Death and taxa: time invariant differences in species duration

Peter D Smits

Committee on Evolutionary Biology, University of Chicago

January 15, 2015

Abstract

1 Introduction

Extinction is one of most important components of diversification, second only to speciation. 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 making predictions about species' vulnerability to anthropogenic impacts. Yet investigations into the factors that affect extinction rate have proved inconclusive and often contradictory results.

I am testing the following hypotheses: is extinction non-random with respect to ecology during times of background extinction, and is extinction taxon age independent among Cenozoic mammals? I use a Bayesian hierarchical survival modeling approach for understanding the effects of organismal and species traits, time of origination, and shared evolutionary history on species duration while also allowing for potentially time–dependent extinction. This is a direct extension of the more traditional dynamic and cohort survival analyses 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 phylogeny 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 means 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 of that effect does not change. For example, geographic range size has been identified as a potentially 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 was the point of Alroy et al. (2000): there was no consistent, time invariant response to climate change. While it is almost certain that selection pressures vary with time, consistent effects reveal fundamental differences in fitness or responses to selection.

Background extinction, or extinction occurring not at a mass extinction, is considered to involve many potential factors influencing 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 for a given sample size. The relationship between extinction risk and other traits, specifically organismal ones, is less well known because the effect size is most likely much smaller which makes 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 taxon's 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 influencing 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 differences in organismal traits (Raup, 1991). However, organismal traits directly related to species–environment interactions may play an important role in determining extinction risk. The inclusion of organismal trait data are necessary to determine if and how both species and organismal traits contribute to differences in extinction risk or not. 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 taxon, 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 paleontological 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 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 better understood from a mechanistic perspective.

Locomotor categories describe the motility of a taxon, plausibility of occurrence in a particular habitat, and dispersal ability. Dispersal ability is important for determining both the extent of a taxon's 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 taxa during the Paleogene will have a greater expected duration than Neogene taxa while the opposite will be true for ground dwelling taxa. In comparison, taxon duration of scansorial taxa is 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

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.

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).

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 these results comparable to previous studies (Jernvall and Fortelius, 2004, Price et al., 2012).

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 bin of occurrence to the last bin of occurrence, inclusive.

Species body size estimates were sourced from a large selection of primary literature and 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 time bin, a bipartite biogeographic network was created between species occurrences and spatial units. Spatial units were defined was 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*. Taxa are not linked to each other, nor are localities linked. 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.

2.1.2 Semi-formal supertree

Because there exists no 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 basis for additional revision. The taxonomy of many species was 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 (Scott Chamberlain and Eduard Szocs, 2013). This was additionally correct using various published phylogenies and taxonomies of fossil mammals (Janis et al., 1998, 2008, Raia et al., 2012), producing a tree that was a series of nested polytomies.

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 an important convenience 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 of interest in this study.

2.2 Survival model

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 to fit 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. 1) with shape α and scale σ parameters.

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

Following standard practice in survival analysis (Klein and Moeschberger, 2003), α was assumed constant and was given a weakly informative half-Cauchy prior. σ was reparameterized as an exponentiated regression model (Eq. 2). Note, the inclusion of α in the denominator of the exponentiated term is a standard parameterization for survival models using a Wiebull distribution (Klein and Moeschberger, 2003).

$$\sigma = \exp\left(\frac{-(h_i + \eta_{j[i]} + \sum \beta^T \mathbf{X}_i)}{\alpha}\right) \tag{2}$$

K species level covariates were included as a $n \times K$ matrix, $\mathbf{X}$. The covariates of interest included the logit of mean relative occupancy and the logarithm of body size (g). The discrete covariate index variables 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. Only $k-1$ indicator variables are necessary as the intercept takes on the remaining value. Finally, a vector of 1-s were 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 was 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 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). 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 were divided by twice their standard deviation (Gelman, 2008). The above model was fit with both unstandardardized and standardized 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 (Gelman and Hill, 2007), where complete pooling is when the differences between groups are ignored while no pooling is where different groups are analyzed separately. By allowing for partial pooling, we are modeling a compromise between these two extremes which allows for better inference. This is done by having all of the groups sharing the same prior with the 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. Hierarchical modeling is analogous to mixed-effects modeling (Gelman and Hill, 2007).

Origination cohort is defined as the group of species which originated during the same 2 My temporal bin. The most recent temporal bin, 0-2 Mya, was excluded, leaving 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 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 relative importance of these two end members.

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

The choice of prior for 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 was assumed known up to a constant variance, σ_p^2 (Housworth et al., 2004, Lynch, 1991). More fully, this is written

7

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

with $\mathbf{V}_{phy}$ being 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.

Both η and h were centered at 0 so that these effects can be interpreted as differences from the mean of the observed.

2.2.2 Censored observations

An important part of survival analysis is the inclusion of "censored" observations (Ibrahim et al., 2001, Klein and Moeschberger, 2003) or observations where the failure time has not been observed. The most common censored observation is right censored, where the point of extinction had not yet been observed in the period of study. In this case, this means 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 this study, taxa occurring in only a single time bin were left censored. Because of the minimum resolution of the record, we cannot observe if these taxa went extinct in less than that single bin or not.

To model censored data, there are a few functions which must first be defined following Klein and Moeschberger (2003). The survival function, $S(t)$, for a Weibull model is

$$S(t|\alpha, \sigma) = \exp\left(-\left(\frac{t}{\sigma}\right)^{\alpha}\right), \tag{3}$$

where σ is defined as above (Eq. 2). $S(t)$ is the probability that an observation will survive longer than a give time, t. The likelihood of uncensored observations is evaluated as as normal using Equation 1 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 cummulative density function and $S(t)$ is equivalent to the complementary cummulative density function (Gelman et al., 2013).

Fully, the likelihood for both uncensored and both right and left censored observations is written

$$L \propto \prod_{i \in C} \text{Weibull}(y_i|\alpha, \sigma) \prod_{j \in R} S(y_j|\alpha, \sigma) \prod_{k \in L} \left(1 - S(y_k|\alpha, \sigma)\right),$$
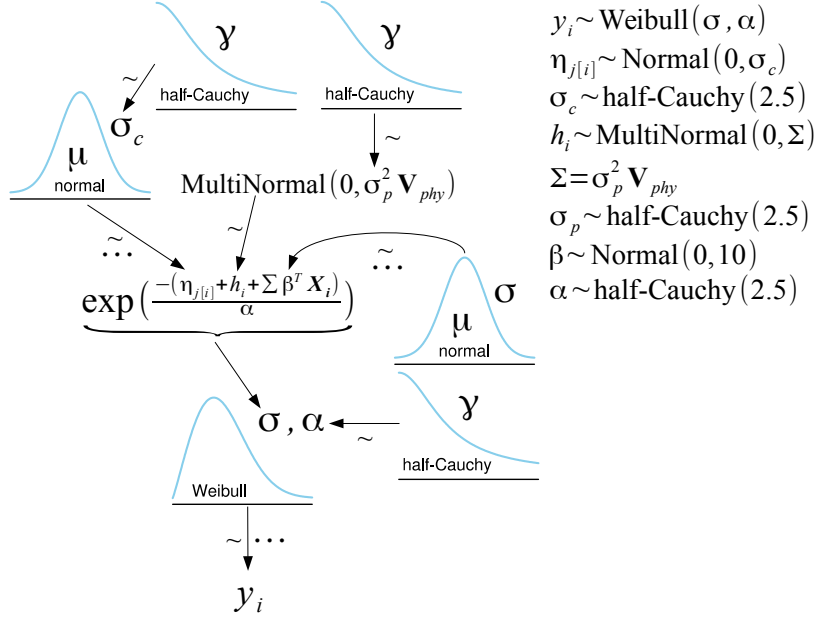
Figure 1: Graphical depiction of full survival model, save censoring, used here. Exact values for the hyperparameters are presented to the right. The observed duration of the ith observation is indicated towards the bottom left as y_i, which is assumed to follow a Weibull distribution.

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 calculations 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. Chain 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).

Both models with and without phylogenetic effects were estimated. Because inverting a large matrix is a memory intense procedure and because the phylogenetic covariance matrix is only assumed known up to a constant, every iteration of the MCMC would involve solving a very large matrix which is not ideal. In order to speed up the MCMC routine, this aspect of the model had to be reparameterized for efficiency purposes. Because of the size of the covariance matrix a custom multivariate sampler was used (see Appendix: Code).

For the model without phylogenetic effect the four MCMC chains ran for 1000 steps, with the first 500 used as warm-up and the last 500 as samples from the posterior. Because of the added complexity of estimating the phylogenetic effect, instead all four chains were run 20000 steps 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 predictors 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 repeated 1000 times and the distribution of y^{rep} was compared with the observed y (Gelman et al., 2013). This was done both graphically and numerically.

An example posterior predictive check used in this study is a graphical comparison of the Kaplan-Meier (K-M) survival curve estimated from the observed data with the survival curves from 1000 simulations. K-M survival curves are non-parametric estimates of the function $S(t)$ or the probability of a species going extinct given that it has lived to time t (Klein and Moeschberger, 2003). Other posterior predictive checks included 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 ordinary linear regression, the residuals are defined as $r = y - y_{est}$. For hierarchical models, this definition is inadequate. For survival analysis, the equivalent values 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. 3), 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

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

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. 1) does not include an estimate of overall variance, as the errors are not normally distributed. The partitioning of the variance between these three components was instead 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. 2) for all w simulations, holding h constant at 0.

3. Calculate υ_c, the Weibull variance (Eq. 4) of each element of σ^{c*} with α drawn from the posterior estimate.

4. Draw a random value from the diagonal of $\mathbf{V}_{phy}$, p_r

5. Simulate w values of h; $h \sim \mathcal{N}\left(0, \sqrt{\sigma_p^2 \times p_r}\right)$

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

7. Calculate υ_p, the Weibull variance (Eq. 4) of each element of σ^{p*} with α drawn from the posterior estimate.

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

9. $\sigma_{c*}^2 = var(\upsilon_c)$ and $\sigma_{p*}^2 = var(\upsilon_p)$

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

$$var(x) = \sigma^2 \left(\Gamma \left(1 + \frac{2}{\alpha} \right) - \left(\Gamma \left(1 + \frac{1}{\alpha} \right) \right)^2 \right), \tag{4}$$

where Γ is the gamma function.

The variance partitioning coefficients are then calculated $VPC_{phylo} = \frac{\sigma^2_{p*}}{\sigma^2_{y*} + \sigma^2_{c*} + \sigma^2_{p*}}$ and similar for the other variance components.

I used ratios of the variances and variance partitioning coefficients (VPC) to measure the amount of partial pooling. The former, when compared to the sample size of the hierarchical effect such as the number of cohorts, is an estimate of the relative amount of pooling, while the latter is a measure of the relative importance of the different variance components (Gelman and Hill, 2007). Phylogenetic heritability, h_p^2 (Housworth et al., 2004), is effectively 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 tree effect.

3   Results

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

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

First, visual examination of the deviance residuals from twelve different sets of posterior predictive values indicate little systematic problems (Fig. 2). The only concern is that the model tends to over estimate observed durations. This can also be observed when comparing the K-M estimate of the empirical survival function to 100 estimated survival functions from posterior simulations (Fig. 3).

Though this bias appears to be concentrated at the tails of the duration distribution, as posterior predictive checks comparing the 25th, 50th, 75th quantiles as well as the mean of the observed durations to the results from simulation indicate adequate model (Fig. 4). Between all of these different posterior predictive checks, 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).

Locomotor 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 dietary 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 55 and 50 My that have a much lower extinction risk than other cohorts of similar age.

Because the estimate of the Weibull shape parameter, α, is greater than 1 this

14

means that extinction risk is expected to increase with taxon age (Table 2). The estimate of α is also rather tightly constrained with 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

$$h(t) = \frac{\alpha}{\sigma} \left(\frac{t}{\sigma} \right)^{\alpha-1}.$$ (5)

$h(t)$ 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 will gradually increase with age (Fig. 8).

Of the three sources of variance present in the model, variation in individual species accounts for at least 70% of unexplained variance (Fig. 9). Both cohort and phylogenetic effects account for approximately 30% of the unexplained 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 accounts for approximately 15% of unexplained variance individually, the total combined effect of these factors indicates that neither can be ignored.

For example, Housworth et al. (2004) state that when the value of 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 "phylogenetic signal" in generic duration (Tomiya, 2013).

4   Discussion

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, P. L. Koch, and J. C. Zachos. Global climate change and North American mammalian evolution. *Paleobiology*, 26(1981):259–288, 2000.
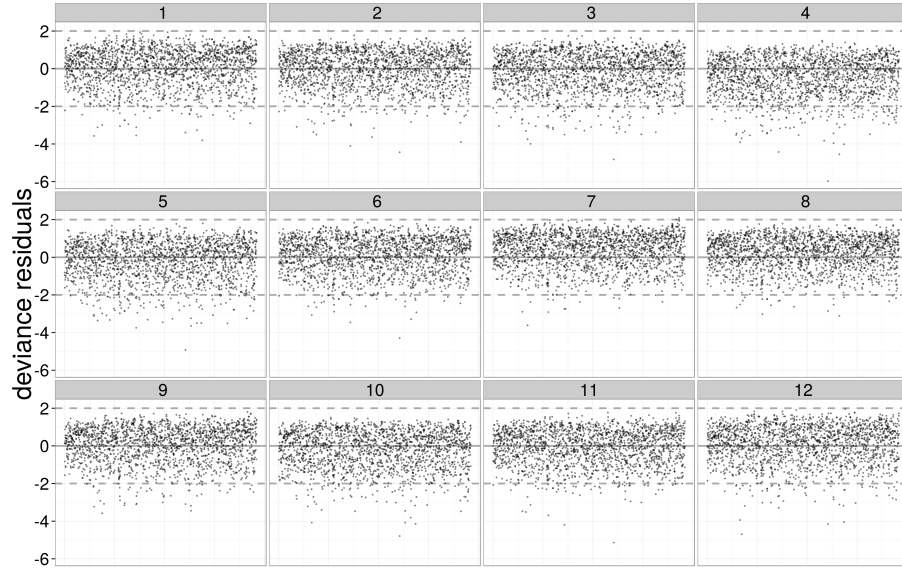
Figure 2: Deviance residuals from the fitted survival model. Each graph depicts the residuals from single draws from the posteriors distributions of all estimated parameters. Positive values indicate an under estimate of the observed duration, while negative values indicate an over estimate of the observed duration. Twelve difference examples are provided here to indicate the lack of individual observation based biases.

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`.

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

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`.

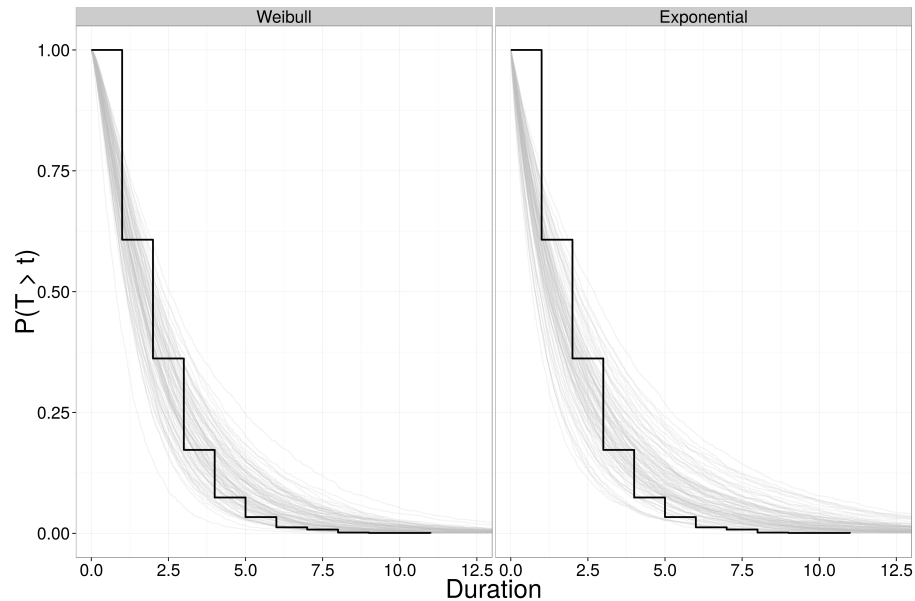J. L. Blois and E. A. Hadly. Mammalian Response to Cenozoic Climatic Change.

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.

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.

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

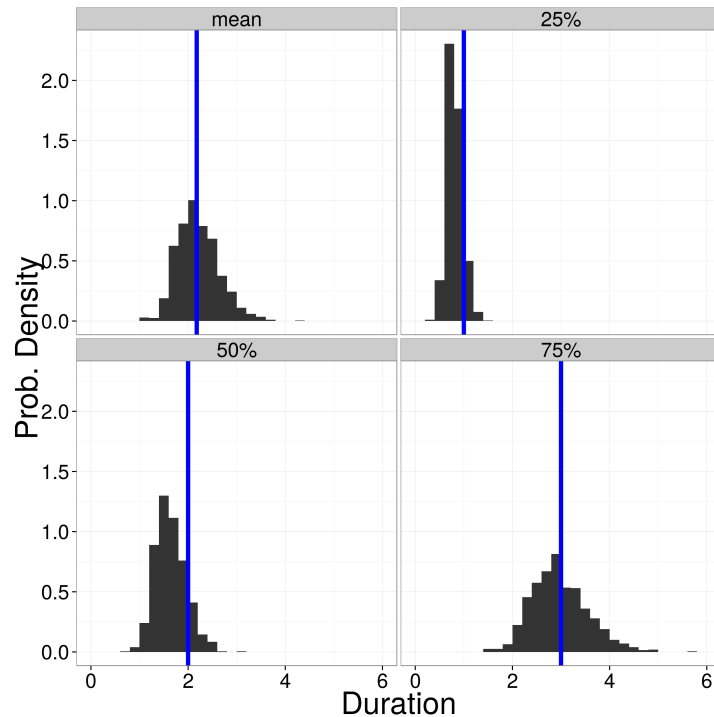J. Damuth. Home range, home range overlap, and species energy use among

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.

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`.
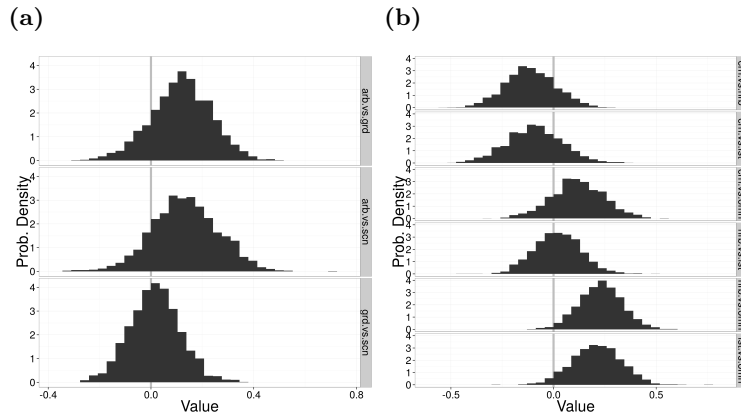
Figure 5: Pairwise differences in effect of the locomotor (a) and dietary categories (b) on expected duration from 1000 samples from the posterior distribution. Comparisons of locomotor categories, from top to bottom (a), are: arboreal versus ground dwelling, arboreal versus scansorial, and ground dwelling 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.

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

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.
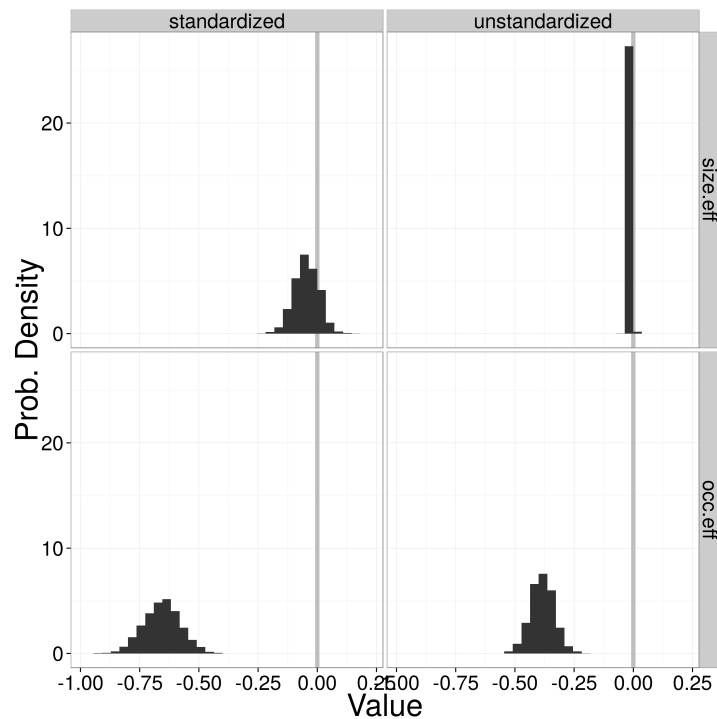
Figure 6

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*, (October), Oct. 2013. ISSN 0962-8452. doi: 10.1098/rspb. 2012.1902. URL `http://rspb.royalsocietypublishing.org/cgi/doi/10.1098/rspb.2012.1902`.

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`.
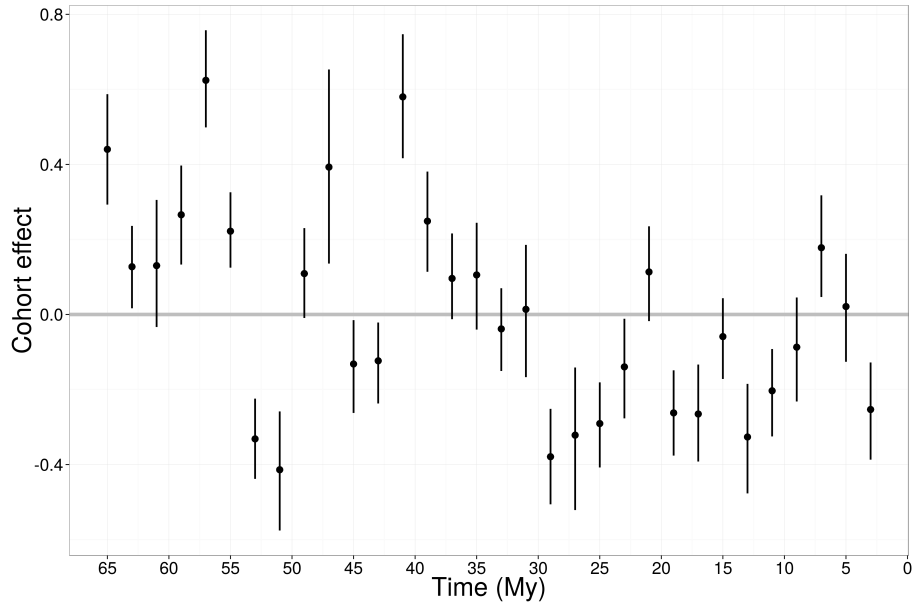
Figure 7: Summaries of estimated individual cohort effect posteriors. Depicted are medians and 80% credible intervals of the estimated posterior distributions. High values correspond to shorter species durations while lower values correspond to greater species durations compared to the mean duration. Lines are placed at the middle of the 2 My origination cohorts.

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

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. *Paleobiology*, 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.
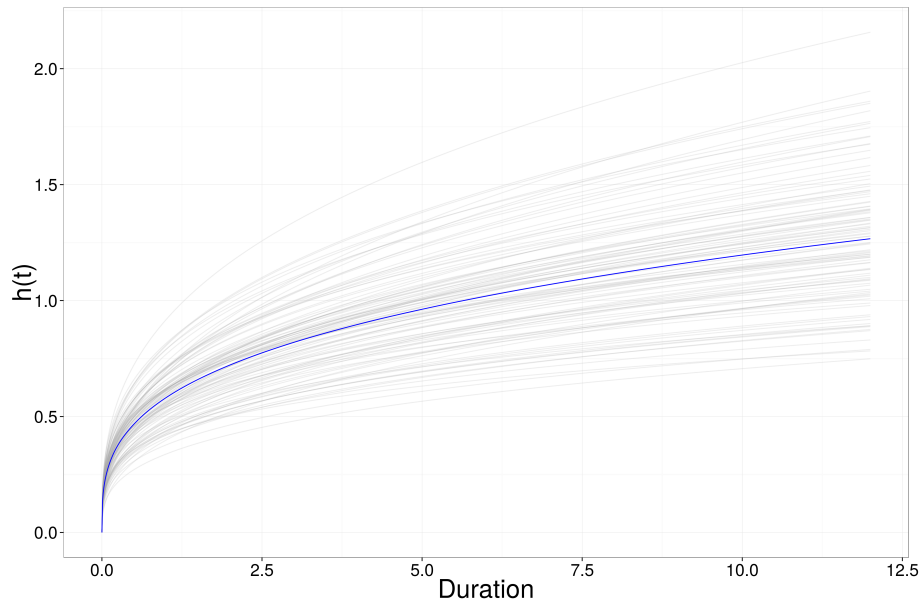
Figure 8: 100 estimates of the hazard function ($h(t)$ for the observed species mean (grey), along with the median estimated hazard function. $h(t)$ is an estimate of the rate at which a species of age t is expected to go extinct. Hazard functions were estimated from random draws from the estimated posterior distributions and evaluated with all covariate information set to 0, which corresponds to the expected duration of the mean species.

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, 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.
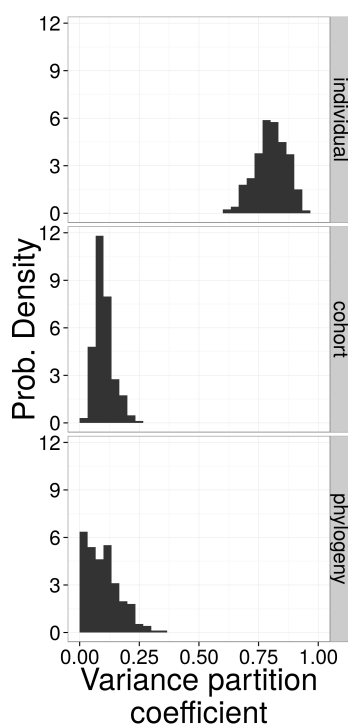
Figure 9: Estimates of the variance partitioning coefficients for the three different sources of variance: species, cohort, and phylogeny. Higher values correspond to greater contribution to total unexplained variance. Each of the estimates is a distribution of 1000 approximating simulations due to the model's non-normality.

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, 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`.

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 National 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`.

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.

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`.

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

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

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

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.

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`.