- 663 Alroy, J. 2010. The Shifting Balance of Diversity Among Major Marine Animal Groups. *Science* 329:1191–1194.
- . 2014. A simple Bayesian method of inferring extinction. *Paleobiology* 40:584–607.
- 666 Baumiller, T. K. 1993. Survivorship analysis of Paleozoic Crinoidea: effect of 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. Inouye. 2015. Unveiling 669 the species-rank abundance distribution by generalizing the Good-Turing sample coverage theory. *Ecology* 96:1189–1201.
- Cooper, W. S. 1984. Expected time to extinction and the concept of fundamental fitness. *Journal 672 of Theoretical Biology* 107:603–629.
- Crampton, J. S., R. A. Cooper, P. M. Sadler, and M. Foote. 2016. Greenhouse–icehouse transition in the Late Ordovician marks a step change in extinction regime in the marine plankton. 675 *Proceedings of the National Academy of Sciences* 113:1498–1503.
- Ezard, T. H. G., P. N. Pearson, T. Aze, and A. Purvis. 2012. The meaning of birth and death (in macroevolutionary birth-death models). *Biology Letters* 8:139–142.
- 678 Felsenstein, J. 1985. Phylogenies and the comparative method. *The American Naturalist* 125:1–15.
- Finnegan, S., N. A. Heim, S. E. Peters, and W. W. Fischer. 2012. Climate change and the selective 681 signature of the Late Ordovician mass extinction. *Proceedings of the National Academy of Sciences* 109:6829–6834.

- Finnegan, S., J. L. Payne, and S. C. Wang. 2008. The Red Queen revisited: reevaluating the age selectivity of Phanerozoic marine genus extinctions. *Paleobiology* 34:318–341.
- 684 Fitzjohn, R. G. 2010. Quantitative Traits and Diversification. *Systematic Biology* 59:619–633.
- Foote, M. 1988. Survivorship analysis of Cambrian and Ordovician Trilobites. *Paleobiology* 14:258–271.
- 687 ———. 2000a. Origination and extinction components of taxonomic diversity: general problems. *Paleobiology* 26:74–102.
- . 2000b. Origination and extinction components of taxonomic diversity: Paleozoic and 690 post-Paleozoic dynamics. *Paleobiology* 26:578–605.
- . 2014. Environmental controls on geographic range size in marine animal genera. *Paleobiology* 40:440–458.
- 693 Foote, M., and A. I. Miller. 2013. Determinants of early survival in marine animal genera. *Paleobiology* 39:171–192.
- Foote, M., and D. M. Raup. 1996. Fossil preservation and the stratigraphic ranges of taxa. *Paleobiology* 22:121–140.
- 696 Fritz, S. A., J. Schnitzler, J. T. Eronen, C. Hof, K. Böhning-Gaese, and C. H. Graham. 2013. Diversity in time and space: wanted dead and alive. *Trends in Ecology & Evolution* 28:509–516.
- 699 Gelman, A., J. B. Carlin, H. S. Stern, D. B. Dunson, A. Vehtari, and D. B. Rubin. 2013. Bayesian data analysis. 3rd ed. Chapman and Hall, Boca Raton, FL.
- Gelman, A., and J. Hill. 2007. Data Analysis using Regression and Multilevel/Hierarchical Models. Cambridge University Press, New York, NY.
- 702 Goldberg, E. E., L. T. Lancaster, and R. H. Ree. 2011. Phylogenetic inference of reciprocal effects between geographic range evolution and diversification. *Systematic Biology* 60:451–465.

- 705 Goldberg, E. E., K. Roy, R. Lande, and D. Jablonski. 2005. Diversity, endemism, and age distributions in macroevolutionary sources and sinks. *The American Naturalist* 165:623–633.
- Harnik, P. G. 2011. Direct and indirect effects of biological factors on extinction risk in fossil bivalves. *Proceedings of the National Academy of Sciences* 108:13594–13599.
- 708 Harnik, P. G., P. C. Fitzgerald, J. L. Payne, and S. J. Carlson. 2014. Phylogenetic signal in extinction selectivity in Devonian terebratulide brachiopods. *Paleobiology* 40:675–692.
- 711 Harnik, P. G., C. Simpson, and J. L. Payne. 2012. Long-term differences in extinction risk among the seven forms of rarity. *Proceedings of the Royal Society B: Biological Sciences* 279:4969–4976.
- Heim, N. A., and S. E. Peters. 2011. Regional environmental breadth predicts geographic range 714 and longevity in fossil marine genera. *PloS one* 6:e18946.
- Hijmans, R. J. 2015. raster: Geographic data analysis and modeling. R package version 2.3-24.
- Hoehn, K. B., P. G. Harnik, and V. L. Roth. 2015. A framework for detecting natural selection 717 on traits above the species level. *Methods in Ecology and Evolution* pages doi: 10.1111/2041—210X.12461.
- Hoffman, M. D., and A. Gelman. 2014. The No-U-Turn Sampler: Adaptively Setting Path Lengths 720 in Hamiltonian Monte Carlo. *Journal of Machine Learning Research* 15:1351–1381.
- Hopkins, M. J., C. Simpson, and W. Kiessling. 2014. Differential niche dynamics among major marine invertebrate clades. *Ecology Letters* 17:314–323.
- 723 Hunt, G., and D. L. Rabosky. 2014. Phenotypic Evolution in Fossil Species: Pattern and Process. *Annual Review of Earth and Planetary Sciences* 42:421–441.
- Hunt, G., K. Roy, and D. Jablonski. 2005. Species-level heritability reaffirmed: a comment on "On 726 the heritability of geographic range sizes". *The American Naturalist* 166:129–135.

- Jablonski, D. 1986. Background and mass extincitons: the alternation of macroevolutionary regimes. *Science* 231:129–133.
- 729 ———. 1987. Heritability at the species level: analysis of geographic ranges of cretaceous mollusks. *Science* 238:360–363.
- . 2007. Scale and hierarchy in macroevolution. *Palaeontology* 50:87–109.
- 732 ———. 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 living molluscs. *Proceedings of the Royal Society B: Biological Sciences* 270:401–406.
- 735 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.
- 738 Klein, J. P., and M. L. Moeschberger. 2003. *Survival Analysis: Techniques for Censored and Truncated Data*. 2nd ed. Springer, New York.
- 741 Kleinbaum, D. G., and M. Klein. 2005. *Survival analysis: a self-learning text*. 2nd ed. Springer, New York, NY.
- Lewandowski, D., D. Kurowicka, and H. Joe. 2009. Generating random correlation matrices based on vines and extended onion method. *Journal of Multivariate Analysis* 100:1989–2001.
- 744 Liow, L. H. 2004. A test of Simpson's "rule of the survival of the relatively unspecialized" using fossil crinoids. *The American Naturalist* 164:431–443.
- 747 ———. 2007. Does versatility as measured by geographic range, bathymetric range and morphological variability contribute to taxon longevity? *Global Ecology and Biogeography* 16:117–128.

- Liow, L. H., M. Fortelius, E. Bingham, K. Lintulaakso, H. Mannila, L. Flynn, and N. C. Stenseth.
750 2008. Higher origination and extinction rates in larger mammals. *Proceedings of the National Academy of Sciences* 105:6097–6102.
- Liow, L. H., and J. D. Nichols. 2010. Estimating rates and probabilities of origination and ex-
753 tinction using taxonomic occurrence data: Capture-mark-recapture (CMR) approaches. Pages
81–94 in J. Alroy and G. Hunt, eds. *Quantitative Methods in Paleobiology*. The Paleontological
Society.
- 756 Liow, L. H., T. B. Quental, and C. R. Marshall. 2010. When can decreasing diversification rates be
detected with molecular phylogenies and the fossil record? *Systematic biology* 59:646–659.
- Maddison, W. P., P. E. Midford, and S. P. Otto. 2007. Estimating a binary character's effect on
759 speciation and extinction. *Systematic Biology* 56:701.
- Miller, A. I., and M. Foote. 2009. Epicontinental seas versus open-ocean settings: the kinetics of
mass extinction and origination. *Science* 326:1106–1109.
- 762 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 inverte-
brates. *Nature Communications* 6:6602.
765
- Palmer, M. E., and M. W. Feldman. 2012. Survivability is more fundamental than evolvability.
PLoS one 7:e38025.
- 768 Payne, J. L., A. M. Bush, E. T. Chang, N. A. Heim, M. L. Knope, and S. B. Pruss. 2016. Extinc-
tion intensity, selectivity and their combined macroevolutionary influence in the fossil record.
Biology Letters 12:20160202.
- 771 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* 104:10506–10511.
- 774 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: Biological Sciences* 281:20133122.
- 777 Peters, S. E. 2007. The problem with the Paleozoic. *Paleobiology* 33:165–181.
- . 2008. Environmental determinants of extinction selectivity in the fossil record. *Nature* 454:626–629.
- 780 Quental, T. B., and C. R. Marshall. 2009. Extinction during evolutionary radiations: reconciling the fossil record with molecular phylogenies. *Evolution* 63:3158–3167.
- . 2013. How the Red Queen Drives Terrestrial Mammals to Extinction. *Science* 341:290–292.
- 783 Rabosky, D. L. 2010. Extinction rates should not be estimated from molecular phylogenies. *Evolution* 64:1816–1824.
- Rabosky, D. L., and A. R. McCune. 2010. Reinventing species selection with molecular phylogenies. *Trends in Ecology & Evolution* 25:68–74.
- 786 Rabosky, D. L., F. Santini, J. Eastman, S. a. Smith, B. Sidlauskas, J. Chang, and M. E. Alfaro. 2013. Rates of speciation and morphological evolution are correlated across the largest vertebrate radiation. *Nature Communications* 4:1–8.
- Raia, P., F. Carotenuto, F. Passaro, D. Fulgione, and M. Fortelius. 2012. Ecological specialization in fossil mammals explains Cope's rule. *The American Naturalist* 179:328–337.
- 792 Raia, P., F. Carotenuto, F. Passaro, P. Piras, D. Fulgione, L. Werdelin, J. Saarinen, and M. Fortelius. 2013. Rapid action in the Palaeogene, the relationship between phenotypic and taxonomic diversification in Coenozoic mammals. *Proceedings of the Royal Society B: Biological Sciences* 280:20122244.

- Raup, D. M. 1975. Taxonomic survivorship curves and Van Valen's Law. *Paleobiology* 1:82–96.
- . 1978. Cohort Analysis of generic survivorship. *Paleobiology* 4:1–15.
- 798 ———. 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.
- 801 Ritterbush, K. A., and M. Foote. 2017. Association between geographic range and initial survival of Mesozoic marine animal genera: Circumventing the confounding effects of temporal and taxonomic heterogeneity. *Paleobiology* 43:209–223.
- 804 Rosindell, J., L. J. Harmon, and R. S. Etienne. 2015. Unifying ecology and macroevolution with individual-based theory. *Ecology Letters* 18:472–482.
- Royle, J. A., and R. M. Dorazio. 2008. *Hierarchical modeling and inference in ecology: the analysis of data from populations, metapopulations and communities*. Elsevier, London.
- 807 Rubin, D. B. 1996. Multiple imputation after 18+ years. *Journal of the American Statistical Association* 91:473–489.
- 810 Sepkoski, J. J. 1975. Stratigraphic biases in the analysis of taxonomic survivorship. *Paleobiology* 1:343–355.
- . 1981. A factor analytic description of the Phanerozoic marine fossil record. *Paleobiology* 7:36–53.
- 813 Sheehan, P. M. 2001. The late Ordovician mass extinction. *Annual Review of Earth and Planetary Sciences* 29:331–364.
- 816 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
819 over the post-Paleozoic. *Paleobiology* 35:631–647.
- Simpson, C., W. Kiessling, H. Mewis, R. C. Baron-Szabo, and J. Müller. 2011. Evolutionary
820 diversification of reef corals: a comparison of the molecular and fossil records. *Evolution*
822 65:3274–3284.
- 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.
- Slater, G. J. 2013. Phylogenetic evidence for a shift in the mode of mammalian body size evolution
825 at the Cretaceous-Palaeogene boundary. *Methods in Ecology and Evolution* 4:734–744.
- . 2015. Iterative adaptive radiations of fossil canids show no evidence for diversity-
828 dependent trait evolution. *Proceedings of the National Academy of Sciences* 112:4897–4902.
- Slater, G. J., L. J. Harmon, and M. E. Alfaro. 2012. Integrating fossils with molecular phylogenies
improves inference of trait evolution. *Evolution* 66:3931–3944.
- Smithson, M., and J. Verkuilen. 2006. A better lemon squeezer? Maximum-likelihood regression
831 with beta-distributed dependent variables. *Psychological Methods* 11:54–71.
- Smits, P. D. 2015. Expected time-invariant effects of biological traits on mammal species duration.
834 *Proceedings of the National Academy of Sciences* 112:13015–13020.
- Solow, A. R., and W. Smith. 1997. On fossil preservation and the stratigraphic ranges of taxa.
Paleobiology 23:271–277.
- Stadler, T. 2011. Mammalian phylogeny reveals recent diversification rate shifts. *Proceedings of
837 the National Academy of Sciences* 108:6187–6192.
- . 2013. Recovering speciation and extinction dynamics based on phylogenies. *Journal of
840 Evolutionary Biology* 26:1203–1219.

- Stadler, T., and F. Bokma. 2013. Estimating speciation and extinction rates for phylogenies of higher taxa. *Systematic biology* 62:220–230.
- 843 Stan Development Team. 2014. Stan: A c++ library for probability and sampling, version 2.5.0.
- Stanley, S. M. 1975. A theory of evolution above the species level. *Proceedings of the National Academy of Sciences* 72:646–650.
- 846 ———. 1979. Macroevolution: pattern and process. W. H. Freeman, San Francisco.
- Team, S. D. 2017. Stan Modeling Language Users Guide and Reference Manual.
- Tomiya, S. 2013. Body Size and Extinction Risk in Terrestrial Mammals Above the Species Level.
- 849 The American Naturalist 182:196–214.
- Van Valen, L. 1973. A new evolutionary law. *Evolutionary Theory* 1:1–30.
- . 1979. Taxonomic survivorship curves. *Evolutionary Theory* 4:129–142.
- 852 Vilhena, D. A., E. B. Harris, C. T. Bergstrom, M. E. Maliska, P. D. Ward, C. A. Sidor, C. A. E. Strömberg, and G. P. Wilson. 2013. Bivalve network reveals latitudinal selectivity gradient at the end-Cretaceous mass extinction. *Scientific Reports* 3:1790.
- 855 Wagner, P. J., and G. F. Estabrook. 2014. Trait-based diversification shifts reflect differential extinction among fossil taxa. *Proceedings of the National Academy of Sciences* 111:16419–16424.
- Wagner, P. J., and J. D. Marcot. 2013. Modelling distributions of fossil sampling rates over time, space and taxa: assessment and implications for macroevolutionary studies. *Methods in Ecology and Evolution* 4:703–713.
- Wang, S. C. 2003. On the continuity of background and mass extinction. *Paleobiology* 29:455–467.
- 861 Wang, S. C., and C. R. Marshall. 2004. Improved confidence intervals for estimating the position of a mass extinction boundary. *Paleobiology* 30:5–18.

Williams, C., C. H. C. Brunton, and S. J. Carlson. 2007. Treatise on invertebrate paleontology. Part
864 H, Brachiopoda. Geological Society of America, Boulder, Colorado.

Tables

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

Category	Parameter	Effect of...	Mean	SD	10%	50%	90%
	μ^i	intercept	-2.34	0.14	-2.52	-2.34	-2.16
	μ^r	geographic	1.19	0.21	0.93	1.19	1.45
Mean		range					
	μ^v	environmental	-0.34	0.20	-0.59	-0.34	-0.07
		preference					
	μ^{v^2}	environmental	2.15	0.30	1.76	2.15	2.53
		preference ²					
	μ^m	body size	0.05	0.11	-0.09	0.05	0.19
	τ^i	intercept	0.51	0.11	0.38	0.50	0.65
	τ^r	geographic	0.77	0.17	0.55	0.75	1.00
Standard deviation		range					
	τ^v	environmental	0.97	0.18	0.76	0.95	1.20
		preference					
	τ^{v^2}	environmental	1.08	0.34	0.66	1.07	1.52
		preference ²					
	τ^m	body size	0.41	0.11	0.27	0.41	0.56
Other	δ	sampling	-2.77	0.15	-2.97	-2.77	-2.58
	α	ageing	1.61	0.05	1.54	1.61	1.67

Note: These parameters are the group-level estimates of the effects of biological traits on brachiopod generic survival, the standard deviation of the between-cohort effects, as well as the estimates of the effect of sampling δ and the Weibull shape parameter α . The mean, standard deviation (SD), 10th, 50th, and 90th quantiles of the marginal posteriors are presented.

Figure legends

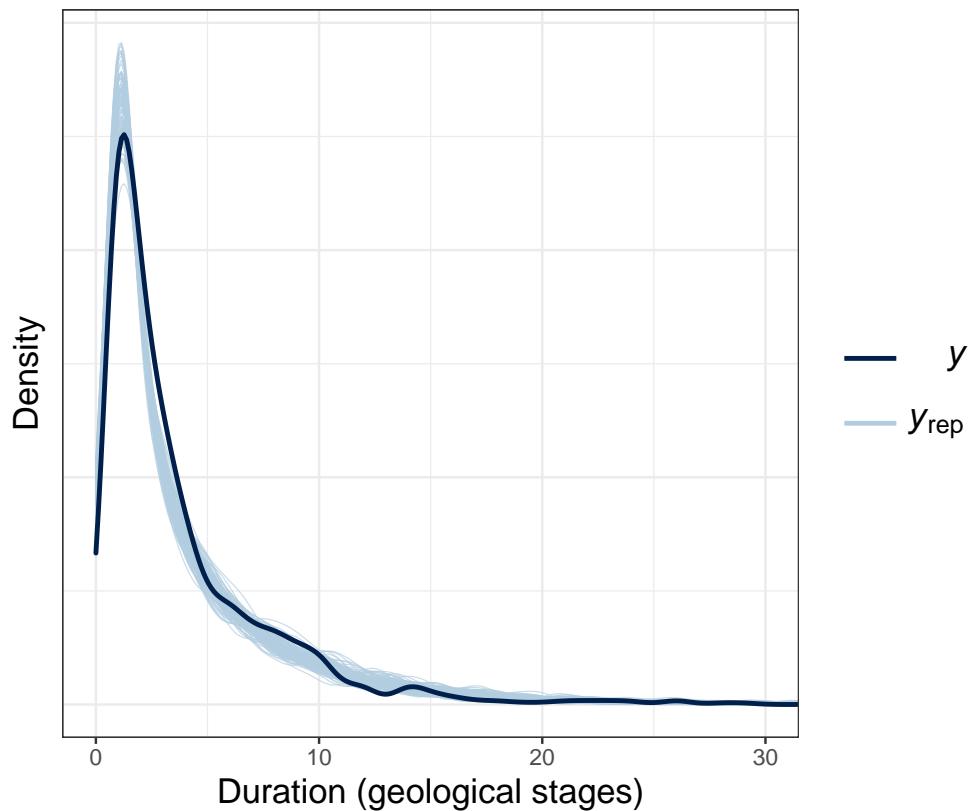


Figure 1: Comparison of the distribution of the observed data (black) to 100 simulated distributions (blue). This is a close-up view of the bulk of the distribution which shows the more subtle aspects of (mis)fit between the data and the model.

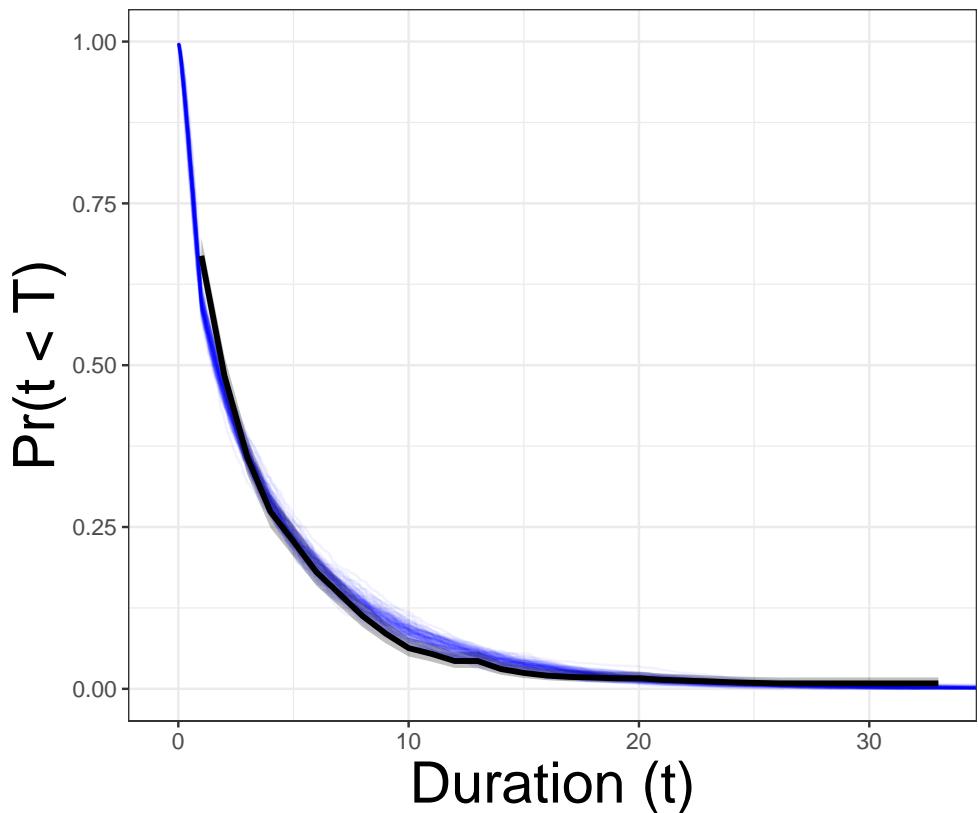


Figure 2: Comparison of the empirical estimate of $S(t)$ (highlighted) versus estimates from 100 posterior predictive data sets (black). $S(t)$ corresponds to the probability that the age of a genus t is less than the genus' ultimate duration T .

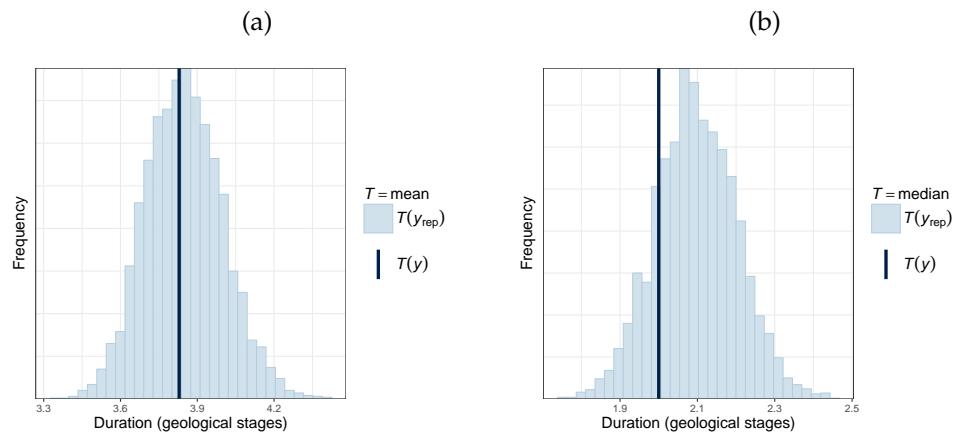


Figure 3: Comparison of the (A) observed mean genus duration (black vertical line) to a distribution of means estimated from 100 simulated datasets (blue), and (B) comparison of the observed median genus duration (black vertical line) to a distribution of medians estimated from 100 simulated datasets (blue). Model fit is evaluated by the similarity between the observed and the estimated, where good fit is demonstrated by the vertical line being “within” the simulated distribution.

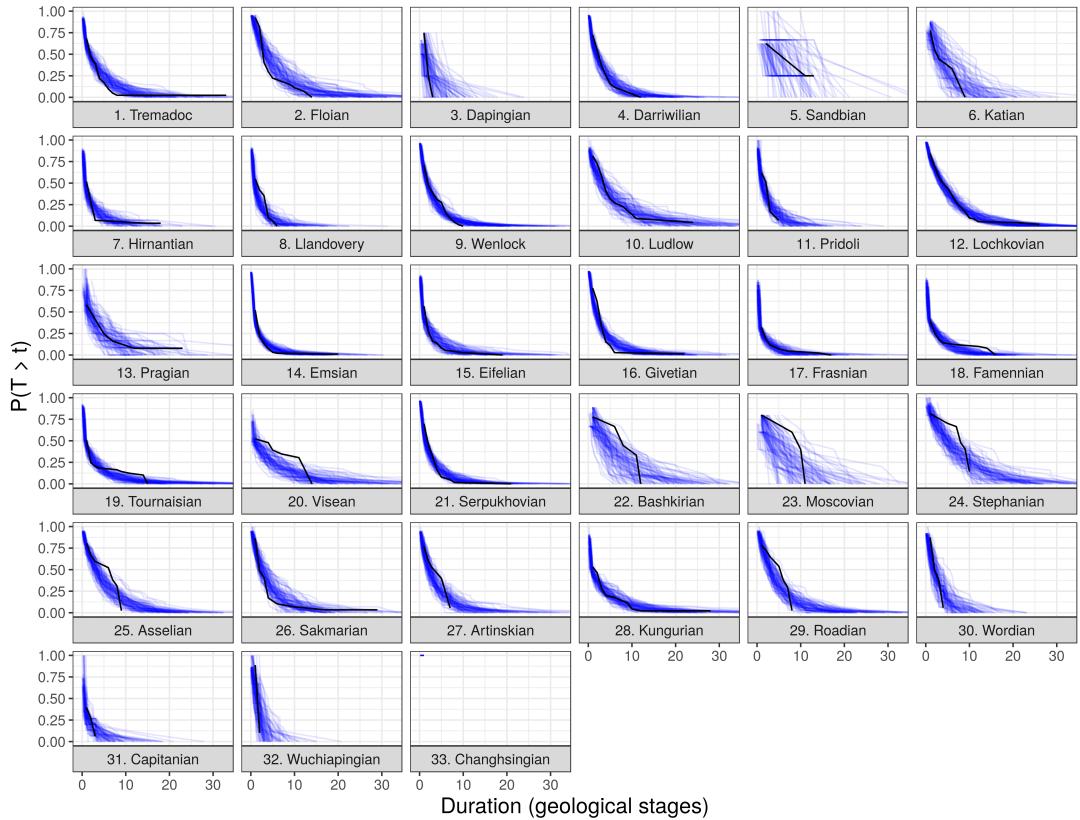


Figure 4: Comparison of the empirical estimate of $S(t)$ (highlighted) versus estimates from 100 posterior predictive data sets (black) for each of the origination cohorts. $S(t)$ corresponds to the probability that the age of a genus t is less than the genus' ultimate duration T . By comparing the fit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

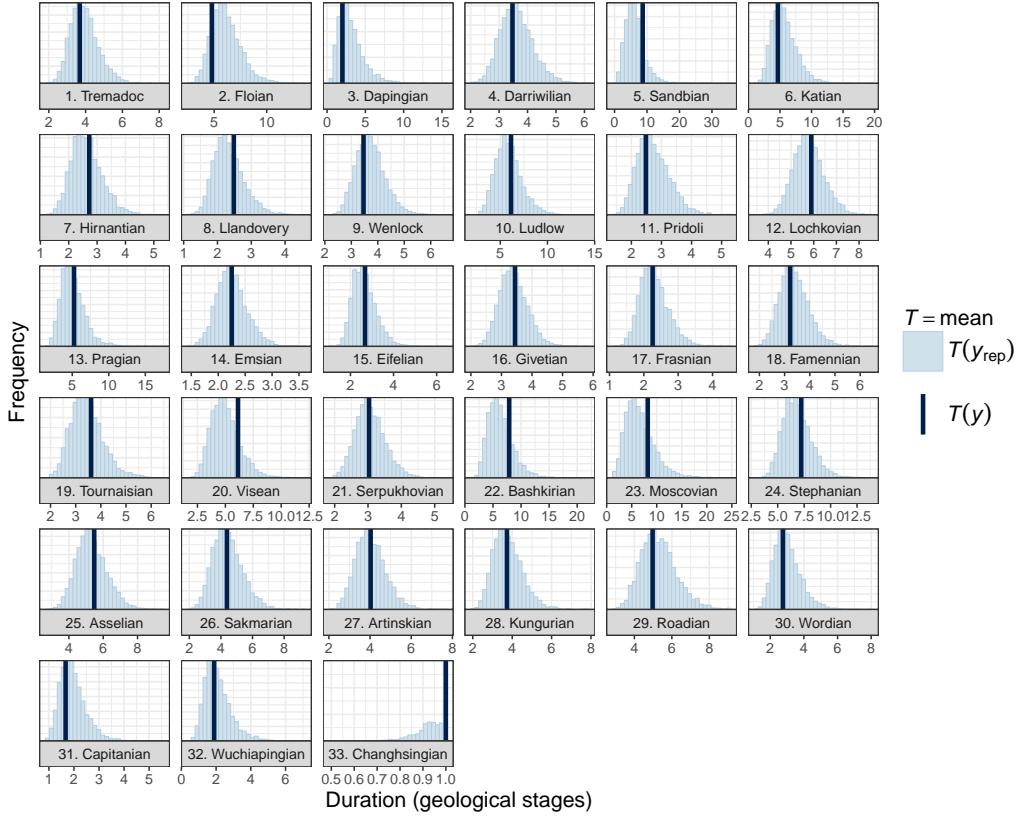


Figure 5: Comparison of the observed mean genus duration (black vertical line) to a distribution of means estimated from 100 simulated datasets (blue) for each of the origination cohorts. Model fit is evaluated by the similarity between the observed and the estimated, where good fit is demonstrated by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

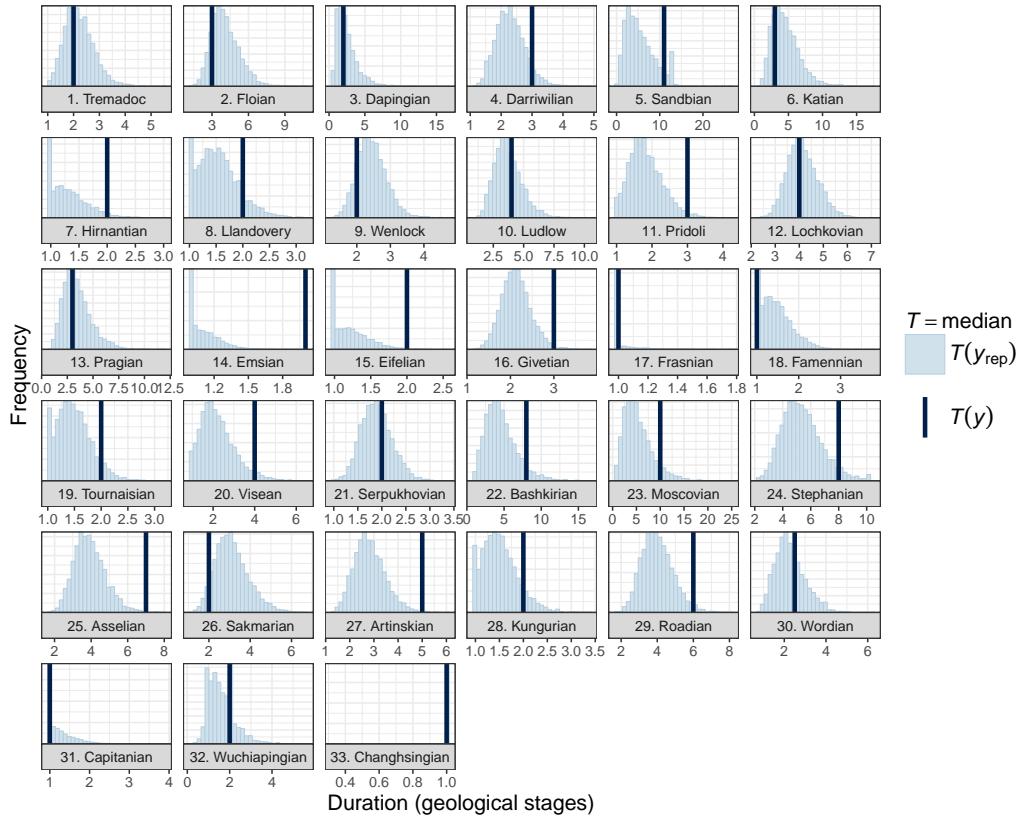


Figure 6: Comparison of the observed median genus duration (black vertical line) to a distribution of medians estimated from 100 simulated datasets (blue) for each of the origination cohorts. Model fit is evaluated by the similarity between the observed and the estimated, where good fit is demonstrated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

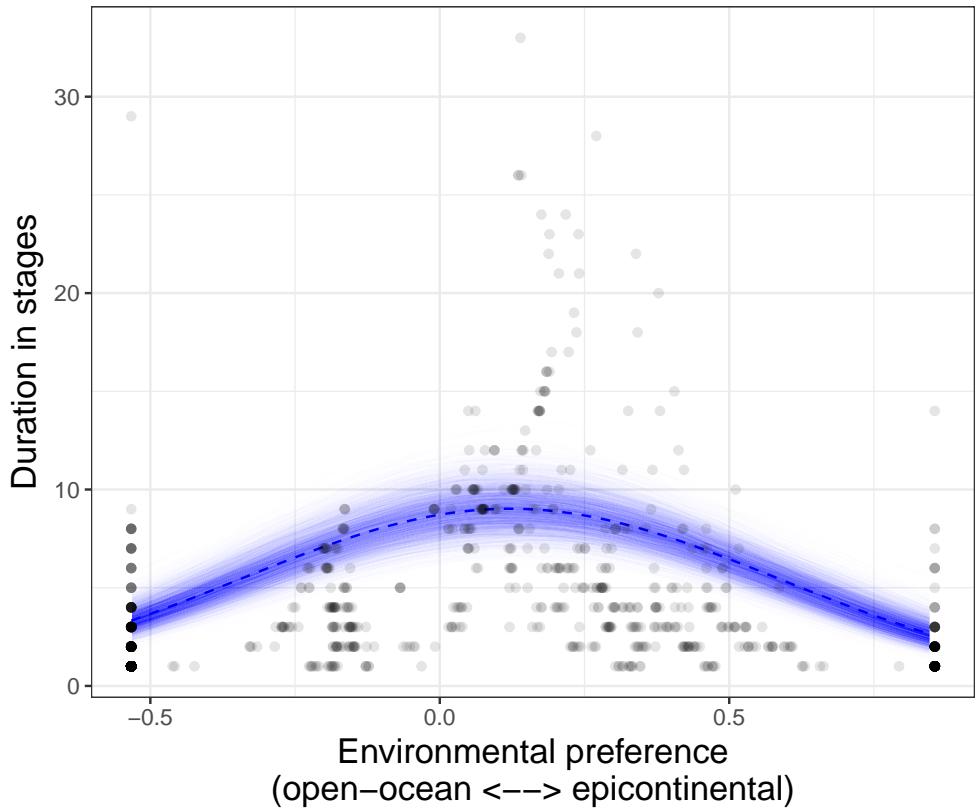


Figure 7: The overall expected relationship between environmental affinity v_i and $\log(\sigma)$ when $r = 0$ and $m = 0$. The 1000 semi-transparent lines corresponds to a single draw from the posterior predictive distribution, while the highlighted line corresponds to the median of the posterior predictive distribution. The overall relationship 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. The tick marks along the bottom of the plot correspond to the (rescaled) observed values of environmental preference.

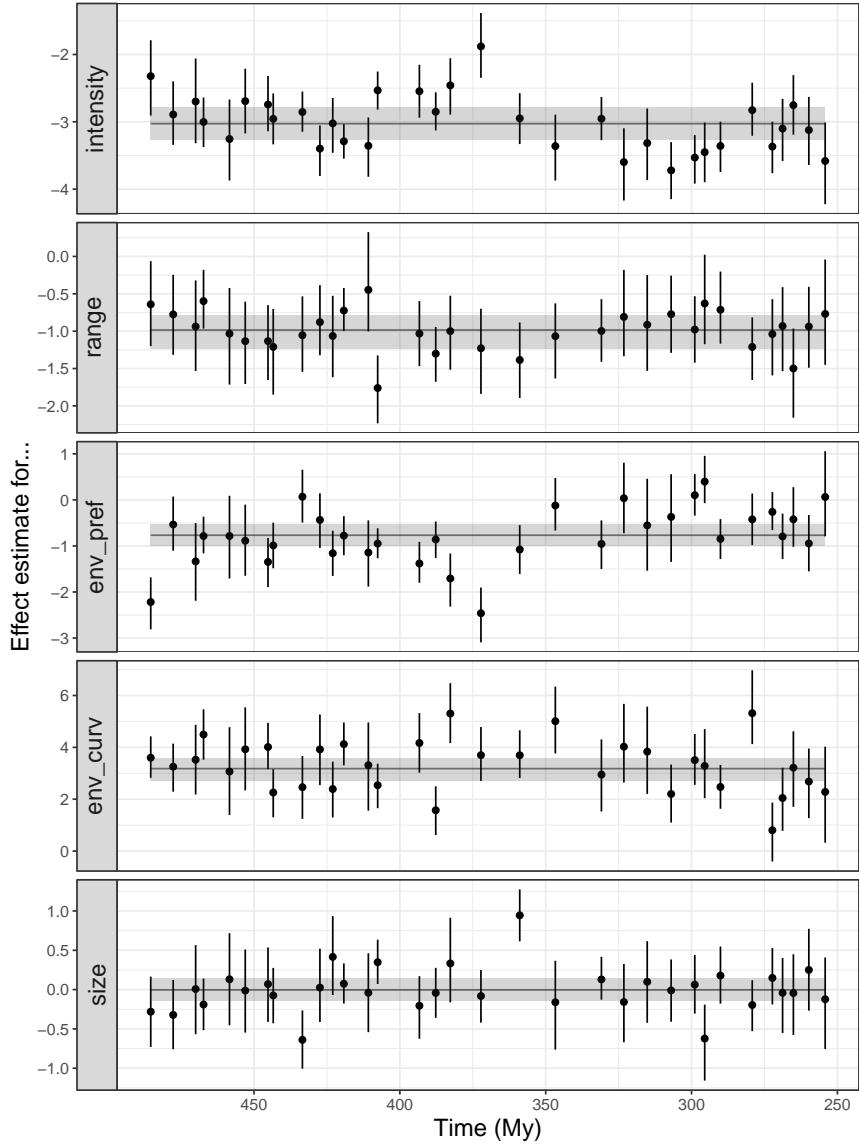


Figure 8: Comparison of cohort-specific estimates of β^0 , the effect of geographic range on extinction risk β^r , the effect of environmental preference β^v and β^{v^2} , and body size β^m . Points correspond to the median of the cohort-specific estimate, along with 80% credible intervals. Points are plotted at the midpoint of the cohorts stage of origination in millions of years before present (My). Black, horizontal lines are the overall estimates of covariate effects along with 80% credible intervals (shaded).

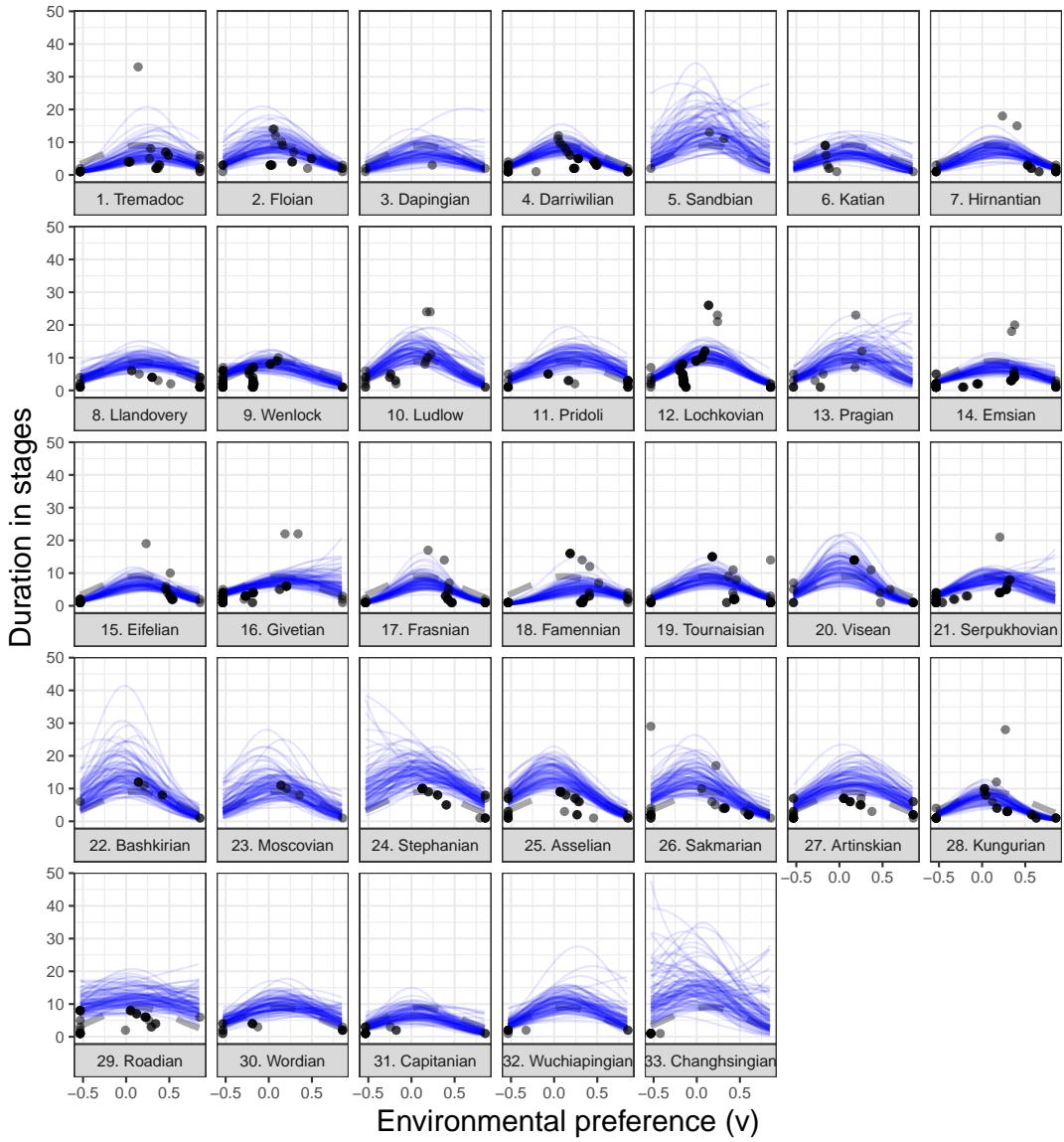


Figure 9: Comparison of origination cohort-specific (posterior predictive) estimates of the effect of environmental preference on $\log(\sigma)$ to the mean overall estimate of the effect of environmental preference. Cohort-specific estimates are from 100 posterior predictive simulations across the range of (transformed and rescaled) observed values of environmental preference. The oldest cohort is in the top-right and younger cohorts proceed left to right, with the youngest cohort being the right-most facet of the last row. Panel names correspond to the name of the stage in which that cohort originated.

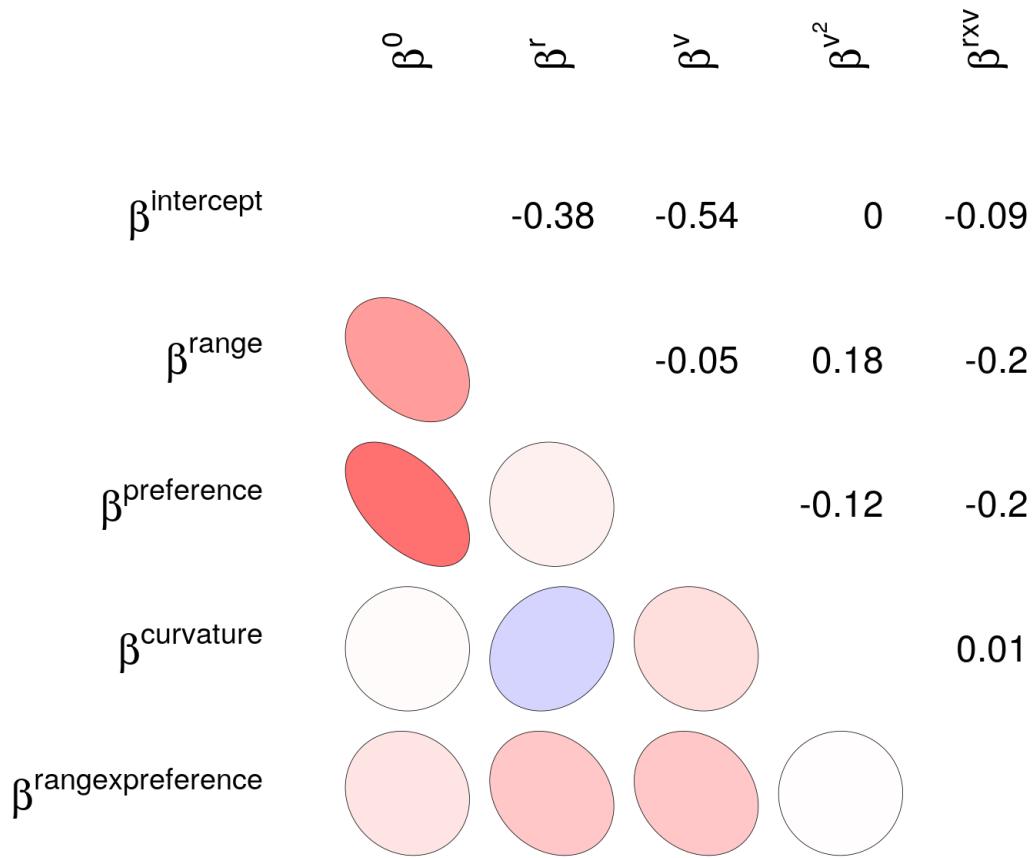


Figure 10: Mixed graphical and numerical representation of the correlation matrix Ω of variation in cohort-specific covariate estimates. These correlations are between the estimates of the cohort-level effects of covariates, along with intercept/baseline extinction risk. The median estimates of the correlations are presented numerically (upper-triangle) and as idealized ellipses representing that much correlation (lower-triangle). The darkness of the ellipse corresponds to the magnitude of the correlation.