

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)

6 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
over time due to speciation, migration, extinction. Local scale processes like resource competition
8 only affect the regional species pool if all communities are affected.

10 How do species pools change over time as species are recruited or go extinct? When are specific
species ecologies enriched or depleted in the environment? How does global and regional
12 environmental context affect the distribution of species ecotypes (e.g. guilds) in a regional species
pool? All of these questions fall under a single umbrella of analysis of ecotypic diversity and
diversification.

14 Guilds are a set of species with similar sets of interactions and interactors (i.e. macroecology)
(Bambach, 1977; Valentine, 1969). Species within a guild are expected to have more similar
16 macroecological dynamics than species in different guilds. Building on the framework of guilds,
(Bush et al., 2007) presented an ecocube for describing the position, motility, and trophic role of
18 marine invertebrates. Unique combinations along the three ecological trait axes represent which
among the possible ecotypes are observed. This approach has proven quite popular as it attempts
20 to operationalize the guild in terms of shared characteristics (Bambach et al., 2007; Bush and
Bambach, 2011; Bush et al., 2007), but the overall utility of this approach is limited due to its
22 condition as just a data type.

24 Analysis of mammal diversity and hypotheses as to the processes that have shaped it tend to fall
into one or more of the following categories: diversity of an entire system (e.g. continent) (Alroy,
1996; Alroy et al., 2000; Figueirido et al., 2012; Liow et al., 2008), guild based (Janis et al., 2004;
26 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
28 (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
30 simultaneously, and instead integration of these diverse observations and hypotheses tends to be
based on coincidence. One of the goals of this study is to present a framework for simultaneously
32 analyzing a diversity of hypotheses by pulling information from multiple levels of organization by
integrating both species traits and environmental factors into a single analysis in order to infer a
34 more holistic picture of the processes which may have shaped mammal species diversity.

In the analyses done here, a few key covariates which describe species' macroecology and
36 environmental context are considered. Because of the complexity inherent in this question and
related analysis in terms of both number of covariates considered and structure of each model, it is
38 possible to consider and test a large number of possible hypotheses. The analytical approach used
here is appropriate for mitigating complications arising from this complexity (e.g. multiple
40 comparisons, garden of forking paths) CITATIONS.

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

There is no previous evidence of any major turnover events in history of North American mammal
48 diversity, unlike the Neogene record European mammals (Alroy, 1996, 2009; Alroy et al., 2000;
Eronen et al., 2015; Janis, 1993). There is also little evidence for simultaneous changes in
50 cross-ecotype or cross-guild diversity. Instead, turnover is distributed through time. It is then
expected then that turnover events or periods of rapid diversification or depletion should not occur
52 simultaneously for all ecotypes.

Translating previous work into hypotheses applicable to this analysis is difficult for a variety of

54 reasons. Taxonomic groupings such as order or family are frequently invoked as an important factor
in many proposed hypotheses for how mammal diversity is structured (Janis, 2008; Janis and
56 Wilhelm, 1993; Pires et al., 2015; Quental and Marshall, 2013; Slater, 2015). Because taxonomic
grouping conflates both species macroecology or guild membership with shared evolutionary history,
58 there are no clear hypotheses as to macroecological change viewed through the lens of species
interactions. Specifically, this issue arises when trying to generalize previous observations from
60 taxonomy-based to ecology-based hypotheses.

Jernvall and Fortelius (2004) found that for the Neogene of Europe the relative abundance of
62 mammal guilds was stable over time even in the face of high turnover rates, though they only
considered large bodied taxa from a small set of mammal orders. Similar results have been observed
64 for some taxonomic groups in North America CITATIONS.

Many discussions of the effects or associates of species ecology and diversity have focused on
66 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).

68 The diversity history of ungulate herbivores is characterized by more recently originating taxa
having longer legs, higher crowned teeth, and a shift from graze-dominated to browse-dominated
70 diets than their earlier originating counterparts (Janis et al., 2004; Janis, 2008; Janis et al., 2000;
Janis and Wilhelm, 1993); all of which have all been attributed to some combination of
72 environmental change and tectonic activity driving that environmental change (Blois and Hadly,
2009; Eronen et al., 2015; Janis, 2008). Additionally, it has been observed that ungulate cursorial
74 forms arose prior to cursorial carnivore forms, an observation attributed to the reorganization of
plant communities towards the end of the Cenozoic and the latter emergence of “modern”
76 environments and communities (Janis and Wilhelm, 1993).

Within the canid guild of North America (e.g. plantigrade and digitigrade carnivores) there is
78 evidence that their diversity is self-regulating or somehow limited. Specifically, it has been proposed
that different clades of “canids” have replaced each other as dominating that macroecological role
80 in the species pool (Silvestro et al., 2015). A pattern of generally constant diversity through time is

also observed within the canid carnivore subguilds of hypercarnivore, hypocarnivore, and
82 mesocarnivores even in the face of constant species turnover is consistent with limited possibility of
increased diversity, even though there was no evidence of diversity-dependence in trait (e.g. body
84 size) evolution (Slater, 2015).

There is some uncertainty as to the effect of species body size on mammal diversity and aspects of
86 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
88 correlated this trait is with other traits such as metabolic rate and home range size CITATIONS. It
is also relatively easy to estimate the body mass of extinct species using proxy measures and
90 regression equations, as was done in this study (see below). However, body size is normally
considered in other studies without reference to other ecological descriptors of the species
92 CITATION, but see (Smits, 2015). Additionally, this high amount of correlation between life history
traits and body size limits process-based inference and hypothesis testing because the actual
94 mechanisms underlying any observed pattern are obscured.

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 constant over time or if there was a change in extinction risk at
100 the Paleogene/Neogene boundary. Each of these possible explanations for the results of Smits
(2015) have clear and testable predictions for this analysis. Specifically, 1) the extinction risk
102 arboreal taxa increased in the Neogene compared to the Paleogene, driving the average extinction
risk of arboreal mammals up and leading to the loss of arboreal taxa from the species pool, or 2) if
104 arboreal taxa have just a generally higher extinction risk than other ecotypes but have maintained
a constant diversity for the Cenozoic.

106 Fundamentally, all species respond differently to climate and environmental change (Blois and
Hadly, 2009). Macroecological patterns are the similarities across species and the emergent

¹⁰⁸ properties of how species react to a similar “stimulus.”

The effect of climate on diversity and the diversification process has been the focus of considerable
¹¹⁰ 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.,
¹¹² 2012). Both temporal and geographic scale of analysis can make a big difference in the
interpretation of results. For example when the mammal fossil record analyzed at small temporal
¹¹⁴ and geographic 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 most of
¹¹⁶ the Cenozoic there is no 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 the
¹¹⁸ correlation between diversity and climate or even reverse direction over time; instead this result
means that there is no single direction of correlation between diversity and climate (Figueirido
¹²⁰ et al., 2012).

In the case of a fluctuating correlation between diversity and climate it is hard to make the
¹²² argument of an actual causal link between the two without modeling the underlying ecological
differences between species; when this analysis is based on diversity or taxonomy alone no
¹²⁴ mechanisms are possible to infer. Taxonomy, like body size, stands in for many important species
traits to the point that mechanistic or process based inference is impossible. While emergent
¹²⁶ patterns might correspond to taxonomic grouping, this itself is an emergent phenomenon. Instead,
by framing hypotheses in terms of species traits and their environmental context, these emergent
¹²⁸ phenomenon are actually being studied as opposed to be assumed.

The Cenozoic is generally characterized by a global cooling trend and the development of polar
¹³⁰ ice-caps during the Neogene, though there are a few notable exceptions to this broad
characterization (Cramer et al., 2011; Zachos et al., 2008, 2001). The Cenozoic of North America is
¹³² additionally characterized by an environmental transition from the closed, partially forested
environments of the Paleogene to the savannah and grasslands environments of the Neogene (Blois
¹³⁴ and Hadly, 2009; Janis, 1993; Janis et al., 2000; Strömberg, 2005).

A lot of the climate and environmental changes observed for North America have been attributed to
136 tectonic activity and uplift (Badgley and Finarelli, 2013; Blois and Hadly, 2009; Eronen et al., 2015;
Janis, 2008) CITATIONS. Additionally, tectonic activity and uplift is considered the causal
138 mechanism secondarily behind both the diversification process and trait evolution (Badgley and
Finarelli, 2013; Blois and Hadly, 2009) CITATIONS. Tectonic uplift changes weather patterns (e.g.
140 rain shadow) and mobilizes grit into the environment. Increased grit in the environment combined
with decreased rain fall is considered the primary reason behind the trend of increased hypsodonty,
142 or high crowned teeth, among herbivore groups over the Cenozoic of both North America and
Europe CITATIONS.

144 The Eocene-Oligocene transition is associated with high extinction amongst ungulate taxa (Janis,
2008). This period is also the transition from the Paleogene to the Neogene and from herbivores
146 being browsing dominated to grazing dominated CITATION. This transition is marked by WHAT?
This transition is associated globally with the appearance of ice caps WHERE? Additionally, this
148 transition is marked by GRASS? 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
150 optimum (Zachos et al., 2008, 2001). The Mid-Miocene climatic optimum is bookended by periods
of temperature decline. We would then expect that, for the Miocene, turnover and other
152 diversification events would most likely be biologically mediated because of the constancy of
climate, and that for groups that are driven primarily by environmental factors, the Miocene would
154 be a period of marked by an absence of major changes to diversity or the diversification process.

The environmental factors included in this study are estimates of global temperature and the
156 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
158 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
160 for inference as to how species ecology mediates environmental context.

Fourth-corner modeling is an approach to explaining the patterns of either species abundance or

¹⁶² presence/absence as a product of species traits, environmental factors, and the interaction between
traits and environment (Brown et al., 2014; Jamil et al., 2013; Pollock et al., 2012; Warton et al.,
¹⁶⁴ 2015) CITATION. In modern ecological studies, what is being modeled is species occurrences at
localities distributed across a region (Jamil et al., 2013; Pollock et al., 2012). In this study, what is
¹⁶⁶ being modeled is the pattern of species occurrence over time for most of the Cenozoic in North
America (Fig. 1). These two approaches, modern and paleontological, are different views of the same
¹⁶⁸ three-dimensional pattern: species at localities over time. The temporal limitations of modern
ecological studies and difficulties with uneven spatial occurrences of fossils in paleontological studies
¹⁷⁰ means that these approaches are complimentary but reveal different patterns of how species are
distributed in time and space.

¹⁷² All observations, paleontological or modern, are made with uncertainty CITATIONS. With
presence/absence data this uncertainty comes from now knowing if an absence is a “true” absence
¹⁷⁴ or just a failure to observe (Foote, 2001; Foote and Sepkoski, 1999; Lloyd et al., 2011; Royle and
Dorazio, 2008; Royle et al., 2014; Wang and Marshall, 2016). For paleontological data, the
¹⁷⁶ incomplete preservation of whatever species were present into fossil form combined with incomplete
sampling of what fossils are present means that the true times of origination or extinction may not
¹⁷⁸ be observed (Foote, 2001; Foote and Sepkoski, 1999; Wang et al., 2016; Wang and Marshall, 2016).

Ultimately, the goal of this analysis are to understand when are unique ecotypes enriched or
¹⁸⁰ depleted in the North American mammal regional species pool and how changes in ecotypic
diversity are related to changes in species’ environmental context.

¹⁸² Materials and Methods

Taxon occurrences and species-level information

¹⁸⁴ All fossil occurrence information used in this analysis was downloaded from the
Paleobiology Database (PBDB). The initial download restricted all occurrences to all Mammalia
¹⁸⁶ observed in North America between the Maastrichtian and Gelasian stages. Occurrences were then

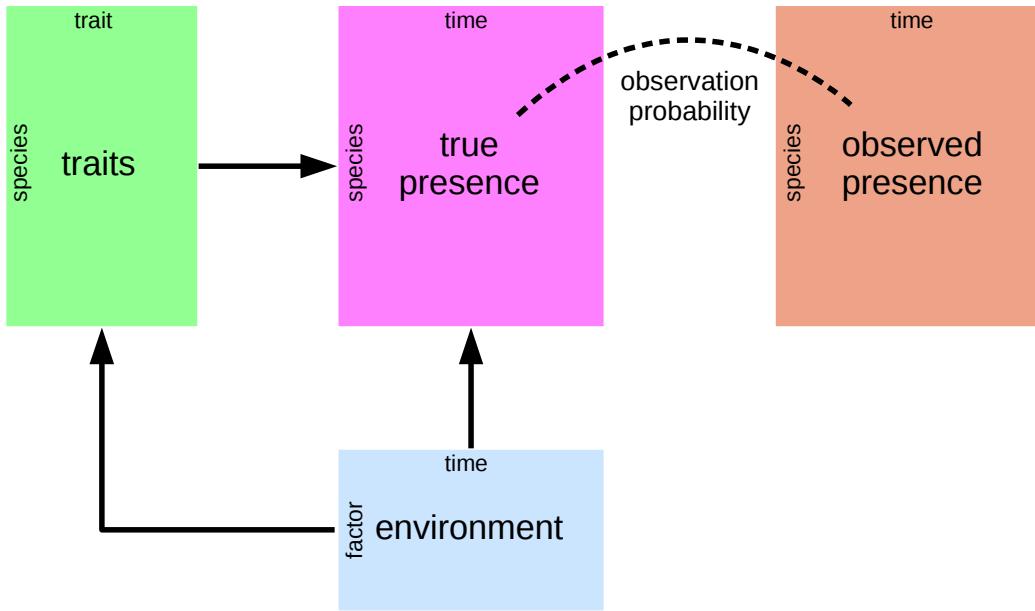


Figure 1: Conceptual diagram of the paleontological fourth corner problem. The observed presence matrix (orange) is the empirical presence/absence pattern for all species for all time points; this matrix is an incomplete observation of the “true” presence/absence pattern (purple). The estimated true 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).

further limited to those occurring between 64 and 2 million years ago (Mya). Taxonomic,
 188 stratigraphic, and ecological metadata for each occurrence was included. A new download for a raw,
 unfiltered PBDB datafile following the same criterion used here is available at
 190 <http://goo.gl/2s1geU>. The raw datafiles used as a part of this study, along with all code for
 filtering and manipulating this download is available at GITHUB.
 192 After download, the raw occurrence data was then sorted, cleaned, and manipulated
 programmatically prior to analysis. Many species taxonomic assignments as present in the raw
 194 PBDB data were updated for accuracy and consistency. For example, species classified in the order

Artiodactyla were reclassified as Cetartiodactyla. These re-assignments follow Smits (2015) which
 196 were based on taxonomies present in the Encyclopedia of Life WEBSITE and Janis et al. (2008,
 198 1998). All taxa who's life habit was classified as either volant (i.e. Chiroptera) or aquatic (e.g.
 198 Cetacea) were excluded from this analysis because of both differences in fossilization potential and
 environmental context as well as their lack of direct applicability to the study of terrestrial species
 200 pools.

Species ecotype is defined as the interaction between life habit and diet categories, the goal of
 202 which is to classify species based on the manner they interact with their environment. Most species
 records in the PBDB have life habit and dietary category assignments. In order to simplify
 204 interpretation, analysis, and per ecotype sample size these classifications were coarsened in a similar
 manner to (Smits, 2015) following Table 1. Additionally for this study, the life history category was
 206 further broken up to better reflect the many locomotor modes employed by mammals. Specifically,
 species were reassigned based on their ankle posture associated with their taxonomic group as
 208 described in Table 2 (Carrano, 1999). Ecotype categories with less than 10 species having ever been
 that combination were excluded, yielding a 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: Ankle posture assignment as based on taxonomy. Assignments are based on (Carrano, 1999). Taxonomic groups are presented alphabetically and without reference for the nestedness of families in orders. As such, do not infer higher-order structure from this table.

Order	Family	Stance
	Ailuridae	plantigrade
	Allomyidae	plantigrade
	Amphicyonidae	plantigrade
	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

Continued on next page

Table 2 – continued from previous page

Order	Family	Stance
	Dipodidae	digitigrade
	Elephantidae	digitigrade
	Entelodontidae	unguligrade
	Eomyidae	plantigrade
	Erethizontidae	plantigrade
	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

Continued on next page

Table 2 – continued from previous page

Order	Family	Stance
	Mammutidae	unguligrade
	Megalonychidae	plantigrade
	Megatheriidae	plantigrade
	Mephitidae	plantigrade
	Merycoidodontidae	digitigrade
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
	Peritychidae	digitigrade
Perissodactyla		unguligrade

Continued on next page

Table 2 – continued from previous page

Order	Family	Stance
	Phenacodontidae	unguligrade
Primates		plantigrade
	Procyonidae	plantigrade
	Proscalopidae	plantigrade
	Protoceratidae	unguligrade
	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

210

Estimates of species mass used as a covariate in this study were sources from multiple databases
212 and papers, especially those focusing on similar macroevolutionary or macrecological questions; this
is similar to what was done in Smits (2015). When species mass was not available, proxy measures
214 were used and then transformed into estimates of mass. For example, given a measurement of a
mammal tooth size, it is possible and routine to estimate its mass given some regression equation.
216 The PBDB has one or more body part measures for many species. These were used as body size

proxies for many species, as was the case in Smits (2015). Mass was log-transformed and then
 218 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;
 220 Gelman and Hill, 2007).

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	?

All fossil occurrences from 64 to 2 million years ago (Mya) were binned into 31 2 million year (My)
 222 bins. This temporal length was chosen because it is approximately the resolution of the North
 American mammal fossil record (Alroy, 1996; Alroy et al., 2000; Marcot, 2014).

224 Environmental and temporal covariates

The environmental covariates used in this study are collectively referred to as group-level covariates
 226 because they predict the response for a “group” of individual-level observations (i.e. species
 occurrences). Additionally, these covariates are defined for temporal bins and not the species
 228 themselves; as such they predict the parts of each species occurrence history. The group-level
 covariates in this study are two global temperature estimates and the Cenozoic “plant phases”
 230 defined by Graham 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
232 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
234 temperature with ice sheet volume and depth/stratification changes. The former is particularly
important to this analysis as the current polar ice-caps appeared and grew over the Cenozoic.
236 These properties make Mg/Ca based temperature estimates preferable for macroevolutionary and
macroecological studies (Ezard et al., 2016).

238 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
240 mean and range were then rescaled by subtracting the mean and then dividing by twice the
standard deviation.

242 The other major set of environmental factors included in this study are the Cenozoic plant phases
defined in Graham CITATION. Graham's plant phases are holistic descriptors of the taxonomic
244 composition of 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
246 CITATION. Graham 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
248 included as a series of "dummy variables" encoding the three phases included in this analysis. This
means that the first phase is synonymous with the intercept and phases

250 Modelling species occurrence

Two different models were used in this study: a pure-presence model and a birth-death model. Both
252 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 probabilities of a
254 species originating or surviving 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.
256 0/1), there are in fact three states in a birth-death model: never having originated, extant, and

		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).

extinct. The last of these is the absorbing state, as once a species has gone extinct it cannot
 258 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 explanations (Tables 6, and
 260 7).

Data augmentation

- 262 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
 264 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
 266 incompleteness of all observations, especially in the case of paleontological data, there are obviously
 many species which were never sampled (Royle and Dorazio, 2008; Royle et al., 2007).
- 268 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 $N^* = M - N$.
 270 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
 272 M is user defined, this approach effectively gives ψ a uniform prior over N to M (Royle and
 Dorazio, 2008). For this study, $M = \lfloor 1.35 \times N \rfloor$.
- 274 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
276 (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
278 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
280 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
282 this study, the body mass distribution is already known ($\mathcal{N}(0, 0.5)$) so the body mass of the
augmented species are generated by simple random draws from this distribution.

284 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
286 grouping indicating their unknown biology. This classification has no biological interpretation.

Observation process

288 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
290 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
292 $\phi = \pi$, 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
294 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
296 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.

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

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

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

298 Pure-presence process

For the pure-presence model there is only a single probability dealing with the presence of a species
 300 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
 302 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 where the intercept varies by ecotype.

304 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 on the
306 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
308 choice of priors for these regression coefficients.

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

314 All parameters not modeled elsewhere were given weakly informative priors (Gelman et al., 2013)
CITATION STAN MANUAL STATISTICAL RETHINKING. Weakly informative means that
316 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
318 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.** The
320 general line of thinking behind this approach is that a result of 0 or “no effect” is more preferable

to a wrong or extremely weak result.

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

³²² Birth-death process

In the birth-death version of the model, $\phi \neq \pi$ and so each of these probabilities are modeled ³²⁴ separately but in a 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 ³²⁶ the full sampling 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 ³²⁸ 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 ³³⁰ priors (Eq. 3).

The intercepts of ϕ and π both vary by species ecotype and those values are themselves the product ³³² of group-level regression using environmental factors as covariates (Eq. 3); this is identical to the

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 $z_{-,t} = 1 z_{-,t-1} = 0$; origination
π	$N \times T - 1$	probability of $z_{-,t} = 1 z_{-,t-1} = 1$; survival
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

pure presence model (Eq. 2).

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

³³⁴ **Posterior inference and model adequacy**

Programs that implement joint posterior inference for the above models (Eqs. 2, 3) were
³³⁶ 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
³³⁸ 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.
³⁴⁰ Instead of implementing a latent discrete parameterization, the posterior probabilities of all possible
states of the latent parameters z were estimated (i.e. marginalized).

³⁴² 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
³⁴⁴ 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
³⁴⁶ the species having only existed between the directly observed FAD and LAD to all possible
intermediaries CITATION STAN MANUAL. This process is identical, language-wise, to assuming
³⁴⁸ range-through and then estimating the possibility of range extension due to incomplete sampling.

The combined size of the dataset and large number of parameters in both models (Eqs. 2, 3),
³⁵⁰ specifically the total number of latent parameters that are the matrix z , means that stochastic
approximate posterior inference is computationally very slow even using HMC. Instead, an
³⁵² approximate Bayesian approach was used: variational inference. A recently developed automatic
variational inference algorithm called “automatic differentiation variational inference” (ADVI) is
³⁵⁴ implemented in Stan and was used here CITATION. ADVI assumes that the posterior is Gaussian
but still yields a true Bayesian posterior; this assumption is similar to quadratic approximation of
³⁵⁶ the likelihood function used in maximum likelihood inference CITATION. The principal limitation
of assuming the joint posterior is Gaussian is that the true topology of the log-posterior isn’t
³⁵⁸ estimated; this is a particular burden for scale parameters which are bound to be positive (e.g.
standard deviation).

³⁶⁰ After fitting both models (Eqs. 2, 3) using ADVI, model adequacy and quality of fit was assessed

	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

using a simple posterior predictive check CITATION CITATION. Because all Bayesian models are
 362 inherently generative, simulations of new data sets is “free” CITATION. By simulating many
 theoretical data sets using the observed covariate information the congruence between predictions
 364 made by the model and the observed empirical data can be assessed. By combining multiple
 posterior predictive tests of congruence between empirical and simulated values of interest, the
 366 holistic adequacy of the model can be analyzed CITATION. The ADVI assumption of a purely
 Gaussian posterior limits the utility and accuracy of the posterior predictive checks because
 368 parameter estimates do not reflect the true posterior distribution and are instead just an
 approximation. Because of this, posterior predictive estimates are themselves only approximate
 370 checks of model adequacy. The posterior predictive check that is used in this study focuses on mean
 occurrence and not to any scale parameters that might be most affected by the ADVI assumptions.
 372 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
- 374 observed value sits in the middle of the distribution than the model is adequate in reproducing the
observed number of occurrences per species.
- 376 Posterior simulations for time series are start with the values at $t = 1$ and then just simulating
forward.
- 378 Given parameter estimates, diversity and diversification rates are estimated through posterior
predictive simulations. Given the observed presence-absence matrix y , estimates of the true
380 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
382 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

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

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

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

- 386 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

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

³⁸⁸ **Results**

The posterior estimates and analysis in this study take one of two forms: direct inspection of
³⁹⁰ parameter estimates from both models, and downstream estimates of diversity and diversification
rates based on posterior predictive simulations from the birth-death model as explained below in
³⁹² the comparison of the models' posterior predictive check results.

³⁹⁴ **Comparing parameter estimates from the pure-presence and birth-death
models**

Comparison of the posterior predictive results from the pure-presence and birth-death models
³⁹⁶ reveals a striking difference in the ability for the model to predict the structure of the underlying
data (Fig. 3). The simulated datasets generated from the birth-death model are clearly able to
³⁹⁸ better reproduce the observed average number of occurrence than the pure-birth model which
greatly underestimates the observed average number of occurrences. This result means that
⁴⁰⁰ inferences based on the birth-death model are more likely to be representative of the underlying
data than inferences based on the pure-presence model. Further inspection of the posterior
⁴⁰² parameter estimates from both models can provide further insight into the reasons for this difference
in posterior predictive results (Gelman et al., 2013).

⁴⁰⁴ Occurrence probabilities estimated from the pure-presence model (Fig. 4) are broadly similar to the
estimates of origination probability from the birth-death model (Fig. 5) as opposed to the estimates
⁴⁰⁶ of survival probability (Fig. 6). This result supports the idea that changes to the North American
regional species pool is more likely due to changes to origination than extinction, a result that is
⁴⁰⁸ returned to later in the discussion of per-capita diversification, origination, and extinction rates.

For most ecotypes, both estimated occurrence probabilities from the pure-presence model (Fig. 4)
⁴¹⁰ and origination probabilities estimated from the birth-death model (Fig. 5) increase with time. This
makes sense given that, over time, all species that have at least one observed occurrence must have
⁴¹² had that occurrence by the last time point, so our certainty in a species occurring must increase

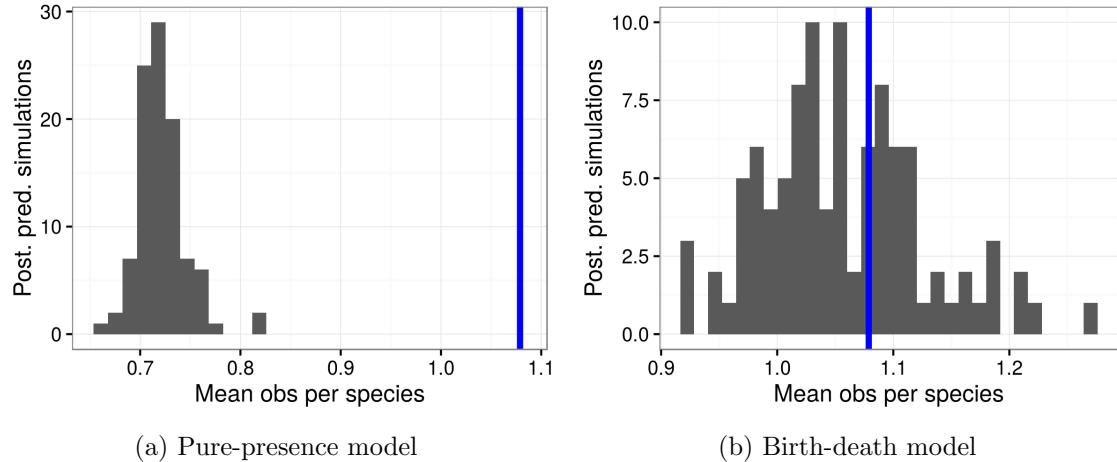


Figure 3: Comparison of the average observed number of occurrences per species (blue line) to the average number of occurrences from 100 posterior predictive datasets using the posterior estimates from the pure-presence and birth-death models.

with time. Notably, ecotypes with arboreal components do not appear to follow a similar pattern; instead, occurrence and origination probabilities appear relatively flat for most of the Cenozoic.

The dramatic differences in the estimates origination and survival probabilities are indicative of how differently these processes affect the diversification process and may also 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 at all points along both time 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. 5, and 6).

For most ecotypes, the estimates of origination probabilities are with less uncertainty than similar estimates of survival probabilities (Fig. 5, and 6). In logistic regression, high uncertainty in the estimates of the underlying log-odds of occurrence, origination, or survival tends to be indicative of extreme rarity or complete absence of the specific ecotype; the latter is called complete separation which occurs when there is no uncertainty in the effect of a covariate on presence/absence, the effect of which has been mitigated by the hierarchical modeling strategy used here (Gelman et al., 2013; Gelman and Hill, 2007) CITATION Statistical Rethinking.

428 The pure-presence and birth-death models also differ in the estimated effect of mass on the

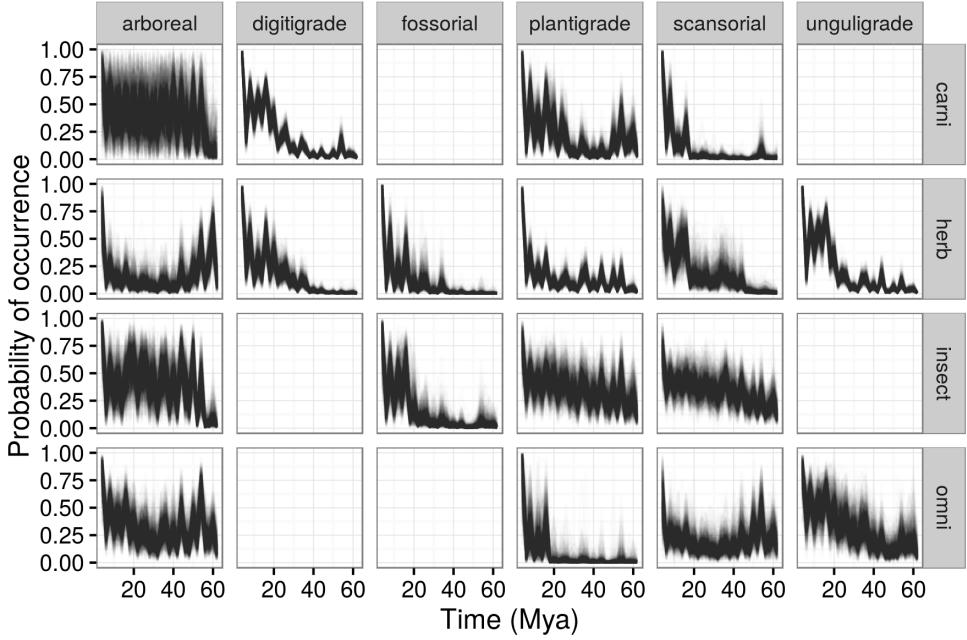


Figure 4: Probability of a mammal ecotype occurring over time as estimated from the pure-presence model. Each panel depicts 100 random samples from the model’s posterior. The columns are by locomotor category and rows by dietary category; their intersections are the observed and analyzed ecotypes. Panels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

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

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

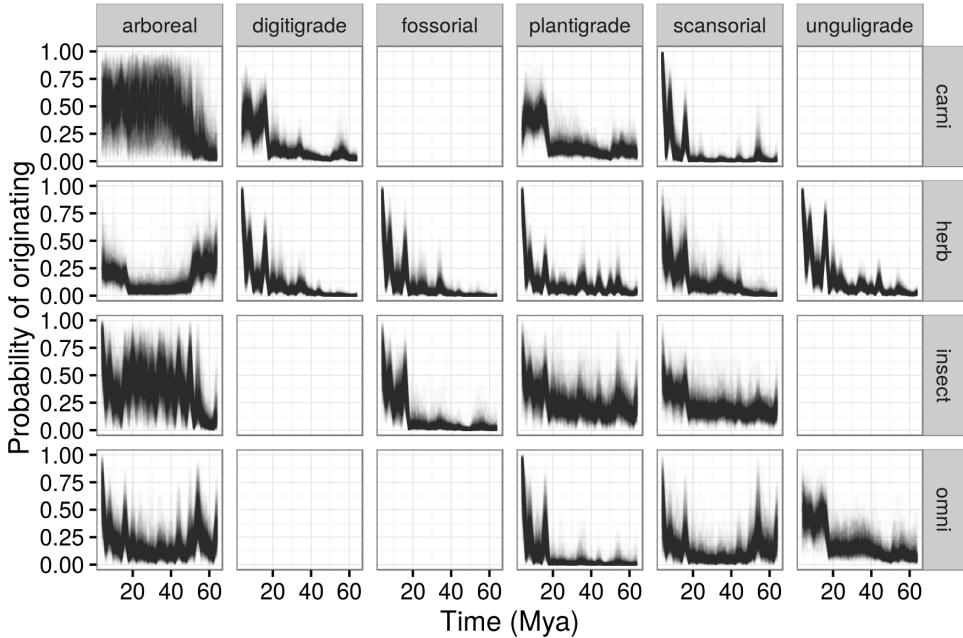


Figure 5: Probability of a mammal ecotype origination probabilities at each time point as estimated from the birth-death model. Each panel depicts 100 random samples from the model’s posterior. The columns are by locomotor category and rows by dietary category; their intersections are the observed and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

442 The average sampling probabilities for both the pure-presence model and birth-death model are
 443 both at the point where (rescaled log) mass equals 0; visual comparison indicates that, on average,
 444 sampling probability has greater posterior estimate in the pure-presence model than the birth-death
 445 model (Fig. 7). The probability that one estimate is different from the other, however, are not
 446 directly calculable as they come from different models; what this tells us is how adding more
 447 information to the model (i.e. replacing occurrence with origination and extinction) changes
 448 parameter estimates in the model.

The effect of species mass on probability of occurrence as estimated from the pure-presence (Fig. 8)
 450 are most similar to the estimated effect of species mass on probability of origination for the
 451 birth-death model (Fig. 9). The striking pattern observable in both sets of estimates is the higher
 452 probability of occurrence for species with body sizes closer to the mean than either extremes. This
 453 result is consistent with the canonically normal distribution of mammal body sizes CITATION; it is
 454 then expected that the most likely to occur species would be those from the middle of the

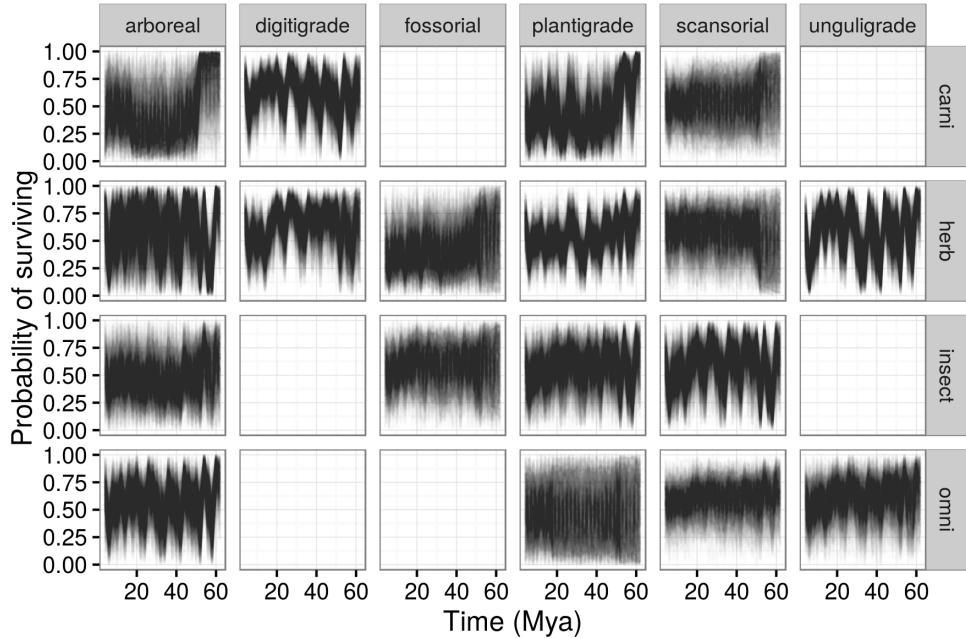


Figure 6: Probability of a mammal ecotype survival probabilities at each time point as estimated from the birth-death model. Each panel depicts 100 random samples from the model’s posterior. The columns are by locomotor category and rows by dietary category; their intersections are the observed and analyzed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

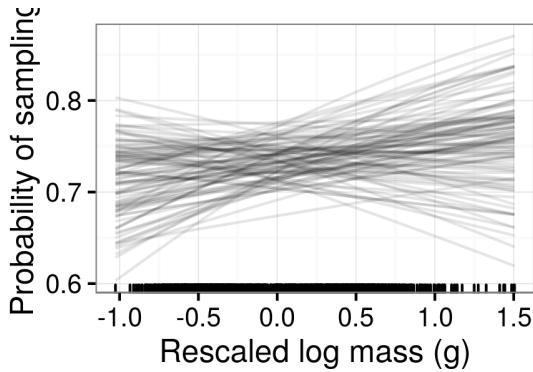
distribution, and that species originating will on average be of average mass, especially considering

456 species shared common ancestry CITATION. Note that all variation in estimates between ecotypes
 (Fig. 9) is due to differences in ecotype-specific survival probability and the associated effects of
 458 plant phase; the effect of mass was considered constant for all ecotypes.

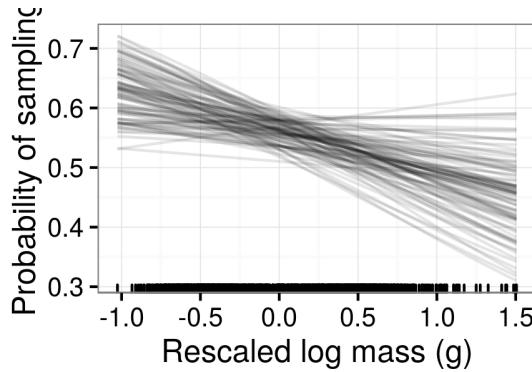
In contrast, the effect of species mass on probability of survival as estimated from the birth-death

460 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
 462 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.

464 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
 466 estimates from the pure-presence and birth-death models, the effects of the group-level covariates in



(a) Pure-presence model



(b) Birth-death model

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.

the pure-presence model (Fig. 11) are more similar to those estimates of the group-level effects on

468 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 model (Fig. 8) with the effect of mass on
470 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.

472 Which ecotypes have major differences in occurrence probabilities by plant phase? When are there
increases vs decreases? Are their similarities along ecotype axes?

474 Which ecotypes are associated with a strong effect of either of the temperature covariates? Are
these relationships positive (inc temp, inc occ) or negative (dec temp, inc occ)? Are their
476 similarities along ecotype axes?

Which ecotypes have major differences in origination probabilities by plant phase? When are there
478 increases vs decreases? Are their similarities along ecotype axes?

Which ecotypes are associated with a strong effect of either of the temperature covariates? Are
480 these relationships positive (inc temp, inc orig) or negative (dec temp, inc orig)? Are their
similarities along ecotype axes?

482 Which ecotypes have major differences in survival probabilities by plant phase? When are there

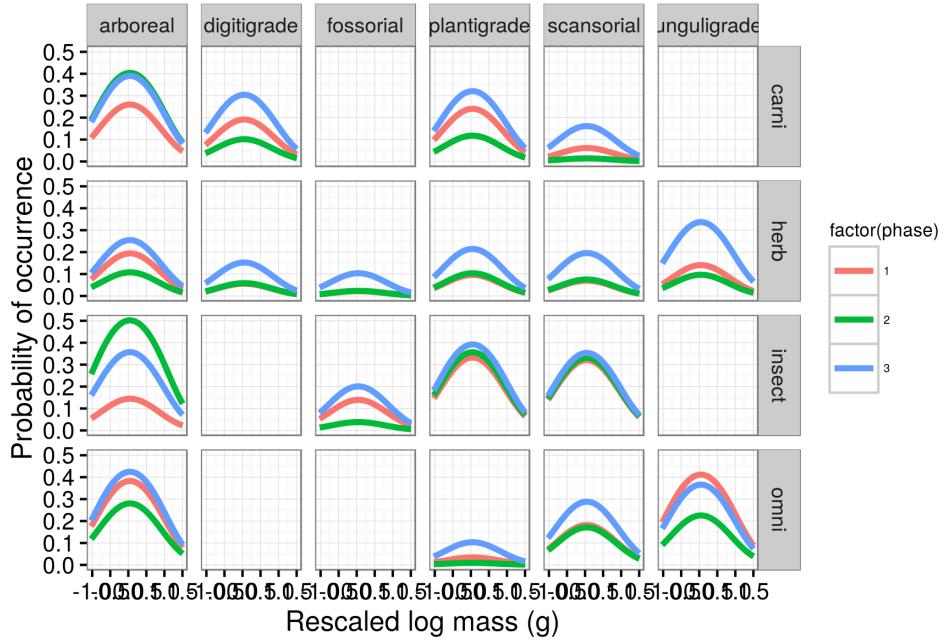


Figure 8: Mean estimate of the effect of species mass on the probability of a species occurrence for each of the three plant phases. The effect of mass is considered constant over time and that the only aspect of the model that changes with plant phase is the intercept of the relationship between mass 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.

increases vs decreases? Are their similarities along ecotype axes?

- 484 Which ecotypes are associated with a strong effect of either of the temperature covariates? Are
 these relationships positive (inc temp, inc surv) or negative (dec temp, inc surv)? Are their
 486 similarities along ecotype axes?

Analysis of diversity

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

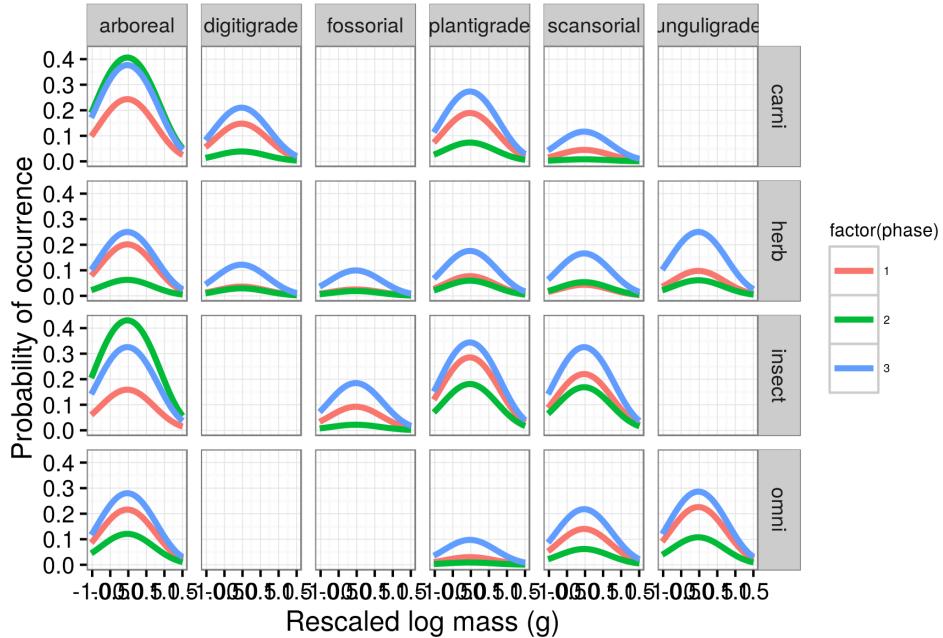


Figure 9: Mean estimate of the effect of species mass on the probability of a species originating for each of the three plant phases. The effect of mass is considered constant over time and that the only aspect of the model that changes with plant phase is the intercept of the relationship between mass 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.

The general pattern of total estimated North American mammal diversity for the Cenozoic is
 492 “stable” in that mean standing diversity does not fluctuate wildly and rapidly over the Cenozoic
 (Fig. 14a). In broad strokes, the first 15 or so million years of the Cenozoic are characterized by a
 494 gradual decline in standing diversity until approximately 45-50 million years ago (early-middle
 Eocene). Following this decline, standing diversity is broadly constant from 45 to 18 Mya (early
 496 Miocene). After this, there is a rapid spike in diversity followed by a slight decline in diversity up to
 the Modern. This characterization of the estimated diversity history is knowingly broad strokes and
 498 diversity time series is not without variation and vagaries.

When viewed through the lens of diversification rate, some of the structure behind the estimated
 500 diversity history begins to take shape (Fig. 14b). For most of the Cenozoic, the diversification rate

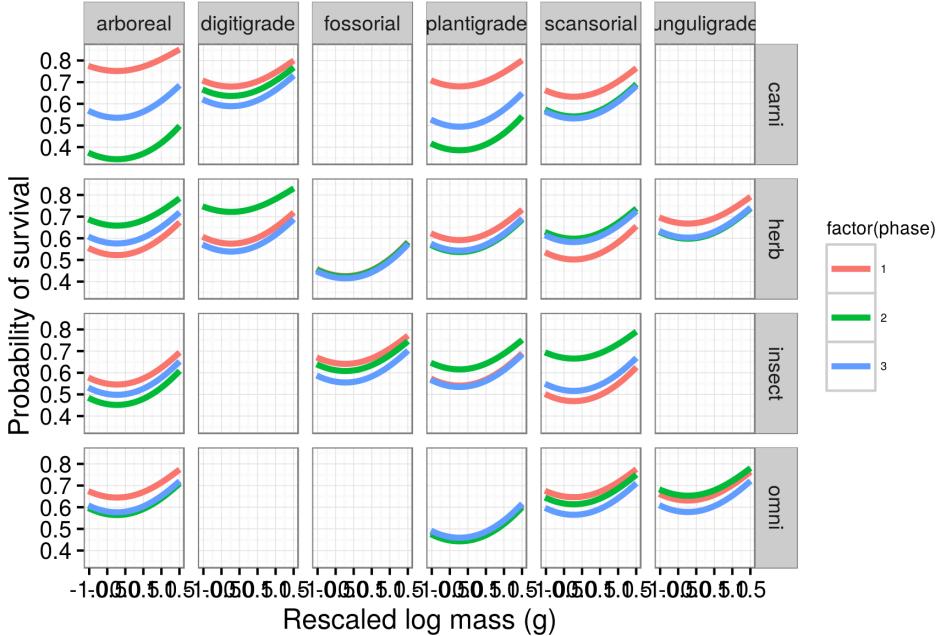


Figure 10: Mean estimate of the effect of species mass on the probability of a species survival for each of the three plant phases. The effect of mass is considered constant over time and that the only aspect of the model that changes with plant phase is the intercept of the relationship between mass 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.

hovers around zero, punctuated by both positive and negative spikes. The largest spike in

502 diversification rate is at 16 Mya, which is early Oligocene (Fig. 14b). Other notable increases in
 503 diversification rate occur at WHEN? Mya (Table 14); other possible increases in diversification rate
 504 are less certain (e.g. 8 Mya). Notable decreases in diversification rate occur at WHEN? (Table 14).

The comparison between per capita origination and extinction rate estimates reveals how

506 diversification rate is formed (Fig. 14c, 14d). As expected given previous inspection of origination
 507 and survival probabilities, diversification rate seems most driven by changes in origination rate as
 508 opposed to extinction rate. Extinction rate, on the other hand, demonstrates an almost saw-toothed
 pattern around a constant mean.

510 Now ask what origin or extinct are doing at the important time points indicated above.

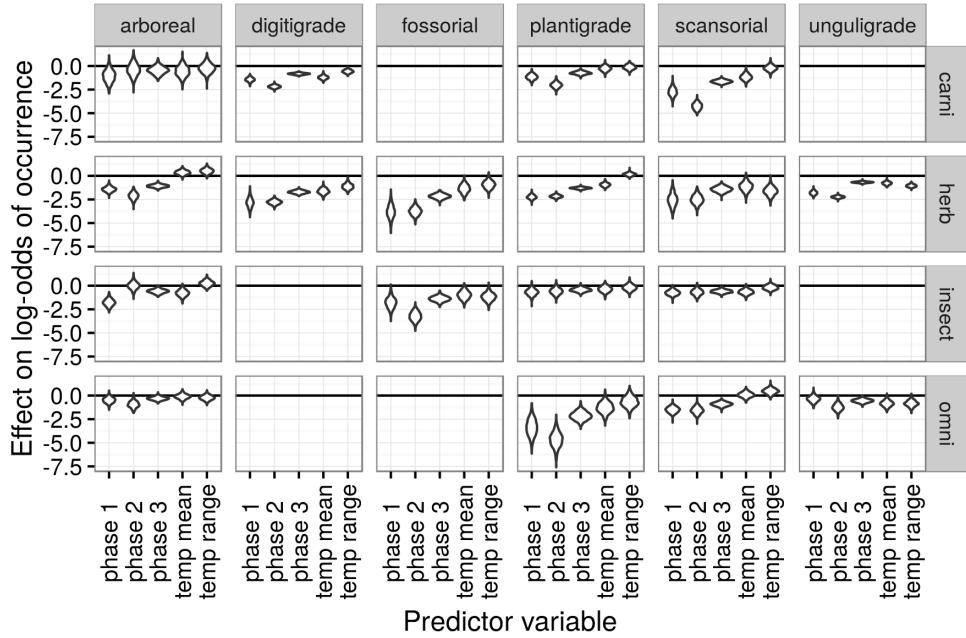


Figure 11: Estimated effects of the group-level covariates describing environmental context on log-odds of species occurrence. These estimates are from the pure-presence model.

Diversity partitioned by ecotype reveals a lot of the complexity behind the pattern of mammal

512 diversity for the Cenozoic (Fig. 15).

Arboreal ecotypes obtain peak diversity early in the Cenozoic and then decline for the rest of the
514 time series, becoming increasingly rare or absent as diversity approaches the Modern (Fig. 15).

Arboreal herbivores and omnivores obtain peak diversity at the beginning of the Cenozoic then go
516 into decline while still possibly remaining a part of the species pool, while arboreal carnivores and
insectivores obtain peak diversity 52-50 Mya and then quickly decline and become extremely rare or
518 absent from the species pool.

The diversity of both digitigrade and unguligrade herbivores increase over the Cenozoic (Fig. 15).

520 In contrast, plantigrade herbivore diversity does not have a single, broad-strokes pattern; instead,
diversity increases, decreases, and may have then increased till the Modern. Contrastingly, fossorial
522 and scansorial herbivores demonstrate a much flatter history of diversity, with a slight increase in
diversity that over time is more pronounced among fossorial taxa than scansorial taxa.

Table 8: Posterior probability estimates of differences in the log-odds of an ecotype occurring based on plant phase. These probabilities are calculated as $P(\text{Phase 1} > \text{2}) = (\gamma_{\text{phase1}} - \gamma_{\text{phase1}} + \gamma_{\text{phase2}})/100$ and similarly for the other comparisons. The divisor is 100 because there were 100 simulated datasets. These estimates are from the pure-presence model.

	P(Phase 1 > Phase 2)	P(Phase 2 > Phase 3)	P(Phase 1 > Phase 3)
arboreal carnivore	0.460	0.776	0.866
digitigrade carnivore	1.000	0.000	1.000
plantigrade carnivore	1.000	0.040	1.000
scansorial carnivore	1.000	0.001	1.000
arboreal herbivore	1.000	0.540	1.000
digitigrade herbivore	1.000	0.995	1.000
fossorial herbivore	1.000	0.920	1.000
plantigrade herbivore	1.000	0.998	1.000
scansorial herbivore	0.999	0.754	1.000
unguligrade herbivore	1.000	0.000	1.000
arboreal insectivore	0.028	1.000	0.999
fossorial insectivore	1.000	0.161	1.000
plantigrade insectivore	0.706	0.774	0.985
scansorial insectivore	0.630	0.937	1.000
arboreal omnivore	0.981	0.165	0.944
plantigrade omnivore	1.000	0.325	1.000
scansorial omnivore	0.987	0.746	1.000
unguligrade omnivore	0.990	0.344	0.997

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

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

Table 9: Posterior probability of the effects of the two temperature covariates on the log-odds of an ecotype occurring. What is estimated is the probability that these estimates are greater than 0; high or low probabilities indicate the “strength” of the covariate in that direction (positive and negative, respectively). These estimates are from the pure-presence model.

	$P(\gamma_{temp\ mean} > 0)$	$P(\gamma_{temp\ range} > 0)$
arboreal carnivore	0.169	0.317
digitigrade carnivore	0.000	0.000
plantigrade carnivore	0.168	0.304
scansorial carnivore	0.000	0.206
arboreal herbivore	0.943	0.969
digitigrade herbivore	0.000	0.000
fossorial herbivore	0.001	0.022
plantigrade herbivore	0.000	0.832
scansorial herbivore	0.009	0.003
unguligrade herbivore	0.000	0.000
arboreal insectivore	0.006	0.783
fossorial insectivore	0.016	0.003
plantigrade insectivore	0.127	0.260
scansorial insectivore	0.009	0.238
arboreal omnivore	0.337	0.191
plantigrade omnivore	0.012	0.120
scansorial omnivore	0.597	0.935
unguligrade omnivore	0.002	0.002

536 scansorial omnivores.

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

Discussion

542 Both the composition of a species pool and its environmental context change over time, though not
 necessarily at the same rate. Local communities, who’s species are drawn from the regional species
 544 pool, have “roles” in their communities defined by their interactions with a host of biotic and abiotic
 interactors (i.e. species niche). For higher level ecological characterizations like ecotypes and guilds,

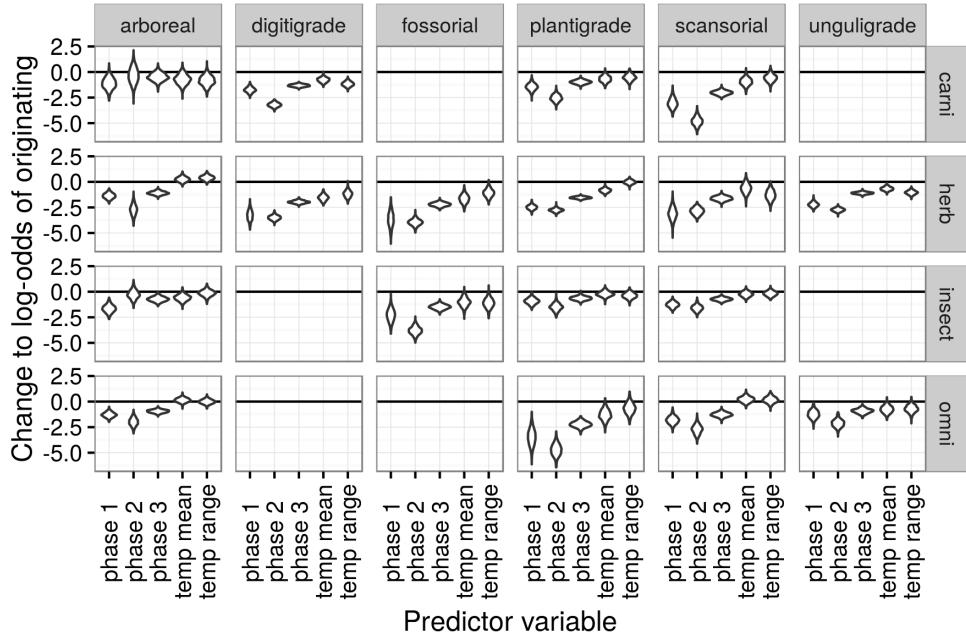


Figure 12: Estimated effects of the group-level covariates describing environmental context on log-odds of species origination. These estimates are from the birth-death model.

- 546 these roles are broadly defined and not defined by specific interactions but the genre of interactions
that species within that grouping participate in. The diversity of species within an ecotype or guild
548 can be stable over millions of years despite constant species turnover (Jernvall and Fortelius, 2004;
Slater, 2015) CITATIONS. This implies that the size and scope of the role of an ecotype or guild is
550 preserved even as the individual interactors change. This also implies the structure of regional
species pools can be constant over time despite a constantly changing set of “players.”
- 552 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 occurrence probability and is
554 instead better modeled as the result of both origination and extinction. Additionally, changes to
ecotypic composition of the North American regional species pool are driven primarily by variation
556 in origination rates. These aspects of how regional species pool diversity is shaped is not observable
from studies of the Modern CITATION.
- 558 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

Table 10: Posterior probability estimates of differences in the log-odds of an ecotype originating based on plant phase. These probabilities are calculated as $P(\text{Phase 1} > 2) = (\gamma_{\text{phase1}} - \gamma_{\text{phase1}} + \gamma_{\text{phase2}})/100$ and similarly for the other comparisons. The divisor is 100 because there were 100 simulated datasets. These estimates are from the birth-death model.

	P(Phase 1 > Phase 2)	P(Phase 2 > Phase 3)	P(Phase 1 > Phase 3)
arboreal carnivore	0.460	0.776	0.866
digitigrade carnivore	1.000	0.000	1.000
plantigrade carnivore	1.000	0.040	1.000
scansorial carnivore	1.000	0.001	1.000
arboreal herbivore	1.000	0.540	1.000
digitigrade herbivore	1.000	0.995	1.000
fossorial herbivore	1.000	0.920	1.000
plantigrade herbivore	1.000	0.998	1.000
scansorial herbivore	0.999	0.754	1.000
unguligrade herbivore	1.000	0.000	1.000
arboreal insectivore	0.028	1.000	0.999
fossorial insectivore	1.000	0.161	1.000
plantigrade insectivore	0.706	0.774	0.985
scansorial insectivore	0.630	0.937	1.000
arboreal omnivore	0.981	0.165	0.944
plantigrade omnivore	1.000	0.325	1.000
scansorial omnivore	0.987	0.746	1.000
unguligrade omnivore	0.990	0.344	0.997

560 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
562 with previous observations (Alroy, 1996; Alroy et al., 2000), however, is that diversity seems to be
most structured by origination than extinction.

564 Arboreal taxa disappear over the Cenozoic, with massive disappearance by the Paleogene-Neogene
transition \sim 22 Mya. This is consistent with one of the two possible patterns that would result in
566 arboreal taxa having a greater extinction risk than other ecotypes: Paleogene-Neogene are different
and while the earliest Cenozoic may have been neutral wrt arboreal taxa, they disappeared quickly
568 over the Cenozoic which may account for their higher extinction risk.

Digitigrade carnivores have a relatively stable diversity history through the Cenozoic and could be
570 characterized as varying around a constant mean diversity. The ecotypie has a large amount of
overlap with the carnivore guild which has been the focus of much research CITATIONS. This

Table 11: Posterior probability of the effects of the two temperature covariates on the log-odds of an ecotype origination. What is estimated is the probability that these estimates are greater than 0; high or low probabilities indicate the “strength” of the covariate in that direction (positive and negative, respectively). These estimates are from the birth-death model.

	$P(\gamma_{temp\ mean} > 0)$	$P(\gamma_{temp\ range} > 0)$
arboreal carnivore	0.086	0.045
digitigrade carnivore	0.001	0.000
plantigrade carnivore	0.013	0.054
scansorial carnivore	0.007	0.062
arboreal herbivore	0.853	0.957
digitigrade herbivore	0.000	0.001
fossorial herbivore	0.000	0.002
plantigrade herbivore	0.000	0.428
scansorial herbivore	0.106	0.003
unguligrade herbivore	0.000	0.000
arboreal insectivore	0.028	0.314
fossorial insectivore	0.010	0.006
plantigrade insectivore	0.188	0.090
scansorial insectivore	0.182	0.224
arboreal omnivore	0.749	0.482
plantigrade omnivore	0.007	0.117
scansorial omnivore	0.765	0.699
unguligrade omnivore	0.016	0.023

572 result is consistent with some form of “control” on the ecotype, such as environmental stability, diversity-dependence, or similar Silvestro et al. (2015); Slater (2015).

574 Both digitigrade and unguligrade herbivores increase in diversity over the Cenozoic. The increase of
 575 these cursorial forms is consistent with the gradual opening up of the North American landscape
 576 CITATION. Only these herbivore from increase in diversity over the Cenozoic which may indicate a
 577 long shift in the interactors associated with those ecotypes leading to increased contribution to the
 578 regional species pool. This result may be comparable to the increasing percentage of hypodont
 579 (high-crowned teeth) mammals in the Neogene of Europe being due to an enrichment of hypodont
 580 taxa and not a depletion of non-hypodont taxa.

What these results support is a gradual change to the ecotypic diversity of the regional species pool
 582 for the Cenozoic. The rapidity of Cenozoic environmental change is worth discussing. If change is
 583 rapid, ecotypic composition of species pool does not seem to track environmental change. If change

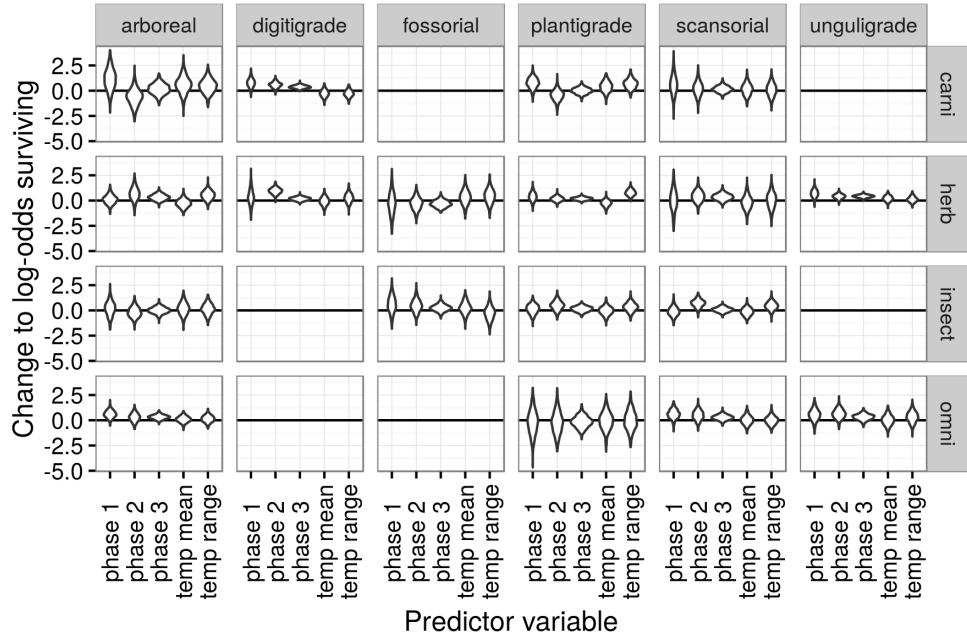


Figure 13: Estimated effects of the group-level covariates describing environmental context on log-odds of species survival. These estimates are from the birth-death model.

- 584 is gradual then there is the possibility that changes to ecotypic composition may be tracking
environmental change.
- 586 If plant phase is associated with differences in ecotype occurrence this is interpreted to mean that
ecotype enrichment or depletion is due to associations between that ecotype and whatever plants
588 are dominate at that time.

Temperature affects very few of the occurrence, origination, or survival probabilities of the mammal
590 ecotypes except for a negative relationship between temperature and the origination probabilities of
digitigrade carnivores, and both digitigrade and unguilgrade herbivores. The origination
592 probabilities and diversity of these three groups all increase over the Cenozoic as average global
temperature decreased. This result coupled with the lack of relationship between temperature and
594 the other ecotypes may be responsible for the continued confusion surrounding the impact of
temperature on mammal diversity and diversification (Alroy, 1996; Alroy et al., 2000; Blois and
596 Hadly, 2009; Figueirido et al., 2012; Janis and Wilhelm, 1993).

Table 12: Posterior probability estimates of differences in the log-odds of an ecotype surviving based on plant phase. These probabilities are calculated as $P(\text{Phase 1} > \text{2}) = (\gamma_{\text{phase1}} - \gamma_{\text{phase1}} + \gamma_{\text{phase2}})/100$ and similarly for the other comparisons. The divisor is 100 because there were 100 simulated datasets. These estimates are from the birth-death model.

	P(Phase 1 > Phase 2)	P(Phase 2 > Phase 3)	P(Phase 1 > Phase 3)
arboreal carnivore	0.904	0.121	0.382
digitigrade carnivore	0.181	0.248	0.004
plantigrade carnivore	0.857	0.195	0.519
scansorial carnivore	0.477	0.438	0.310
arboreal herbivore	0.278	0.510	0.140
digitigrade herbivore	0.001	0.978	0.175
fossorial herbivore	0.480	0.723	0.816
plantigrade herbivore	0.558	0.192	0.111
scansorial herbivore	0.444	0.286	0.133
unguligrade herbivore	0.548	0.022	0.002
arboreal insectivore	0.691	0.359	0.492
fossorial insectivore	0.334	0.488	0.221
plantigrade insectivore	0.189	0.677	0.308
scansorial insectivore	0.017	0.918	0.375
arboreal omnivore	0.549	0.196	0.074
plantigrade omnivore	0.528	0.537	0.618
scansorial omnivore	0.326	0.442	0.125
unguligrade omnivore	0.191	0.487	0.145

What is the comparative size of the effects of plant phase and temperature? Both seem of “equal”

598 importance in the sense that they have similar effect sizes on the ecotypes. Perhaps focusing on
 599 temperature and not considering other measures of environmental context has been a mistake and
 600 perhaps led to increasing confusion in discussions of how “environment” effects mammal diversity
 601 and diversification. The environment or climate is not just global or regional temperature, it is the
 602 set of all possible biotic and abiotic interactions. By including more descriptors of species’
 603 environmental context a more complete “picture” of the diversification process is inferred.

604 The effect of species mass on either occurrence or origination and extinction was not allowed to
 605 vary by ecotype or environmental context even though it is not known if this is the case or not
 606 CITATION. The primary reason for this modeling choice was that this study focuses on ecotypic
 607 based differences in either occurrence, or origination and extinction. Allowing the effect of body size
 608 to vary by ecotype, time, and environmental factors would increase the overall complexity of the

Table 13: Posterior probability of the effects of the two temperature covariates on the log-odds of an ecotype survival. What is estimated is the probability that these estimates are greater than 0; high or low probabilities indicate the “strength” of the covariate in that direction (positive and negative, respectively). These estimates are from the birth-death model.

	$P(\gamma_{temp\ mean} > 0)$	$P(\gamma_{temp\ range} > 0)$
arboreal carnivore	0.777	0.745
digitigrade carnivore	0.236	0.211
plantigrade carnivore	0.763	0.929
scansorial carnivore	0.596	0.554
arboreal herbivore	0.261	0.878
digitigrade herbivore	0.438	0.720
fossorial herbivore	0.676	0.731
plantigrade herbivore	0.215	0.997
scansorial herbivore	0.377	0.535
unguligrade herbivore	0.768	0.655
arboreal insectivore	0.614	0.610
fossorial insectivore	0.673	0.337
plantigrade insectivore	0.470	0.787
scansorial insectivore	0.364	0.879
arboreal omnivore	0.620	0.645
plantigrade omnivore	0.476	0.484
scansorial omnivore	0.514	0.494
unguligrade omnivore	0.513	0.729

model, something that I felt was not necessary because the overall scope of the study. Instead, body

size was included in order to control for its possible underlying effects CITATION. A control means
 610 that if there is variation due to body mass, having a term to “absorb” that effect is better than
 ignoring it which may affect other parameter estimates. Additionally, the effect of body size was
 612 allowed to have a second-order polynomial form and no higher order polynomials were considered;
 614 this was done because it is hard to conceive of a more complex third- or higher-order relationship
 between body size and the other parameters. Additionally, nonlinearity is rarely if ever considered
 616 in the first place, so the simple act of estimating a potential second-order relationship is an
 opportunity to test more complex hypotheses of the effects of body size on macroevolutionary and
 618 macroecological processes.

The only covariate allowed to affect sampling probability was mass and only as a linear predictor.

620 Other covariates, such as the environmental factors considered here, could have affected the

underlying preservation process that limits sampling probability. Their exclusion as covariates of
622 sampling/observation was the product of a few key decisions: model complexity, model
interpretability, the scope of this study, and a lack of good hypotheses related to these covariates to
624 warrant their inclusion. It should be noted that in other similar studies that use a hidden
birth-death model to handle simultaneous estimation of sampling, origination, and extinction have
626 not considered the possible effects of covariates, both species traits and environmental factors, on
sampling CITATION.

628 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
630 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
632 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.

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

644 Phylogenetic comparative community ecology and phylogenetic comparative biogeography also
discusses how the macroevolutionary processes helps structure an observed community, though it is
646 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,

Acknowledgements

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

References

- 656 Allen, L. J. S. 2011. An introduction to stochastic processes with applications to biology. 2nd ed.
 Chapman and Hall/CRC, Boca Raton, FL.
- 658 Alroy, J. 1996. Constant extinction, constrained diversification, and uncoordinated stasis in North
 American mammals. *Palaeogeography, Palaeoclimatology, Palaeoecology* 127:285–311.
- 660 ———. 2009. Speciation and extinction in the fossil record of North American mammals. Pages
 302–323 *in* R. K. Butlin, J. R. Bridle, and D. Schlüter, eds. *Speciation and patterns of diversity*.
 662 Cambridge University Press, Cambridge.
- Alroy, J., P. L. Koch, and J. C. Zachos. 2000. Global climate change and North American
 664 mammalian evolution. *Paleobiology* 26:259–288.
- Badgley, C., and J. A. Finarelli. 2013. Diversity dynamics of mammals in relation to tectonic and
 666 climatic history: comparison of three Neogene records from North America. *Paleobiology*
 39:373–399.
- 668 Bambach, R. K. 1977. Species richness in marine benthic habitats through the Phanerozoic.
 Paleobiology 3:152–167.

- 670 Bambach, R. K., A. M. Bush, and D. H. Erwin. 2007. Autecology and the filling of ecospace: Key
metazoan radiations. *Palaeontology* 50:1–22.
- 672 Blois, J. L., and E. A. Hadly. 2009. Mammalian Response to Cenozoic Climatic Change. *Annual
Review of Earth and Planetary Sciences* 37:181–208.
- 674 Brown, A. M., D. I. Warton, N. R. Andrew, M. Binns, G. Cassis, and H. Gibb. 2014. The
fourth-corner solution - using predictive models to understand how species traits interact with
676 the environment. *Methods in Ecology and Evolution* 5:344–352.
- Bush, A. M., and R. K. Bambach. 2011. Paleoeologic Megatrends in Marine Metazoa, vol. 39.
- 678 Bush, A. M., R. K. Bambach, and G. M. Daley. 2007. Changes in theoretical ecospace utilization in
marine fossil assemblages between the mid-Paleozoic and late Cenozoic. *Paleobiology* 33:76–97.
- 680 Carrano, M. T. 1999. What, if anything, is a cursor? Categories versus continua for determining
locomotor habit in mammals and dinosaurs. *Journal of Zoology* 247:29–42.
- 682 Clyde, W. C., and P. D. Gingerich. 1998. Mammalian community response to the latest Paleocene
thermal maximum: an isotaphonomic study in the northern Bighorn Basin, Wyoming. *Geology*
684 26:1011–1014.
- Cramer, B. S., K. Miller, P. Barrett, and J. Wright. 2011. Late Cretaceous-Neogene trends in deep
686 ocean temperature and continental ice volume: Reconciling records of benthic foraminiferal
geochemistry ($\delta^{18}\text{O}$ and Mg/Ca) with sea level history. *Journal of Geophysical Research: Oceans*
688 116:1–23.
- Eronen, J. T., C. M. Janis, C. P. Chamberlain, and A. Mulch. 2015. Mountain uplift explains
690 differences in Palaeogene patterns of mammalian evolution and extinction between North
America and Europe. *Proceedings of the Royal Society B: Biological Sciences* 282:20150136.
- 692 Ezard, T. H. G., A. Purvis, and H. Morlon. 2016. Environmental changes define ecological limits to
species richness and reveal the mode of macroevolutionary competition. *Ecology Letters*
694 19:899–906.

- Figueirido, B., C. M. Janis, J. A. Pérez-Claros, M. De Renzi, and P. Palmqvist. 2012. Cenozoic
696 climate change influences mammalian evolutionary dynamics. *Proceedings of the National
Academy of Sciences* 109:722–727.
- 698 Foote, M. 2001. Inferring temporal patterns of preservation, origination, and extinction from
taxonomic survivorship analysis. *Paleobiology* 27:602–630.
- 700 Foote, M., and J. J. Sepkoski. 1999. Absolute measures of the completeness of the fossil record.
Nature 398:415–7.
- 702 Gelman, A. 2008. Scaling regression inputs by dividing by two standard deviations. *Statistics in
Medicine* pages 2865–2873.
- 704 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.
- 706 Gelman, A., and J. Hill. 2007. Data Analysis using Regression and Multilevel/Hierarchical Models.
Cambridge University Press, New York, NY.
- 708 Harnik, P. G., P. C. Fitzgerald, J. L. Payne, and S. J. Carlson. 2014. Phylogenetic signal in
extinction selectivity in Devonian terebratulide brachiopods. *Paleobiology* 40:675–692.
- 710 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.
- 712 Janis, C., J. Damuth, and J. M. Theodor. 2004. The species richness of Miocene browsers, and
714 implications for habitat type and primary productivity in the North American grassland biome.
Palaeogeography, Palaeoclimatology, Palaeoecology 207:371–398.
- 716 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.
- 718 ———. 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.

- 720 Janis, C. M., J. Damuth, and J. M. Theodor. 2000. Miocene ungulates and terrestrial primary
productivity: where have all the browsers gone? *Proceedings of the National Academy of Sciences*
722 97:7899–904.
- 724 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
Press, Cambridge.
- 726 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
728 University Press, Cambridge.
- 730 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.
- 732 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.
- 734 Legendre, S. 1986. Analysis of mammalian communities from the Late Eocene and Oligocene of
Southern France. *Paleovertebrata* 16:191–212.
- 736 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
Academy of Sciences* 105:6097–6102.
- 738 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.
- 740 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.
- 742 Marcot, J. D. 2014. The fossil record and macroevolutionary history of North American ungulate
ungulate mammals: standardizing variation in intensity and geography of sampling. *Paleobiology*
744 40:237–254.

- Pires, M. M., D. Silvestro, and T. B. Quental. 2015. Continental faunal exchange and the
746 asymmetrical radiation of carnivores. *Proceedings of the Royal Society B: Biological Sciences*
282:20151952.
- 748 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.
- 750 Quental, T. B., and C. R. Marshall. 2013. How the Red Queen Drives Terrestrial Mammals to
Extinction. *Science* 341:290–292.
- 752 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.
- 754 ———. 2012. Parameter-expanded data augmentation for Bayesian analysis of capture-recapture
models. *Journal of Ornithology* 152:521–537.
- 756 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.
- 758 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.
- 760 Rubin, D. B. 1996. Multiple imputation after 18+ years. *Journal of the American Statistical
Assocaition* 91:473–489.
- 762 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
764 the United States of America* 112:8684–9.
- Silvestro, D., J. Schnitzler, L. H. Liow, A. Antonelli, and N. Salamin. 2014. Bayesian estimation of
766 speciation and extinction from incomplete fossil occurrence data. *Systematic biology* 63:349–67.
- Slater, G. J. 2015. Iterative adaptive radiations of fossil canids show no evidence for
768 diversity-dependent trait evolution. *Proceedings of the National Academy of Sciences*
112:4897–4902.

- 770 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.
- 772 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
774 the United States of America 102:11980–4.
- Tomiya, S. 2013. Body Size and Extinction Risk in Terrestrial Mammals Above the Species Level.
776 The American Naturalist 182:196–214.
- Valentine, J. W. 1969. Patterns of taxonomic and ecological structure of the shelf benthos during
778 Phanerozoic time. Paleontology 12:684–709.
- Wang, S. C., P. J. Everson, H. J. Zhou, D. Park, and D. J. Chudzicki. 2016. Adaptive credible
780 intervals on stratigraphic ranges when recovery potential is unknown. Paleobiology 42:240–256.
- Wang, S. C., and C. R. Marshall. 2016. Estimating times of extinction in the fossil record. Biology
782 Letters 12:20150989.
- Warton, D. I., B. Shipley, and T. Hastie. 2015. CATS regression - a model-based approach to
784 studying trait-based community assembly. Methods in Ecology and Evolution 6:389–398.
- Zachos, J. C., G. R. Dickens, and R. E. Zeebe. 2008. An early Cenozoic perspective on greenhouse
786 warming and carbon-cycle dynamics. Nature 451:279–283.
- Zachos, J. C., M. Pagani, L. Sloan, E. Thomas, and K. Billups. 2001. Trends, rhythms, and
788 aberrations in global climate 65 Ma to present. Science 292:686–693.

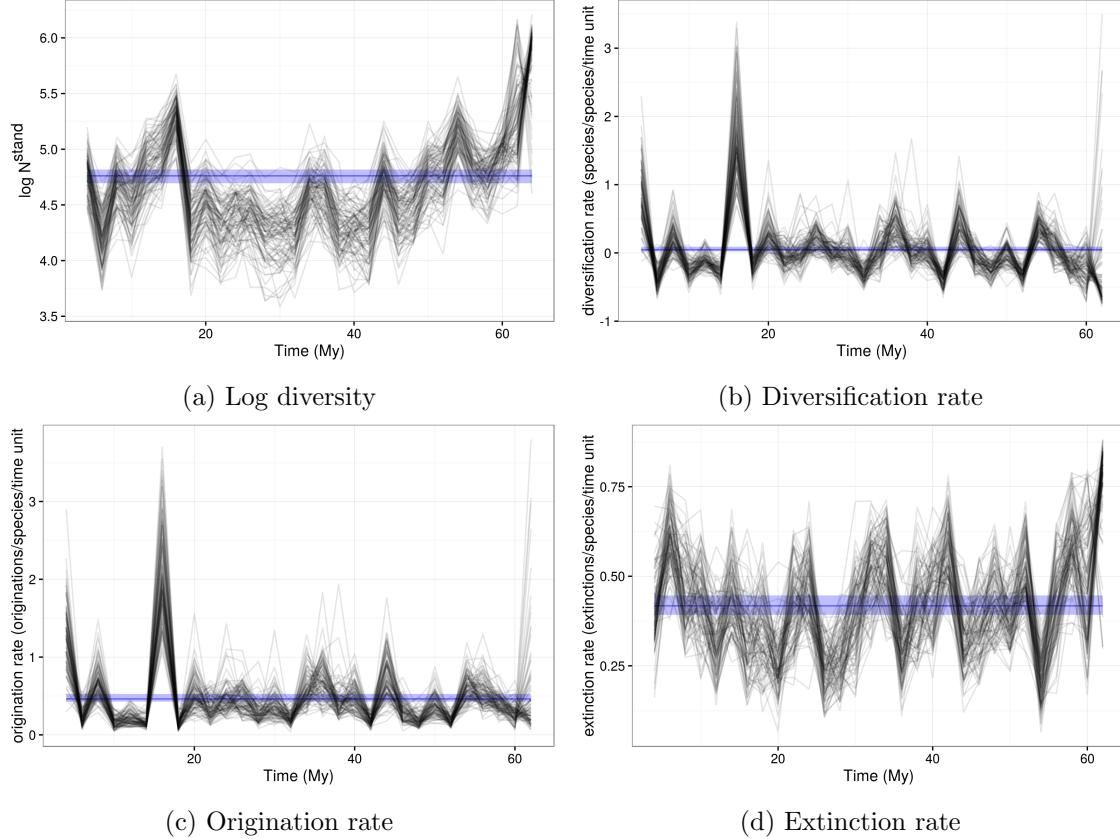


Figure 14: Posterior estimates of the time series of Cenozoic North American mammal diversity and its characteristic macroevolutionary rates; all estimates are from the birth-death model and 100 posterior draws are plotted to indicate the uncertainty in these estimates. The blue horizontal strip corresponds to the 80% credible interval of estimated mean standing diversity, diversification rate, origination rate, and extinction rate respectively; the median estimate is also indicated. What is also plotted is the 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.

Table 14: Posterior probabilities of diversity N_t^{stand} or diversification rate D_t^{rate} being greater than average standing diversity \bar{N}^{stand} or average diversification rate \bar{D}^{rate} for the whole Cenozoic. The “Time” column corresponds to the top of each of the temporal bins. Diversification rate can not be estimated for the last time point because it is unknown how many more species originated or went extinct following this temporal bin. The estimates are from the birth-death model.

Time (Mya)	$P(N_t^{stand} > \bar{N}^{stand})$	$P(D_t^{rate} > \bar{D}^{rate})$
64.00	0.99	0.18
62.00	0.93	0.15
60.00	0.93	0.04
58.00	0.53	0.59
56.00	0.72	0.99
54.00	0.99	0.00
52.00	0.59	0.45
50.00	0.57	0.01
48.00	0.05	0.27
46.00	0.04	0.92
44.00	0.53	0.00
42.00	0.01	0.44
40.00	0.00	0.37
38.00	0.01	0.94
36.00	0.23	0.46
34.00	0.22	0.01
32.00	0.00	0.31
30.00	0.00	0.33
28.00	0.00	0.83
26.00	0.03	0.32
24.00	0.02	0.25
22.00	0.01	0.89
20.00	0.15	0.02
18.00	0.02	1.00
16.00	1.00	0.00
14.00	0.83	0.11
12.00	0.67	0.01
10.00	0.11	0.79
8.00	0.40	0.02
6.00	0.00	0.98
4.00	0.59	

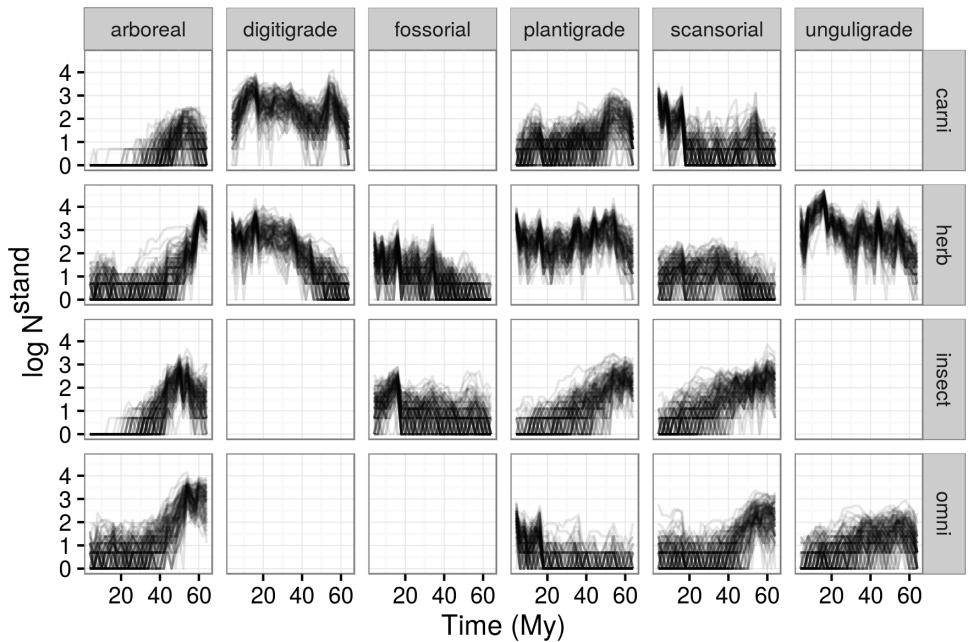


Figure 15: Posterior of standing log-diversity of North American mammals by ecotype for the Cenozoic as estimated from the birth-death model; 100 posterior draws are plotted to indicate the uncertainty in these estimates and what is technically plotted is log of diversity plus 1.