

Expected time-invariant effects of biological traits on mammal species duration

Peter D. Smits¹

Committee on Evolutionary Biology, University of Chicago, Chicago, IL 60637

Edited by Mark A. McPhee, Dartmouth College, Hanover, NH, and accepted by the Editorial Board September 3, 2015 (received for review May 29, 2015)

Determining which biological traits influence differences in extinction risk is vital for understanding the differential diversification of life and for making predictions about species' vulnerability to anthropogenic impacts. Here I present a hierarchical Bayesian survival model of North American Cenozoic mammal species durations in relation to species-level ecological factors, time of origination, and phylogenetic relationships. I find support for the survival of the unspecialized as a time-invariant generalization of trait-based extinction risk. Furthermore, I find that phylogenetic and temporal effects are both substantial factors associated with differences in species durations. Finally, I find that the estimated effects of these factors are partially incongruous with how these factors are correlated with extinction risk of the extant species. These findings parallel previous observations that background extinction is a poor predictor of mass extinction events and suggest that attention should be focused on mass extinctions to gain insight into modern species loss.

macroevolution | macroecology | extinction | paleobiology | Bayesian

Why extinction risk varies among species remains one of the most fundamental questions in paleobiology and conservation biology (1–5). To address this issue, I test for similarities in associations between extinction risk and multiple species-level traits during times of background extinction and in the modern world; which traits have time-invariant effects on species duration; and whether extinction is age independent. I approach these questions together by using a model of species duration whose parameter estimates act as direct tests of these questions. Cenozoic mammals are an ideal focus for this study because their fossil record is well sampled and well resolved both temporally and spatially and because individual species ecology and taxonomic position are generally understood (1, 4, 6–10).

Time-invariant factors are those that have a constant directional effect even if their magnitude varies. Because change in the magnitude of extinction risk is not necessarily the best indicator of a shift from background to mass extinction (11), it is better to look for changes in either the direction of selection, the loss of a selective pressure, or the appearance of novel selective pressures (12).

The species-level traits studied here are bioprovince occupancy, body mass, and dietary and locomotor categories. These traits are related to aspects of a species' adaptive zone such as population density, expected range size, potential prey, and dispersal ability (8, 13) and are a combination of aggregate and emergent traits (14). It is expected that species with larger geographic ranges have lower extinction rates than species with smaller geographic ranges (12, 15); however, how traits more directly related to species–environment interactions may affect species extinction risk is more nebulous.

Body size is a complex trait related to many life history characteristics. There are three general hypotheses of how body size may effect extinction risk: (i) positive effect where an increase in body size causes an increase in extinction risk, potentially due to associated decrease in reproductive rate or other similar life history traits (7, 16); (ii) negative effect where an increase in body size causes a decrease in extinction risk because of an expected positive

relationship between body size and geographic range; and (iii) no effect of body size on extinction risk (9).

The strongest expectation for the effects of dietary category on extinction risk is that omnivores will have the lowest extinction risk of all species. This expectation is based on the long standing “survival of the unspecialized” hypothesis where more generalist species (e.g., omnivores) have greater survival than specialist species (e.g., carnivores/herbivores) (1, 17). It has also been observed that both carnivores and herbivores have greater diversification rates than omnivores, with herbivores diversifying faster than carnivores (18). How this result translates into differences in extinction risk is currently unknown (19). In modern taxa, higher trophic levels (e.g., carnivores vs. herbivores) have been associated with greater extinction risk, most likely because of human extermination of top predators (16, 20).

Similarly, there are few expectations of how locomotor category may effect extinction risk. During the Cenozoic, there was a shift at the Paleogene/Neogene boundary from predominately closed to predominately open environments (21, 22). Based on this observation, a prediction is that arboreal taxa will have the greatest extinction risk of all, with both scansorial and ground dwelling taxa having lower extinction risks.

I use a hierarchical Bayesian survival model of species duration as predicted by the covariates of interest along with species' temporal and phylogenetic context. Species duration, in 2-My bins, was modeled as realizations from either an exponential or Weibull distribution-based hierarchical model (23). The exponential distribution corresponds to the Law of Constant Extinction, which states that extinction is age independent (2). Note that the exponential is a special case of the Weibull when its shape parameter, α , is 1. The Weibull distribution allows for extinction to be

Significance

Determining which biological traits influence differences in extinction risk is vital for understanding differential diversification of life and for making predictions about species' vulnerability to human impacts. The durations of fossil species are a rich source of information for estimating systematic differences in extinction risk. I analyzed Cenozoic North American fossil mammal species durations and their relationship with multiple ecological traits, their time of origination, and their shared evolutionary histories. I find support for generalists having a lower extinction risk than specialists as a general statement. When these results are compared with risk factors associated with living mammals, I find some incongruities that may indicate that the current biodiversity crisis is akin to the great mass extinctions in Earth's history.

Author contributions: P.D.S. designed research, performed research, analyzed data, and wrote the paper.

The author declares no conflict of interest.

This article is a PNAS Direct Submission. M.A.M. is a guest editor invited by the Editorial Board.

¹Email: psmits@uchicago.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1510482112/-DCSupplemental.

taxon age dependent, where values of α greater than 1 corresponds to increasing risk with age and values less than 1 corresponds to decreasing risk with age. Origination cohort and phylogenetic position were modeled as independent effects. Phylogenetic effect was modeled assuming species duration may have evolved via a Brownian motion-like process (24, 25). The results from the Weibull model are detailed here because this model has a better fit to the data the exponential [Weibull widely applicable information criterion (WAIC), 6,140.37; exponential WAIC, 16,697.35; Fig. 1 and *SI Appendix*, Figs. S1 and S2].

Results

A summary of the posterior distributions for the most relevant parameter estimates is presented in Table 1. All posterior inference is based on these estimates. For the results from the posterior predictive checks and discussion of the estimation of α , please see *SI Appendix*. Additionally, see *SI Appendix* for discussion surrounding use of Paleobiology Database and accompanying data quality concerns.

Species with greater bioprovince occupancy are found to be associated with lower extinction risk than taxa with smaller bioprovince occupancy ($\beta_{\text{occupancy}}$ mean = -0.53 , SD = 0.08). This is consistent with previous findings. Body size has nearly zero association with expected duration (β_{size} mean = -0.05 , SD = 0.05), a similar result to some previous studies (9). However, previous studies were performed at the generic level and were unable to determine how body size may effect species-level extinction, as the effect of either extinction or speciation cannot be distinguished (7, 9).

Some clear patterns emerge from the pairwise differences in effect of each dietary category on expected duration (Fig. 2). Consistent with expectations from the “survival of the unspecialized” hypothesis (1, 17), omnivory appears to be associated with the lowest expected extinction risk. Carnivory is associated a greater expected duration than either herbivory or insectivory, but a greater expected extinction risk than omnivory.

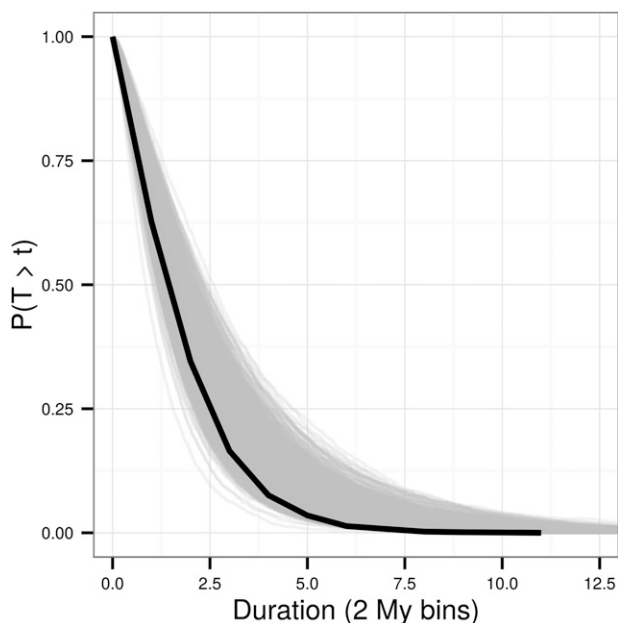


Fig. 1. Weibull-based model estimates (gray) from 1,000 posterior predictive datasets of the empirical survival function (black). The survival function is the probability that a species with duration t will not have gone extinct. Simulated datasets were generated by drawing parameter values randomly from their estimated posteriors and using the observed covariate information to estimate durations for all of the observed species.

Finally, herbivory and insectivory have approximately equal effects on expected duration. Given previous results, these results imply that carnivores have a greater origination rate than omnivores (18). These results also imply that herbivores, which have the greatest extinction risk, must also have a very high origination rate to have the greatest diversification rate among these three categories (18).

For locomotor category, both scansoriality and ground-dwelling life habitat are associated with a greater expected duration than arboreality (Fig. 2). Scansorial and ground-dwelling life habits also have approximately equal expected effects on extinction risk. This result is consistent with the expectation that arboreality will confer greater extinction risk due to the loss of associated environment with the shift from open to closed habitat at the Paleogene/Neogene boundary (21). However, there are two possible processes that could lead to the observed pattern: arboreality confers an intrinsic difference in extinction risk or it might not be that arboreal taxa have an intrinsically higher risk but were instead “hit harder” by the environmental shift than other taxa. This analysis cannot distinguish between these two processes. Note that, although this is a study of North American Cenozoic mammals, for European Cenozoic mammals this transitional period corresponds to the Vallesian, which was a sudden shift in species demography away from arboreality (26, 27).

Of the three sources of variance present in the model, individual species variance accounts for $\sim 80\%$ of the observed unmodeled variance (see Fig. 4). Note that the individual variance was approximated using a simulation approach (28) because the Weibull distribution does not have a variance term. Both cohort and phylogenetic effects account for the other 20% of the observed variance. This result means that extinction risk has both temporal and phylogenetic aspects, as both contribute greater than 0% of the observed variability in the data (25).

The estimates for the individual cohort effects show a weak pattern of greater extinction risk in older Cenozoic cohorts compared with younger cohorts (Fig. 3). This potential slowdown in extinction risk is consistent with previous analyses of marine invertebrates (29, 30) and mammals (31, 32). There are two prevailing hypotheses as to the cause of this slowdown: (i) extinction risk is constant within, but varies between, clades so over time clades with low extinction rates increases in proportion of total diversity, thus bringing down expected extinction risk; or (ii) over time taxa increase in mean fitness and thus decrease in expected extinction risk (29). The observed decrease in extinction risk with age, along with the variance partitioning results (Fig. 4), are consistent with both of these hypotheses with neither being more important than the other.

Interestingly, the shift from older cohorts with a higher extinction risk to younger cohorts with lower extinction risk is approximately at the Paleogene–Neogene boundary. Given the association with arboreality and increased extinction risk (Fig. 2), the decrease in expected extinction risk over time might relate to the preferential loss of arboreal taxa over the Cenozoic. However, because the model used here does not allow for time-varying effects, I cannot identify whether this boundary is associated with a shift in the direction or magnitude of the expected effect of arboreality on extinction risk.

Discussion

My results indicate that Cenozoic North American mammal generalists are expected to have a lower extinction risk than specialists, which implies that the diversification of specialized taxa would have required either a driven trend away from generality (33) or an increase in speciation rate relative to extinction rate (34). For specialist traits to increase in frequency, they would have to be associated with factors that may increase in speciation rate, such as increased niche partitioning among descendant taxa or increasing habitat heterogeneity leading to an increase in both

Table 1. Marginal posterior estimates for the parameters of interest

Parameter	Effect	Mean	SD	2.5%	25%	50%	75%	97.5%	\hat{R}
α	"Age"	1.29	0.03	1.23	1.27	1.29	1.31	1.36	1.00
β_0	Arboreal/carnivore	-0.78	0.14	-1.05	-0.87	-0.78	-0.68	-0.51	1.00
β_{occ}	Occupancy	-0.53	0.08	-0.69	-0.59	-0.53	-0.48	-0.38	1.00
β_{size}	Body size	-0.05	0.05	-0.14	-0.08	-0.05	-0.01	0.05	1.00
β_g	Ground dwelling	-0.28	0.10	-0.47	-0.34	-0.28	-0.21	-0.09	1.00
β_s	Scansorial	-0.22	0.11	-0.43	-0.29	-0.22	-0.14	-0.00	1.00
β_h	Herbivore	0.09	0.09	-0.09	0.03	0.09	0.14	0.27	1.00
β_i	Insectivore	0.10	0.11	-0.11	0.03	0.10	0.17	0.31	1.00
β_o	Omnivore	-0.12	0.11	-0.33	-0.19	-0.12	-0.05	0.09	1.00
σ_c	SD cohort	0.33	0.06	0.23	0.29	0.33	0.37	0.48	1.00
σ_p	SD phylogeny	0.11	0.05	0.03	0.07	0.10	0.14	0.23	1.03

The intercept β_0 can also be interpreted as the estimate for the mean observed species. The remaining β values can be interpreted as the effect of a trait on the expected species duration as expressed as deviation from the mean. The categorical variables are binary index variables where an observation is of that category or not. See Eq. 6 for an explanation of the effect of α on extinction risk. \hat{R} values of less than 1.1 indicate approximate chain convergence for the posterior samples. Values are based on 1,000 posterior samples.

speciation and extinction. For example, descendant species of omnivores may divide available prey items more finely or arboreal taxa may increase in both extinction and speciation rates via increases in habitat heterogeneity. Possible evidence to support this hypothesis would be to demonstrate differences in speciation rate associated with those traits analyzed here or other similar traits.

When these results are compared with factors contributing to increased extinction risk in extant mammals, there are some incongruities. As expected, large range size is consistently associated with lower extinction risk in the modern world (16, 20, 35, 36). Although my analysis found body size to have almost no time-invariant effect on extinction risk, in extant mammals this is not necessarily the case as increased body size is associated with increased extinction risk (16, 20). However, this pattern is partially clade dependent (35). As stated earlier, higher trophic levels have been found to be associated with greater extinction risk in extant mammals (16, 20). In contrast, I found that omnivores and carnivores have a lower expected extinction risk than either insectivores or herbivores (Fig. 2). Finally, phylogeny has been found to be a good predictor of differences in extinction risk in extant mammals as certain clades are at much higher risks than others (36). This effect seems much greater in the Recent than for the whole Cenozoic, implying that current extinction risk is more phylogenetically concentrated than during times of background extinction levels during the Cenozoic.

Whether these incongruities are within the standard range of time-variant effects is unknown, although my comparisons do imply that current processes are different from those studied here. However, this is not a model of what makes taxa vulnerable during mass extinctions and that may account for these incongruities, assuming mass extinctions are qualitatively different from background extinction (12). These results would also be inapplicable if the current biodiversity crisis is qualitatively different from either background or mass extinction as preserved in the fossil record.

By modeling how different ecologies and historical factors affect a species' expected extinction risk, it is possible to better understand what processes may have driven the resulting pattern of selection (i.e., diversity) while also providing a baseline for evaluating the current biodiversity crisis. This analysis finds support for the survival of the unspecialized hypothesis (1, 17) as a time-invariant generalization about extinction risk. I also find that there are substantial effects of both cohort and phylogeny on extinction risk, which supports the idea that the decrease in extinction risk (29) over time has both temporal and phylogenetic components. Additionally, I found evidence of increasing ex-

inction risk with species age, the cause of which is unknown. These results show that, like prior mass extinction events in the fossil record, the current biodiversity crisis is qualitatively different from the previous period of background extinction in the fossil record (12).

Materials and Methods

Species Occurrence and Covariate Information. Fossil occurrence information was downloaded from the Paleobiology Database (PBDB; <https://paleobiodb.org/#/>). Occurrence, taxonomic, stratigraphic, and biological information was downloaded for all North American mammals. This dataset was filtered so that only occurrences were 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 from the sample.

Species dietary and locomotor category assignments were done using the assignments in the PBDB, which were reassigned into coarser categories (*SI Appendix, Table S1*) in order to improve interpretability, increase sample size per category, and make results comparable to previous studies (13, 18).

All individual fossil occurrences were assigned to 2-My bins ranging through the entire Cenozoic. Taxon duration was measured as the number of 2-My 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 (6, 10, 32). Species originating in the youngest cohort, 0–2 My, were excluded from analysis because every species duration would be both left and right censored, which is illogical.

Species body size estimates in grams were sourced from a large selection of primary literature and database compilations. Databases used include the PBDB, PanTHERIA (37), and the Neogene Old World Mammal database (NOW; www.helsinki.fi/science/now/). Major sources of additional compiled body size estimates include refs. 9 and 38–42, which were then supplemented with an additional literature search to try and fill in the remaining gaps. In many cases, species body mass was estimated using various published regression equations based on tooth or skull measurements (*SI Appendix, Table S2*). If multiple specimens were measured, I used the mean of specimen measures as the species mean. See *Dataset S1* for a complete list of mass estimates and sources.

Biogeographic Network. Species geographic extent was measured as the mean of the relative number of bioprovinces occupied by a species for each 2-My bin in which the species was present. Bioprovinces were identified using a network-theoretic approach that has previously been applied to paleontological data (43, 44). This approach relies on defining a biogeographic bipartite network of taxa and localities. In this study, taxa were defined as species and localities were grid cells from a regular lattice on a global equal-area cylinder map projection. The regular lattice was defined as a 70 × 34 global grid, where each cell corresponds to ~250,000 km². An advantage of this approach is that this approach reduces to occupancy when all localities are independent and to a single bioprovince when all localities are identical.

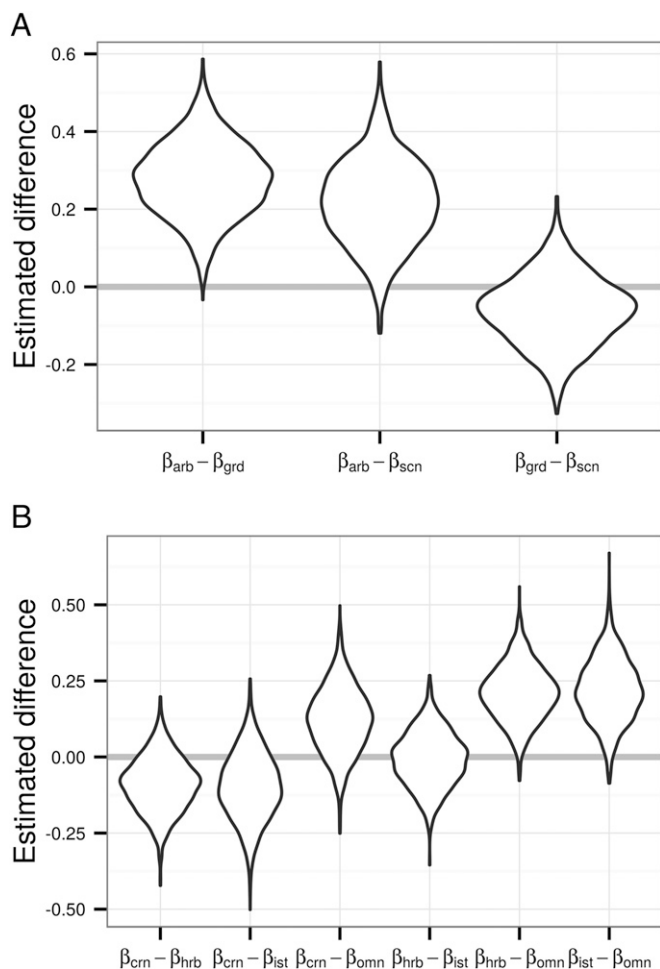


Fig. 2. Pairwise differences in effect of the locomotor (A) and dietary categories (B) on expected duration from 1,000 samples from the posterior distribution. Comparisons of locomotor categories, from top to bottom (A) are as follows: arboreal ($\beta_{arb} = \beta_0$) vs. ground dwelling ($\beta_{grd} = \beta_0 + \beta_g$), arboreal vs. scansorial ($\beta_{scn} = \beta_0 + \beta_s$), and ground dwelling vs. scansorial. For dietary category, from top to bottom (B): carnivore ($\beta_{crn} = \beta_0$) vs. herbivore ($\beta_{hrb} = \beta_0 + \beta_h$), carnivore vs. insectivore ($\beta_{ist} = \beta_0 + \beta_i$), carnivore vs. omnivore ($\beta_{omn} = \beta_0 + \beta_o$), herbivore vs. insectivore, herbivore vs. omnivore, and insectivore vs. omnivore. Negative values indicate that the first category is expected to have a greater duration than the second, whereas positive values indicate that the first category is expected to have a shorter duration.

A biogeographic network was constructed for each of the 2-My bins used in this study. Emergent bioprovinces were then identified using the map Eq (45, 46), as has been done before (43, 44, 47). These bioprovinces correspond to taxa and localities that are more interconnected with each other than with other nodes.

The map projection and regular lattice were made using shape files from www.naturalearthdata.com/ and the raster package for R (48). Bioprovince identification was done using the map equation as implemented in the igraph package for R (49).

Supertree. As there is no single, combined formal phylogenetic hypothesis of all Cenozoic fossils mammals from North America, it was necessary to construct a semiformal supertree. This tree was formed by combining taxonomic information for all of the observed species and a few published phylogenies using matrix representation parsimony (50). For further explanation of the methodology used to construct this supertree, please see *SI Appendix*.

Survival Model. Presented here is the model development process used to formulate the two survival models used in this study. First, define y as a vector of length n where the i th element is the duration of species i , where $i = 1, \dots, n$.

The simplest survival model where durations are assumed to follow an exponential distribution with a single rate or inverse-scale parameter λ (51). This is written as

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

The exponential distribution corresponds to situations where extinction risk is independent of age. To understand this, we need to define two functions: the survival function $S(t)$ and the hazard function $h(t)$. $S(t)$ is the probability that a species having existed for t 2-My bins will not have gone extinct, whereas $h(t)$ corresponds to the instantaneous extinction rate for some taxon age t (51). For an exponential model, $S(t)$ is

$$S(t) = \exp(-\lambda t), \quad [2]$$

and $h(t)$ is defined

$$h(t) = \lambda. \quad [3]$$

The choice of the exponential distribution corresponds directly to the Law of Constant Extinction (2), because the right side of Eq. 3 does not depend on species age t .

The current sampling statement (Eq. 1) 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 $\lambda_i = \exp(\sum \beta^T X_i)$, where i is indexing a given observation and its covariates, β is a vector of regression coefficients, and X is a matrix of covariates. Using a standard regression approach, one column of X is all 1 – s and its corresponding β coefficient is the intercept.

X is an $n \times K$ matrix of species-level covariates. Three of the covariates of interest are the logit of mean relative occupancy, and the logarithm of body size (grams). 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' category, with k being the number of categories of the index variable (3 and 4, respectively). Only $k - 1$ indicator variables are necessary as the intercept takes on the remaining value. For example, the difference in effect of arboreality vs. scansoriality on extinction risk, given that arboreality is the reference category, is the coefficient for the scansorial indicator variable as that is the difference between between the effect of arboreality (the intercept β_0) and scansoriality (the intercept + scansorial effect β_s) (Fig. 2). Finally, a vector of 1 – s was included in the matrix X whose corresponding β coefficient is the intercept, making K equal 8.

β is the vector of regression coefficients. The intercept term was given a weak normal prior, $\beta_0 \sim \mathcal{N}(0, 10)$, whereas all of these other coefficients were slightly more informative priors, e.g., $\beta_{mass} \sim \mathcal{N}(0, 5)$. 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 (52).

Regression coefficients are not directly comparable without first standardizing the input variables to have equal SDs. Standardization was accomplished by subtracting the mean of the covariate from all values and

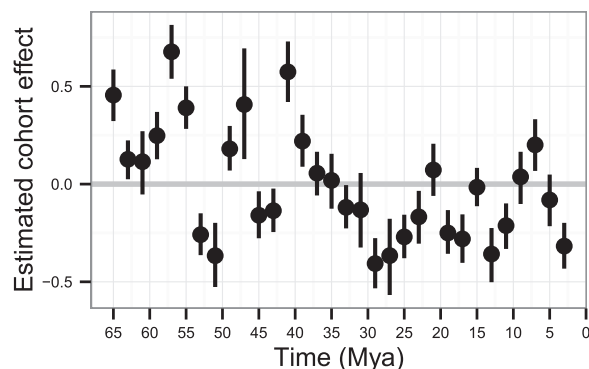


Fig. 3. Summaries of posterior estimates of individual cohort effect depicted as medians and 80% credible intervals. High values correspond to shorter species durations, whereas lower values correspond to greater species durations compared with the mean duration. Lines are placed at the middle of the 2-My origination cohorts.

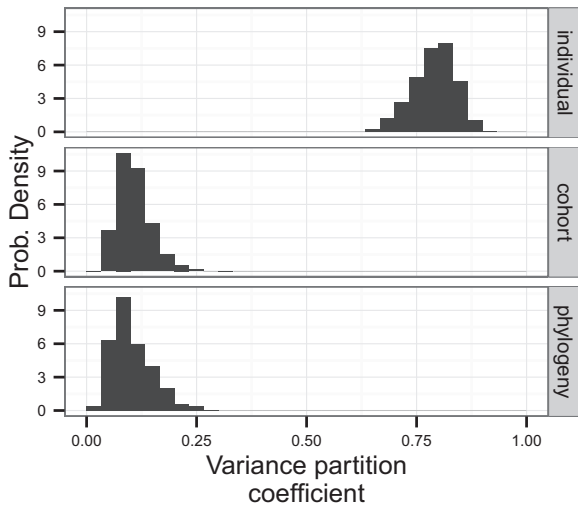


Fig. 4. Estimates of the variance partitioning coefficients for the three different sources of variance: species, cohort, and phylogeny. Higher values correspond to greater contribution to total observed variance. Each of the estimates is a distribution of 1,000 approximating simulations due to the model's nonnormally distributed errors.

then dividing by the SD, resulting in a variable with mean of 0 and an SD of 1. This linear transform greatly improves the interpretability of the coefficients as expected change in mean duration given a difference of 1 SD in the covariate (53). Additionally, this makes the intercept directly interpretable as the estimate of mean (transformed) σ (Eq. 7). However, because the expected SD for a random binary variable is 0.5, to make comparisons between the binary and continuous variables, the continuous inputs were divided by twice their SD (54).

The origination cohort is defined as the group of species that all originated during the same 2-My temporal bin. Because the most recent temporal bin, 0–2 My, was excluded, there are 32 total cohorts. The effect of origination cohort j was modeled with each group being a sample from a common cohort effect, η_j , which was considered normally distributed with mean of 0 and SD of 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 (55–59), with σ_c acting as a measure of compromise between these two end members. The choice of the half-Cauchy prior for σ_c follows (60)

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

$$\sigma_c \sim C^+(0, 2.5).$$

The impact of shared evolutionary history, or phylogeny, was modeling as an individual effect where each observation, i , is modeled as a multivariate normal, h , where the covariance matrix Σ is known up to a constant, σ_p^2 (24, 25). This is written

$$h \sim \text{MVN}(0, \Sigma)$$

$$\Sigma = \sigma_p^2 \mathbf{V}_{phy}$$

$$\sigma_p \sim C^+(0, 2.5).$$

\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. Note that because the phylogeny used here is primarily based on taxonomy, estimates of σ_p represent minimum estimates (24, 25). Improved phylogenetic estimates of all fossil Cenozoic mammals would greatly improve this estimate.

To relax the assumption of age-independent extinction of the Law of Constant Extinction, the Weibull distribution is substituted for the exponential (51). The Weibull distribution has a shape parameter α and scale parameter σ . Conceptually, σ is the inverse of λ . α modifies the impact of taxon age on extinction risk. When $\alpha > 1$, then $h(t)$ is a monotonically increasing function, but

when $\alpha < 1$, then $h(t)$ is a monotonically decreasing function. When $\alpha = 1$, then the Weibull distribution is equivalent to the exponential.

The Weibull distribution and sampling statement were defined as

$$p(y|\alpha, \sigma) = \frac{\alpha}{\sigma} \left(\frac{y}{\sigma}\right)^{\alpha-1} \exp\left[-\left(\frac{y}{\sigma}\right)^\alpha\right]$$

$$y \sim \text{Weibull}(\alpha, \sigma). \quad [4]$$

The corresponding $S(t)$ and $h(t)$ functions are defined as

$$S(t) = \exp\left[-\left(\frac{t}{\sigma}\right)^\alpha\right], \quad [5]$$

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

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

$$\sigma_i = \exp\left(\frac{-\beta}{\alpha}\right). \quad [7]$$

Given the above, the survival model was then fit in a Bayesian context using both exponential and Weibull distributions. The Weibull's α parameter was assumed constant across species, which is standard practice in survival analysis (51). α was given a weakly informative half-Cauchy (C^+) prior. σ was reparameterized as an exponentiated regression model (Eq. 7). Eq. 8 was further expanded to allow for two hierarchical factors as discussed above. This is written

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

where an equivalent statement for the exponential distribution is defined as

$$\lambda_i = \exp\left(h_i + \eta_{j|i} + \sum \beta^T \mathbf{X}_i\right). \quad [9]$$

An important part of survival analysis is the inclusion of censored observations where the failure time has not been observed (51, 61). 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 present in the most recent time bin (0–2 My). Left censored observations, on the other hand, correspond to observations that went extinct any time between 0 and some known point. To account for this uncertainty, the probability of a left censored observation is found by integrating over all possible durations between the 0 and 1 time bins. For an explanation of how censored observations are included in the sampling statement, please see *SI Appendix*.

Estimation. Parameter posteriors were approximated using a Markov-chain Monte Carlo (MCMC) routine implemented in the Stan programming language (62). Stan implements a version of Hamiltonian Monte Carlo called the No-U-Turn sampler (63). Posterior approximation was done using four parallel MCMC chains run for 30,000 steps, thinned to every 30th sample, split evenly between warmup and sampling. 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 (23).

Posterior Evaluation. The most basic assessment of model fit is that simulated data generated given the 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 1,000 times, and the distribution of y^{rep} was compared with the observed (23). For results from the posterior predictive tests used in this study, please see *SI Appendix*.

The exponential and Weibull models were compared for out-of-sample predictive accuracy using the WAIC (64). Because the Weibull model reduces to the exponential model when $\alpha = 0$, our interest is not in choosing between these models. Instead comparison of WAIC values is useful for better understanding the effect of model complexity on out-of-sample predictive

accuracy. An explanation of how WAIC is calculated is presented in *SI Appendix* following the recommended WAIC 2 formulation (23).

There are three different variance components in this model: sample component, cohort σ_c^2 , and phylogenetic σ_p^2 . Partitioning the variance between these sources allows the relative amount of unexplained variance of the sample to be compared. The sample component is similar to the residual variance from a normal linear regression. However, the Weibull-based model used here (Eq. 4) does not include an estimate of the variance similar to the squared scale term of a normal distribution. Instead, the sample component was approximated via a simulation approach modified from ref. 28. For explanation of this method, please see *SI Appendix*.

I used variance partitioning coefficients (VPCs) to estimate the relative importance of the different variance components (52). Phylogenetic heritability, h_p^2

(24, 25), is identical to the VPC of the phylogenetic effect. Phylogenetic heritability is a measure of how shared evolutionary history impacts differences in individual species trait values (e.g., duration), a broad sense definition of “heritability” as it combines both genetic inheritance and other, nongenetic shared history factors. Importantly, because phylogenetic effect was estimated using a principally taxonomy-based tree, the estimates derived here can be considered minimum estimates of the phylogenetic effect.

ACKNOWLEDGMENTS. I thank M. Foote, K. Angielczyk, R. Ree, and P. D. Polly for discussion; N. Pierrehumbert, E. Sander, and L. Southcott for draft comments; and J. Alroy and the Fossilworks/Paleobiology Database for data accumulation, management, and availability. This is Paleobiology Database publication number 239.

1. Simpson GG (1944) *Tempo and Mode in Evolution* (Columbia Univ Press, New York).
2. Van Valen L (1973) A new evolutionary law. *Evol Theory* 1(1):1–30.
3. Raup DM (1994) The role of extinction in evolution. *Proc Natl Acad Sci USA* 91(15):6758–6763.
4. Quental TB, Marshall CR (2013) How the red queen drives terrestrial mammals to extinction. *Science* 341(46):290–292.
5. Wagner PJ, Estabrook GF (2014) Trait-based diversification shifts reflect differential extinction among fossil taxa. *Proc Natl Acad Sci USA* 111(46):16419–16424.
6. Alroy J (2009) Speciation and extinction in the fossil record of North American mammals. *Speciation and Patterns of Diversity*, eds Butlin RK, Bridle JR, Schluter D (Cambridge Univ Press, Cambridge, UK), pp 302–323.
7. Liow LH, et al. (2008) Higher origination and extinction rates in larger mammals. *Proc Natl Acad Sci USA* 105(16):6097–6102.
8. Smith FA, Brown J, Haskell J, Lyons S (2004) Similarity of mammalian body size across the taxonomic hierarchy and across space and time. *Am Nat* 163(5):672–691.
9. Tomiya S (2013) Body size and extinction risk in terrestrial mammals above the species level. *Am Nat* 182(6):196–214.
10. Marcot JD (2014) 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.
11. Wang SC (2003) On the continuity of background and mass extinction. *Paleobiology* 29(4):455–467.
12. Jablonski D (1986) Background and mass extinctions: The alternation of macroevolutionary regimes. *Science* 231(4734):129–133.
13. Jernvall J, Fortelius M (2004) Maintenance of trophic structure in fossil mammal communities: site occupancy and taxon resilience. *Am Nat* 164(5):614–624.
14. Jablonski D (2008) Species selection: Theory and data. *Annu Rev Ecol Syst* 39:501–524.
15. Roy K, Hunt G, Jablonski D, Krug AZ, Valentine JW (2009) A macroevolutionary perspective on species range limits. *Proc Biol Sci* 276(1661):1485–1493.
16. Liow LH, Fortelius M, Lintulaakko K, Mannila H, Stenseth NC (2009) Lower extinction risk in sleep-or-hide mammals. *Am Nat* 173(2):264–272.
17. Liow LH (2004) A test of Simpson’s “rule of the survival of the relatively un-specialized” using fossil crinoids. *Am Nat* 164(4):431–443.
18. Price SA, Hopkins SSB, Smith KK, Roth VL (2012) Tempo of trophic evolution and its impact on mammalian diversification. *Proc Natl Acad Sci USA* 109(18):7008–7012.
19. Rabosky DL (2010) Extinction rates should not be estimated from molecular phylogenies. *Evolution* 64(6):1816–1824.
20. Purvis A, Gittleman JL, Cowlshaw G, Mace GM (2000) Predicting extinction risk in declining species. *Proc Biol Sci* 267(1456):1947–1952.
21. Blois JL, Hadly EA (2009) Mammalian response to Cenozoic climatic change. *Annu Rev Earth Planet Sci* 37:181–208.
22. Janis CM (1993) Tertiary mammal evolution in the context of changing climates, vegetation, and tectonic events. *Annu Rev Ecol Syst* 24:467–500.
23. Gelman A, et al. (2013) *Bayesian Data Analysis* (Chapman and Hall, Boca Raton, FL), 3rd Ed, p 675.
24. Lynch M (1991) Methods for the analysis of comparative data in evolutionary biology. *Evolution* 45(5):1065–1080.
25. Housworth EA, Martins P, Lynch M (2004) The phylogenetic mixed model. *Am Nat* 163(1):84–96.
26. Agustí J, Cabrera L, Garcés M (2013) The Vallesian mammal turnover: A Late Miocene record of decoupled land-ocean evolution. *Geobios* 46(1–2):151–157.
27. Moyà-Solà ICVS, Agustí J, Köhler M (2005) The geography of a faunal turnover: Tracking the Vallesian Crisis. *Migration of Organisms: Climate, Geography, Ecology*, ed Elewa AMT (Springer, Berlin), pp 247–300.
28. Goldstein H, Browne W, Rasbash J (2002) Partitioning variation in multilevel models. *Understat* 1(4):1–12.
29. Raup DM, Sepkoski JJ (1982) Mass extinctions in the marine fossil record. *Science* 215(4539):1501–1503.
30. Foote M (2003) Origination and extinction through the Phanerozoic: A new approach. *J Geol* 111(2):125–148.
31. Alroy J (2010) Fair sampling of taxonomic richness and unbiased estimation of origination and extinction rates. *Quantitative Methods in Paleobiology*, eds Alroy J, Hunt G (Yale University, New Haven, CT), pp 55–80.
32. Alroy J, Koch PL, Zachos JC (2000) Global climate change and North American mammalian evolution. *Paleobiology* 26(4):259–288.
33. McShea DW (1994) Mechanisms of large-scale evolutionary trends. *Evolution* 48(6):1747–1763.
34. Stanley SM (1975) A theory of evolution above the species level. *Proc Natl Acad Sci USA* 72(2):646–650.
35. Fritz SA, Bininda-Emonds ORP, Purvis A (2009) Geographical variation in predictors of mammalian extinction risk: Big is bad, but only in the tropics. *Ecol Lett* 12(6):538–549.
36. Fritz SA, Purvis A (2010) Selectivity in mammalian extinction risk and threat types: a new measure of phylogenetic signal strength in binary traits. *Conserv Biol* 24(4):1042–1051.
37. Jones KE, et al. (2009) PanTHERIA: A species-level database of life history, ecology, and geography of extant and recently extinct mammals. *Ecology* 90(9):2648.
38. Brook BW, Bowman DMJS (2004) The uncertain blitzkrieg of Pleistocene megafauna. *J Biogeogr* 31(4):517–523.
39. Freudenthal M, Martín-Suárez E (2013) Estimating body mass of fossil rodents. *Scripta Geol* 145(1):1–130.
40. McKenna RT (2011) Potential for speciation in mammals following vast, late Miocene volcanic interruptions in the Pacific Northwest. Masters thesis (Portland State University, Portland, OR).
41. Raia P, Carotenuto F, Passaro F, Fulgione D, Fortelius M (2012) Ecological specialization in fossil mammals explains Cope’s rule. *Am Nat* 179(3):328–337.
42. Smith FA, et al. (2004) Similarity of mammalian body size across the taxonomic hierarchy and across space and time. *Am Nat* 163(5):672–691.
43. Sidor CA, et al. (2013) Provincialization of terrestrial faunas following the end-Permian mass extinction. *Proc Natl Acad Sci USA* 110(20):8129–8133.
44. Vilhena DA, et al. (2013) Bivalve network reveals latitudinal selectivity gradient at the end-Cretaceous mass extinction. *Sci Rep* 3:1790.
45. Rosvall M, Bergstrom CT (2008) Maps of random walks on complex networks reveal community structure. *Proc Natl Acad Sci USA* 105(4):1118–1123.
46. Rosvall M, Axelsson D, Bergstrom C (2009) The map equation. *Eur Phys J Spec Top* 178(1):13–24.
47. Vilhena DA (2013) Boundaries and dynamics of biomes. PhD thesis (University of Washington, Seattle).
48. Hijmans RJ (2015) Raster: Geographic Data Analysis and Modeling R Package, Version 2.3-24. Available at <https://CRAN.R-project.org/package=raster>.
49. Csardi G, Nepusz T (2006) The igraph software package for complex network research. *InterJournal*, Complex Systems, 1695.
50. Bininda-Emonds ORP, et al. (2007) The delayed rise of present-day mammals. *Nature* 446(7135):507–512.
51. Klein JP, Moeschberger ML (2003) *Survival Analysis: Techniques for Censored and Truncated Data* (Springer, New York), 2nd Ed.
52. Gelman A, Hill J (2007) *Data Analysis Using Regression and Multilevel/Hierarchical Models* (Cambridge Univ Press, New York), p 625.
53. Schielzeth H (2010) Simple means to improve the interpretability of regression coefficients. *Methods Ecol Evol* 1(2):103–113.
54. Gelman A (2008) Scaling regression inputs by dividing by two standard deviations. *Stat Med* 27(15):2865–2873.
55. Foote M (1988) Survivorship analysis of Cambrian and Ordovician trilobites. *Paleobiology* 14(3):258–271.
56. Raup DM (1978) Cohort analysis of generic survivorship. *Paleobiology* 4(1):1–15.
57. Raup DM (1975) Taxonomic survivorship curves and Van Valen’s Law. *Paleobiology* 1(1):82–96.
58. Van Valen L (1979) Taxonomic survivorship curves. *Evol Theory* 4(3):129–142.
59. Baumiller TK (1993) Survivorship analysis of Paleozoic Crinoida: Effect of filter morphology on evolutionary rates. *Paleobiology* 19(3):304–321.
60. Gelman A (2006) Prior distributions for variance parameters in hierarchical models. *Bayesian Anal* 1(3):515–533.
61. Ibrahim JG, Chen MH, Sinha D (2001) *Bayesian Survival Analysis* (Springer, New York), p 480.
62. Stan Development Team (2014) Stan: A C++ Library for Probability and Sampling, Version 2.5.0. Available at mc-stan.org.
63. Hoffman MD, Gelman A (2011) The no-U-turn sampler: Adaptively setting path lengths in Hamiltonian Monte Carlo. arXiv:1111.
64. Watanabe S (2010) Asymptotic equivalence of Bayes cross validation and widely applicable information criterion in singular learning theory. *J Mach Learn Res* 11:3571–3594.