

Species occurrence as a function of both emergent biological traits
and environmental context

Peter D. Smits^{1,*}

1. University of Chicago, Chicago, Illinois 60637.

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

Manuscript elements:

Keywords:

Manuscript type: Article

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

Introduction

² All the world's a stage, And all the men and women merely players; They have their exits and their entrances...

⁴ (Shakespeare, *As You Like It*, Act II, Scene VII)

How do species pools change over time as species are recruited or go extinct? When are ecotypes enriched or depleted? How does global and regional environmental context affect the distribution of species ecotypes (e.g. guilds) in a regional species pool?

8 A regional species pool is the set of species which form communities in a specific region; local
communities are subsets of the regional pool. The composition of a regional species pool changes
10 over time due to speciation, migration, extinction. Local scale processes like resource competition
only affect the regional species pool if all communities are affected.

¹² Valentine and Bambach how they presented guilds in paleobiology which is taxa united by similarity of their macroecology (Bambach, 1977; Valentine, 1969). Bush and Bambach presented
¹⁴ an ecocube to describe what marine invertebrates partition space and resources (Bambach et al., 2007; Bush and Bambach, 2011; Bush et al., 2007). Unique combinations represent what
¹⁶ possible ecotypes are observable. The distribution of ecocube occupancy is then normally analyzed as raw counts of unique combinations or using ordination methods and the change in disparity over
¹⁸ time is estimated (Bambach et al., 2007; Bush and Bambach, 2011; Bush et al., 2007).

Analysis of mammal diversity and hypotheses as to the processes that have shaped it tend to fall under a few categories: diversity of the whole system (Alroy, 1996; Alroy et al., 2000; Figueirido et al., 2012; Liow et al., 2008), guild based (Janis et al., 2004; Janis, 2008; Janis et al., 2000; Janis and Wilhelm, 1993; Jernvall and Fortelius, 2004; Pires et al., 2015), clade based (Quental and Marshall, 2013; Silvestro et al., 2015; Slater, 2015), climate based (Blois and Hadly, 2009; Janis, 1993; Janis and Wilhelm, 1993), and location based (Badgley and Finarelli, 2013; Eronen et al., 2015). Rarely are more than two of these categories considered simultaneously, and instead integration of these diverse observations and hypotheses tends to be based on coincidence. The goal

of this study is to pool information from multiple levels of organization by integrating both species
28 and climate data into a single analysis in order to provide a more holistic interpretation of the
processes which may have shaped mammal species diversity.

30 Fourth-corner modeling is an approach to explaining the patterns of either species abundance or
presence/absence as a product of species traits, environmental factors, and the interaction between
32 traits and environment (Brown et al., 2014; Jamil et al., 2013; Pollock et al., 2012; Warton et al.,
2015) CITATION. In modern ecological studies, what is being modeled is species occurrences at
34 localities distributed across a region (Jamil et al., 2013; Pollock et al., 2012). In this study, what is
being modeled is the pattern of species occurrence over time for most of the Cenozoic in North
36 America (Fig. 1). These two approaches, modern and paleontological, are different views of the same
three-dimensional pattern: species at localities over time. The temporal limitations of modern
38 ecological studies and difficulties with uneven spatial occurrences of fossils in paleontological studies
means that these approaches are complimentary but reveal different patterns of how species are
40 distributed in time and space.

One of the greatest challenges with analyzing species occurrence data is the inherent incompleteness
42 of any sample (Foote, 2001; Foote and Sepkoski, 1999; Lloyd et al., 2011; Royle and Dorazio, 2008;
Royle et al., 2014; Wang and Marshall, 2016). In the modern, only presences are certain as an
44 absence can be caused by both the species being truly absent or the species never having been
sampled (Royle and Dorazio, 2008; Royle et al., 2014). For paleontological data in the context of
46 this study, the incomplete preservation of fossil communities combined with the incomplete
sampling of what fossils there are means that the true times of origination or extinction may not be
48 observed (Foote, 2001; Foote and Sepkoski, 1999; Wang et al., 2016; Wang and Marshall, 2016).

In the analyses done here, a few key covariates which describe species' macroecology and
50 environmental context are considered. Because of the complexity of fourth-corner analyses in terms
of both number of covariates considered and structure of each model, it is possible to consider and
52 test a large number of possible hypotheses. Presented here are the species traits and related
hypotheses, followed by the environmental factors and related hypotheses.

54 The principle species trait considered in this study is a species' ecotype, defined here as the unique
combination of species dietary category and locomotor category (e.g. arboreal omnivore versus
56 unguligrade herbivore). This classification is analogous to the marine invertebrate ecocube discussed
above (Bush and Bambach, 2011; Bush et al., 2007; ?). Species mass was also included as a species
58 trait, but is mostly included in order to control for that effect on species observation and occurrence.

There are no observed massive cross-taxonomic turnover events in the North American record,
60 unlike the Neogene record Europe (Alroy, 1996, 2009; Alroy et al., 2000; Eronen et al., 2015; Janis,
1993). What does this mean for analysis of ecotype dynamics? Well, it means that we probably
62 shouldn't find one.

Translating previous work into hypotheses applicable to this analysis is difficult. Taxonomic
64 grouping is frequently invoked in many proposed hypotheses for how mammal diversity is
structured (Janis, 2008; Janis and Wilhelm, 1993; Pires et al., 2015; Quental and Marshall, 2013;
66 Slater, 2015). However, this practice is problematic because taxonomic grouping conflates shared
evolutionary history and similarities in species macroecology which means that whatever aspects of
68 species biology are important for the processes underlying species diversity are obscured.

Jernvall and Fortelius (2004) found that for the Neogene of Europe the relative abundance of
70 mammal guilds was stable over time even in the face of high turnover rates.

Many discussions of the effects or associates of species ecology and diversity have focused on
72 ungulate herbivores (Janis et al., 2004; Janis, 2008; Janis et al., 2000; Janis and Wilhelm, 1993) and
carnivores (Janis and Wilhelm, 1993; Pires et al., 2015; Silvestro et al., 2015; Slater, 2015).

74 The diversity history of ungulate herbivores is characterized by younger originating taxa having
longer legs, higher crowned teeth, and a preference for grazing over browsing than their earlier
76 originating counterparts (Janis et al., 2004; Janis, 2008; Janis et al., 2000; Janis and Wilhelm,
1993); all of which have all been attributed to environmental change or tectonic activity driving
78 climate and environmental change (Blois and Hadly, 2009; Eronen et al., 2015; Janis, 2008).

Additionally, it has been observed that ungulate cursorial forms arose before carnivore cursorial
80 forms, an observation attributed to the reorganization of plant communities towards the end of the

- Cenozoic (Janis and Wilhelm, 1993).
- 82 Within the canid guild of North America (e.g. plantigrade and digitigrade carnivores) there is
evidence that their diversity, or at least the structure of that diversity, is self-regulating in the sense
84 that species from different clades are “competing” for niche or guild space in the North American
regional species pool (Silvestro et al., 2015).
- 86 A pattern of generally constant diversity through time is also observed within the canid carnivore
subguilds of hypercarnivore, hypocarnivore, and mesocarnivores though there was no evidence of
88 diversity-dependence in trait (e.g. body size) evolution (Slater, 2015). The general pattern of
constant diversity in the face of turnover is however consistent with a possibly self-regulating
90 system.
- There is some uncertainty as to the effect of species body size on mammal diversity and aspects of
92 the diversification processes, specifically extinction (Liow et al., 2008, 2009; Smits, 2015; Tomiya,
2013). Species body size is frequently framed as an important biological descriptor because of how
94 correlated this trait is with other traits such as metabolic rate and home range size CITATIONS.
Smits (2015) found several systematic differences in mammal species durations associated with
96 various species traits. Omnivorous taxa were found to have, on average, a greater duration than
other dietary categories. Additionally, arboreal taxa were found to have a shorter duration than
98 other locomotor categories.
- An unresolved question from Smits (2015) is whether the greater extinction risk faced by arboreal is
100 constant over time or if there was a change in extinction risk at the Paleogene/Neogene boundary.
Specifically, the question is whether the extinction risk of arboreal taxa increased in the Neogene,
102 driving the loss of arboreal taxa and average extinction risk of arboreal taxa down.
- Fundamentally, all species respond differently to climate and environmental change (Blois and
104 Hadly, 2009). Macroecological patterns are the similarities across species, the emergent properties
of how species react to a similar “stimulus.”
- 106 The Cenozoic is generally characterized by a global cooling trend and the development of polar

ice-caps during the Neogene, with a few notable exceptions (Cramer et al., 2011; Zachos et al., 2008,
108 2001). The Cenozoic of North America is additionally characterized by an environmental transition
from the closed, partially forested environments of the Paleogene to the savannah and grasslands
110 environments of the Neogene (Blois and Hadly, 2009; Janis, 1993; Janis et al., 2000; Strömberg,
2005).

112 With respect to North America specifically, much of regional climatic and environmental change has
been attributed to tectonic activity and uplift (Badgley and Finarelli, 2013; Blois and Hadly, 2009;
114 Eronen et al., 2015; Janis, 2008) CITATIONS. Additionally, tectonic activity and uplift is
considered the driving causal mechanism behind both changes to diversity and trait evolution
116 (Badgley and Finarelli, 2013; Blois and Hadly, 2009) CITATIONS.

The Eocene-Oligocene transition is associated with high extinction amongst ungulate taxa (Janis,
118 2008). This period is also the transition from the Paleogene to the Neogene and from herbivores
being browsing dominated to grazing dominated CITATION.

120 There is an observed stability in estimates of global temperature from the E/O transition till the
end of the Miocene; this is called the Mid-Miocene climatic optimum (Zachos et al., 2008, 2001).
122 The Mid-Miocene climatic optimum is bookended by periods of temperature decline.

The environmental factors included in this study include estimates of global temperature and the
124 changing floral groups present in North America across the Cenozoic. These covariates were chosen
because they provide high level characterizations of the environmental context of the entire North
126 American regional species pool for most of the Cenozoic. Importantly, the effects of a species
ecotype on diversity are themselves modeled as functions of environmental factors (Fig. 1) allowing
128 for inference as to how species ecology mediates environmental context.

The effect of climate on diversity and the diversification process has been the focus of considerable
130 research with many analyses favoring diversification being more biologically-mediated than
climate-mediated (Alroy, 1996; Alroy et al., 2000; Clyde and Gingerich, 1998; Figueirido et al.,
132 2012). Scale of analysis makes a big difference in interpretation of results, both temporal and
geographic. For example when the mammal fossil record analyzed at small temporal and geographic

¹³⁴ scales a correlation between diversity and climate are observable (Clyde and Gingerich, 1998). However, when the record is analyzed at the scale of the continent and the Cenozoic there is no
¹³⁶ correlation with diversity and climate (Alroy et al., 2000). This results, however, does not go against the idea that there may be short periods of correlation and that this correlation change or
¹³⁸ reverse direction over time; instead this result means that there is no single direction of correlation between diversity and climate (Figueirido et al., 2012).

¹⁴⁰ In the case of a fluctuating correlation between diversity and climate it is hard to make the argument of an actual causal link between the two without understanding the ecological differences
¹⁴² in mammalian fauna over time; when this analysis is based on diversity or taxonomy alone no mechanisms are possible to infer. After all, taxonomy conflates many potential factors that could
¹⁴⁴ affect diversification into a single variable; by separating the effects of shared common ancestry (i.e. phylogeny) from species ecology the subtle differences in the diversification process can be observed
¹⁴⁶ (Smits, 2015).

Ultimately, the goal of this analysis are to understand when are unique ecotypes enriched or
¹⁴⁸ depleted in the North American mammal regional species pool and how changes in ecotypic diversity are related to changes in species' environmental context.

¹⁵⁰ Materials and Methods

Taxon occurrences and species-level information

¹⁵² All fossil occurrence information was downloaded from the Paleobiology Database. Occurrences (PBDB) were restricted to all Mammalia sampled in North America between the
¹⁵⁴ Maastrichtian and Gelasian stages. Taxonomic, stratigraphic, and ecological metadata for each occurrence was included. The raw data is available for download at <http://goo.gl/2s1geU>.
¹⁵⁶ This raw data was then sorted, cleaned, and manipulated programmatically prior to analysis. Species taxonomic assignments given by the PBDB were updated for accuracy and consistency. For

158 example, species classified in the order Artiodactyla were reclassified as Cetartiodactyla. These
 159 re-assessments follow Smits (2015) and were Janis et al. (2008, 1998) and the Encyclopedia of Life
 160 WEBSITE. Additionally, Taxa who's life habit was classified as either volant (i.e. Chiroptera) or
 161 aquatic (e.g. Cetacea) were excluded from this analysis because of both differences in fossilization
 162 potential and applicability to the study of terrestrial species pools.

The life habit and dietary categories provided through the PBDB where coarsened to increase per
 164 ecotype sample size; this coarsening follows the same procedure as Smits (2015). Additionally, life
 165 habit category was further modified to break-up the vague “ground-dwelling” category;
 166 re-classifying these species by ankle posture gives more precise information about that species'
 167 environmental context. Ground-dwelling taxa were reassigned following ? by species taxonomic
 168 context. Species ecotype is defined as the interaction between life habit and diet categories. Ecotype
 169 categories with less than 10 species having ever been that combination were excluded, yielding a
 170 total of 18 of 24 possible ecotypes.

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.

Table 2: Posture assignment based on taxonomy

Order	Family	Stance
	Ailuridae	plantigrade
	Allomyidae	plantigrade
	Amphicyonidae	plantigrade

Continued on next page

Table 2 – continued from previous page

Order	Family	Stance
	Amphilemuridae	plantigrade
	Anthracotheriidae	digitigrade
	Antilocapridae	unguligrade
	Apheliscidae	plantigrade
	Aplodontidae	plantigrade
	Apternodontidae	scansorial
	Arctocyonidae	unguligrade
	Barbourofelidae	digitigrade
	Barylambdidae	plantigrade
	Bovidae	unguligrade
	Camelidae	unguligrade
	Canidae	digitigrade
	Cervidae	unguligrade
	Cimolodontidae	scansorial
	Coryphodontidae	plantigrade
	Cricetidae	plantigrade
	Cylindrodontidae	plantigrade
	Cyriacotheriidae	plantigrade
	Dichobunidae	unguligrade
Dinocerata		unguligrade
	Dipodidae	digitigrade
	Elephantidae	digitigrade
	Entelodontidae	unguligrade
	Eomyidae	plantigrade
	Erethizontidae	plantigrade

Continued on next page

Table 2 – continued from previous page

Order	Family	Stance
	Erinaceidae	plantigrade
	Esthonychidae	plantigrade
	Eutypomyidae	plantigrade
	Felidae	digitigrade
	Florentiamyidae	plantigrade
	Gelocidae	unguligrade
	Geolabididae	plantigrade
	Glyptodontidae	plantigrade
	Gomphotheriidae	unguligrade
	Hapalodectidae	plantigrade
	Heteromyidae	digitigrade
	Hyaenidae	digitigrade
	Hyaenodontidae	digitigrade
	Hypertragulidae	unguligrade
	Ischyromyidae	plantigrade
	Jimomyidae	plantigrade
Lagomorpha		digitigrade
	Leptictidae	plantigrade
	Leptochoeridae	unguligrade
	Leptomerycidae	unguligrade
	Mammutidae	unguligrade
	Megalonychidae	plantigrade
	Megatheriidae	plantigrade
	Mephitidae	plantigrade
	Merycoidodontidae	digitigrade

Continued on next page

Table 2 – continued from previous page

Order	Family	Stance
Mesonychia		unguligrade
	Mesonychidae	digitigrade
	Micropternodontidae	plantigrade
	Mixodectidae	plantigrade
	Moschidae	unguligrade
	Muridae	plantigrade
	Mustelidae	plantigrade
	Mylagaulidae	fossorial
	Mylodontidae	plantigrade
	Nimravidae	digitigrade
	Nothrotheriidae	plantigrade
Notoungulata		unguligrade
	Oromerycidae	unguligrade
	Oxyaenidae	digitigrade
	Palaeomerycidae	unguligrade
	Palaeoryctidae	plantigrade
	Pampatheriidae	plantigrade
	Pantolambdidae	plantigrade
	Periptychidae	digitigrade
Perissodactyla		unguligrade
	Phenacodontidae	unguligrade
Primates		plantigrade
	Procyonidae	plantigrade
	Proscalopidae	plantigrade
	Protoceratidae	unguligrade

Continued on next page

Table 2 – continued from previous page

Order	Family	Stance
	Reithroparamyidae	plantigrade
	Sciuravidae	plantigrade
	Sciuridae	plantigrade
	Simimyidae	plantigrade
	Soricidae	plantigrade
	Suidae	digitigrade
	Talpidae	fossorial
	Tayassuidae	unguligrade
	Tenrecidae	plantigrade
	Titanoideidae	plantigrade
	Ursidae	plantigrade
	Viverravidae	plantigrade
	Zapodidae	plantigrade

¹⁷² Species mass information was gathered from multiple different sources where a plurality of the body size estimates are from the PBDB. Body part measurements for many species are also available
¹⁷⁴ through the PBDB. Just as with Smits (2015), these measurements and corresponding regression equations were used to get mass estimates for more species. Additional mass estimates and body
¹⁷⁶ part measurements were sourced from numerous publications and the Neogene Old World Database; see the supplementary material to Smits (2015) for details. Mass was log-transformed and then
¹⁷⁸ mean-centered and rescaled by dividing by two-times its standard deviation; this insures that the magnitude of effects for both continuous and discrete covariates are comparable (Gelman, 2008;
¹⁸⁰ Gelman and Hill, 2007).

All fossil occurrences from 64 to 2 million years ago (Mya) were binned into 31 2 million year (My)

Table 3: Regression equations used in this study for estimating body size. Equations are presented with reference to taxonomic grouping, part name, and reference.

Group	Equation	log(Measurement)	Source
General	$\log(m) = 1.827x + 1.81$	lower m1 area	Legendre (1986)
General	$\log(m) = 2.9677x - 5.6712$	mandible length	?
General	$\log(m) = 3.68x - 3.83$	skull length	?
Carnivores	$\log(m) = 2.97x + 1.681$	lower m1 length	?
Insectivores	$\log(m) = 1.628x + 1.726$	lower m1 area	?
Insectivores	$\log(m) = 1.714x + 0.886$	upper M1 area	?
Lagomorph	$\log(m) = 2.671x - 2.671$	lower toothrow area	Tomiya (2013)
Lagomorph	$\log(m) = 4.468x - 3.002$	lower m1 length	Tomiya (2013)
Marsupials	$\log(m) = 3.284x + 1.83$	upper M1 length	?
Marsupials	$\log(m) = 1.733x + 1.571$	upper M1 area	?
Rodentia	$\log(m) = 1.767x + 2.172$	lower m1 area	Legendre (1986)
Ungulates	$\log(m) = 1.516x + 3.757$	lower m1 area	?
Ungulates	$\log(m) = 3.076x + 2.366$	lower m2 length	?
Ungulates	$\log(m) = 1.518x + 2.792$	lower m2 area	?
Ungulates	$\log(m) = 3.113x - 1.374$	lower toothrow length	?

¹⁸² bins. This temporal length was chosen because it is approximately the resolution of the North American mammal fossil record.

¹⁸⁴ Environmental and temporal covariates

The group-level covariates in this study are descriptors of species' environmental context, specifically global temperature estimates and Graham's floral intervals CITATION. Global temperature across most of the Cenozoic was calculated from Mg/Ca isotope record from deep sea carbonates (Cramer et al., 2011). Mg/Ca based temperature estimates are preferable to the frequently used $\delta^{18}\text{O}$ temperature proxy (Alroy et al., 2000; Figueirido et al., 2012; Zachos et al., 2008, 2001) because Mg/Ca estimates do not conflate temperature with ice sheet volume and depth/stratification changes; this makes it preferable as an estimate of global temperature for macroevolutionary and macroecological studies (Ezard et al., 2016).

Two aspects of the Mg/Ca-based temperature curve were included in this analysis: mean and range. Both were calculated as the mean of all respective estimates for each 2 My temporal bins. Both mean and range were then rescaled as above: subtract mean, divide by twice the standard deviation.

¹⁹⁶ The other major set of environmental factors included in this study are Graham's Cenozoic plant
phases CITATION. Graham's plant phases are holistic descriptors of the taxonomic composition of
¹⁹⁸ 12 ecosystem types, which plants are present at a given time, and the relative modernity of those
plant groups with younger phases representing increasingly modern taxa CITATION. Graham
²⁰⁰ CITATION defines four intervals from the Cretaceous to the Pliocene, though only three of these
intervals are included in this analysis. Graham's plant phases CITATION was included as a series
²⁰² of "dummy variables" encoding the three phases included in this analysis. This means that the first
phase is synonymous with the intercept and phases

²⁰⁴ Modelling species occurrence

Two different models were used in this study: a pure-presence model and a birth-death model. Both
²⁰⁶ models at their core are hidden Markov model where the latent aspect of the process has an
absorbing state (Allen, 2011). The difference between these two models is if the probability of a
²⁰⁸ species origination and survival are considered equal or different (Table 4). Something that is
important to realize is that while there are only two state "codes" in a presence-absence matrix (i.e.
²¹⁰ 0/1), there are in fact three states in a birth-death model: never having originated, extant, and
extinct. The last of these is the absorbing state, as once a species has gone extinct it cannot
²¹² re-originate (Allen, 2011); this is made obvious in the transition matrices as the probability of an
extinct species changing states is 0 (Table 4). See below for parameter explainations (Tables 6, and
²¹⁴ 7).

Data augmentation

²¹⁶ All presence/absence observations are incomplete. The hidden Markov model at the core of this
analysis allows for observed absences to be used meaningfully to estimate the number of unobserved
²¹⁸ species. Of specific concern in this analysis is the unknown "true" size of the dataset; how many
species could have actually been observed? While many species have been observed, the natural
²²⁰ incompleteness of all observations, especially in the case of paleontological data, there are obviously

		State at $t + 1$		
		0_{never}	1	$0_{extinct}$
State at t	0_{never}	$1 - \theta$	θ	0
	1	0	θ	$1 - \theta$
	$0_{extinct}$	0	0	1

(a) Pure-presence

		State at $t + 1$		
		0_{never}	1	$0_{extinct}$
State at t	0_{never}	$1 - \phi$	ϕ	0
	1	0	π	$1 - \pi$
	$0_{extinct}$	0	0	1

(b) Birth-death

Table 4: Transition matrices for the pure-presence (4a) and birth-death (4b) models. Both of these models share the core machinery of discrete-time birth-death processes but make distinct assumptions about the equality of originating and surviving (Eq. 2, and 3). Note also that while there are only two state “codes” (0, 1), there are in fact three states: never having originated 0_{never} , present 1, extinct $0_{extinct}$ (Allen, 2011).

many species which were never sampled (Royle and Dorazio, 2008; Royle et al., 2007).

- 222 Let N by the total number of observed species, M be the upper limit of possible species that could have existed given a model of species presence, and N^* is the all-zero histories where $N^* = M - N$.
- 224 This approach assumes that $\hat{N} \sim \text{Binomial}(M, \psi)$ where \hat{N} is the estimated “true” number of species and ψ is the probability that any augmented species should actually be “present.” Because
- 226 M is user defined, this approach effectively gives ψ a uniform prior over N to M (Royle and Dorazio, 2008). For this study, $M = N + \lfloor N/4 \rfloor$.
- 228 Data imputation is the process of estimating missing data for partially observed covariates (Gelman and Hill, 2007; Rubin, 1996), this is simple in a Bayesian context because data are also parameters
- 230 (Gelman et al., 2013). Augmented species also have no known mass so a mass estimate must be imputed for each possible species (Royle and Dorazio, 2012). This procedure assumes that mass
- 232 values for augmented species are from the same distribution as observed species. The distribution of observed mass values is estimated as part of the model, and new mass values are then generated
- 234 from this distribution. This approach is an example of imputing data missing completely at random (Gelman and Hill, 2007; Royle and Dorazio, 2012). Because log mass values are rescaled as a part of
- 236 this study, the body mass distribution is already known ($\mathcal{N}(0, 0.5)$); augmented species body mass just simply drawn from this distribution.
- 238 In addition to body mass information, the augmented species need an ecotype classification. Because these species are completely unknown, they were all classified as “augmented,” an additional

Table 5: Observation parameters

Parameter	dimensions	explanation
y	$N \times T$	observed species presence/absence
z	$N \times T$	“true” species presence/absence
p	T	probability of observing a species that is present at time t
m	N	species log mass, rescaled
α_0	1	average log-odds of p
α_1	1	change in average log-odds of p per change mass
r	T	difference from α_0 associated with time t
σ	1	standard deviation of r

²⁴⁰ grouping indicating their unknown biology. This classification has no biological interpretation.

Observation process

²⁴² The type of hidden Markov model used in this study has three characteristic probabilities:
probability p of observing a species given that it is present, probability ϕ of a species surviving from
²⁴⁴ one time to another, and probability π of a species first appearing (Royle and Dorazio, 2008). In
this formulation, the probability of a species going extinct is $1 - \pi$. For the pure-presence model
²⁴⁶ $\phi = \pi$, while for the birth-death model $\phi \neq \pi$.

The probability of observing a species that is present p is modeled as a logistic regression was a
²⁴⁸ time-varying intercept and species mass as a covariate. The effect of species mass on p was assumed
linear and constant over time and given a prior reflecting a possible positive relationship; these
²⁵⁰ assumptions are reflected in the structure of the model Equation 1. The parameters associated with
this part of the model are described in Table 5.

$$\begin{aligned}
y_{i,t} &\sim \text{Bernoulli}(p_{i,t} z_{i,t}) \\
p_{i,t} &= \text{logit}^{-1}(\alpha_0 + \alpha_1 m_i + r_t) \\
r_t &\sim \mathcal{N}(0, \sigma)
\end{aligned} \tag{1}$$

Table 6: Parameters for the model of presence in the pure-presence model

Parameter	dimensions	explanation
z	$N \times T$	“true” species presence/absence
θ	$N \times T - 1$	probability of $z = 1$
a	$T - 1 \times D$	ecotype-varying intercept; mean value of log-odds of θ
m	N	species log mass, rescaled
b_1	1	effect of species mass on log-odds of θ
b_2	1	effect of species mass, squared, on log-odds of θ
U	$T \times D$	matrix of group-level covariates
γ	$U \times D$	matrix of group-level regression coefficients
Σ	$D \times D$	covariance matrix of a
Ω	$D \times D$	correlation matrix of a
τ	D	vector of standard deviations for each ecotype a_d

252 Pure-presence process

For the pure-presence model there is only a single probability dealing with the presence of a species
254 θ (Table 4a). This probability was modeled as multi-level logistic regression with both species-level
and group-level covariates (Gelman et al., 2013; Gelman and Hill, 2007). The parameters associated
256 with pure-presence model are presented in Table 6 and the full sampling statement in Equation 2.

The species-level of the model (Eq. 2) is a logistic regression with varying-intercept that varies by
258 ecotype. Additionally, species mass was included as a covariate associated with two regression
coefficients allowing a quadratic relationship with log-odds of occurrence. This assumption is based
260 on the known distribution of mammal body masses where species with intermediate mass values are
more common than either small or large bodied species. These assumptions are also reflected in the
262 choice of priors for these regression coefficients.

The values of each ecotype’s intercept are themselves modeled as regressions using the group-level
264 covariates associated with environmental context. Each of these regressions has an associated
variance of possible values of each ecotype’s intercept (Gelman and Hill, 2007). In addition, the
266 covariances between ecotype intercepts, given this group-level regression, are modeled (Gelman and
Hill, 2007).

268 All parameters not modeled elsewhere were given weakly informative priors (Gelman et al., 2013)

CITATION STAN MANUAL STATISTICAL RETHINKING. Weakly informative means that
 270 priors do not necessarily encode actual prior information but instead help regularize or weakly
 constrain posterior estimates. These priors have a concentrated probability density around and near
 272 zero; this has the effect of tempering our estimates and help prevent overfitting the model to the
 data (Gelman et al., 2013) CITATION STAN MANUAL STATISTICAL RETHINKING.

$$\begin{aligned}
 y_{i,t} &\sim \text{Bernoulli}(p_{i,t} z_{i,t}) & \alpha_0 &\sim \mathcal{N}(0, 1) \\
 p_{i,t} &= \text{logit}^{-1}(\alpha_0 + \alpha_1 m_i + r_t) & \alpha_1 &\sim \mathcal{N}(1, 1) \\
 r_t &\sim \mathcal{N}(0, \sigma) & \sigma &\sim \mathcal{N}^+(1) \\
 z_{i,1} &\sim \text{Bernoulli}(\rho) & b_1 &\sim \mathcal{N}(0, 1) \\
 z_{i,t} &\sim \text{Bernoulli}(\theta_{i,t}) & b_2 &\sim \mathcal{N}(-1, 1) \\
 \theta_{i,t} &= \text{logit}^{-1}(a_{t,j[i]} + b_1 m_i + b_2 m_i^2) & \gamma &\sim \mathcal{N}(0, 1) \\
 a &\sim \text{MVN}(u\gamma, \Sigma) & \tau &\sim \mathcal{N}^+(1) \\
 \Sigma &= \text{diag}(\tau)\Omega\text{diag}(\tau) & \Omega &\sim \text{LKJ}(2)
 \end{aligned} \tag{2}$$

274 **Birth-death process**

In the birth-death model, $\phi \neq \pi$ and so each of these probabilities are modeled separately but in a
 276 similar manner to how θ is modeled in the pure-presence model (Eq. 2, Table 4b). The parameters
 associated with the birth-death presence model are presented in Table 7 and the full sampling
 278 statement, including observation (Eq. 1), is described in Equation 3.

Similar to the pure-presence model, both ϕ and π are modeled as logistic regressions with
 280 varying-intercept and one covariate associated with two parameters. The possible relationships
 between mass and both ϕ and π are reflected in the parameterization of the model and choice of
 282 priors (Eq. 3).

The intercepts of ϕ and π both vary by species ecotype and those values are themselves the product

Table 7: Parameters for the model of presence in the pure-presence model

Parameter	dimensions	explanation
z	$N \times T$	“true” species presence/absence
ϕ	$N \times T$	probability of $z_{-,t} = 1 z_{-,t-1} = 0$
π	$N \times T - 1$	probability of $z_{-,t} = 1 z_{-,t-1} = 1$
a^ϕ	$T - 1 \times D$	ecotype-varying intercept; mean value of log-odds of θ
a^π	$T - 1 \times D$	ecotype-varying intercept; mean value of log-odds of θ
m	N	species log mass, rescaled
b_1^ϕ	1	effect of species mass on log-odds of ϕ
b_1^π	1	effect of species mass on log-odds of π
b_2^ϕ	1	effect of species mass, squared, on log-odds of ϕ
b_2^π	1	effect of species mass, squared, on log-odds of π
U	$T \times D$	matrix of group-level covariates
γ^ϕ	$U \times D$	matrix of group-level regression coefficients
γ^π	$U \times D$	matrix of group-level regression coefficients
Σ^ϕ	$D \times D$	covariance matrix of a^ϕ
Σ^π	$D \times D$	covariance matrix of a^π
Ω^ϕ	$D \times D$	correlation matrix of a^ϕ
Ω^π	$D \times D$	correlation matrix of a^π
τ^ϕ	D	vector of standard deviations for each ecotype a_d^ϕ
τ^π	D	vector of standard deviations for each ecotype a_d^π

284 of group-level regression using environmental factors as covariates (Eq. 3); this is identical to the

pure presence model (Eq. 2).

$$\begin{aligned}
y_{i,t} &\sim \text{Bernoulli}(p_{i,t} z_{i,t}) & \Sigma^\phi &= \text{diag}(\tau^\phi) \Omega^\phi \text{diag}(\tau^\phi) \\
p_{i,t} &= \text{logit}^{-1}(\alpha_0 + \alpha_1 m_i + r_t) & \Sigma^\pi &= \text{diag}(\tau^\pi) \Omega^\pi \text{diag}(\tau^\pi) \\
r_t &\sim \mathcal{N}(0, \sigma) & \rho &\sim U(0, 1) \\
\alpha_0 &\sim \mathcal{N}(0, 1) & b_1^\phi &\sim \mathcal{N}(0, 1) \\
\alpha_1 &\sim \mathcal{N}(1, 1) & b_1^\pi &\sim \mathcal{N}(0, 1) \\
\sigma &\sim \mathcal{N}^+(1) & b_2^\phi &\sim \mathcal{N}(-1, 1) \\
z_{i,1} &\sim \text{Bernoulli}(\phi_{i,1}) & b_2^\pi &\sim \mathcal{N}(-1, 1) \\
z_{i,t} &\sim \text{Bernoulli} \left(z_{i,t-1} \pi_{i,t} + \sum_{x=1}^t (1 - z_{i,x}) \phi_{i,t} \right) & \gamma^\phi &\sim \mathcal{N}(0, 1) \\
&& \gamma^\pi &\sim \mathcal{N}(0, 1) \\
\phi_{i,t} &= \text{logit}^{-1}(a_{t,j[i]}^\phi + b_1^\phi m_i + b_2^\phi m_i^2) & \tau^\phi &\sim \mathcal{N}^+(1) \\
\pi_{i,t} &= \text{logit}^{-1}(a_{t,j[i]}^\pi + b_1^\pi m_i + b_2^\pi m_i^2) & \tau^\pi &\sim \mathcal{N}^+(1) \\
a^\phi &\sim \text{MVN}(U \gamma^\phi, \Sigma^\phi) & \Omega^\phi &\sim \text{LKJ}(2) \\
a^\pi &\sim \text{MVN}(U \gamma^\pi, \Sigma^\pi) & \Omega^\pi &\sim \text{LKJ}(2)
\end{aligned} \tag{3}$$

286 Posterior inference and model adequacy

Programs that implement joint posterior inference for the above models (Eqs. 2, 3) were
288 implemented in the probabilistic programming language Stan CITATION. The models used here
both feature latent discrete parameters in the large matrix z (Tables 5, 6, 7; Eqs. 1, 2, 3). All
290 methods for posterior inference implemented in Stan are derivative based which causes
complications for actually implementing the above models because integers do not have derivatives.
292 Instead of implementing a latent discrete parameterization, the posterior probabilities of all possible
states of the latent parameters z were estimated (i.e. marginalized).
294 Species durations at minimum range-through from the FAD to the LAD, but the incompleteness of

all observations means that the actual time of origination or extinction is unknown. The
296 marginalization approach used here means that the probabilities all possible histories for a species
are calculated, from the end members of the species having existed for the entire study interval and
298 the species having only existed between the directly observed FAD and LAD to all possible
intermediaries CITATION STAN MANUAL.

300 The combined size of the dataset and large number of parameters in both models (Eqs. 2, 3),
specifically the total number of latent parameters that are the matrix z , means that stochastic
302 approximate posterior inference is computationally very slow even using HMC. Instead, an
approximate Bayesian approach was used: variational inference. A recently developed automatic
304 variational inference algorithm called “automatic differentiation variational inference” (ADVI) is
implemented in Stan and was used here CITATION. ADVI assumes that the posterior is Gaussian
306 but still yields a true Bayesian posterior; this assumption is similar to quadratic approximation of
the likelihood function used in maximum likelihood inference CITATION. The principal limitation
308 of assuming the joint posterior is Gaussian is that the true topology of the log-posterior isn’t
estimated; this is a particular burden for scale parameters which are bound to be positive (e.g.
310 standard deviation).

After fitting both models (Eqs. 2, 3) using ADVI, model adequacy and quality of fit was assessed
312 using a series of posterior predictive checks CITATION CITATION. Because all Bayesian models
are inherently generative, simulations of new data sets is “free” CITATION. By simulating many
314 theoretical data sets using the observed covariate information the congruence between predictions
made by the model and the observed empirical data can be assessed. By combining multiple
316 posterior predictive tests of congruence between empirical and simulated values of interest, the
holistic adequacy of the model can be analyzed CITATION.

318 An example posterior predictive check used in this study was comparing the observed average
number of observations per species to a distribution of simulated averages; if the empirically
320 observed value sits in the middle of the distribution than the model is adequate in reproducing the
observed number of occurrences per species.

- 322 Posterior simulations for time series are start with the values at $t = 1$ and then just simulating forward.
- 324 Given parameter estimates, diversity and diversification rates are estimated through posterior predictive simulations. Given the observed presence-absence matrix y , estimates of the true
- 326 presence-absence matrix z can be simulated and the distribution of possible occurrence histories can be analyzed. This is conceptually similar to marginalization where the probability of each
- 328 possible occurrence history is estimated (Fig. 2).

The posterior distribution of z gives the estimate of standing diversity N_t^{stand} for all time points as

$$N_t^{stand} = \sum_{i=1}^M z_{i,t}. \quad (4)$$

- 330 Given estimates of N^{stand} for all time points, the estimated number of originations O_t are be estimated as

$$O_t = \sum_{i=1}^M z_{i,t} = 1 | z_{i,t-1} = 0 \quad (5)$$

- 332 and number of extinctions E_t estimated as

$$E_t = \sum_{i=1}^M z_{i,t} = 0 | z_{i,t-1} = 1. \quad (6)$$

Per-captia growth D^{rate} , origination O^{rate} and extinction E^{rate} rates are then calculated as

$$\begin{aligned} O_t^{rate} &= \frac{O_t}{N_{t-1}^{stand}} \\ E_t^{rate} &= \frac{E_t}{N_{t-1}^{stand}} \\ D_t^{rate} &= O_t^{rate} - E_t^{rate}. \end{aligned} \quad (7)$$

³³⁴ **Results**

Posterior results take one of two forms: direct inspection of parameter estimates, and downstream
³³⁶ estimates of diversity and diversification rates. For the former, both the pure-presence and birth-death models (Eq. 2, and 3 are inspected. For the latter, only posterior estimates from the
³³⁸ birth-death model are considered; the reason for this is explained below in the comparison of the models' posterior predictive check results.

³⁴⁰ **Comparing parameter estimates from the pure-presence and birth-death models**

³⁴² Comparison of the posterior predictive performance of the pure-presence and birth-death models reveals a striking difference in quality of the models' fits to the data (Fig. 3a and 3b). The
³⁴⁴ birth-death model is clearly able to reproduce the observed average number of occurrence, in contrast to the pure-birth model which greatly underestimates the ovserved average number of
³⁴⁶ occurrences. The interpretation of these results is that the results of the birth-death model are more representative of the data than the pure-presence model, though further inspection of the
³⁴⁸ posterior parameter estimates can provide further insight into why these models give different posterior predictive results (Gelman et al., 2013). However, it is expected that downstream analyses
³⁵⁰ from the birth-death model will be more reliable than that from the pure-presence model.

Occurrence probabilities estimated from the pure-presence model (Fig. 4) are broadly similar to the
³⁵² estimates of origination probability from the birth-death model (Fig. 5) as opposed to the estimates of survival probability (Fig. 6). This result supports the idea that changes to the North American
³⁵⁴ regional species pool is more likely due to changes to origination than extinction.

For most ecotypes, both estimated occurrence probabilities from the pure-presence model (Fig. 4)
³⁵⁶ and origination probabilities estimated from the birth-death model (Fig. 5) increase with time. Notably, ecotypes with arboreal components do not appear to follow a similar pattern; instead,
³⁵⁸ occurrence and origination probabilities appear relatively flat for most of the Cenozoic.

The dramatic differences between origination and survival probabilities indicate how different these
360 processes are, and may be responsible for the better posterior predictive performance of the birth-death model over the pure-presence model (Fig. 3a, and 3b). While the estimates of both time
362 series have high variance, what is striking is how mean origination probability changes over time while most ecotype survival probabilities have relatively stable means for the entire Cenozoic (Fig.
364 5, and 6).

For most ecotypes, the estimates of origination probabilities are with less uncertainty than similar
366 estimates of survival probabilities (Fig. 5, and 6). High uncertainty in the estimates of the underlying log-odds of occurrence, origination, or survival tends to be indicative of extreme rarity
368 or complete absence of the specific ecotype; the latter is called complete separation, the effect of which has been mitigated by the hierarchical modeling strategy used here (Gelman et al., 2013;
370 Gelman and Hill, 2007) CITATION Statistical Rethinking.

The pure-presence and birth-death models differ in estimated effect of mass on the probability of
372 sampling a species that is present (Fig. 7). For the pure-presence model, mass is estimated to have not have a great effect on the probability of sampling a species that is presence (Fig. 7a).
374 Contrastingly, for the birth-death model mass is found to have a negative relationship with observation such that larger species are less likely to be observed if present than smaller species
376 (Fig. 7b).

The result from the birth-death model is unexpected given that it is generally assumed that larger
378 mammals are more likely to have been collected than smaller mammals CITATION. However, collection is not preservation; similarities in preservation rate indicate similarities in how gap-filled
380 species records are. What this result means is that the record of large bodied species is expected on average to be more gap-filled and less consistent from time point to time point than smaller bodied
382 species. Additionally, this is presence/absence data, so higher preservation and collection in terms of individual specimens at a location or a single temporal horizon does not necessarily translate to
384 high preservation over time.

The average probabilities of sampling for both the pure-presence model and birth-death model are

³⁸⁶ both at the point where (rescaled log) mass equals 0; visual comparison indicates that, on average,
sampling probability has greater posterior estimate in the pure-presence model than the birth-death
³⁸⁸ model (Fig. 7).

The effect of species mass on probability of occurrence as estimated from the pure-presence (Fig. 8)
³⁹⁰ are most similar to the estimated effect of species mass on probability of origination for the
birth-death model (Fig. 9). The striking pattern observable in both sets of estimates is the higher
³⁹² probability of occurrence for species with body sizes closer to the mean than either extremes. This
result is consistent with the canonically normal distribution of mammal body sizes CITATION; it is
³⁹⁴ then expected that the most likely to occur species would be those from the middle of the
distribution, and that species originating will on average be of average mass, especially considering
³⁹⁶ species shared common ancestry CITATION. Note that all variation between ecotypes (Fig. 9) is
due to differences in ecotype-specific survival probability and the associated effects of plant phase;
³⁹⁸ the effect of mass was considered constant for all ecotypes.

In contrast, the effect of species mass on probability of survival as estimated from the birth-death
⁴⁰⁰ model (Fig. 10) indicates little effect of mass on extinction; this is consistent with previous findings
from the North American mammal fossil record (Smits, 2015; Tomiya, 2013). Note that all variation
⁴⁰² between ecotypes (Fig. 10) is due to differences in ecotype-specific survival probability and the
associated effects of plant phase; the effect of mass was considered constant for all ecotypes.

⁴⁰⁴ Similarities in parameters estimates between ecotypes may be due to similar response to
environmental factors (Fig. 11, 12, and 13). As with previous comparisons between posterior
⁴⁰⁶ estimates from the pure-presence and birth-death models, the effects of the group-level covariates in
the pure-presence model (Fig. 11) are more similar to those estimates of the group-level effects on
⁴⁰⁸ origination (Fig. 12) as opposed to survival (Fig. 13).

As demonstrated in the comparisons of the effect of mass on occurrence from the pure-presence
⁴¹⁰ model (Fig. 8) with the effect of mass on origination and survival from the birth-death model (Fig.
9, and 10), there is considerable variation in the effect of plant phases on ecotype-specific estimates.

⁴¹² Plant phase is estimated to structure ecotype occurrence probability, specifically at least one phase

has a very different estimates from the others, for non-arboreal carnivores, arboreal and unguligrade
414 herbivores, arboreal and fossorial insectivores, and arboreal and unguligrade omnivores (Fig. 11).

For the other ecotypes, plant phase does not correspond to major differences in diversity over time.

416 The temperature covariates do not appear to strongly structure occurrence history for most
ecotypes (Fig. 11). Ecotypes for which at least one temperature covariate is estimated to have
418 strong effect on occurrence are digitigrade carnivores (mean only), scansorial carnivores (mean only),
and non-arboreal herbivores. For the other ecotypes neither of the temperature covariates are
420 expected to have strong effects on occurrence history.

Plant phase is estimated to at least partially structure ecotype origination probability for
422 non-arboreal carnivores, non-fossorial or plantigrade herbivores, arboreal and fossorial insectivores,
and plantigrade and scansorial omnivores (Fig. 12). In the case of the temperature covariates, at
424 least one of them is estimated to have strong effects on origination history for the following
ecotypes: digitigrade carnivores, and both digitigrade and unguligrade herbivores (Fig. 12). Neither
426 of the temperature covariates are estimate to have strong effects for the other ecotypes.

In contrast to both of the above (Fig. 11, and 12) group-level covariates are estimate to have almost
428 no effect on survival for all ecotypes (Fig. 13); this is the case for both the plant phases and
temperature coviates.

430 Analysis of diversity

All of the following analyses of diversification and macroevolutionary rates has been done using
432 only the birth-death model; this is because of the models better posterior predictive check
performance (Fig. 3a, and 3b).

434 The general pattern of total estimated North American mammal diversity for the Cenozoic is
“stable” meaning that mean standing diversity does not fluctuate wildly over the Cenozoic (Fig.
436 14a). In broad strokes, the first 15 or so million years of the Cenozoic are characterized by a gradual
decline in standing diversity until approximately 45-50 million years ago (early-middle Eocene).
438 Following this decline, standing diversity is broadly constant from 45 to 18 Mya (early Miocene).

- After this, there is a rapid spike in diversity followed by a slight decline in diversity up to the
440 Modern. This characterization of the estimated diversity history is knowingly broad strokes and
diversity time series is not without variation and vagaries.
- 442 When viewed through the lens of diversification rate, some of the structure behind the estimated
diversity history begins to take shape (Fig. 14b). For most of the Cenozoic, the diversification rate
444 hovers around zero, punctuated by both positive and negative spikes. The largest spike in
diversification rate is at 16 Mya, which is early Oligocene (Fig. 14b). Other notable increases in
446 diversification rate occur at 54, 44, 36, 26, and 20 Mya; other possible increases in diversification
rate are less certain (e.g. 8 Mya). Notable decreases in diversification rate occur at 52, 48, 42, 32,
448 14, 10, and 6 Mya.

The comparison between per capita origination and extinction rate estimates reveals how
450 diversification rate is formed (Fig. 14c, 14d). Diversification rate seems most driven by changes in
origination rate as opposed to extinction rate. Extinction rate, on the other hand, demonstrates an
452 almost saw-toothed pattern around a constant mean.

- Now ask what origin or extinct are doing at the important time points indicated above.
- 454 Increases in diversification rate at 54, 44, 36, 26, 20, 16
- Decreases in diversification rate at 52, 48, 42, 32, 14, 10, 6
- 456 Diversity partitioned by ecotype reveals a lot of the complexity behind the pattern of mammal
diversity for the Cenozoic (Fig. 15). There are many possible orders to present these results; I've
458 chosen to focus on those with particular relevance to observations and hypotheses presented in the
introduction.
- 460 Arboreal ecotypes obtain peak diversity early in the Cenozoic and then decline for the rest of the
time series, becoming increasingly rare or absent as diversity approaches the Modern (Fig. 15).
- 462 Arboreal herbivores and omnivores obtain peak diversity at the beginning of the Cenozoic then go
into decline while still possibly remaining a part of the species pool, while arboreal carnivores and
464 insectivores obtain peak diversity 52-50 Mya and then quickly decline and become extremely rare or

absent from the species pool.

466 The diversity of both digitigrade and unguligrade herbivores increase over the Cenozoic (Fig. 15).
In contrast, plantigrade herbivore diversity does not have a single, broad-strokes pattern; instead,
468 diversity increases, decreases, and may have shallowly increased again. Contrastingly, fossorial and
scansorial herbivores demonstrate a much flatter history of diversity, with a slight increase in
470 diversity that is more pronounced in the diversity fossorial taxa than scansorial taxa.

Digitigrade carnivores have a multi-modal diversity history, with peaks 54-52 and 12-10 Mya
472 (Fig.15). Between these two peaks digitigrade carnivore diversity dips below average diversity
following the first peak and then grows slowly until the second peak. Plantigrade carnivores obtain
474 peak diversity in the early Cenozoic and then maintain a relatively stable diversity until another
peak at the end of the Cenozoic.

476 There are some broad similarities in diversity histories of insectivorous and omnivorous taxa. The
diversity histories of arboreal, plantigrade, and scansorial insectivorous taxa all demonstrate a
478 decreasing pattern with time, while fossorial insectivores have a flat diversity history with a rapid
peak approximately 10 Mya (Fig. 15). Arboreal and scansorial omnivores decrease in diversity from
480 their initial peaks early in the Cenozoic, and plantigrade omnivores have a generally flat diversity
history with a sudden peak in diversity late in the Cenozoic (Fig. 15). Unguligrade omnivores also
482 demonstrate a possible decrease in diversity over the Cenozoic, but not as clearly as arboreal and
scansorial omnivores.

484 Many of the estimated ecotype specific diversity histories share a similar increases in diversity to
one degree or another at the late Cenozoic 16-14 Mya (Fig. 15); these increases are either sustained
486 or temporary: digitigrade carnivores, plantigrade carnivores, scansorial carnivores, unguligrade
herbivores, fossorial insectivores, and plantigrade omnivores.

⁴⁸⁸ Discussion

Both species pools and environmental context change over time, though not at the same rate. Local
⁴⁹⁰ communities, who's species are drawn from this pool, have “roles” in their communities as their interact with a host of biotic and abiotic interactors. For higher level ecological characterizations
⁴⁹² like ecotypes and guilds, these roles are broadly defined and not about specific interactions but the genre of interactions covered by that grouping.

⁴⁹⁴ It has been observed that the diversity of an ecotype or guild can be stable over millions of years despite constant species turnover (Jernvall and Fortelius, 2004; Slater, 2015) CITATIONS. This
⁴⁹⁶ implies that the size and scope of the role of an ecotype or guild is preserved even as the individual interactors change.

⁴⁹⁸ What is the pace of environmental change in North America over the Cenozoic? Is it a constant process or a pulsed one?

⁵⁰⁰ Comparison of the pure-presence model to the birth-death model support the conclusion that regional species pool dynamics cannot simply be described by a single probability of occurrence and
⁵⁰² is instead the product of both origination and extinction. Additionally, changes to ecotypic composition of the North American regional species pool are driven primarily by variation in
⁵⁰⁴ origination rates. This aspect of how regional species pool diversity is shaped is not observable from studies of the Modern CITATION.

⁵⁰⁶ The time scale available with paleontological data is much greater than that obtainable from modern ecological studies, even long running observations CITATION. Specifically, the temporal
⁵⁰⁸ scale of paleontological data allows for the complete turnover of a species pool to be observed, something that is impossible in “real time.” However, paleontological data is very limited in its
⁵¹⁰ spatial resolution, so the analysis of how the ecotypic diversity local communities change over time and how that is also the product of larger scale regional turnover remains unanswered.

⁵¹² Phylogenetic comparative community ecology and phylogenetic comparative biogeography also discusses how the macroevolutionary processes helps structure an observed community, though it is

514 not necessarily phrased that way. However, that community did not form in isolation but it the
result of many factors interacting over time including incumbency, competition, limiting similarity,
516 etc.

How do the results line up with previous observations and hypotheses discussed in the introduction?

518 Extinction rate for the entire regional species pool through time is highly variable and demonstrates
a saw-toothed pattern around an apparently constant mean. While a constant mean extinction rate
520 is consistent with previous observation (Alroy, 1996; Alroy et al., 2000), the degree to which
extinction rate is actually variable may not have been equally appreciated. What is most consistent
522 with previous observations (Alroy, 1996; Alroy et al., 2000), however, is that diversity seems to be
most structured by origination than extinction.

524 Plant phase always affects plantigrade ecotypes. Does it affect all plantigrade taxa?

Temperature affects very little in general with a few major exceptions: origination probabilities of
526 digitigrade carnivores, and digitigrade and unguligrade herbivores. Why?

Arboreal taxa disappear over the Cenozoic, with massive disappearance by the Paleogene-Neogene
528 barrier. This is consistent with one of the possible explanations presented: Paleogene-Neogene are
different and while the earliest Cenozoic may have been neutral wrt arboreal taxa, they disappeared
530 quickly which may account for their higher extinction risk.

Digitigrade carnivores have a relatively stable diversity history through the Cenozoic and could be
532 characterized as varying around a constant mean diversity. This result is consistent with similar
observations in Silvestro et al. (2015); Slater (2015).

534 Both digitigrade and unguligrade herbivores increase in diversity over the Cenozoic. The increase of
these cursorial forms is consistent with the gradual opening up of the North American landscape
536 CITATION.

What these results support is a gradual change to the ecotypic diversity of the regional species pool
538 for the Cenozoic.

The rapidity of Cenozoic environmental change is worth discussing. If change is rapid, ecotypic
540 composition of species pool does not seem to track environmental change. If change is gradual then
there is the possibility that changes to ecotypic composition may be tracking environmental change.

542 The effects of phylogeny on origination and extinction are not directly considered in this analysis.
While a birth-death process approximates the speciation-extinction process (Silvestro et al., 2014)
544 this is not same as considering how the similarity between closely related species may affect
estimates for the effects of species traits or response to environmental on both origination and
546 extinction (Harnik et al., 2014; Smits, 2015). One of the principle barriers to the inclusion of the
effect of phylogeny in either the pure-presence or birth-death models is computational; with over
548 1000 tips, the calculation of the scale parameter defining phylogenetic effect would be very slow and
further increase the already slow computation time necessary for both the marginalization of the
550 discrete occurrence histories and data augmentation already included in both models.

The effect of species mass on either occurrence or origination and extinction was not allowed to
552 vary by ecotype even though there may be difference amoungst those ecotypes CITATION. The
primary reason for this modeling choice was this studies focus on ecotypic differences in occurrence,
554 or origination and extinction. Allowing the effect of this covariate to vary by ecotype, time, and
environmental factors would increase the overall complexity of the model, something that may not
556 be necessary because the covariate is not the focus of this study. Instead, this covariate was
included in order to control for its possible underlying effects CITATION. Additionally, body size
558 was allowed to have a second-order polynomial form and no higher order polynomials were
considered; this was done BECAUSE

560 The only covariate allowed to affect sampling probability is mass and only as a linear predictor.
Other covariates, such as the environmental factors considered here, may have affected the
562 underlying preservation process that limits sampling probability. It should be noted that in other
similar studies that use a hidden birth-death model to handle simultaneous estimation of sampling,
564 origination, and extinction have not considered the possible effects of covariates, both species traits
and environmental factors, on sampling CITATION.

⁵⁶⁶ An ideal system would have the temporal scale of the fossil record combined with the spatial scale afforded in studies of extant systems.

⁵⁶⁸ **Acknowledgements**

I would like to thank K. Angielczyk, M. Foote, P. D. Polly, and R. Ree for helpful discussion and
⁵⁷⁰ advice. This entire study would not have been possible without the Herculean effort of the many contributors to the Paleobiology Database. In particular, I would like to thank J. Alroy and
⁵⁷² M. Uhen for curating most of the mammal occurrences recorded in the PBDB. This is Paleobiology Database publication XXX.

⁵⁷⁴ **References**

- Allen, L. J. S. 2011. An introduction to stochastic processes with applications to biology. 2nd ed.
⁵⁷⁶ Chapman and Hall/CRC, Boca Raton, FL.
- Alroy, J. 1996. Constant extinction, constrained diversification, and uncoordinated stasis in North
⁵⁷⁸ American mammals. *Palaeogeography, Palaeoclimatology, Palaeoecology* 127:285–311.
- . 2009. Speciation and extinction in the fossil record of North American mammals. Pages
⁵⁸⁰ 302–323 *in* R. K. Butlin, J. R. Bridle, and D. Schluter, eds. *Speciation and patterns of diversity*. Cambridge University Press, Cambridge.
- ⁵⁸² Alroy, J., P. L. Koch, and J. C. Zachos. 2000. Global climate change and North American
mammalian evolution. *Paleobiology* 26:259–288.
- ⁵⁸⁴ Badgley, C., and J. A. Finarelli. 2013. Diversity dynamics of mammals in relation to tectonic and
climatic history: comparison of three Neogene records from North America. *Paleobiology*
⁵⁸⁶ 39:373–399.
- Bambach, R. K. 1977. Species richness in marine benthic habitats through the Phanerozoic.
⁵⁸⁸ *Paleobiology* 3:152–167.

- Bambach, R. K., A. M. Bush, and D. H. Erwin. 2007. Autecology and the filling of ecospace: Key
590 metazoan radiations. *Palaeontology* 50:1–22.
- Blois, J. L., and E. A. Hadly. 2009. Mammalian Response to Cenozoic Climatic Change. Annual
592 Review of Earth and Planetary Sciences 37:181–208.
- Brown, A. M., D. I. Warton, N. R. Andrew, M. Binns, G. Cassis, and H. Gibb. 2014. The
594 fourth-corner solution - using predictive models to understand how species traits interact with
the environment. *Methods in Ecology and Evolution* 5:344–352.
- 596 Bush, A. M., and R. K. Bambach. 2011. Paleoeologic Megatrends in Marine Metazoa, vol. 39.
- Bush, A. M., R. K. Bambach, and G. M. Daley. 2007. Changes in theoretical ecospace utilization in
598 marine fossil assemblages between the mid-Paleozoic and late Cenozoic. *Paleobiology* 33:76–97.
- Clyde, W. C., and P. D. Gingerich. 1998. Mammalian community response to the latest Paleocene
600 thermal maximum: an isotaphonomic study in the northern Bighorn Basin, Wyoming. *Geology*
26:1011–1014.
- 602 Cramer, B. S., K. Miller, P. Barrett, and J. Wright. 2011. Late Cretaceous-Neogene trends in deep
ocean temperature and continental ice volume: Reconciling records of benthic foraminiferal
604 geochemistry ($\delta^{18}\text{O}$ and Mg/Ca) with sea level history. *Journal of Geophysical Research: Oceans*
116:1–23.
- 606 Eronen, J. T., C. M. Janis, C. P. Chamberlain, and A. Mulch. 2015. Mountain uplift explains
differences in Palaeogene patterns of mammalian evolution and extinction between North
608 America and Europe. *Proceedings of the Royal Society B: Biological Sciences* 282:20150136.
- Ezard, T. H. G., A. Purvis, and H. Morlon. 2016. Environmental changes define ecological limits to
610 species richness and reveal the mode of macroevolutionary competition. *Ecology Letters*
19:899–906.
- 612 Figueirido, B., C. M. Janis, J. A. Pérez-Claros, M. De Renzi, and P. Palmqvist. 2012. Cenozoic

- climate change influences mammalian evolutionary dynamics. *Proceedings of the National Academy of Sciences* 109:722–727.
- Foote, M. 2001. Inferring temporal patterns of preservation, origination, and extinction from taxonomic survivorship analysis. *Paleobiology* 27:602–630.
- Foote, M., and J. J. Sepkoski. 1999. Absolute measures of the completeness of the fossil record. *Nature* 398:415–7.
- Gelman, A. 2008. Scaling regression inputs by dividing by two standard deviations. *Statistics in Medicine* pages 2865–2873.
- Gelman, A., J. B. Carlin, H. S. Stern, D. B. Dunson, A. Vehtari, and D. B. Rubin. 2013. Bayesian data analysis. 3rd ed. Chapman and Hall, Boca Raton, FL.
- Gelman, A., and J. Hill. 2007. Data Analysis using Regression and Multilevel/Hierarchical Models. Cambridge University Press, New York, NY.
- Harnik, P. G., P. C. Fitzgerald, J. L. Payne, and S. J. Carlson. 2014. Phylogenetic signal in extinction selectivity in Devonian terebratulide brachiopods. *Paleobiology* 40:675–692.
- Jamil, T., W. A. Ozinga, M. Kleyer, and C. J. F. Ter Braak. 2013. Selecting traits that explain species-environment relationships: A generalized linear mixed model approach. *Journal of Vegetation Science* 24:988–1000.
- Janis, C., J. Damuth, and J. M. Theodor. 2004. The species richness of Miocene browsers, and implications for habitat type and primary productivity in the North American grassland biome. *Palaeogeography, Palaeoclimatology, Palaeoecology* 207:371–398.
- Janis, C. M. 1993. Tertiary mammal evolution in the context of changing climates, vegetation, and tectonic events. *Annual Review of Ecology and Systematics* 24:467–500.
- . 2008. An evolutionary history of browsing and grazing ungulates. Pages 21–45 in I. J. Gordon and H. H. T. Prins, eds. *The Ecology of Browsing and Grazing*. Springer-Verlag.
- Janis, C. M., J. Damuth, and J. M. Theodor. 2000. Miocene ungulates and terrestrial primary

- 638 productivity: where have all the browsers gone? *Proceedings of the National Academy of Sciences*
97:7899–904.
- 640 Janis, C. M., G. F. Gunnell, and M. D. Uhen. 2008. Evolution of Tertiary mammals of North
America. Vol. 2. Small mammals, xenarthrans, and marine mammals. Cambridge University
642 Press, Cambridge.
- 644 Janis, C. M., K. M. Scott, and L. L. Jacobs. 1998. Evolution of Tertiary mammals of North
America. Vol. 1. Terrestrial carnivores, ungulates, and ungulate-like mammals. Cambridge
University Press, Cambridge.
- 646 Janis, C. M., and P. B. Wilhelm. 1993. Were there mammalian pursuit predators in the tertiary?
Dances with wolf avatars. *Journal of Mammalian Evolution* 1:103–125.
- 648 Jernvall, J., and M. Fortelius. 2004. Maintenance of trophic structure in fossil mammal
communities: site occupancy and taxon resilience. *The American Naturalist* 164:614–624.
- 650 Legendre, S. 1986. Analysis of mammalian communities from the Late Eocene and Oligocene of
Southern France. *Paleovertebrata* 16:191–212.
- 652 Liow, L. H., M. Fortelius, E. Bingham, K. Lintulaakso, H. Mannila, L. Flynn, and N. C. Stenseth.
2008. Higher origination and extinction rates in larger mammals. *Proceedings of the National
654 Academy of Sciences* 105:6097–6102.
- 656 Liow, L. H., M. Fortelius, K. Lintulaakso, H. Mannila, and N. C. Stenseth. 2009. Lower Extinction
Risk in SleeporHide Mammals. *The American Naturalist* 173:264–272.
- 658 Lloyd, G. T., J. R. Young, and A. B. Smith. 2011. Taxonomic Structure of the Fossil Record is
Shaped by Sampling Bias. *Systematic Biology* 61:80–89.
- 660 Pires, M. M., D. Silvestro, and T. B. Quental. 2015. Continental faunal exchange and the
asymmetrical radiation of carnivores. *Proceedings of the Royal Society B: Biological Sciences*
282:20151952.

- 662 Pollock, L. J., W. K. Morris, and P. A. Vesk. 2012. The role of functional traits in species distributions revealed through a hierarchical model. *Ecography* 35:716–725.
- 664 Quental, T. B., and C. R. Marshall. 2013. How the Red Queen Drives Terrestrial Mammals to Extinction. *Science* 341:290–292.
- 666 Royle, J. A., and R. M. Dorazio. 2008. Hierarchical modeling and inference in ecology: the analysis of data from populations, metapopulations and communities. Elsevier, London.
- 668 ———. 2012. Parameter-expanded data augmentation for Bayesian analysis of capture-recapture models. *Journal of Ornithology* 152:521–537.
- 670 Royle, J. A., R. M. Dorazio, and W. a. Link. 2007. Analysis of Multinomial Models With Unknown Index Using Data Augmentation. *Journal of Computational and Graphical Statistics* 16:67–85.
- 672 Royle, J. A., J. D. Nichols, M. Kéry, E. Ranta, and M. Kery. 2014. detection is of species when Modelling occurrence and abundance imperfect 110:353–359.
- 674 Rubin, D. B. 1996. Multiple imputation after 18+ years. *Journal of the American Statistical Association* 91:473–489.
- 676 Silvestro, D., A. Antonelli, N. Salamin, and T. B. Quental. 2015. The role of clade competition in the diversification of North American canids. *Proceedings of the National Academy of Sciences of the United States of America* 112:8684–9.
- 680 Silvestro, D., J. Schnitzler, L. H. Liow, A. Antonelli, and N. Salamin. 2014. Bayesian estimation of speciation and extinction from incomplete fossil occurrence data. *Systematic biology* 63:349–67.
- 682 Slater, G. J. 2015. Iterative adaptive radiations of fossil canids show no evidence for diversity-dependent trait evolution. *Proceedings of the National Academy of Sciences* 112:4897–4902.
- 684 Smits, P. D. 2015. Expected time-invariant effects of biological traits on mammal species duration. *Proceedings of the National Academy of Sciences* 112:13015–13020.
- 686 Strömberg, C. A. E. 2005. 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* 102:11980–4.
- 688
- Tomiya, S. 2013. Body Size and Extinction Risk in Terrestrial Mammals Above the Species Level. *The American Naturalist* 182:196–214.
- 690
- Valentine, J. W. 1969. Patterns of taxonomic and ecological structure of the shelf benthos during Phanerozoic time. *Paleontology* 12:684–709.
- 692
- Wang, S. C., P. J. Everson, H. J. Zhou, D. Park, and D. J. Chudzicki. 2016. Adaptive credible intervals on stratigraphic ranges when recovery potential is unknown. *Paleobiology* 42:240–256.
- 694
- Wang, S. C., and C. R. Marshall. 2016. Estimating times of extinction in the fossil record. *Biology Letters* 12:20150989.
- 696
- Warton, D. I., B. Shipley, and T. Hastie. 2015. CATS regression - a model-based approach to studying trait-based community assembly. *Methods in Ecology and Evolution* 6:389–398.
- 698
- Zachos, J. C., G. R. Dickens, and R. E. Zeebe. 2008. An early Cenozoic perspective on greenhouse warming and carbon-cycle dynamics. *Nature* 451:279–283.
- 700
- Zachos, J. C., M. Pagani, L. Sloan, E. Thomas, and K. Billups. 2001. Trends, rhythms, and aberrations in global climate 65 Ma to present. *Science* 292:686–693.
- 702

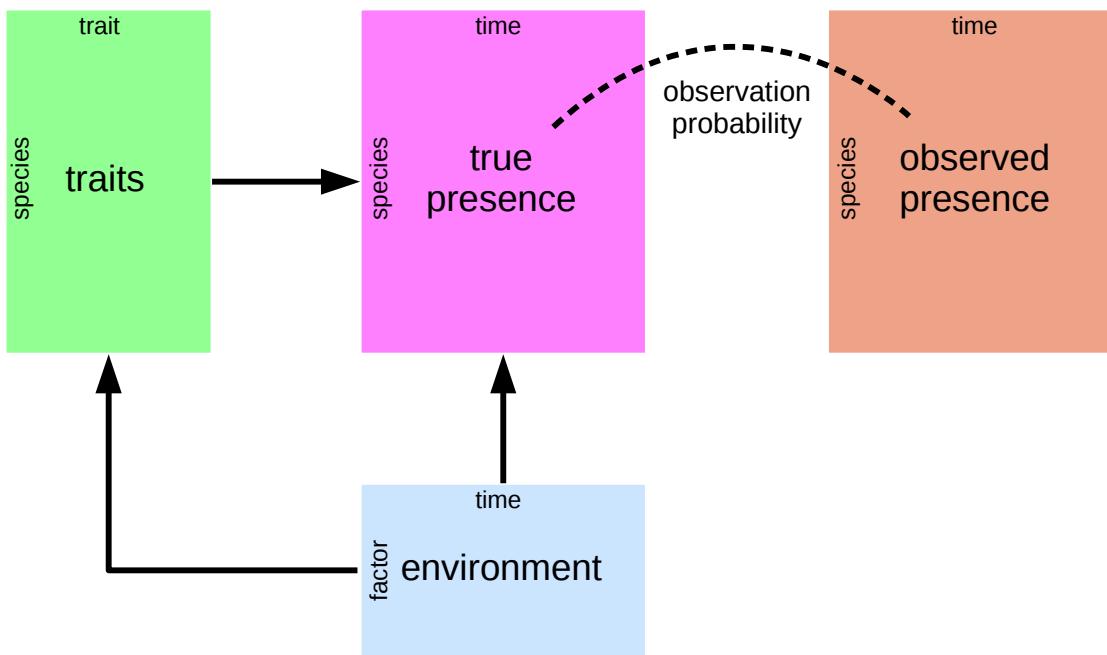


Figure 1: Conceptual diagram of the paleontological fourth corner problem. The observed presence matrix (orange) is the empirical presence/absence pattern for all species for all time points; this matrix is an incomplete observation of the “true” presence/absence pattern (purple). The estimated true presence matrix is modeled as a function of both environmental factors over time (blue) and multiple species traits (green). Additionally, the affect of environmental factors on species traits are also modeled as traits are expected to mediate the effects of a species environmental context. This diagram is based partially on material presented in Brown et al. (2014) and Warton et al. (2015).

	Time Bin							
	1	2	3	4	5	6	7	8
Observed	0	0	0	1	0	1	1	0
Certain	?	?	?	1	1	1	1	?
Potential	0	0	0	1	1	1	1	0
Potential	0	0	1	1	1	1	1	0
Potential	0	1	1	1	1	1	1	0
Potential	1	1	1	1	1	1	1	0
Potential	0	0	0	1	1	1	1	1
Potential	0	0	1	1	1	1	1	1
Potential	0	1	1	1	1	1	1	1
Potential	1	1	1	1	1	1	1	1

Figure 2: Conceptual figure of all possible occurrence histories for an observed species. The first row represents the observed presence/absence pattern for a single species at eight time points. The second row corresponds to the known aspects of the “true” occurrence history of that species. The remaining rows correspond to all possible occurrence histories that are consistent with the observed data. The process of parameter marginalization described in the text

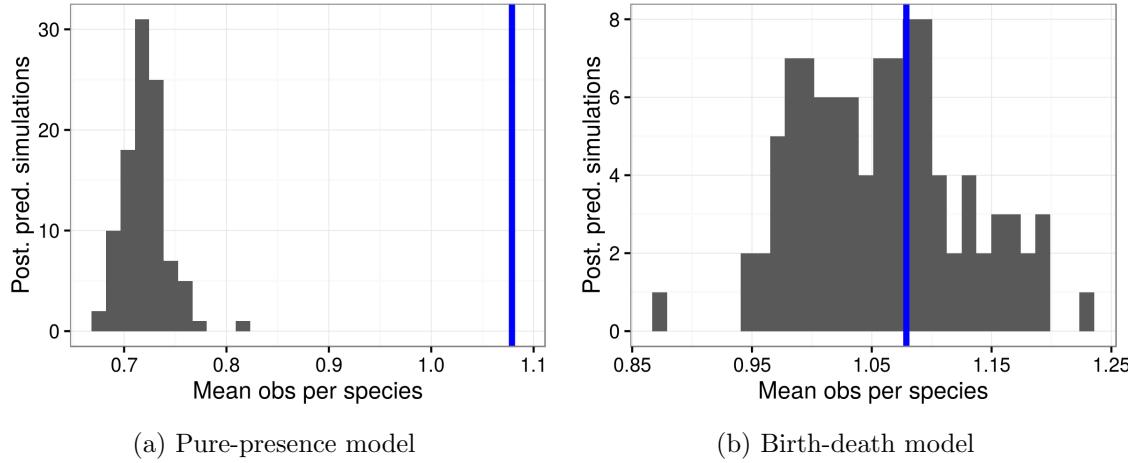


Figure 3: Comparison of the average observed number of occurrences per species (blue line) to the average number of occurrences from 100 posterior predictive datasets using the posterior estimates from the pure-presence and birth-death models.

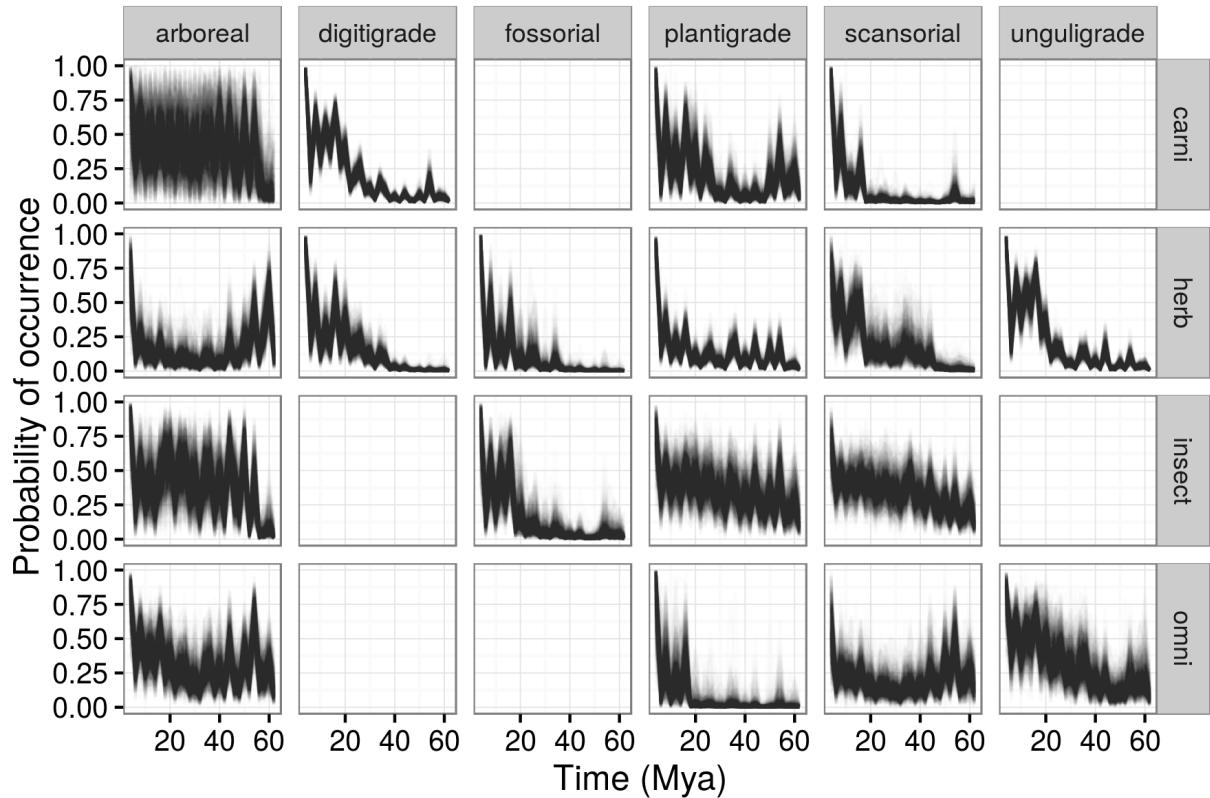


Figure 4: Probability of a mammal ecotype occurring over time as estimated from the pure-presence model. Each panel depicts 100 random samples from the model's posterior. The columns are by locomotor category and rows by dietary category; their intersections are the observed and analyzed ecotypes. Panels with no lines are ecotypes not observed in the dataset.

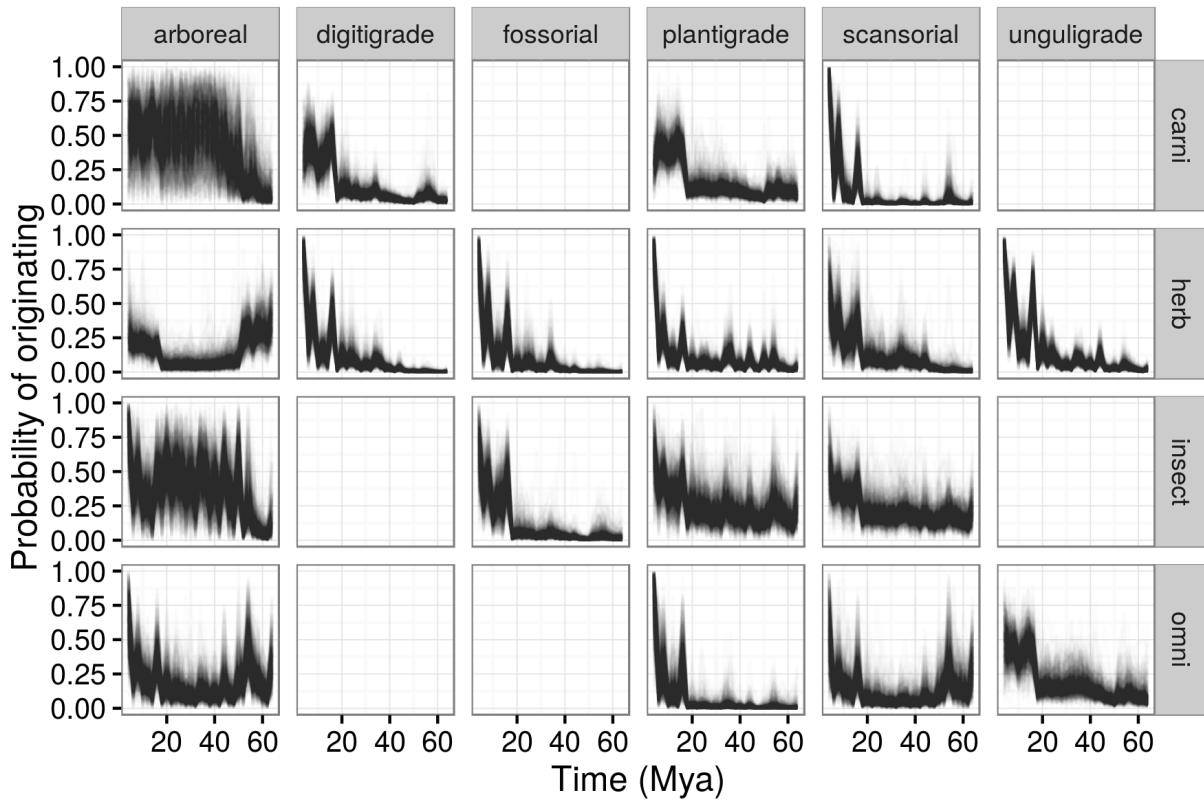


Figure 5: Probability of a mammal ecotype origination probabilities at each time point as estimated from the birth-death model. Each panel depicts 100 random samples from the model's posterior. The columns are by locomotor category and rows by dietary category; their intersections are the observed and analyzed ecotypes. Panels with no lines are ecotypes not observed in the dataset.

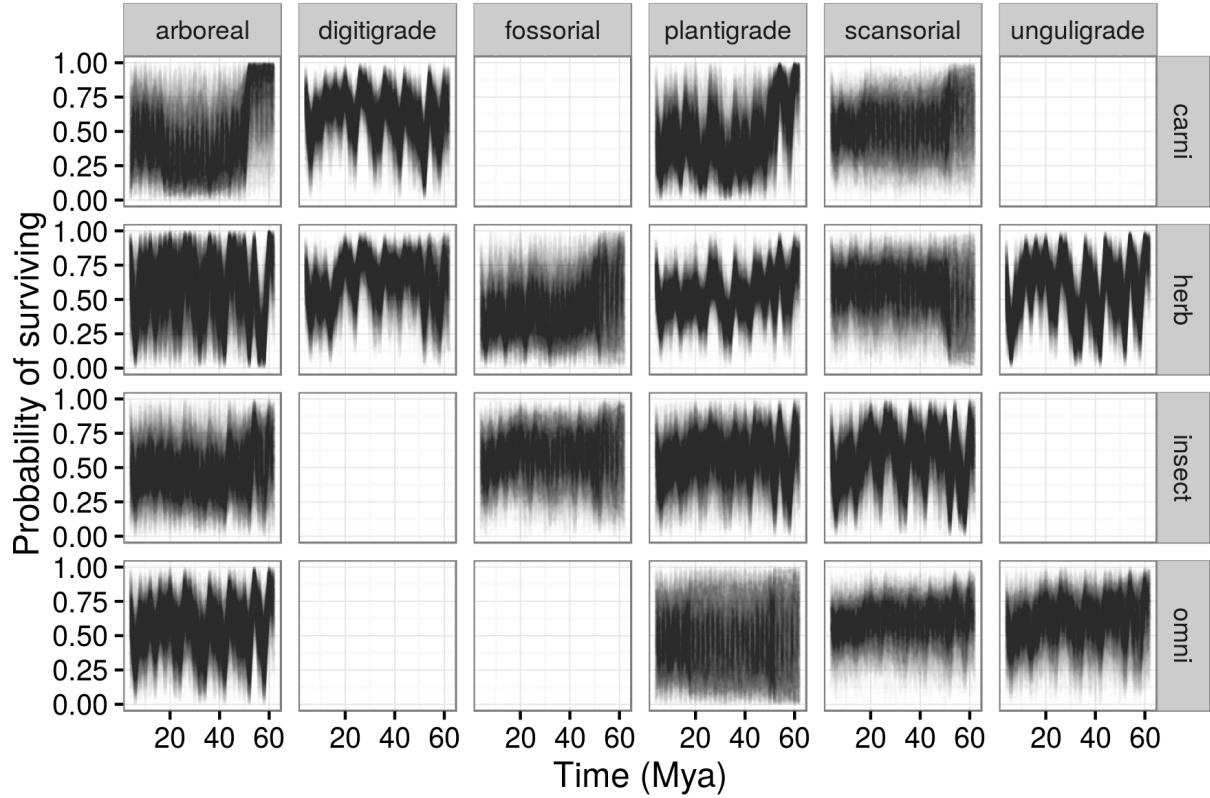


Figure 6: Probability of a mammal ecotype survival probabilities at each time point as estimated from the birth-death model. Each panel depicts 100 random samples from the model’s posterior. The columns are by locomotor category and rows by dietary category; their intersections are the observed and analyzed ecotypes. Panels with no lines are ecotypes not observed in the dataset.

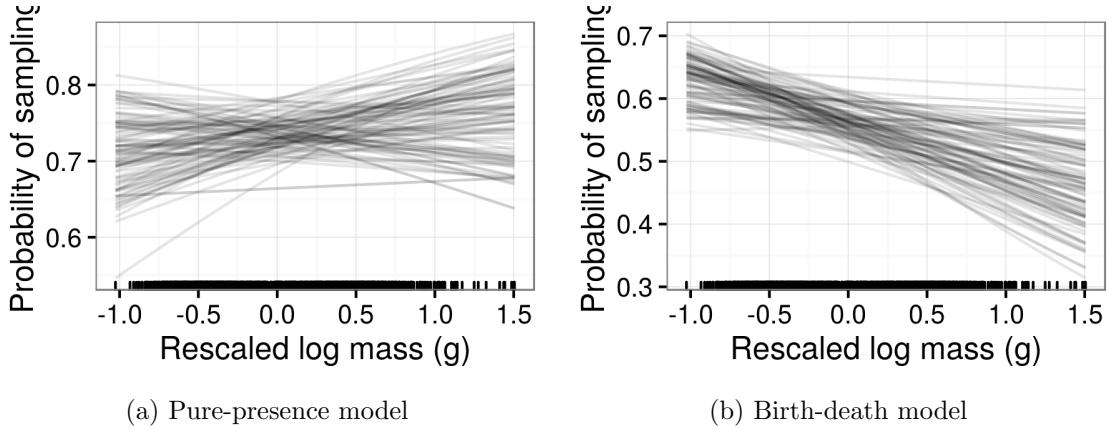


Figure 7: Estimates of the effect of species mass on probability of sampling a present species (p). Mass has been log-transformed, centered, and rescaled; this means that a mass of 0 corresponds to the mean of log-mass of all observed species and that mass is in standard deviation units. Estimates are from both the pure-presence and birth-death models.

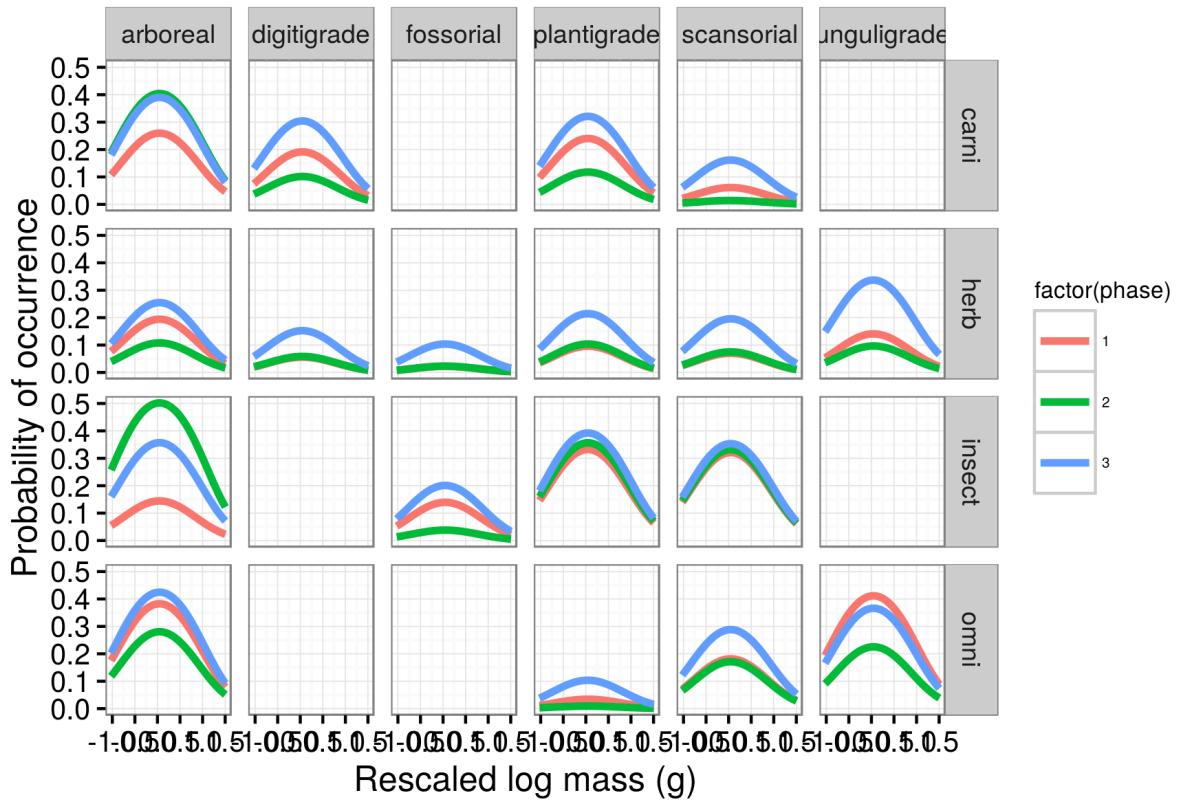


Figure 8: Mean estimate of the effect of species mass on the probability of a species occurrence for each of the three plant phases. The effect of mass is considered constant over time and that the only aspect of the model that changes with plant phase is the intercept of the relationship between mass and occurrence. The three plant phases are indicated by the color of the line. Mass has been log-transformed, centered, and rescaled; this means that a mass of 0 corresponds to the mean of log-mass of all observed species and that mass is in standard deviation units. Only the mean estimates of the effects of both mass and plant phase are plotted for clarity; these estimates are obviously made with uncertainty.

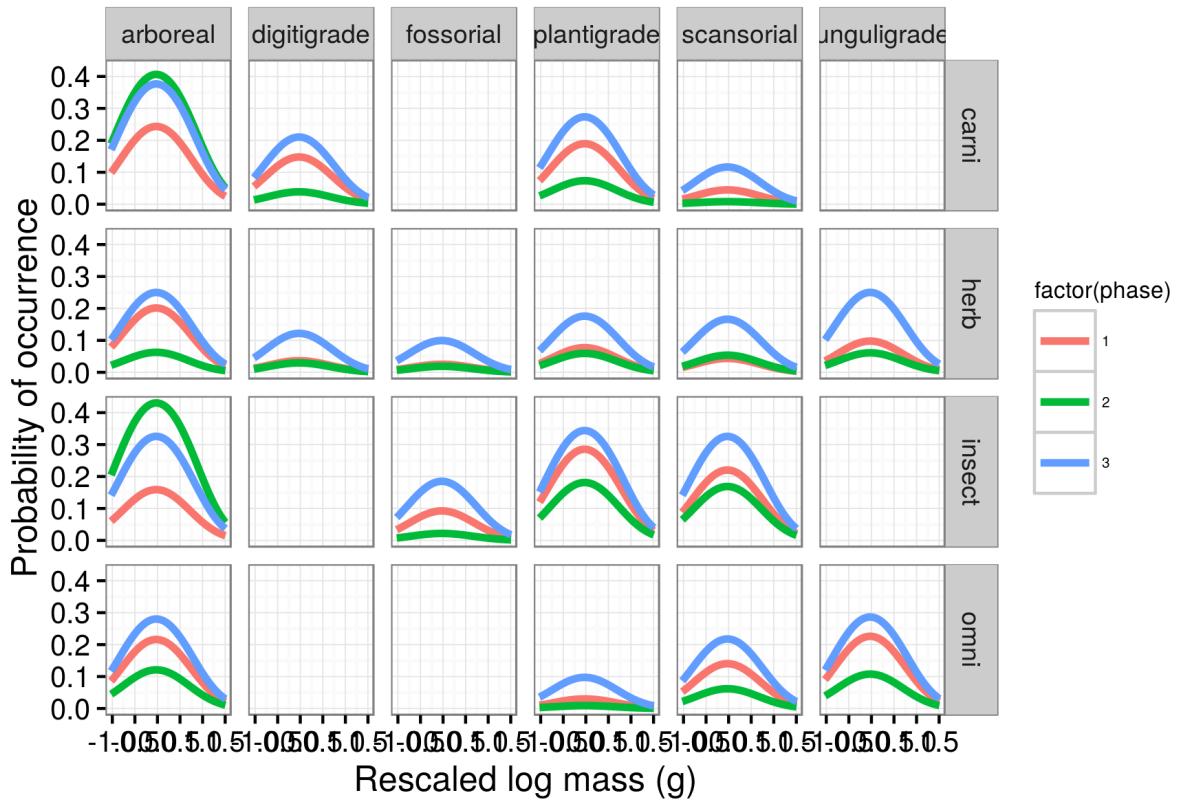


Figure 9: Mean estimate of the effect of species mass on the probability of a species originating for each of the three plant phases. The effect of mass is considered constant over time and that the only aspect of the model that changes with plant phase is the intercept of the relationship between mass and origination. The three plant phases are indicated by the color of the line. Mass has been log-transformed, centered, and rescaled; this means that a mass of 0 corresponds to the mean of log-mass of all observed species and that mass is in standard deviation units. Only the mean estimates of the effects of both mass and plant phase are plotted for clarity; these estimates are obviously made with uncertainty.

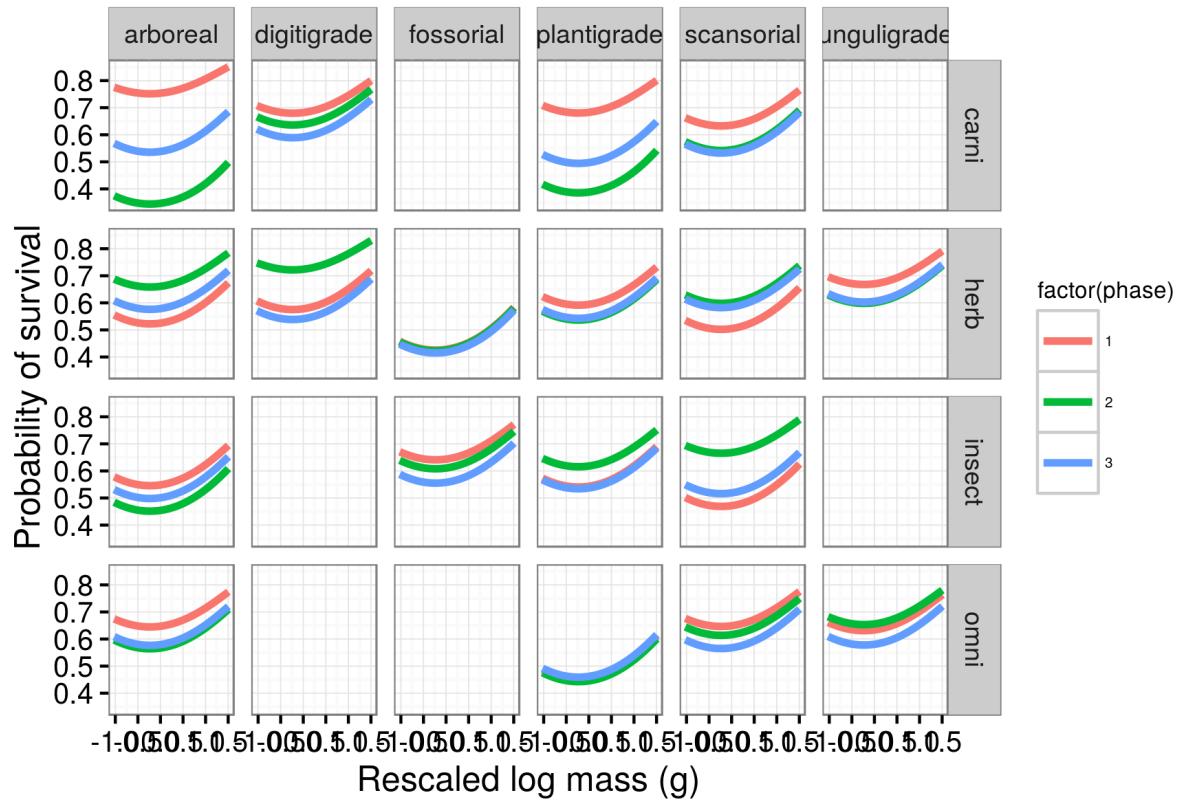


Figure 10: Mean estimate of the effect of species mass on the probability of a species survival for each of the three plant phases. The effect of mass is considered constant over time and that the only aspect of the model that changes with plant phase is the intercept of the relationship between mass and survival. The three plant phases are indicated by the color of the line. Mass has been log-transformed, centered, and rescaled; this means that a mass of 0 corresponds to the mean of log-mass of all observed species and that mass is in standard deviation units. Only the mean estimates of the effects of both mass and plant plant are plotted for clarity; these estimates are obviously made with uncertainty.

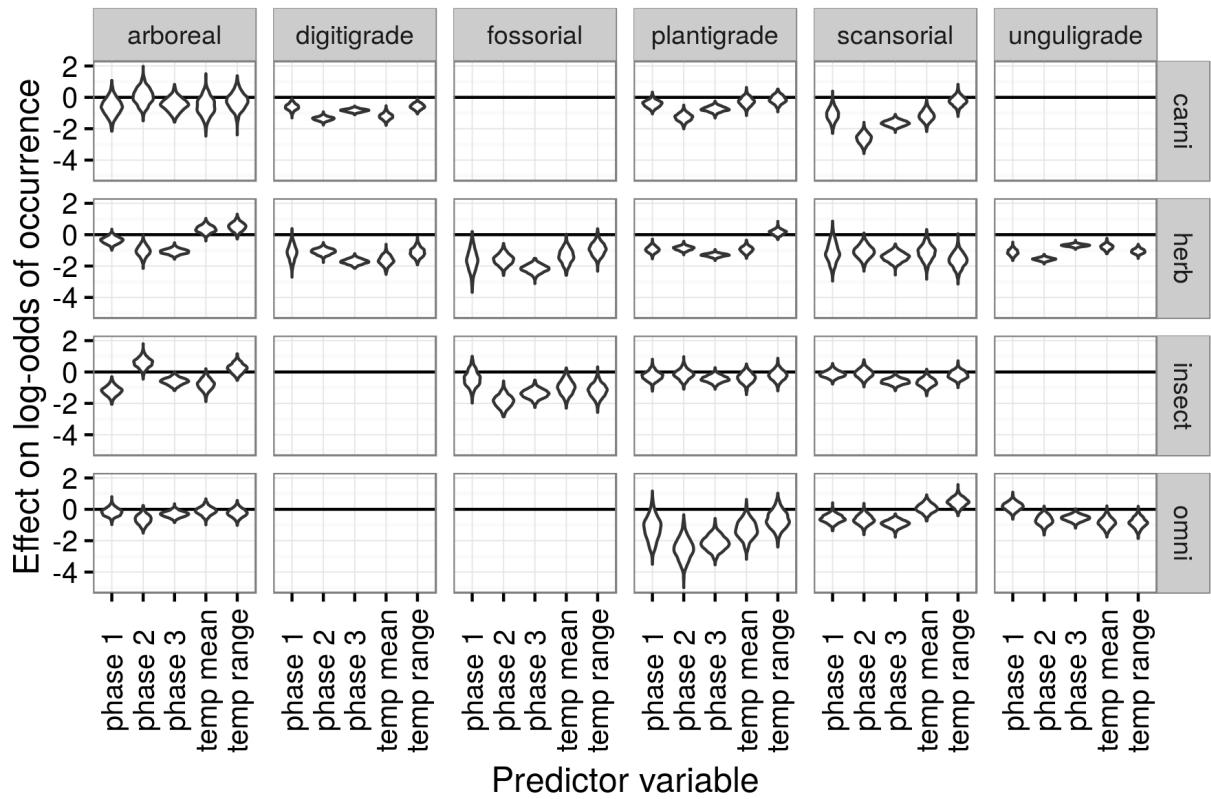


Figure 11: Estimated effects of the group-level covariates describing environmental context on log-odds of species occurrence. These estimates are from the pure-presence model.

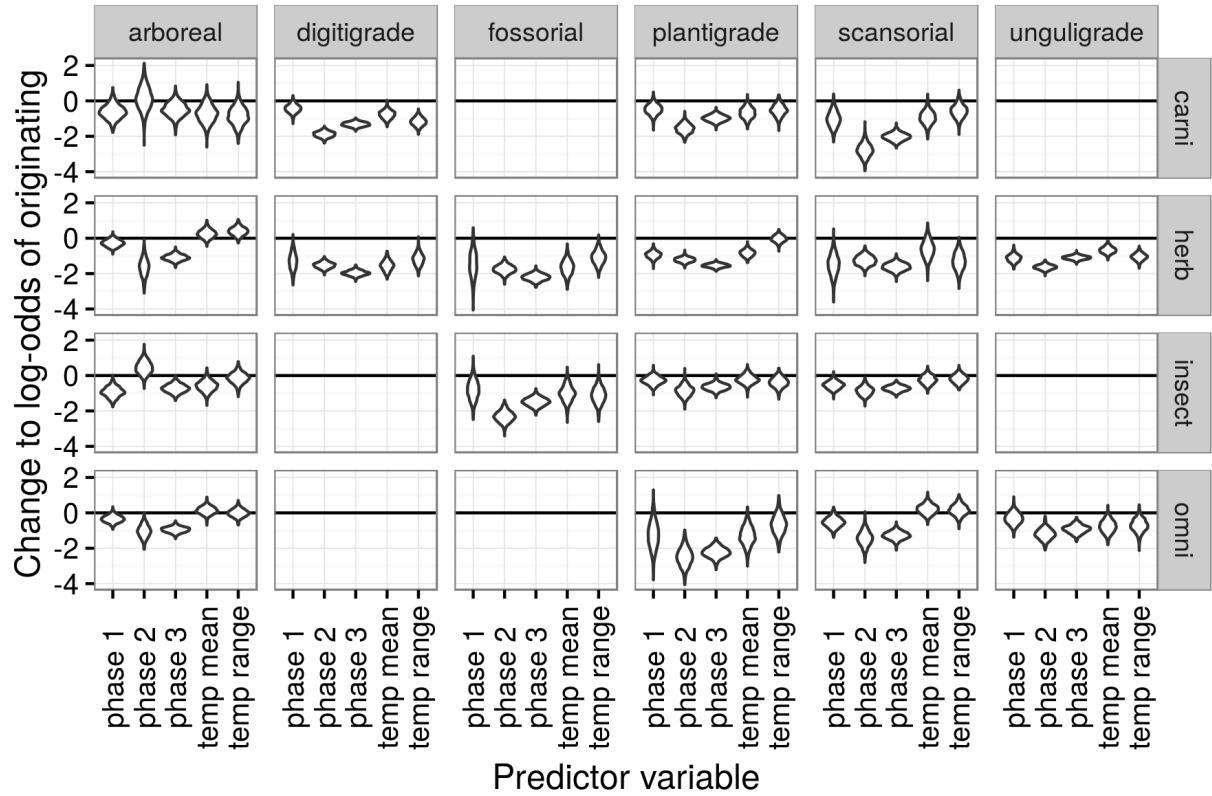


Figure 12: Estimated effects of the group-level covariates describing environmental context on log-odds of species origination. These estimates are from the birth-death model.

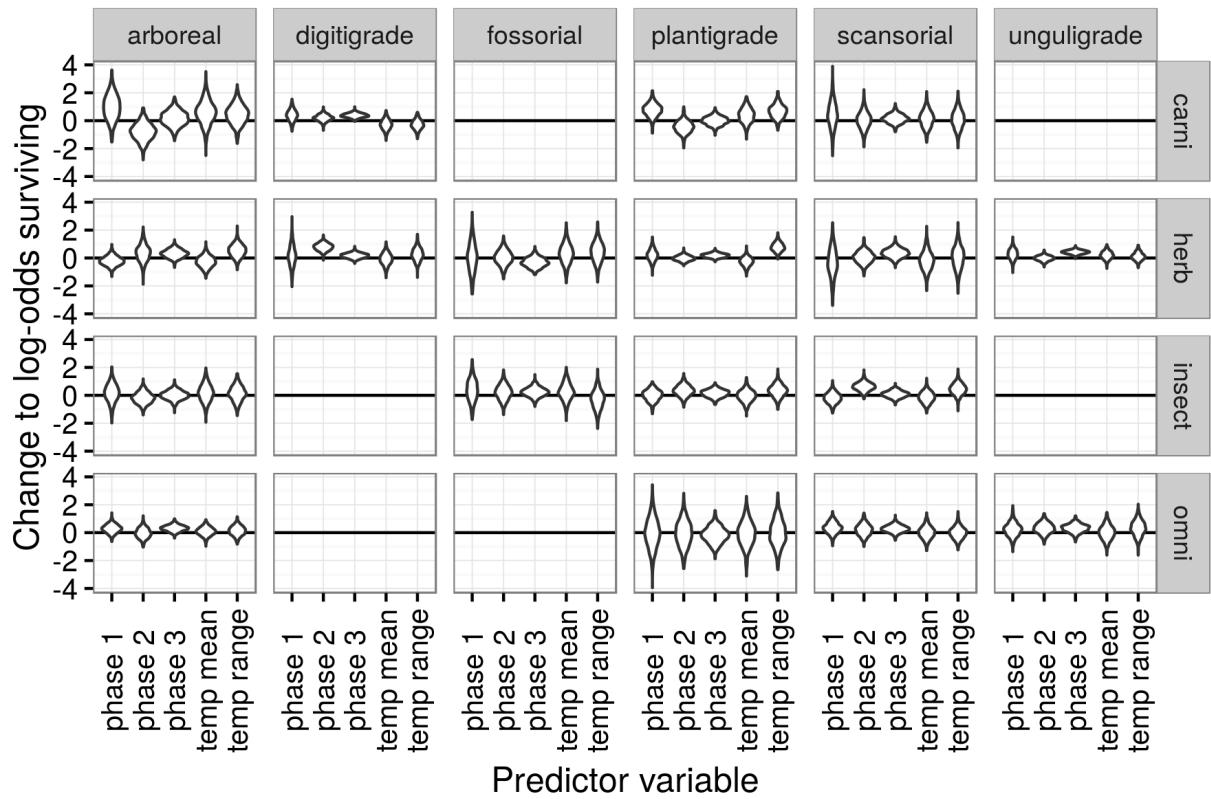


Figure 13: Estimated effects of the group-level covariates describing environmental context on log-odds of species survival. These estimates are from the birth-death model.

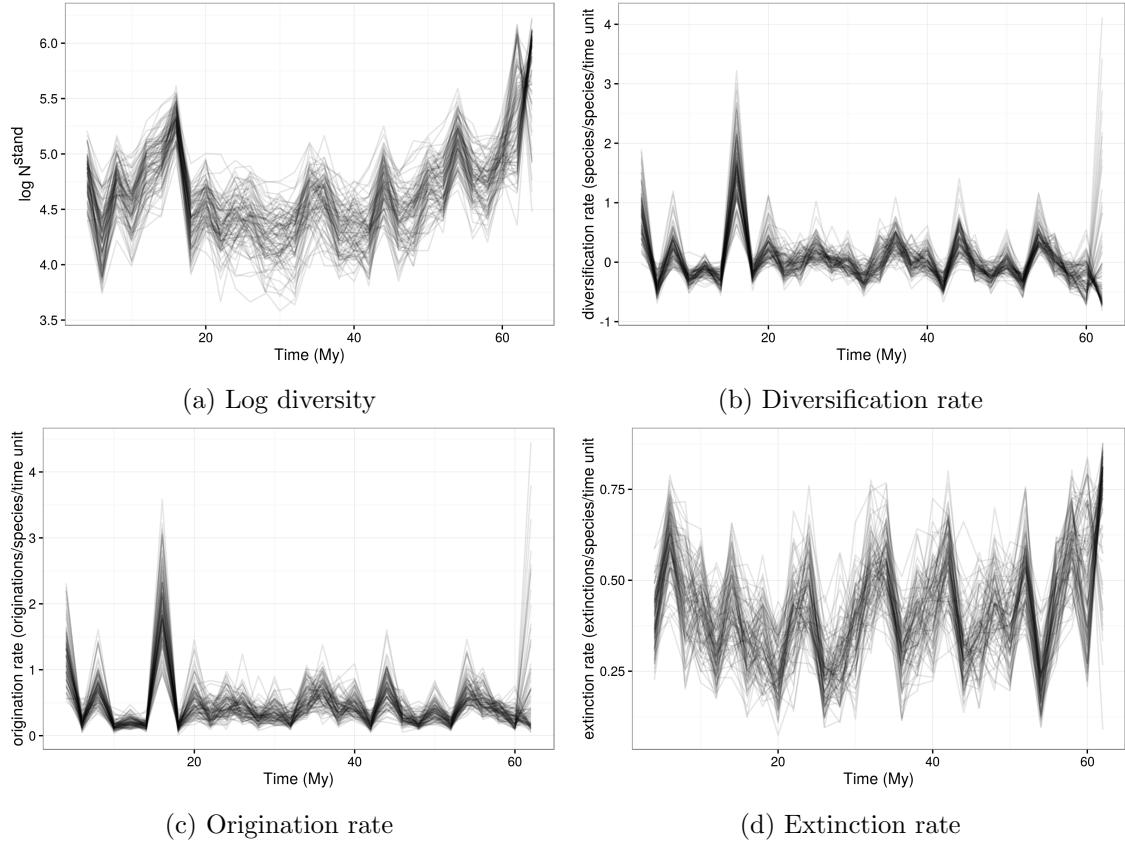


Figure 14: Posterior estimates of the time series of Cenozoic North American mammal diversity and its characteristic macroevolutionary rates; all estimates are from the birth-death model and 100 posterior draws are plotted to indicate the uncertainty in these estimates. The dramatic differences between diversity estimates at the first and second time points and the penultimate and last time points in this series are caused by well known edge effects in discrete-time birth-death models caused by $p_{-,t=1}$ and $p_{-,t=T}$ being partially unidentifiable (Royle and Dorazio, 2008); the hierarchical modeling strategy used here helps mitigate these effects but they are still present (Gelman et al., 2013; Royle and Dorazio, 2008). Diversification rate is in units of species gained per species present per time unit (2 My), origination rate is in units of species originating per species present per time unit, and extinction rate is in units of species becoming extinct per species present per time unit.

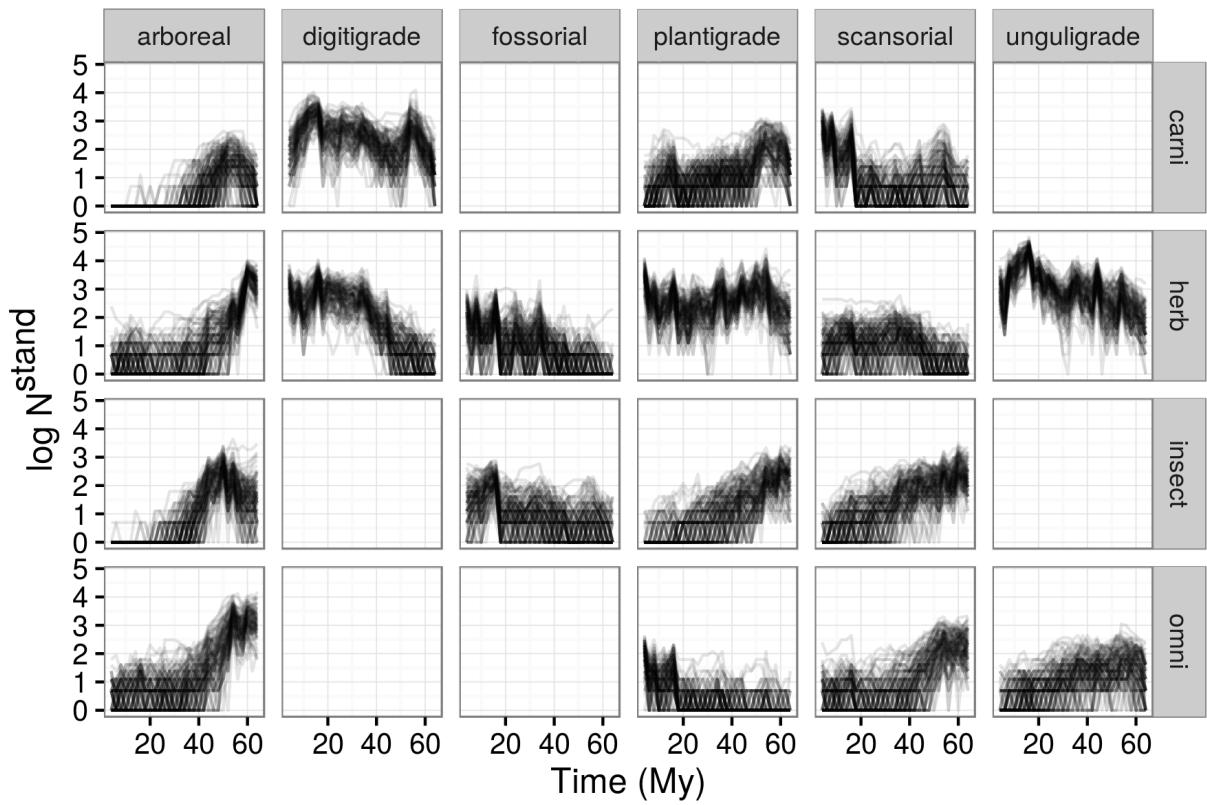


Figure 15: Posterior of standing log-diversity of North American mammals by ecotype for the Cenozoic as estimated from the birth-death model; 100 posterior draws are plotted to indicate the uncertainty in these estimates and what is technically plotted is log of diversity plus 1.