Death and taxa: time-invariant differences in mammal species duration

Peter D Smits

Committee on Evolutionary Biology, University of Chicago

February 18, 2015

Abstract

1   Introduction

Why species go extinct at different rates remains one of the most fundamental questions in paleobiology (Jablonski, 2005, Kitchell et al., 1986, Payne and Finnegan, 2007, Quental and Marshall, 2013, Raup, 1991, 1994, Simpson, 1944, Van Valen, 1973, Wagner and Estabrook, 2014). Determining which and how biological traits influence extinction risk is vital for understanding the differential diversification of life during the Phanerozoic and for making predictions about species' vulnerability to anthropogenic impacts.

I am testing if extinction is non-random with respect to organismal- and species-level traits during times of background extinction, if and which traits have time-invariant effects on species duration, and if extinction is taxon-age independent among Cenozoic mammals? I use a Bayesian hierarchical survival modeling approach to estimate the effects of organismal and species traits, time of origination, and shared evolutionary history on species duration while allowing for potentially age–dependent extinction. This is a direct extension and statistical unification of the dynamic and cohort survival approaches from paleontology (Baumiller, 1993, Foote, 1988, Raup, 1975, 1978, Sepkoski, 1975, Simpson, 1944, Van Valen, 1973, 1979). Cenozoic mammals represent an ideal group and time period because their fossil record is well sampled, well resolved both temporally and spatially, and the ecology and taxonomy of individual species are generally understood (Alroy, 1996, 1998, 2009, Alroy et al., 2000, Blois and Hadly, 2009, Jernvall and Fortelius, 2002, Liow et al., 2008, Marcot, 2014, Quental and Marshall, 2013, Simpson, 1944, Smith et al., 2004a, Tomiya, 2013).

This study focuses on identifying the time-invariant effects of organismal and species traits on the expected duration of a species. A time-invariant effect is that

when comparing taxa over a long time period there is a consistent effect that is generalizable over the entire period of interest. While the strength of that effect may vary over time, the direction does not change. For example, geographic range size has been identified as a generally time-invariant factor throughout the Phanerozoic, where large geographic range is associated with a decreased expected extinction risk (Payne and Finnegan, 2007). In many ways, this is a similar goal as Alroy et al. (2000): they found was no consistent, time-invariant evolutionary response to climate change. While it is almost certain that selection pressures vary with time, consistent effects reveal fundamental differences in fitness. Also, a shift in time-invariant factors may be an indicator of a change in "macroevolutionary regime" (Jablonski, 1986) or a "tipping point" (Barnosky et al., 2011, 2012), where the pattern of selection or "rules" are fundamentally changed or different between two time periods.

Background extinction, extinction not occurring at a mass extinction event, is considered to involve a mix of different factors which influence the instantaneous extinction risk of any given species (Harnik et al., 2013, Jablonski, 1986, Kitchell et al., 1986, Nürnberg and Aberhan, 2013, Payne and Finnegan, 2007, Wang, 2003). Factors such as geographic range have well known effects on survival (Jablonski, 1987, Payne and Finnegan, 2007) because the effect size is large and thus easy to identify. The relationship between extinction risk and other traits, specifically organismal ones, is less well known because the effect size is most likely much smaller, making inference difficult.

Periods of background extinction provide a great opportunity to study how traits are related to survival because they represent the majority of geologic time, remain relatively predictable, and change slowly (Jablonski, 1986, Raup and Boyajian, 1988). The Law of Constant Extinction (Van Valen, 1973) states that a taxon's extinction risk within a given adaptive zone is age-independent (memoryless). This law is the foundation for the Red Queen hypothesis as well as most approaches for quantifying extinction. However there is some evidence contrary to this law (Drake, 2014, Finnegan et al., 2008, Raup, 1975, Sepkoski, 1975). By analyzing survival patterns within adaptive zones during periods of background extinction, it should be possible to both estimate the effects of various ecological strategies on survival and determine if extinction is age-independent or dependent.

The organismal and species traits studied here are both dietary and locomotor categories, bioprovince occupancy, and body mass. Each of these traits describe different aspects of a species' adaptive zone such as energetic cost, population density, expected home range size, set of potential prey items, and dispersal ability (Damuth, 1979, 1981, Jernvall and Fortelius, 2004, Lyons, 2005, Lyons et al., 2010, Smith et al., 2004a, 2008). This is a mixture of well established factors that potentially influence extinction risk (i.e. occupancy and body mass) and less well understood ones (i.e. dietary and locomotor categories). It is expected that species with larger geographic ranges have lower extinction rates than species with smaller geographic ranges (Harnik et al., 2013, Jablonski, 1986,

Jablonski and Roy, 2003, Nürnberg and Aberhan, 2013, Roy et al., 2009), though this pattern may be random with respect to other traits (Raup, 1991). However, organismal traits directly related to species–environment interactions may play an important role in determining extinction risk. By modeling extinction via traits related to environmental preference, the relative importance of species- and organismal-level properties can be elucidated.

Dietary category roughly describes the trophic relations of a species, a central component of its biotic environment. The categories used here are coarse groupings of similar ecologies: carnivore, herbivore, omnivore, and insectivore. The first three of these represent commonly used groupings of mammals in paleobiological and macroevolutionary studies (Jernvall and Fortelius, 2004, Price et al., 2012), while the fourth is a biologically important grouping. Price et al. (2012) found that mammalian herbivores and carnivores have a greater diversification rates than omnivores which may indicate that these traits are better for survival. An increase in diversification can be due to either an increase in speciation relative to extinction or a decrease in extinction relative to speciation. Which scenario occurred, however, is currently impossible to determine from a phylogeny of only extant organisms (Rabosky, 2010). By analyzing the fossil record of extinct organisms, the results of Price et al. (2012) can be decomposed into the relative contributions of speciation and extinction.

Locomotor categories describe the motility of a species, plausibility of occurrence in a particular habitat, and dispersal ability. Dispersal ability is important for determining both the extent of a species' geographic range and ability to track changing environments (Birand et al., 2012, Gaston, 2009, Jablonski and Hunt, 2006) which then affects both extinction risk and community similarity. Here, mammals are categorized as either arboreal, ground dwelling, or scansorial. With the transition from primarily closed to open environments during the Cenozoic (Blois and Hadly, 2009, Janis, 1993, Strömberg, 2005, Strömberg et al., 2013), it is expected that arboreal species during the Paleogene will have a greater expected duration than Neogene species while the opposite will be true for ground dwelling species. In comparison, the durations of scansorial species are expected to remain relatively similar between the two time periods because it represents a mixed environmental preference that may be viable in either closed or open environments.

Body size, here defined as mass, has an associated energetic cost in order to maintain homeostasis which in turn necessitates a supply of prey items. Many life history traits are associated with body size such as reproductive rate, metabolic rate, and home range size (Brown and Maurer, 1987, Damuth, 1979, Peters, 1983, Smith et al., 2004a). Body size may affect extinction risk because as body size increases, home range size increases (Damuth, 1979). If individual home range size scales up to reflect a species geographic range, this would mean that extinction risk would decrease. Alternatively, it could be argued that as body size increases, reproductive rate decreases (Johnson, 2002), populations get smaller (White et al., 2007), and generations get longer (Martin and Palumbi,

1993), all of which increase extinction risk. A negative relationship between mammal body size and duration of genera has been observed (Davidson et al., 2012, Liow et al., 2008) though this is inconsistent between continents (Liow et al., 2008, Tomiya, 2013).

I analyze the effect of each of these species traits along with shared origination cohort and shared evolutionary history (i.e. phylogeny) in the context of a Bayesian survival model. Survival analysis is concerned with the time till an event (Klein and Moeschberger, 2003), in this case the time from origination to extinction. The model used here is parameterized in a hierarchical, or mixed effects, context so that known structure in the data can be modeled along with the traits of interest. In that way, the importance and contribution to variance by species traits can be compared with these hierarchical effects.

The purpose of this study is the estimate the time-invariant effects of organismal- and speices-level traits upon expected duration while also estimating the effect of taxon age upon extinction risk. The Bayesian model used here allows for comparison of the marginal posterior estimates of each of the parameters acts as tests of the multiple trait effect hypotheses. All of these traits are analyzed with respect to the individual species origination cohort and phylogenetic position. From there the relative contribution species, cohort, and phylogeny to the unexplained variance can be estimated. Finally, the effect of time on extinction risk is explictly modeled allowing for inference about the applicability of the Law of Constant Extinction (Van Valen, 1973).

2 Methods

2.1 Species information

Fossil occurrence information was downloaded from the Paleobiology Database (PBDB; `http://paleodb.org/`). Occurrence, taxonomic, stratigraphic, and biological information was downloaded for all North American mammals. This data set was filtered so that only occurrences identified to the species-level, excluding all "sp."-s. All aquatic and volant taxa were also excluded. Additionally, all occurrences without latitude and longitude information were excluded.

Species dietary and locomotor category assignments were done using the assignments in initial the PBDB which were then reassigned into coarser categories (Table 1). This was done to improve interpretability, increase sample size per category, and make results comparable to previous studies (Jernvall and Fortelius, 2004, Price et al., 2012).

All individual fossil occurrences were assigned to 2 My bins ranging through the entire Cenozoic. Taxon duration was measured as the number of bins from the first occurrence to the last occurrence, inclusive. This bin size was chosen because it approximately reflects the resolution of the North American Cenozoic

mammal fossil record (Alroy, 2009, Alroy et al., 2000, Marcot, 2014). The youngest cohort, 0-2 My, was excluded from analysis.

Species body size estimates were sourced from a large selection of primary literature and data base compilations, principally the PBDB, PanTHERIA (Jones et al., 2009), the Neogene Old World Mammal database (Now; `http://www.helsinki.fi/science/now/`), and other large scale data collection efforts (Brook and Bowman, 2004, Freudenthal and Martín-suárez, 2013, McKenna, 2011, Raia et al., 2012, Smith et al., 2004b). In many cases, species body mass was estimated from anatomical dimensions such as tooth size. These estimates were made using a variety of published regression equations. See Appendix: Data for a complete list of individual sources and equations.

2.1.1 Bioprovince occupancy

For each 2 My bin, a bipartite biogeographic network was created between species occurrences and spatial units, defined as 2x2 latitude–longitude grid cells from an azimuthal equal-area map projection. In these bipartite networks, taxa can only be linked to localities and *vice versa*. Emergent bioprovinces within the biogeographic occurrence network were identified using the map equation (Rosvall and Bergstrom, 2008, Rosvall et al., 2009). A bioprovince is a set of species–locality connections that are more interconnected within the group than without. This was done for each bin's biogeographic network using the `igraph` package for R (Csardi and Nepusz, 2006, R Core Team, 2014). The relative number of bioprovinces occupied per time bin was then determined for each species. This approach to understanding biogeographic patterns of species occurrence was first used by Sidor et al. (2013) and Vilhena et al. (2013).

2.1.2 Informal phylogeny

As there is no single, combined formal phylogenetic hypothesis of all Cenozoic fossils mammals from North America, it was necessary to construct a semi-formal supertree. This was done by combining taxonomic information for all the observed species and a few published phylogenies.

The taxonomic information from the PBDB served as the initial phylogenetic placement. This information was then updated using the Encyclopedia of Life (`http://eol.org/`), which collects and collates taxonomic information in a single database. This was done programatically using the `taxize` package for R (Chamberlain and Szocs, 2013). This was modified using a published taxonomy of fossil mammals (Janis et al., 1998, 2008). This taxonomy based tree was combined with a selection of species level phylogenies Bininda-Emonds et al. (2007), Raia et al. (2012) into a supertree via matrix representation parsimony as implemented in the R package `phytools` (Revell, 2012).

Polytomies were resolved in order of species first appearance. The resulting tree was then time scaled using the `paleotree` package via the "minimum branch length" approach with a minimum length of 0.1 My (Bapst, 2012). The minimum length is necessary to avoid zero-length branches which cause the phylogenetic covariance matrix not be positive definite, which is important for computation (see below). While other time scaling approaches are possible (Bapst, 2013, Hedman, 2010) this method was chosen for it's simplicity and not requiring additional information about diversification rates which are the interest of this study.

2.2 Survival model

First, define y as a vector of i species durations where $i = 1, \cdots, n$. The simplest survival model is when durations are considered drawn from an exponential distribution with a single rate or inverse-scale parameter λ (Eq. 1).

$$p(y|\lambda) = \lambda \exp(-\lambda y)$$
$$y \sim \text{Exp}(\lambda). \tag{1}$$

In survival analysis, two quantities of interest that are derived directly from the sampling statement (e.g. Eq. 1) are the survival function $S(t)$ and hazard function $h(t)$ where t is a given duration. $S(t)$ corresponds to the probability that a species having existed for t My will not have gone extinct while $h(t)$ corresponds to the instantaneous extinction rate for some taxon age t (Klein and Moeschberger, 2003). For an exponential model, $S(t)$ is defined

$$S(t) = \exp(-\lambda t) \tag{2}$$

and $h(t)$ is defined

$$h(t) = \lambda \tag{3}$$

The choice of the exponential distribution corresponds directly to the Law of Constant Extinction (Van Valen, 1973) as extinction risk ($h(t)$) does not depend on species age (Eq. 3).

The current sampling statement (Eq. 1) currently assumes that all species share the same rate parameter with no variation. To allow for variation in λ associated with relevant covariate information like species body size, λ is reparameterized as $\lambda_i = \exp(\beta^T \mathbf{X}_i)$ with i indexing a given observation and its covariates, β is a vector of regression coefficients, and $\mathbf{X}$ is a matrix of covariates. This is a standard regression formulation, where one column of $\mathbf{X}$ is all 1-s and its corresponding β coefficient is the intercept. This approach is essentially a generalized linear model (GLM) approach where instead of normally distributed errors there are exponentially distributed errors (Klein and Moeschberger, 2003).

To relax the assumption of age-independent extinction or the Law of Constant Extinction we substitute the Weibull distribution for the exponential (Klein and

6

Moeschberger, 2003). σ is the inverse of λ. The Weibull distribution has two parameters: shape α and scale σ and is written

$$p(y|\alpha,\sigma) = \frac{\alpha}{\sigma}\left(\frac{y}{\sigma}\right)^{\alpha-1}\exp\left(-\left(\frac{y}{\sigma}\right)^{\alpha}\right)$$
$$y \sim \text{Weibull}(\alpha,\sigma). \tag{4}$$

$S(t)$ and $h(t)$ for the Weibull distribution are defined

$$S(t) = \exp\left(-\left(\frac{t}{\sigma}\right)^{\alpha}\right) \tag{5}$$

$$h(t) = \frac{\alpha}{\sigma}\left(\frac{t}{\sigma}\right)^{\alpha-1}. \tag{6}$$

The Weibull is useful because when $\alpha = 1$, the distribution is equivalent to the exponential. When $\alpha > 1$ then $h(t)$ (Eq. 6) is a monotonically increasing function and when $\alpha < 1$ $h(t)$ is a monotonically decreasing function.

To allow for σ to vary with y's covariate information it reparameterized in a similar fashion to λ with a few key differences. Because $\sigma = 1/\lambda$ and to preserve interpretation of β, along with taking into account the α, σ is reparameterized as

$$\sigma_i = \exp\left(\frac{-(\beta^T \mathbf{X}_i)}{\alpha}\right). \tag{7}$$

The statistical model described here was the final model at the end of a continuous model development framework where the sampling and prior distributions were iteratively modified to best reflect theory, knowledge of the data, the inclusion of important covariates, and the fit to the data. This follows the approach described in Gelman and Hill (2007) and Gelman et al. (2013). A survival model was fit in a Bayesian context where species duration were assumed to be drawn from a Wiebull distribution (Eq. 4) with shape α and scale σ parameters. α was assumed constant, which is standard practice in survival analysis (Klein and Moeschberger, 2003). α was given a weakly informative half-Cauchy prior. σ was reparameterized as an exponentiated regression model (Eq. 7). This was further expanded (Eq. 8) to allow for two hierarchical factors as discussed below.

$$\sigma_i = \exp\left(\frac{-(h_i + \eta_{j[i]} + \sum \beta^T \mathbf{X}_i)}{\alpha}\right) \tag{8}$$

K species-level covariates were included as a $n \times K$ matrix, $\mathbf{X}$. Two of the covariates of interest are the logit of mean relative occupancy and the logarithm of body size (g). The discrete covariate index variables of dietary and locomotor category were transformed into $n \times (k-1)$ matrices where each column is an indicator variable (0/1) for that species's category, k being the number of categories of the index variable (3 and 4, respectively). Only $k-1$ indicator

7

variables are necessary as the intercept takes on the remaining value. Finally, a vector of 1-s was included in the matrix $\mathbf{X}$ whose corresponding β coefficient is the intercept, making K equal eight.

β is a vector of regression coefficients, where each element is given a unique, weakly informative Normally distributed prior. These priors were chosen because it is expected that the effect size of each variable on duration will be small, as is generally the case with binary covariates.

Regression coefficients are not directly comparable without first standardizing the input variables to have equal standard deviations. This is accomplished by subtracting the mean of the covariate from all values and then dividing by the standard deviation, resulting in a variable with mean of zero and a standard deviation of one. This linear transform greatly improves the interpretability of the coefficients as expected change in mean duration given a difference of one standard deviation in the covariate (Schielzeth, 2010). Additionally, this makes the intercept directly interpretable as the estimate of mean (transformed) σ (Eq. 7). However, because the expected standard deviation for a binary variable is 0.5, in order to make comparisons between the binary and continuous variables, the continuous inputs must instead by divided by twice their standard deviation (Gelman, 2008). The above model was fit with both unstandardardized and standardized continuous inputs for illustrative purposes.

2.2.1 Hierarchical effects

The two hierarchical effects of interest in this study are origination cohort and shared evolutionary history, or phylogeny. Hierarchical modeling can be considered an intermediate between complete and no pooling of groups (Gelman and Hill, 2007), where complete pooling is when the differences between groups are ignored and no pooling is where different groups are analyzed separately. By allowing for partial pooling, we are modeling the appropriate compromise between these two extremes, allowing for better and potentially more informative overall inference. This is done by having all of the groups share the same Normal prior with mean 0 and a scale parameter estimated from the data, which then acts as an indicator of the amount of pooling. A scale of 0 and ∞ indicate complete and no pooling, respectively. The choice of mean 0 allows for the individual groups estimates to be interpreted as deviations from the intercept. Hierarchical modeling is analogous to mixed-effects modeling (Gelman and Hill, 2007).

Origination cohort is defined as the group of species which all originated during the same 2 My temporal bin. Because the most recent temporal bin, 0-2 Mya was excluded, there are 32 different cohorts. The effect of origination cohort j was modeled with each group being a sample from a common cohort effect, η, which was considered Normally distributed with mean 0, and standard deviation σ_c. The value of σ_c was then estimated from the data itself, corresponding

to the amount of pooling in the individual estimates of η_j. This approach is a conceptual and statistical unification between dynamic and cohort survival analysis in paleontology (Baumiller, 1993, Foote, 1988, Raup, 1975, 1978, Van Valen, 1979), with σ_c acting as a measure of compromise between these two end members.

$$\eta_j \sim \mathcal{N}(0, \sigma_c)$$
$$\sigma_c \sim \text{halfCauchy}(0, 2.5)$$

The choice of the half-Cauchy prior on σ_c follows Gelman (2006)

The impact of shared evolutionary history, or phylogeny, was modeling as an individual effect where each observation, i, is distributed as a multivariate normal, h, where the covariance matrix Σ is known up to a constant, σ_p^2 (Housworth et al., 2004, Lynch, 1991). This is written

$$h \sim \text{Multi}\mathcal{N}(0, \boldsymbol{\Sigma})$$
$$\boldsymbol{\Sigma} = \sigma_p^2 \mathbf{V}_{phy}$$
$$\sigma_p \sim \text{halfCauchy}(0, 2.5).$$

$\mathbf{V}_{phy}$ is the phylogenetic covariance matrix defined as an $n \times n$ matrix where the diagonal elements are the distance from root to tip, in branch length, for each observation and the off-diagonal elements are the amount of shared history, measured in branch length, between observations i and j. σ_p was given a weakly informative half-Cauchy hyperprior.

2.2.2 Censored observations

An important part of survival analysis is the inclusion of censored observations where the failure time has not been observed (Ibrahim et al., 2001, Klein and Moeschberger, 2003). The most common censored observation is right censored, where the point of extinction had not yet been observed in the period of study, such as taxa that are still extant. Left censored observations, on the other hand, correspond to observations that went extinct any time between 0 and some known point. In order to account for the minimum resolution of the fossil record encountered here, taxa that occurred in only a single time bin were left censored.

Censored data is modeled using the survival function of the distribution, $S(t)$, defined earlier for the Weibull distribution (Eq. 5) with σ is defined as above (Eq. 8). $S(t)$ is the probability that an observation will survive longer than a given time t. The likelihood of uncensored observations is evaluated as normal using Equation 4 while right censored observations are evaluated at $S(t)$ and

left censored observations are evaluated at $1 - S(t)$. Note, $1 - S(t)$ is equivalent to the cumulative density function and $S(t)$ is equivalent to the complementary cumulative density function (Gelman et al., 2013).

The full likelihood for both uncensored and both right and left censored observations is written

$$L \propto \prod_{i \in C} \text{Weibull}(y_i | \alpha, \sigma) \prod_{j \in R} S(y_j | \alpha, \sigma) \prod_{k \in L} \left(1 - S(y_k | \alpha, \sigma)\right),$$

where C is the set of uncensored observations, R is the set of right censored observations, and L is the set of left censored observations.

A summary of the entire model, save for the formulation for censored observations, along with the exact priors for every estimated parameter is presented in Figure 1.

2.2.3 Estimation

Parameter posteriors were approximated using a Markov-chain Monte Carlo (MCMC) routine implemented in the Stan programming language (Stan Development Team, 2014). Stan implements a Hamiltonian Monte Carlo using a No-U-Turn sampler (Hoffman and Gelman, 2011). Posterior approximation was done using four parallel MCMC chains where convergence was evaluated using the scale reduction factor, $\hat{R}$. Values of $\hat{R}$ close to 1, or less than or equal to 1.1, indicate approximate convergence. Convergence means that the chains are approximately stationary and the samples are well mixed (Gelman et al., 2013).

In order to speed up the poisterior approximation, a custom multivariate normal sampler was used to estimate the unknown constant term in the covariance matrix. This is necessary because inverting and solving the complete covariance matrix on every iteration is a memory intense procedure. The custom sampler limits the necessary number of operations and matrix inversions per iteration.

Four different MCMC chains were run for 20000 steps and were thinned to every twentieth sample split evenly between warm-up and sampling.

2.3 Posterior predictive checks

The most basic assessment of model fit is that simulated data generated using the fitted model should be similar to the observed. This is the idea behind posterior predictive checks. Using the covariates from each of the observed durations, and randomly drawn parameter estimates from their marginal posteriors, a simulated data set y^{rep} was generated. This process was repeated 1000 times and the distribution of y^{rep} was compared with the observed (Gelman et al., 2013).

An example posterior predictive check used in this study is a graphical comparison between the Kaplan-Meier (K-M) survival curve estimated from the observed

data and the K-M survival curves estimated from 1000 simulation sets. K-M survival curves are non-parametric estimates of $S(t)$ (Klein and Moeschberger, 2003). Other posterior predictive checks used here include comparison of the mean and quantiles of the observed durations to the distributions of the same quantities from the simulations, and inspection of the deviance residuals, defined below.

In standard linear regression, residuals are defined as $r_i = y_i - y_i^{est}$. For the model used here, this definition is inadequate. The equivalent values for survival analysis are deviance residuals. To define how deviance residuals are calculated, we first define the cummulative hazard function (Klein and Moeschberger, 2003). Given $S(t)$ (Eq. 5), we define the cumulative hazard function as

$$\Lambda(t) = -log\left(S\left(t\right)\right).$$

Next, we define martingale residuals, m, which are defined in relation to the inclusion vector I

$$m_i = I_i - \Lambda(t_i).$$

I is as a vector of length n, where $I_i = 1$ means the observation is completely observed and $I_i = 0$ means the observation is censored.

Martingale residuals have a mean of 0 and ranges between 1 and $-\infty$ and can be viewed as the difference between the observed number of deaths between 0 and t_i and the expected number of deaths based on the model. However, martingale residuals are difficult to interpret, can be asymmetrically distributed, and are not equivalent to standard residuals.

The solution to this is to use the deviance residuals, D. This is defined as a function of martingale residuals and takes the form

$$D_i = \text{sign}(m_i)\sqrt{-2[m_i + I_i log(I_i - m_i)]}.$$

Deviance residuals have a mean of 0 and a standard deviation of 1 by definition.

2.4   Variance partitioning

There are three different variance components in this model (Fig. 1): sample σ_y^2, cohort σ_c^2, and phylogenetic σ_p^2. The sample variance, σ_y^2, is similar to the residual variance from a normal linear regression. Partitioning the variance between these sources allows the relative amount of unexplained variance of the sample to be compared. However, the model used here (Eq. 4) does not include an estimate of the sample variance, σ_y^2. Partitioning the variance between these three components was approximated via a simulation approach modified from Goldstein et al. (2002).

The procedure is as follows:

1. Simulate w (50,000) values of η; $\eta \sim \mathcal{N}(0, \sigma_c)$.

2. For a given value of $\beta^T \mathbf{X}$, calculate σ^{c*} (Eq. 7) for all w simulations, holding h constant at 0.

3. Calculate v_c, the Weibull variance (Eq. 9) of each element of σ^{c*} with α drawn from the posterior estimate.

4. Simulate w values of h; $h \sim \mathcal{N}(0, \sigma_p)$.

5. For a given value of $\beta^T \mathbf{X}$, calculate σ^{p*} (Eq. 7) for all w simulations, holding η constant at 0.

6. Calculate v_p, the Weibull variance (Eq. 9) of each element of σ^{p*} with α drawn from the posterior estimate.

7. $\sigma_{y*}^2 = \frac{1}{2} \left(\left(\frac{1}{w} \sum_i^w v_{pi} \right) + \left(\frac{1}{w} \sum_j^w v_{cj} \right) \right)$.

8. $\sigma_{c*}^2 = var(v_c)$ and $\sigma_{p*}^2 = var(v_p)$.

The simulated values of h were drawn from a univariate normal distribution because each simulated value is in isolation, so there are is no concern of phylogenetic autocorrelation. The chosen value for $\beta^T \mathbf{X}$ was a draw from the posterior estimate of the intercept. Because input variables were standardized prior to model fitting, the intercept corresponds to the estimated effect on survival of the sample mean.

Weibull variance is calculated as

$$var(x) = \sigma^2 \left(\Gamma \left(1 + \frac{2}{\alpha} \right) - \left(\Gamma \left(1 + \frac{1}{\alpha} \right) \right)^2 \right), \tag{9}$$

where Γ is the gamma function.

The variance partitioning coefficients are then calculated, for example, as $VPC_{phylo} = \frac{\sigma_{p*}^2}{\sigma_{y*}^2 + \sigma_{c*}^2 + \sigma_{p*}^2}$ and similarly for the other components.

I used variance partitioning coefficients (VPC) to estimate of the relative size importance of the different variance components (Gelman and Hill, 2007). Phylogenetic heritability, h_p^2 (Housworth et al., 2004, Lynch, 1991), is identical to the VPC of the phylogenetic effect. Additionally, because phylogenetic effect was estimated using a principally taxonomy based tree the estimates derived here can be considered minimum estimates of the phylogenetic effect.

3   Results

With all marginal posterior estimates having converged ($\hat{R} < 1.1$) it is possible be examine the quality of model fit (Table 2). If the model is an adequate

Table 1: Species trait assignments in this study are a coarser version of the information available in the PBDB. Information was coarsened to improve per category sample size and uniformity and followed this table.

This study		PBDB categories
Diet	Carnivore	Carnivore
	Herbivore	Browser, folivore, granivore, grazer, herbivore.
	Insectivore	Insectivore.
	Omnivore	Frugivore, omnivore.
Locomotor	Arboreal	Arboreal.
	Ground dwelling	Fossorial, ground dwelling, semifossorial, saltatorial.
	Scansorial	Scansorial.

$$y_i \sim \text{Weibull}(\sigma, \alpha)$$
$$\eta_{j[i]} \sim \text{Normal}(0, \sigma_c)$$
$$\sigma_c \sim \text{half-Cauchy}(2.5)$$
$$h_i \sim \text{MultiNormal}(0, \Sigma)$$
$$\Sigma = \sigma_p^2 \mathbf{V}_{phy}$$
$$\sigma_p \sim \text{half-Cauchy}(2.5)$$
$$\beta \sim \text{Normal}(0, 10)$$
$$\alpha \sim \text{half-Cauchy}(2.5)$$

Figure 1: Graphical depiction of full survival model, save censoring, used here. Exact values for the hyperparameters are presented to the right. The observed duration of the ith observation is indicated towards the bottom left as y_i, which is assumed to follow a Weibull distribution.

descriptor of the observed data, then relatively confident inference can be made (Gelman et al., 2013).

Visual examination of the deviance residuals from twelve different sets of posterior predictive simulations indicates few systematic problems (Fig. 2). The only concern is that the residuals are slightly skewed, however this bias appears very small. This is confirmed by comparing the K-M estimate of the empirical survival function to 100 estimated survival functions from posterior simulations (Fig. 3).

Comparisons of the observed 25th, 50th, 75th quantiles, and mean durations to the results from the posterior predictive simulations indicate adequate model fit (Fig. 4). Because all the different posterior predictive checks seem to agree, the inferred model appears adequate at capturing the observed variation.

Given that the model appears adequate, it is possible to interpret the parameter estimates with some degree of confidence. The estimates for diet and locomotor categories were inferred as contrasts between the intercept and one of the $k-1$ other states (Table 2). In order to interpret these estimates, I compared the differences between each of the different states to get an estimate of whether either of two traits was associated with a greater mean duration or not. This was done for all pairwise comparisons for diet and locomotor category separately (Fig. 5).

Dietary category has a large amount of variation in the pairwise differences of the effect on expected duration (Fig. 5a). Carnivory appears to be associated with a greater expected duration than herbivory or insectivory, while approximately equal to or less than the expected duration of an omnivore. Omnivory is associated with greater expected duration than either herbivory or insectivory. Finally, herbivory and insectivory are associated with approximately equal effects on expected duration.

For locomotor category, arboreality appears to be associated with a lower expected duration than either scansoriality or a ground dwelling life habit (Fig. 5b). Scansoriality and a ground dwelling life habit have approximately equal expected durations.

The effects of both body size and bioprovince occupancy on expected duration are best interpreted from estimates from the model fit standardized data (Fig. 6). Of all the traits of interest, bioprovince occupancy has the largest effect size with larger occupancy associated with a longer expected duration. Body size has a very small, near zero effect on expected duration, similar to the lack of relationship between body size and generic duration (Tomiya, 2013).

The estimates for the individual cohort effects show a weak pattern of increased extinction risk in older Cenozoic cohorts and decreased extinction risk in younger cohorts (Fig. 7). However, this pattern is not very strong as there is a large amount of variation, particularity for older cohorts. For example, note the two cohorts between 50 and 55 My that have a much lower extinction risk than other cohorts of similar age. However, it is interesting to note that the apparent shift

from older cohorts with a higher extinction risk to younger cohorts with lower extinction risk is approximately 30 Mya or the Paleogene–Neogene boundary. This transition is marked by the opening up of the landscape and the rise of grazers and the decline heavily forested environments.

Because the estimate of the Weibull shape parameter, α, is greater than 1 extinction risk is expected to increase with taxon age (Table 2). The estimate of α is also rather tightly constrained, having a small posterior standard deviation. α is related to the strength of time on extinction risk and is a key parameter in the hazard function $h(t)$, calculated for a Weibull distribution as

$$h(t) = \frac{\alpha}{\sigma} \left(\frac{t}{\sigma} \right)^{\alpha - 1}. \tag{10}$$

$h(t)$ can be interpreted as the rate, or approximate probability, of an individual of age t going extinct. As the value of α is between 1 and 1.5, extinction risk for a given species only gradually increases with age (Fig. 8).

Of the three sources of variance present in the model, variation in individual species accounts for approximately 70% of observed variance (Fig. 9). Both cohort and phylogenetic effects account for the other 30% of the observed variance. Both of these factors are related to some aspect of the relationship between taxa, either temporal and through shared evolutionary history. While both of these effects are the source of approximately 15% of observed variance individually, the total combined effect of these factors indicates that neither can be ignored. For example, Housworth et al. (2004) state that when phylogenetic heritability is greater than 0, phylogenetic effect cannot be ignored. As VPC_{phylo} is equivalent to phylogenetic heritability and is greater than 0, the effect of shared evolutionary history on duration is non-ignorable.

The low amount of variance of phylogenetic effect estimated from the model (Table 2) may explain previously found lack of "significant phylogenetic signal" in generic duration (Tomiya, 2013).

Table 2: Summaries of 1000 samples drawn from the marginal posteriors for the principle parameters of interest. Because of variable standardization, the intercept can be interpreted as the estimate for the mean observed species. As such, the other values are expected effects of trait values expressed as deviation from the mean. The categorical variables are binary index variables where an observation is of that category or not. $\hat{R}$ values of less than 1.1 indicate chain convergence for the posterior samples.

	mean	sd	2.5%	25%	50%	75%	97.5%	$\hat{R}$
alpha	1.31	0.03	1.25	1.29	1.31	1.34	1.38	1.01
intercept	-0.82	0.18	-1.17	-0.94	-0.82	-0.70	-0.48	1.00
ground dwelling	-0.12	0.12	-0.34	-0.20	-0.12	-0.04	0.12	1.00
scansorial	-0.13	0.13	-0.38	-0.22	-0.13	-0.05	0.12	1.00
herbivore	0.11	0.12	-0.13	0.03	0.12	0.19	0.35	1.00
insectivore	0.09	0.13	-0.17	0.00	0.09	0.19	0.36	1.00
omnivore	-0.12	0.13	-0.37	-0.20	-0.12	-0.04	0.14	1.00
logit(occupancy)	-0.66	0.08	-0.81	-0.71	-0.65	-0.60	-0.50	1.00
log(size)	-0.05	0.05	-0.15	-0.08	-0.05	-0.01	0.05	1.00
sd cohort	0.33	0.07	0.22	0.29	0.33	0.37	0.48	1.00
sd phylogeny	0.21	0.10	0.07	0.14	0.19	0.26	0.46	1.05
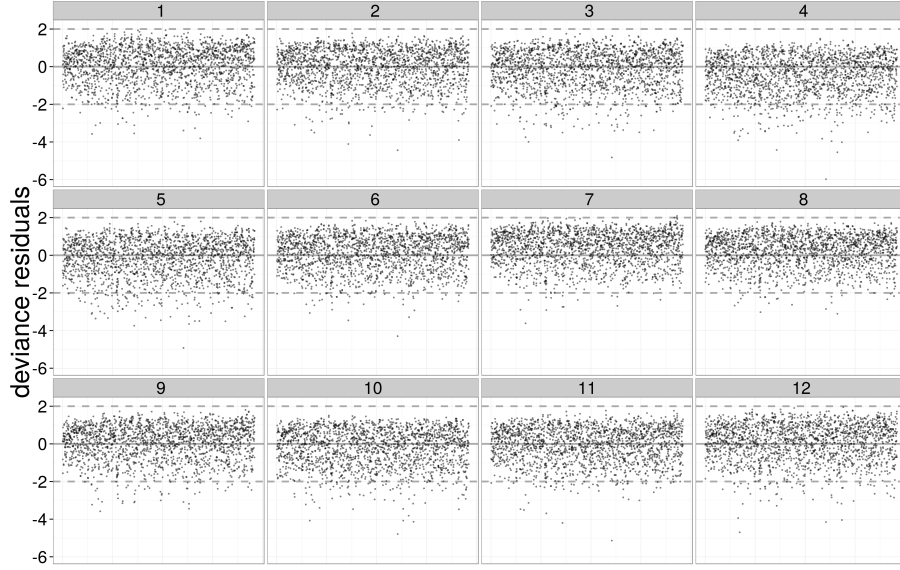
Figure 2: Deviance residuals from the fitted survival model. Each graph depicts the residuals from single draws from the posteriors distributions of all estimated parameters. Positive values indicate an under estimate of the observed duration, while negative values indicate an over estimate of the observed duration. Twelve difference examples are provided here to indicate the lack of individual observation based biases.

4   Discussion

The number and structure of the parameters from the fitted model allows for many different hypotheses to be addressed and inferences to be made. First, I interpret the results in terms of hypotheses of trait effects on survival. Second, I address implications for future paleobiological modeling. Finally, I discuss additional model improvements and complexities that can be used to address further hypotheses of species survival.

As expected, bioprovince occupancy has the largest effect on expected species duration/extinction risk. The other factors all appear to have small effect sizes which is also consistent with expectations.

I found that carnivores have a lower extinction risk than herbivores and insecti-vores, while ominivores have approximately equal or lower extinction risk than carnivores. Omnivorous taxa are also associated with a lower extinction risk than both herbivores and insectivores. These results sit nicely beside those of
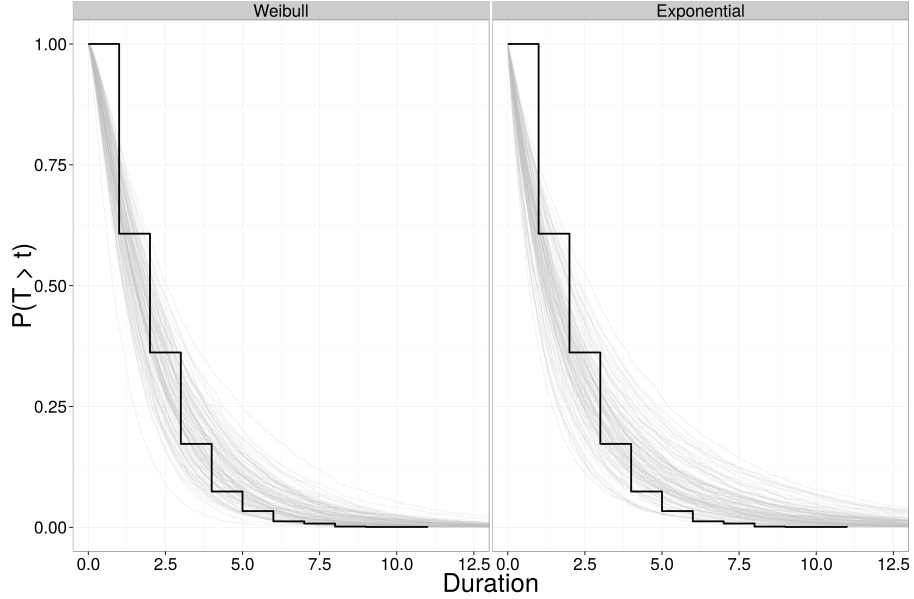
Figure 3: Comparison between K-M estimate of survival function (black) from the observed versus K-M estimates from 100 simulated data sets using the fitted model (dark grey). Simulated data sets were generated by drawing parameter values randomly from their estimated posteriors and using the observed covariate information to estimate durations for all the observed species. On the left are the results from the full survival model (Fig. 1), while on the right are the results from a simplified model where duration follows an exponential distribution and there is no phylogenetic effect.

Price et al. (2012).

Given that carnivores and omnivores have approximately equal extinction risk, and that Price et al. (2012) found that carnivores have a greater diversification rate than omnivores, this implies that carnivores have a greater origination rate than omnivores. In comparison, given the results of this study and Price et al. (2012), this also implies that herbivores which have the greatest extinction risk must also have a very high origination rate in order to have the greatest diversification rate of these three categories.

The large difference in time-invariant extinction risk between omnivores and both herbivores and insectivores is most likely related to the concept of "survival of the unspecialized" where less specialized taxa have lower extinction risks (Liow, 2004, Simpson, 1944). Because larger effects are easier to identify in coarse or small data sets, the magnitude of this effect also explains both the early identification and origin of this hypothesis of time-invariant effects on survival (Simpson, 1944).
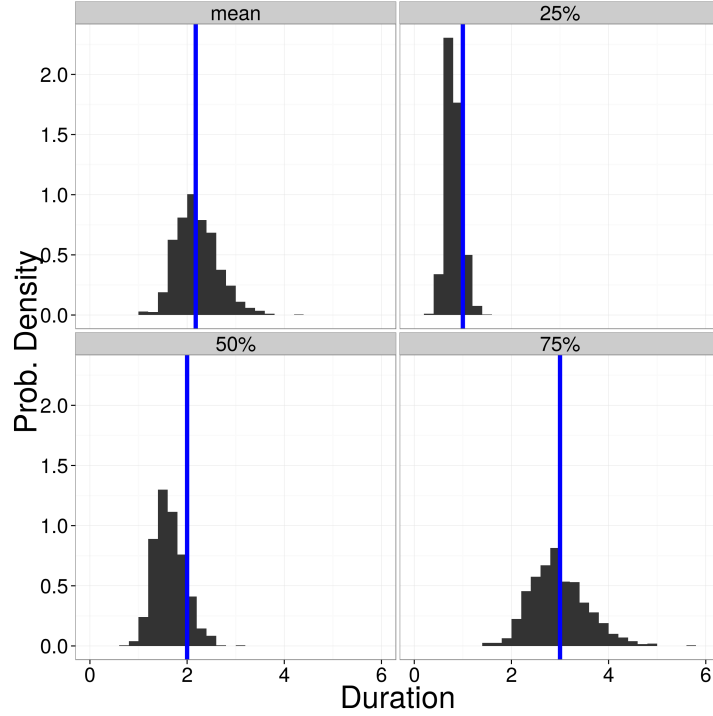
Figure 4: The results of additional posterior predictive checks for four summaries of the observed durations, as labeled. Blue vertical indicate the observed value. None of the observed are significantly different from the posterior predictive distributions.

The comparisons between the effects of locomotor category on expected species extinction risk are consistent with the hypotheses that arboreality is associated with a greater expected extinction risk than either with scansoriality and ground dwelling taxa. Importantly, scansoriality appears to not influence any difference in extinction risk when compared with ground dwelling taxa. This can be interpreted that arboreal taxa, which require a specific kind of environment, may be more prone to extinction because the lack of permanency of those environments may prevent species persistence.

The very weak if non-existent effect of body size on extinction risk, which can be considered not "significant", is consistent with Tomiya (2013). The direction/sign of the modal estimate of effect is not consistent with the prediction of increase in extinction risk associated with increase in body size (Liow et al., 2008). Importantly, however, the other studies were performed at the generic-level which may or may not involve different processes that are not included in this species-level model (Liow et al., 2008, Tomiya, 2013).
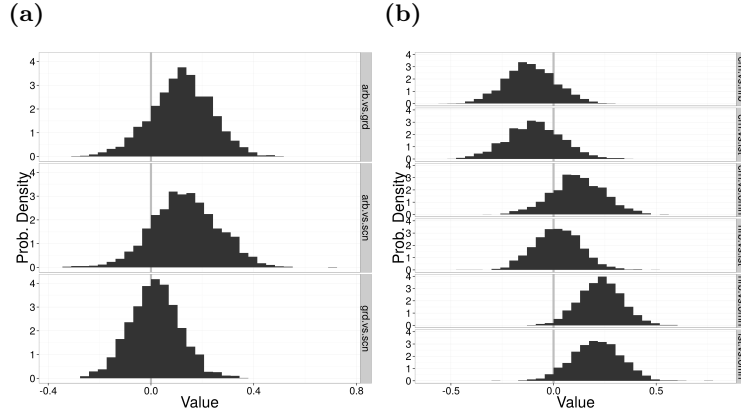
Figure 5: Pairwise differences in effect of the locomotor (a) and dietary categories (b) on expected duration from 1000 samples from the posterior distribution. Comparisons of locomotor categories, from top to bottom (a), are: arboreal versus ground dwelling, arboreal versus scansorial, and ground dwelling versus scansorial. For dietary category, from top to bottom (b): carnivore versus herbivore, carnivore versus insectivore, carnivore versus omnivore, herbivore versus insectivore, herbivore versus omnivore, and insectivore versus omnivore. Values to the left indicate that the first category is expected to have a greater duration than the second, while values to the right indicate that the first category is expected to have a shorter duration.

While explicit phylogenetic relations between taxa are frequently not modeled in paleobiological studies of diversity (Alroy, 2009, Alroy et al., 2000, Foote and Miller, 2013, Hunt, 2007, Jablonski and Hunt, 2006, Jernvall and Fortelius, 2002, 2004, Liow et al., 2008, Marcot, 2014, Payne and Finnegan, 2007), there have been more recent studies which analyze fossil diversification in an explicitly phylogenetic context (Harnik et al., 2014, Simpson et al., 2011, Slater et al., 2012, Slater, 2013, Tomiya, 2013). The partitioning of the different sources of variance involved in this model indicate that phylogeny or shared evolutionary history accounts for approximately 10-15% of the unexplained variance each. Because VPC of phylogeny is greater than 0, it is not appropriate to ignore phylogeny when modeling survival in paleontological studies (Housworth et al., 2004). An addition 10-15% of unexplained variance was due to shared origination cohort. Between these two sources of variance, it is clear that a shared evolutionary history and temporal occurrence are non-ignorable in paleontological studies or survival. Modeling phylogeny as a hierarchical effect, as done here, is a very simple and interpretable means of integrating phylgoenetic information into any and all regression models (Housworth et al., 2004, Lynch, 1991) and is most likely ideal for most paleontological studies.

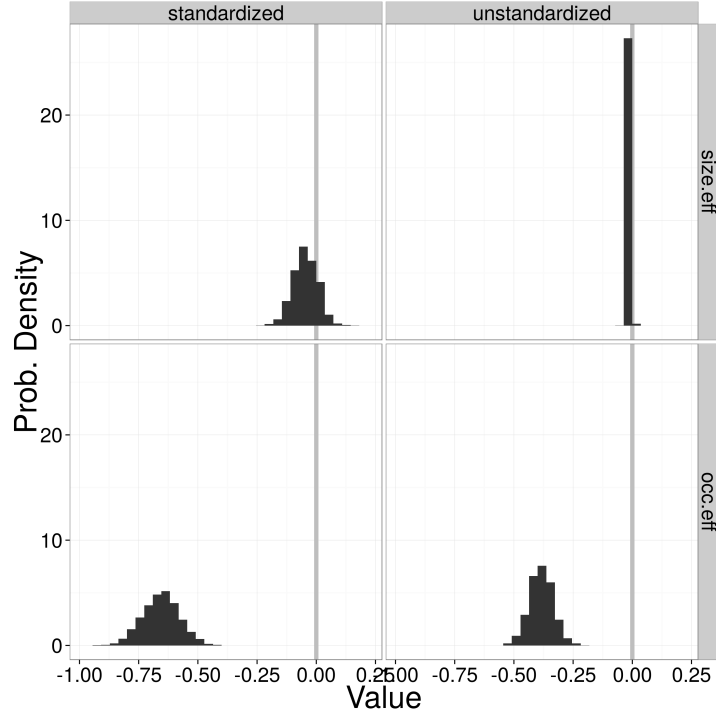The inferred increasing extinction risk with species duration (Fig. 8) has two

Figure 6

possible explinations: (1) older taxa being aged out or out competed by younger taxa, or (2) as an artifact of the minimum resolution of the fossil record.

The observed pattern of younger cohorts have a lower extinction risk than older cohorts is consistent with the idea that younger taxa outcompete or replace older taxa (Fig. 7). Additionally, the hypothesis that older taxa are being outcompeted or replaced by younger taxa is also consistent with the results of Wagner and Estabrook (2014) and Quental and Marshall (2013), both of which require that older taxa have a greater extinction risk than younger taxa, and the earlier density-dependent kinetic model hypothesized by Sepkoski Jr. (1978, 1979, 1984).

An important note is that while Quental and Marshall (2013) argue that their result is consistent with the Red Queen hypothesis, this cannot actually be true as their model of the rise and fall of taxa requires increasing extinction risk with species age in order to obtain the observed symetric distribution of family richness with time. A requirement which is directly opposed by the Law of Constant Extinction which is the original theoretical basis of the Red Queen Hypothesis (Van Valen, 1973). While the meaning of the Red Queen Hypothesis has apparently shifted over time to represent only biotic–biotic interactions
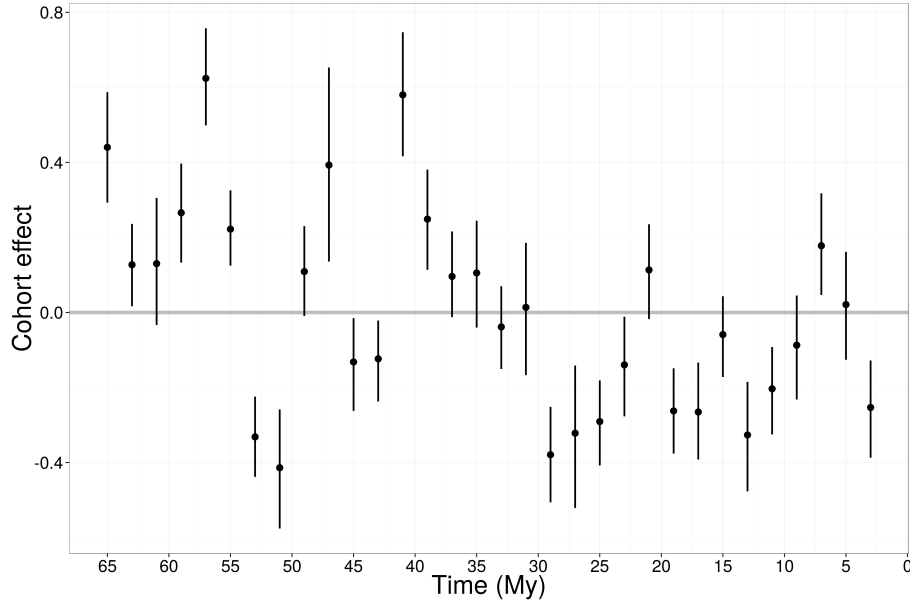
Figure 7: Summaries of estimated individual cohort effect posteriors. Depicted are medians and 80% credible intervals of the estimated posterior distributions. High values correspond to shorter species durations while lower values correspond to greater species durations compared to the mean duration. Lines are placed at the middle of the 2 My origination cohorts.

or competition (Barnosky, 2001) this is possibly due to a misreading of the original argument of Van Valen (1973), who actually argued that the Red Queen Hypothesis is due to the deterioration of a species adaptive zone with time, where an adaptive zone is defined as the set of biotic–biotic and biotic–abiotic interactions experienced by a species (Liow et al., 2011, Simpson, 1944).

The other possible explination for the inferred increase in extinction risk with species age is the minimum resolution which might cause an upward bias in estimates of the Weibull shape parameter α (Sepkoski, 1975), an effect which can be observed by the initial plateau in the K-M estimate of $S(t)$ for the observed (Fig. 3). This plateau is a hallmark of the original survival analyses of Van Valen (1973) which Sepkoski (1975) identified as partially a product of minimum resolution of the fossil records of the different studied groups.

Given this known biasing factor and the results of Wagner and Estabrook (2014) and Quental and Marshall (2013), I hypothesize that the inferred pattern is most likely a combination of these two explinations working on concert. In order to determine the relative importance of these two explinations, more work is required into approaches for directly modeling the minimium resolution of the fossil record.
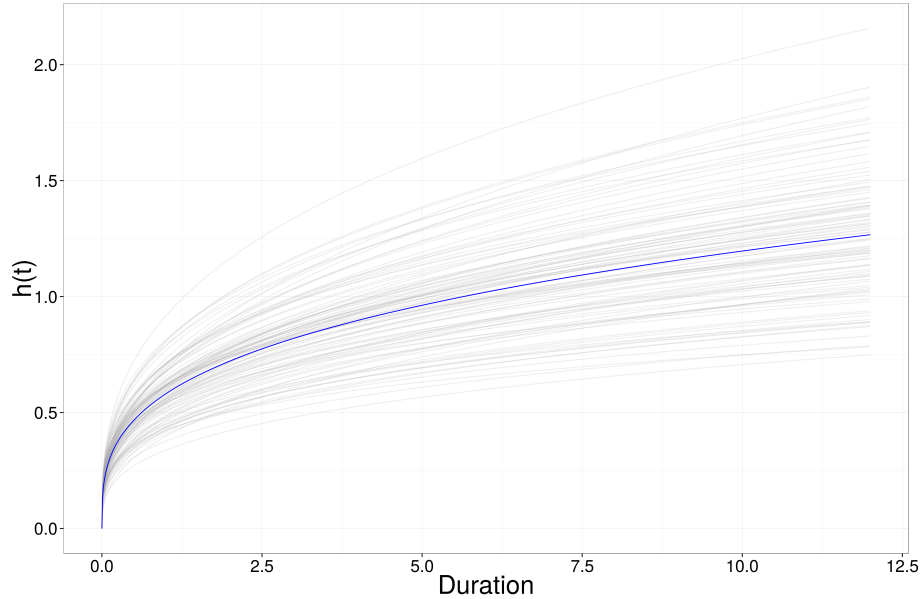
Figure 8: 100 estimates of the hazard function ($h(t)$ for the observed species mean (grey), along with the median estimated hazard function. $h(t)$ is an estimate of the rate at which a species of age t is expected to go extinct. Hazard functions were estimated from random draws from the estimated posterior distributions and evaluated with all covariate information set to 0, which corresponds to the expected duration of the mean species.

While the observed pattern of older cohorts having increased extinction risk compared to younger cohorts is weak, it is notable because the shift from older, higher risk cohorts to younger, lower risk cohorts coincides approximately with the Paleogene–Neogene boundary. The shift from the Paleogene to the Neogene is marked by the transition from a principally closed, forested environment to a principally open, grassland environment (Janis, 1993, Janis et al., 2000, Strömberg, 2005). This shift may underly the inferred increased extinction risk associated with arboreal species compared to ground dwelling or scansorial species (Fig. 5a). However, because the model used here does not allow for change in time-invariant effects, I cannot identify this transition as a tipping point or shift in selective regime (Barnosky et al., 2011, 2012).

It is possible to think of the observed time-invariant effects modeled here as the signature of the Cenozoic background extinction macroevolutionary regime (Jablonski, 1986). One of the open questions in paleobiology and macroecology is whether the current biodiversity crisis qualified as a mass extinction (Alroy, 2010, Barnosky et al., 2011, 2012). Because change in the magnitude of extinction risk is not necessarily the best indicator of a shift from background to mass extinction
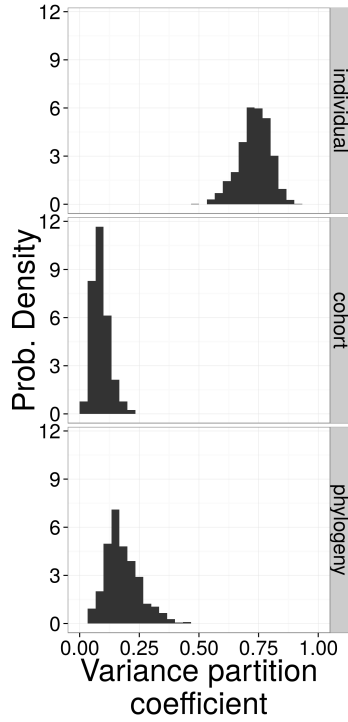
Figure 9: Estimates of the variance partitioning coefficients for the three different sources of variance: species, cohort, and phylogeny. Higher values correspond to greater contribution to total observed variance. Each of the estimates is a distribution of 1000 approximating simulations due to the model's non-normality.

(Wang, 2003), it is more fruitful to look for changes in the direction selection, loss of a selective pressure, or appearance of novel selective pressures. Comparison of the estimated effects of organismal- and species-level traits analyzed here with previous studies demonstrates a mixture of congruence and incongruence.

As expected, large range size is always associated with lower extinction risk in the Recent (Fritz and Purvis, 2010, Fritz et al., 2009, Liow et al., 2009, Purvis et al., 2000). The primacy of geographic range as a time-invariant factor influencing extinction risk has been found for marine invertebrates across the Phanerozoic and in particular at the Cretaceous mass extinction event (Jablonski, 1986, Payne and Finnegan, 2007).

While I found that body size has no time-invariant effect on extinction risk, large body size is associated with increased extinction risk in the Recent, though this is variable across environments and clades (Fritz et al., 2009, Liow et al., 2009, Purvis et al., 2000).

A higher trophic level (e.g. carnivory versus herbivory) is associated with greater

extinction risk in Primates and Carnivora (Purvis et al., 2000) which is not congruous with the results found here that carnivores have lower extinction risk than herbivores.

Finally, phylogeny has been found to be a factor underlying current mammal species extinction risk, though this effect seems much greater in the Recent than for the whole Cenozoic (Fritz and Purvis, 2010). Note that the phylogeny of Recent mammals is much better than the primarily taxonomy based phylogeny used here, which may partially account for the difference in effect.

How many of these incongruities are within the standard range of time-variant effects is unknown, though these comparisons across multiple factors do point to our arrival at a tipping point (Barnosky et al., 2011, 2012) and potentially a shift in macroevolutionary regime (Jablonski, 1986).

There are a few data quality concerns in this study which are also inherent to almost any paleontological study. Almost all of the body mass estimates were obtained using published regression equations that estimate mass from some other body part (e.g. tooth). These estimates are known with error, which was not included in the model. If the standard deviation of the residuals from each of these regression equations was known, it would be possible to directly model this as measurement error (Gelman et al., 2013).

A similar situation occurs with species bioprovince occupancy. Depending on the structure of the biogeographic network, there can be a range in the number of emergent bioprovinces. By estimating the standard deviation of both the number of bioprovinces and the number of occupied bioprovinces, it might be possible in propegate this measurement error correctly through the model (Gelman et al., 2013).

These model improvements were not done here for a variety of reasons: adequacy of current model fit, lack of residual standard deviation information from regression equations, and for improved tractability. Both of these measurement error models involve estimating the actual value given the observed and some known amout of error. Because of this, by allowing two covariates to be known with error approximately 4000 more (nuisance) parameter values would need to be estimated. Because of the combination of these factors, no measurement error in body size or bioprovince occupancy was included.

The phylogeny used here is only a coarse, baseline estimate of the actual species relationships. Because of this, the analysis of phylogenetic effect on survival represents a minimum or rough estimate. With improved topology and resolution, it would be possible to more accurately estimate the effect of shared evolutionary history. As it stand, given the rough minimum estimate of phylogenetic heritability, these results point to the importance of including shared evolutionary history in diversification models.

There are many processes encompassed by background extinction and identifying the exact cause of any one species' reason for extinction is extremely difficult.

By focusing on estimating the effects of different ecologies and historical factors on average extinction risk, it is possible to better understand what processes may have driven the resulting pattern of selection (i.e. diversity). Here, I focused on time-invariant factors and their relation to biological selectivity of extinction, possible reasons for the observed time-invariant effects, and the effects of taxon-age on extinction risk. I found that some organismal- and species-level traits such as omnivory and large geographic range size have time-invariant effects on mammal species extinction risk. I also found that there are small but non-ignorable effects of cohort and phylogeny. Finally, I found putative evidence of increasing extinction risk with species age, though this result may be partially due to the minimum resolution of the fossil record itself (Sepkoski, 1975).

References

J. Alroy. Constant extinction, constrained diversification, and uncoordinated stasis in North American mammals. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 127:285–311, 1996.

J. Alroy. Cope's rule and the dynamics of body mass evolution in North American fossil mammals. *Science*, 280:731–734, 1998. doi: 10.1126/science.280.5364.731.

J. Alroy. Speciation and extinction in the fossil record of North American mammals. In R. K. Butlin, J. R. Bridle, and D. Schluter, editors, *Speciation and patterns of diversity*, pages 302–323. Cambridge University Press, Cambridge, 2009.

J. Alroy. The Shifting Balance of Diversity Among Major Marine Animal Groups. *Science*, 329(5996):1191–1194, 2010.

J. Alroy, P. L. Koch, and J. C. Zachos. Global climate change and North American mammalian evolution. *Paleobiology*, 26(1981):259–288, 2000.

D. W. Bapst. paleotree: an R package for paleontological and phylogenetic analyses of evolution. *Methods in Ecology and Evolution*, 3:803–807, 2012. doi: 10.1111/j.2041-210X.2012.00223.x. URL `http://doi.wiley.com/10.1111/j.2041-210X.2012.00223.xpapers2://publication/doi/10.1111/j.2041-210X.2012.00223.x`.

D. W. Bapst. A stochastic rate-calibrated method for time-scaling phylogenies of fossil taxa. *Methods in Ecology and Evolution*, 4(8):724–733, Aug. 2013. ISSN 2041210X. doi: 10.1111/2041-210X.12081. URL `http://doi.wiley.com/10.1111/2041-210X.12081`.

A. D. Barnosky. Distinguishing the effects of the Red Queen Court Jester on Micene Mammal Evolution in the Northern Rocky Mountains. *Journal of Vertebrate Paleontology*, 21(1):172–185, 2001.

A. D. Barnosky, N. Matzke, S. Tomiya, G. O. U. Wogan, B. Swartz, T. B.

Quental, C. Marshall, J. L. Mcguire, E. L. Lindsey, K. C. Maguire, B. Mersey, and E. A. Ferrer. Has the Earths sixth mass extinction already arrived? *Nature*, 470(7336):51–57, 2011. URL `http://dx.doi.org/10.1038/nature09678papers2://publication/doi/10.1038/nature09678`.

A. D. Barnosky, E. a. Hadly, J. Bascompte, E. L. Berlow, J. H. Brown, M. Fortelius, W. M. Getz, J. Harte, A. Hastings, P. a. Marquet, N. D. Martinez, A. Mooers, P. Roopnarine, G. Vermeij, J. W. Williams, R. Gillespie, J. Kitzes, C. Marshall, N. Matzke, D. P. Mindell, E. Revilla, and A. B. Smith. Approaching a state shift in Earth's biosphere. *Nature*, 486 (7401):52–8, June 2012. ISSN 1476-4687. doi: 10.1038/nature11018. URL `http://www.ncbi.nlm.nih.gov/pubmed/22678279`.

T. K. Baumiller. Survivorship analysis of Paleozoic Crinoidea: effect of filter morphology on evolutionary rates. *Paleobiology*, 19(3):304–321, 1993.

O. R. P. Bininda-Emonds, M. Cardillo, K. E. Jones, R. D. E. Macphee, R. M. D. Beck, R. Grenyer, S. A. Price, R. A. Vos, J. L. Gittleman, and A. Purvis. The delayed rise of present-day mammals. *Nature*, 446(7135):507–512, 2007. URL `papers2://publication/doi/10.1038/nature05634`.

A. Birand, A. Vose, and S. Gavrilets. Patterns of species ranges, speciation, and extinction. *The American naturalist*, 179(1):1–21, Jan. 2012. ISSN 1537-5323. doi: 10.1086/663202. URL `http://www.ncbi.nlm.nih.gov/pubmed/22173457`.

J. L. Blois and E. A. Hadly. Mammalian Response to Cenozoic Climatic Change. *Annual Review of Earth and Planetary Sciences*, 37(1):181–208, May 2009. ISSN 0084-6597. doi: 10.1146/annurev.earth.031208.100055.

B. W. Brook and D. M. J. S. Bowman. The uncertain blitzkrieg of Pleistocene megafauna. *Journal of Biogeography*, 31(4):517–523, Apr. 2004. ISSN 03050270. doi: 10.1046/j.1365-2699.2003.01028.x. URL `http://doi.wiley.com/10.1046/j.1365-2699.2003.01028.x`.

J. H. Brown and B. A. Maurer. Evolution of species assembalges: effects of energetic constraints and species dynamics on the diversification of the North American avidauna. *The American Naturalist*, 130(1):1–17, 1987.

S. Chamberlain and E. Szocs. taxize - taxonomic search and retrieval in r. *F1000Research*, 2013. URL `http://f1000research.com/articles/2-191/v2`.

G. Csardi and T. Nepusz. The igraph software package for complex network research. *InterJournal*, Complex Systems:1695, 2006. URL `http://igraph.org`.

J. Damuth. Home range, home range overlap, and species energy use among herbivorous mammals. *Biological Journal of the Linnean Society*, 15:185–193, 1979.

J. Damuth. Population density and body size in mammals. *Nature*, 290:699–700, 1981. URL `http://www.jstor.org/stable/10.2307/2461771http://www.nature.com/nature/journal/v290/n5808/abs/290699a0.html`.

A. D. Davidson, A. G. Boyer, H. Kim, S. Pompa-Mansilla, M. J. Hamilton, D. P. Costa, G. Ceballos, and J. H. Brown. Drivers and hotspots of extinction risk in marine mammals. *Proceedings of the National Academy of Sciences*, 109(9):3395–400, Feb. 2012. ISSN 1091-6490. doi: 10.1073/pnas.1121469109. URL `http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3295301&tool=pmcentrez&rendertype=abstract`.

J. M. Drake. Tail probabilities of extinction time in a large number of experimental populations. *Ecology*, page 140206083444001, Feb. 2014. ISSN 0012-9658. doi: 10.1890/13-1107.1. URL `http://www.esajournals.org/doi/abs/10.1890/13-1107.1`.

S. Finnegan, J. L. Payne, and S. C. Wang. The Red Queen revisited: reevaluating the age selectivity of Phanerozoic marine genus extinctions. *Paleobiology*, 34(3):318–341, Sept. 2008. ISSN 0094-8373. doi: 10.1666/07008.1. URL `http://www.bioone.org/doi/abs/10.1666/07008.1`.

M. Foote. Survivorship analysis of Cambrian and Ordovician Trilobites. *Paleobiology*, 14(3):258–271, 1988.

M. Foote and A. I. Miller. Determinants of early survival in marine animal genera. *Paleobiology*, 39(2):171–192, Mar. 2013. ISSN 0094-8373. doi: 10.1666/12028. URL `http://www.bioone.org/doi/abs/10.1666/12028`.

M. Freudenthal and E. Martín-suárez. Estimating body mass of fossil rodents. *Scripta Geologica*, 145:1–130, 2013.

S. a. Fritz and A. Purvis. Selectivity in mammalian extinction risk and threat types: a new measure of phylogenetic signal strength in binary traits. *Conservation biology : the journal of the Society for Conservation Biology*, 24(4):1042–51, Aug. 2010. ISSN 1523-1739. doi: 10.1111/j.1523-1739.2010.01455.x. URL `http://www.ncbi.nlm.nih.gov/pubmed/20184650`.

S. a. Fritz, O. R. P. Bininda-Emonds, and A. Purvis. Geographical variation in predictors of mammalian extinction risk: big is bad, but only in the tropics. *Ecology letters*, 12(6):538–49, June 2009. ISSN 1461-0248. doi: 10.1111/j.1461-0248.2009.01307.x. URL `http://www.ncbi.nlm.nih.gov/pubmed/19392714`.

K. J. Gaston. Geographic range limits: achieving synthesis. *Proceedings. Biological sciences / The Royal Society*, 276(1661):1395–406, Apr. 2009. ISSN 0962-8452. doi: 10.1098/rspb.2008.1480. URL `http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2677218&tool=pmcentrez&rendertype=abstract`.

A. Gelman. Prior distributions for variance parameters in hierarchical models. *Bayesian Analysis*, 1(3):515–533, 2006.

A. Gelman. Scaling regression inputs by dividing by two standard deviations. *Statistics in Medicine*, pages 2865–2873, 2008. doi: 10.1002/sim.

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

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

H. Goldstein, W. Browne, and J. Rasbash. Partitioning variation in multilevel models. *Understanding Statistics*, 1(4):1–12, 2002. URL `http://www.tandfonline.com/doi/abs/10.1207/S15328031US0104_02`.

P. G. Harnik, C. Simpson, and J. L. Payne. Long-term differences in extinction risk among the seven forms of rarity. *Proceedings of the Royal Society B: Biological Sciences*, 282(1801), Oct. 2013. ISSN 0962-8452. doi: 10.1098/rspb. 2012.1902. URL `http://rspb.royalsocietypublishing.org/cgi/doi/10.1098/rspb.2012.1902`.

P. G. Harnik, P. C. Fitzgerald, J. L. Payne, and S. J. Carlson. Phylogenetic signal in extinction selectivity in Devonian terebratulide brachiopods. *Paleobiology*, 40(4):675–692, June 2014. ISSN 0094-8373. doi: 10.1666/14006. URL `http://www.bioone.org/doi/abs/10.1666/14006`.

M. M. Hedman. Constraints on clade ages from fossil outgroups. *Paleobiology*, 36(1):16–31, 2010.

M. D. Hoffman and A. Gelman. The no-U-turn sampler: Adaptively setting path lengths in Hamiltonian Monte Carlo. *arXiv*, 1111(4246), 2011. URL `http://arxiv.org/abs/1111.4246`.

E. A. Housworth, P. Martins, and M. Lynch. The Phylogenetic Mixed Model. *The American Naturalist*, 163(1):84–96, 2004.

G. Hunt. The relative importance of directional change, random walks, and stasis in the evolution of fossil lineages. In *Proceedings of the National Academy of ...*, 2007. URL `http://www.pnas.org/content/104/47/18404.shortpapers2://publication/uuid/DE43904B-DCEC-49C7-A107-93353C9A7ABE`.

J. G. Ibrahim, M.-H. Chen, and D. Sinha. *Bayesian Survival Analysis*. Springer, New York, 2001.

D. Jablonski. Background and mass extincitons: the alternation of macroevolutionary regimes. *Science*, 231(4734):129–133, 1986.

D. Jablonski. Heritability at the species level: analysis of geographic ranges of cretaceous mollusks. *Science*, 238(4825):360–363, Oct. 1987. ISSN 0036-8075. doi: 10.1126/science.238.4825.360. URL `http://www.ncbi.nlm.nih.gov/pubmed/17837117`.

D. Jablonski. Mass extinctions and macroevolution. *Paleobiol-*

ogy, 31(sp5):192–210, 2005. URL `papers2://publication/uuid/F296080D-6609-4C57-9514-C74E5E726340`.

D. Jablonski and G. Hunt. Larval ecology, geographic range, and species survivorship in Cretaceous mollusks: organismic versus species-level explanations. *The American naturalist*, 168(4):556–64, Oct. 2006. ISSN 1537-5323. doi: 10.1086/507994. URL `http://www.ncbi.nlm.nih.gov/pubmed/17004227`.

D. Jablonski and K. Roy. Geographical range and speciation in fossil and living molluscs. *Proceedings. Biological sciences / The Royal Society*, 270(1513):401–6, Feb. 2003. ISSN 0962-8452. doi: 10.1098/rspb.2002. 2243. URL `http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1691247&tool=pmcentrez&rendertype=abstract`.

C. M. Janis. Tertiary mammal evolution in the context of changing climates, vegetation, and tectonic events. *Annual Review of Ecology and Systematics*, 24:467–500, 1993.

C. M. Janis, K. M. Scott, and L. L. Jacobs. *Evolution of Tertiary mammals of North America. Vol. 1. Terrestrial carnivores, ungulates, and ungulatelike mammals*. Cambridge University Press, Cambridge, 1998.

C. M. Janis, J. Damuth, and J. M. Theodor. Miocene ungulates and terrestrial primary productivity: where have all the browsers gone? *Proceedings of the National Academy of Sciences*, 97(14):7899–904, July 2000. ISSN 0027-8424. URL `http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=16642&tool=pmcentrez&rendertype=abstract`.

C. M. Janis, G. F. Gunnell, and M. D. Uhen. *Evolution of Tertiary mammals of North America. Vol. 2. Small mammals, xenarthrans, and marine mammals*. Cambridge University Press, Cambridge, 2008.

J. Jernvall and M. Fortelius. Common mammals drive the evolutionary increase of hypsodonty in the Neogene. *Nature*, 417(6888):538–40, May 2002. ISSN 0028-0836. doi: 10.1038/417538a.

J. Jernvall and M. Fortelius. Maintenance of trophic structure in fossil mammal communities: site occupancy and taxon resilience. *American Naturalist*, 164 (5):614–624, Nov. 2004. ISSN 1537-5323. doi: 10.1086/424967.

C. N. Johnson. Determinats of loss of mammal species during the Late Quaternary 'megafauna' extinctions: life history and ecology, but not body size. *Proceedings of the Royal Society B: Biological Sciences*, 269:2221–2227, 2002. doi: 10.1098/rspb.2002.2130.

K. E. Jones, J. Bielby, M. Cardillo, S. A. Fritz, J. O'Dell, C. D. L. Orme, K. Safi, W. Sechrest, E. Boakes, C. Carbone, C. Connolly, M. J. Cutts, J. K. Foster, R. Grenyer, M. Habib, C. A. Plaster, S. A. Price, E. A. Rigby, J. Rist, A. Teacher, O. R. P. Bininda-Emonds, J. L. Gittleman, G. M. Mace, and A. Purvis. PanTHERIA : a species-level database of life history , ecology ,

and geography of extant and recently extinct mammals. *Ecology*, 90(9):2648, 2009.

J. A. Kitchell, D. L. Clark, and A. M. Gombos. Biological selectivity of extinction: a link between background and mass extinction. *Palaios*, 1(5):504–511, 1986.

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

L. H. Liow. A test of Simpson's "rule of the survival of the relatively unspecialized" using fossil crinoids. *The American naturalist*, 164(4):431–43, Oct. 2004. ISSN 1537-5323. doi: 10.1086/423673. URL `http://www.ncbi.nlm.nih.gov/pubmed/15459876`.

L. H. Liow, M. Fortelius, E. Bingham, K. Lintulaakso, H. Mannila, L. Flynn, and N. C. Stenseth. Higher origination and extinction rates in larger mammals. *Proceedings of the National Academy of Sciences*, 105(16):6097, 2008. URL `papers2://publication/uuid/5A93DDA3-204F-4D9C-AFAC-A15722A18C61`.

L. H. Liow, M. Fortelius, K. Lintulaakso, H. Mannila, and N. C. Stenseth. Lower Extinction Risk in SleeporHide Mammals. *The American Naturalist*, 173(2): 264–272, 2009. URL `papers2://publication/doi/10.1086/595756`.

L. H. Liow, L. Van Valen, and N. C. Stenseth. Red Queen: from populations to taxa and communities. *Trends in ecology & evolution*, 26(7):349–58, July 2011. ISSN 0169-5347. doi: 10.1016/j.tree.2011.03.016. URL `http://www.ncbi.nlm.nih.gov/pubmed/21511358`.

M. Lynch. Methods for the analysis of comparative data in evolutionary biology. *Evolution*, 45(5):1065–1080, 1991.

S. K. Lyons. A quantitative model for assessing community dynamics of pleistocene mammals. *The American naturalist*, 165(6):E168–85, June 2005. ISSN 1537-5323. doi: 10.1086/429699. URL `http://www.ncbi.nlm.nih.gov/pubmed/15937741`.

S. K. Lyons, P. J. Wagner, and K. Dzikiewicz. Ecological correlates of range shifts of Late Pleistocene mammals. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 365 (1558):3681–93, Nov. 2010. ISSN 1471-2970. doi: 10.1098/rstb.2010.0263. URL `http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2981999&tool=pmcentrez&rendertype=abstract`.

J. D. Marcot. The fossil record and macroevolutionary history of North American ungulate ungulate mammals: standardizing variation in intensity and geography of sampling. *Paleobiology*, 40(2):237–254, Feb. 2014. ISSN 0094-8373. doi: 10.1666/13052. URL `http://www.bioone.org/doi/abs/10.1666/13052`.

A. P. Martin and S. R. Palumbi. Body size, metabolic rate, generation time, and the molecular clock. *Proceedings of the Na-*

tional Academy of Sciences, 90(9):4087–91, May 1993. ISSN 0027-8424. URL `http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=46451&tool=pmcentrez&rendertype=abstract`.

R. T. McKenna. *Potential for Speciation in Mammals Following Vast , Late Miocene Volcanic Interruptions in the Pacific Northwest*. Masters, Portland State University, 2011.

S. Nürnberg and M. Aberhan. Habitat breadth and geographic range predict diversity dynamics in marine Mesozoic bivalves. *Paleobiology*, 39(3):360–372, Apr. 2013. ISSN 0094-8373. doi: 10.1666/12047. URL `http://www.bioone.org/doi/abs/10.1666/12047`.

J. L. Payne and S. Finnegan. The effect of geographic range on extinction risk during background and mass extinction. *Proceedings of the National Academy of Sciences of the United States of America*, 104(25):10506–11, June 2007. ISSN 0027-8424. doi: 10.1073/pnas.0701257104. URL `http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1890565&tool=pmcentrez&rendertype=abstract`.

R. H. Peters. *The ecological implications of body size*. Cambridge University Press, Cambridge, 1983.

S. A. Price, S. S. B. Hopkins, K. K. Smith, and V. L. Roth. Tempo of trophic evolution and its impact on mammalian diversification. *Proceedings of the National Academy of Sciences of the United States of America*, 109(18):7008–12, May 2012. ISSN 1091-6490. doi: 10.1073/pnas.1117133109. URL `http://www.ncbi.nlm.nih.gov/pubmed/22509033`.

a. Purvis, J. L. Gittleman, G. Cowlishaw, and G. M. Mace. Predicting extinction risk in declining species. *Proceedings. Biological sciences / The Royal Society*, 267(1456):1947–52, Oct. 2000. ISSN 0962-8452. doi: 10.1098/rspb.2000.1234. URL `http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1690772&tool=pmcentrez&rendertype=abstract`.

T. B. Quental and C. R. Marshall. How the Red Queen Drives Terrestrial Mammals to Extinction. *Science*, 341(6143):290–292, Sept. 2013. ISSN 0022-1376. doi: 10.1086/627905. URL `http://www.journals.uchicago.edu/doi/abs/10.1086/627905`.

R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2014. URL `http://www.R-project.org/`.

D. L. Rabosky. Extinction rates should not be estimated from molecular phylogenies. *Evolution*, 64(6):1816–24, June 2010. ISSN 1558-5646. doi: 10.1111/j.1558-5646.2009.00926.x. URL `http://www.ncbi.nlm.nih.gov/pubmed/20030708`.

P. Raia, F. Carotenuto, F. Passaro, D. Fulgione, and M. Fortelius. Ecological specialization in fossil mammals explains Cope's rule. *The American naturalist*,

179(3):328–37, Mar. 2012. ISSN 1537-5323. doi: 10.1086/664081. URL `http://www.ncbi.nlm.nih.gov/pubmed/22322221`.

D. M. Raup. Taxonomic survivorship curves and Van Valen's Law. *Paleobiology*, 1(1):82–96, Jan. 1975. ISSN 0036-8075. doi: 10.1126/science.49.1254.50. URL `http://www.ncbi.nlm.nih.gov/pubmed/17777225`.

D. M. Raup. Cohort Analysis of generic survivorship. *Paleobiology*, 4(1):1–15, 1978.

D. M. Raup. *Extinction: Bad Genes or Bad Luck?* Norton, New York, 1991.

D. M. Raup. The role of extinction in evolution. *Proceedings of the National Academy of Sciences*, 91(July):6758–6763, 1994.

D. M. Raup and G. E. Boyajian. Patterns of generic extinction in the fossil record. *Paleobiology*, 14(2):109–125, 1988.

L. J. Revell. phytools: An r package for phylogenetic comparative biology (and other things). *Methods in Ecology and Evolution*, 3:217–223, 2012.

M. Rosvall and C. T. Bergstrom. Maps of random walks on complex networks reveal community structure. *Proceedings of the National Academy of Sciences*, 105(4):1118–23, Jan. 2008. ISSN 1091-6490. doi: 10.1073/pnas. 0706851105. URL `http://www.pubmedcentral.nih.gov/articlerender. fcgi?artid=2234100&tool=pmcentrez&rendertype=abstract`.

M. Rosvall, D. Axelsson, and C. Bergstrom. The map equation. *The European Physical Journal Special Topics*, 178(14):13–24, 2009. URL `http://www. springerlink.com/index/H8193132U6432363.pdf`.

K. Roy, G. Hunt, D. Jablonski, A. Z. Krug, and J. W. Valentine. A macroevolutionary perspective on species range limits. *Proceedings. Biological sciences / The Royal Society*, 276(1661):1485–93, Apr. 2009. ISSN 0962-8452. doi: 10.1098/rspb.2008.1232. URL `http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid= 2677224&tool=pmcentrez&rendertype=abstract`.

H. Schielzeth. Simple means to improve the interpretability of regression coefficients. *Methods in Ecology and Evolution*, 1(2):103–113, Feb. 2010. ISSN 2041210X. doi: 10.1111/j.2041-210X.2010.00012.x. URL `http://doi.wiley. com/10.1111/j.2041-210X.2010.00012.x`.

J. J. Sepkoski. Stratigraphic biases in the analysis of taxonomic survivorship. *Paleobiology*, 1(4):343–355, 1975.

J. J. Sepkoski Jr. A kinetic model of Phanerozoic taxonomic diversity. I. Analysis of marine orders. *Paleobiology*, 4(3):223–251, 1978.

J. J. Sepkoski Jr. A kinetic model of Phanerozoic taxonomic diversity. II. Early Phanerozoic families and multiple equilibria. *Paleobiology*, 5(3):222–251, 1979.

J. J. Sepkoski Jr. A kinetic model of Phanerozoic taxonomic diversity. III. Post-Paleozoic families and mass extinctions. *Paleobiology*, 10(2):246–267, 1984.

C. A. Sidor, D. A. Vilhena, K. D. Angielczyk, A. K. Huttenlocker, S. J. Nesbitt, B. R. Peecook, J. S. Steyer, R. M. H. Smith, and L. A. Tsuji. Provincialization of terrestrial faunas following the end-Permian mass extinction. *Proceedings of the National Academy of Sciences*, 110(20):8129–33, May 2013. ISSN 1091-6490. doi: 10.1073/pnas.1302323110.

C. Simpson, W. Kiessling, H. Mewis, R. C. Baron-Szabo, and J. Müller. Evolutionary diversification of reef corals: a comparison of the molecular and fossil records. *Evolution; international journal of organic evolution*, 65(11):3274–84, Nov. 2011. ISSN 1558-5646. doi: 10.1111/j.1558-5646.2011.01365.x. URL `http://www.ncbi.nlm.nih.gov/pubmed/22023591`.

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

G. Slater, L. Harmon, and M. Alfaro. Integrating fossils with molecular phylogenies improves inference of trait evolution. *Evolution*, 2012. doi: 10.5061/dryad.q96d7. URL `http://onlinelibrary.wiley.com/doi/10.1111/j.1558-5646.2012.01723.x/full`.

G. J. Slater. Phylogenetic evidence for a shift in the mode of mammalian body size evolution at the Cretaceous-Palaeogene boundary. *Methods in Ecology and Evolution*, 4(8):734–744, Aug. 2013. ISSN 2041210X. doi: 10.1111/2041-210X.12084. URL `http://doi.wiley.com/10.1111/2041-210X.12084`.

F. A. Smith, J. Brown, J. Haskell, and S. Lyons. Similarity of mammalian body size across the taxonomic hierarchy and across space and .... *The American Naturalist*, 2004a. URL `http://www.journals.uchicago.edu/doi/abs/10.1086/382898papers2://publication/uuid/D5606802-FD91-49EB-BE2F-E2D314A5E71D`.

F. A. Smith, J. H. Brown, J. P. Haskell, S. K. Lyons, J. Alroy, E. L. Charnov, T. Dayan, B. J. Enquist, S. K. M. Ernest, E. A. Hadly, K. E. Jones, D. M. Kaufman, P. A. Marquet, B. A. Maurer, K. J. Niklas, W. P. Porter, B. Tiffney, and M. R. Willig. Similarity of Mammalian Body Size across the Taxonomic Hierarchy and across Space and Time. *The American Naturalist*, 163(5):672–691, 2004b.

F. A. Smith, S. K. Lyons, S. Morgan Ernest, and J. H. Brown. Macroecology: more than the division of food and space among species on continents. *Progress in Physical Geography*, 32(2):115–138, Apr. 2008. ISSN 0309-1333. doi: 10.1177/0309133308094425. URL `http://ppg.sagepub.com/cgi/doi/10.1177/0309133308094425`.

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

C. A. E. Strömberg. Decoupled taxonomic radiation and ecological expansion of open-habitat grasses in the Cenozoic of North America. *Proceedings of the National Academy of Sciences of the United States of America*, 102(34):11980–4, Aug. 2005. ISSN 0027-8424. doi: 10.1073/pnas. 0505700102. URL `http://www.pubmedcentral.nih.gov/articlerender. fcgi?artid=1189350&tool=pmcentrez&rendertype=abstract`.

C. A. E. Strömberg, R. E. Dunn, R. H. Madden, M. J. Kohn, and A. A. Carlini. Decoupling the spread of grasslands from the evolution of grazer-type herbivores in South America. *Nature communications*, 4:1478, Jan. 2013. ISSN 2041-1723. doi: 10.1038/ncomms2508. URL `http://www.ncbi.nlm.nih.gov/ pubmed/23403579`.

S. Tomiya. Body Size and Extinction Risk in Terrestrial Mammals Above the Species Level. *The American Naturalist*, pages E000–E000, Sept. 2013. ISSN 00030147. doi: 10.1086/673489. URL `http://www.jstor.org/stable/info/ 10.1086/673489`.

L. Van Valen. A new evolutionary law. *Evolutionary Theory*, 1:1–30, 1973. URL `http://ci.nii.ac.jp/naid/10011264287/`.

L. Van Valen. Taxonomic survivorship curves. *Evolutionary Theory*, 4:129–142, 1979.

D. A. Vilhena, E. B. Harris, C. T. Bergstrom, M. E. Maliska, P. D. Ward, C. A. Sidor, C. A. E. Strömberg, and G. P. Wilson. Bivalve network reveals latitudinal selectivity gradient at the end-Cretaceous mass extinction. *Scientific reports*, 3:1790, May 2013. ISSN 2045-2322. doi: 10.1038/srep01790.

P. J. Wagner and G. F. Estabrook. Trait-based diversification shifts reflect differential extinction among fossil taxa. *Proceedings of the National Academy of Sciences of the United States of America*, pages 1–6, Oct. 2014. ISSN 1091-6490. doi: 10.1073/pnas.1406304111. URL `http://www.ncbi.nlm.nih. gov/pubmed/25331898`.

S. C. Wang. On the continuity of background and mass extinction. *Paleobiology*, 29(4):455–467, Dec. 2003. ISSN 0094-8373. doi: 10.1666/0094-8373(2003) 029⟨0455:OTCOBA⟩2.0.CO;2. URL `http://www.bioone.org/doi/abs/10. 1666/0094-8373%282003%29029%3C0455%3AOTCOBA%3E2.0.CO%3B2`.

E. P. White, S. K. M. Ernest, A. J. Kerkhoff, and B. J. Enquist. Relationships between body size and abundance in ecology. *TRENDS in Ecology and Evolutioncology & evolution*, 22(6):323–30, June 2007. ISSN 0169-5347. doi: 10.1016/j.tree.2007.03.007. URL `http://www.ncbi.nlm.nih.gov/pubmed/ 17399851`.