

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: macroecology, macroevolution, paleobiology, species selection, species pool, community assembly

Manuscript type: Article

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

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

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

4 **Abstract**

The set of species in a region changes over time as new species enter through speciation or
6 immigration, and species leave the system through extinction. Changes in to the functional
composition of regional species pool represents average change across all local communities
8 drawn from that species pool. While a species being present in a local community is due to the
availability of the necessary a biotic-biotic or biotic-abiotic interactions for that species to
10 co-exist, a species being present in a regional species pool just requires the possibility that is at
least one local community that has that set of necessary interactions. How a regional species
12 pool changes over time is the product of many processes acting across multiple levels of
organization. Here, I analyze the diversity history of North American mammals ecotypes for
14 most of the Cenozoic (the last 65 million years). This analysis frames mammal diversity in terms
of both there means of interacting with both the biotic and abiotic environment (i.e. ecotype) as
16 well as their regional and global environmental context, such as changes to the major groups of
plants in North America over time and global temperature. The goal of this analysis is to
18 understand when, and possibly for what reasons, are mammal ecotypes enriched or depleted
relative to their average diversity. Using two hierarchical Bayesian hidden Markov models of
20 diversity, I find that changes to mammal diversity is driven more by the influx of new species
than selective extinction. I also find that the only ecotypes which experience near constant
22 increase in diversity over time are digitigrade and unguligrade herbivores, while arboreal species
become increasingly rare and in many cases disappear from the species pool over the Cenozoic.
24 Additionally, I find that global temperature only affects the origination of some mammal
ecotypes but most likely does not affect ecotype extinction differences. The clear and direct
26 translation of research question to statistical model allows for precise and better contextualized
results. By taking into account more of the complexity surrounding and contributing to species
28 diversity and the diversification process, the idiosyncrasies of ecotype diversification histories are
more clearly contextualized than when only diversity is analyzed.

30 Introduction

A regional species pool is the set of species which form communities in a specific region. Local scale
32 processes like resource competition only affect the regional species pool if all communities are affected. The taxonomic and functional composition of a regional species pool changes over time
34 due to speciation, migration, extinction. How do species pools change over time as species are recruited or go extinct? When are specific species ecologies enriched or depleted in the species pool?
36 How does global and regional environmental context affect the set of species ecotypes (e.g. guilds) in a regional species pool? All of these questions fall under a single umbrella of analysis of ecotypic
38 diversity and diversification.

Functional diversity is frequently broken into or thought of as a set of guilds, which are a set of
40 species with similar sets of interactions and interactors (i.e. macroecology) (Bambach, 1977; Brown and Maurer, 1989; Simberloff and Dayan, 1991; Valentine, 1969; Wilson, 1999). Species within a
42 guild are expected to have more similar macroecological dynamics than species in different guilds. Building on the concept of guilds and a macroecological niche, Bush et al. (2007) presented a
44 three-dimensional construct, or ecocube, for describing the macroecological role of a marine invertebrate species by their physical position (i.e. tiering), motility, and trophic role. Unique
46 combinations along the three ecological trait axes indicate which among the possible ecotypes are observed. This approach has proven quite popular as it attempts to operationalize the guild concept
48 in terms of shared characteristics that are indicative of the type of interactions experience by species of that macroecology (Bambach et al., 2007; Bush and Bambach, 2011; Bush et al., 2007;
50 Bush and Novack-Gottshall, 2012; Novack-Gottshall, 2007; Villéger et al., 2011), but the overall utility of this approach is limited due to its condition as just a data type.

52 Previous analysis of mammal diversity and hypotheses as to the processes that have shaped it tend to be through one or more of the following lenses: diversity of an entire system (e.g. continent)
54 (Alroy, 1996; Alroy et al., 2000; Figueirido et al., 2012; Liow et al., 2008), guild based (Janis et al., 2004; Janis, 2008; Janis et al., 2000; Janis and Wilhelm, 1993; Jernvall and Fortelius, 2004; Pires et al., 2015), clade based (Cantalapiedra et al., 2017; Fraser et al., 2015; Quental and Marshall,

2013; Silvestro et al., 2015; Slater, 2015), and environment based (Badgley and Finarelli, 2013;
58 Badgley et al., 2017; Blois and Hadly, 2009; Eronen et al., 2015; Fraser et al., 2015; Janis, 1993;
Janis and Wilhelm, 1993). Rarely are more than two of these lenses considered simultaneously, and
60 integration across the resulting diversity of observations and hypotheses tends to be based on
coincidence. One of the goals of this study is to present a framework for simultaneously analyzing a
62 diversity of hypotheses by integrating both species traits and environmental factors into a single
model in order to infer a more holistic multi-level picture of the processes which may have shaped
64 mammal species diversity and diversification.

The principle species trait considered in this study is a species' ecotype, defined here as the unique
66 combination of species dietary category and locomotor category (e.g. arboreal omnivore versus
unguligrade herbivore). These classifications can be considered analogous to guilds or unique
68 ecocube combinations 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 its inclusion is principally to
70 control for that effect on the other covariates that are the focus of this study.

Translating previous work into hypotheses applicable to this analysis is difficult for a variety of
72 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
74 Wilhelm, 1993; Pires et al., 2015; Quental and Marshall, 2013; Slater, 2015). Because taxonomic
grouping conflates both species macroecology with shared evolutionary history, there are few clear
76 ways to translate and operationalize these hypotheses in terms of macroecological change viewed
through the lens of species interactions. Hypotheses as to macroecological change viewed through
78 the lens of species interactions. Specifically, this issue arises when trying to generalize previous
observations from taxonomy-based framework to ecology-based one.

80 There is little convincing evidence of any major or sudden cross-ecotypic or cross-taxonomic
turnover events in history of North American mammal diversity, unlike the Neogene record
82 European mammals (Alroy, 1996, 2009; Alroy et al., 2000; Eronen et al., 2015; Janis, 1993). Instead
of being concentrated in time, turnover has been found to be distributed through time. It is then

84 expected then that, for this analysis, turnover events or periods of rapid diversification or depletion
should not occur simultaneously for all ecotypes.

86 Jernvall and Fortelius (2004) found that for the Neogene of Europe the relative abundance of
mammal guilds was stable over time even in the face of high turnover rates, though they only
88 considered large bodied taxa from a small set of mammal orders. Similar results have been observed
for some taxonomic groups in North America CITATIONS. These results imply that there the
90 types of interactions happening in local communities observed over a region are constant over time
even if the interactors are constantly changing. MORE ABOUT DIVERSITY DEPENDENCE
92 HERE. WHAT DO PEOPLE THINK THE MAMMAL DIVERSITY CURVE REPRESENTS?
CAN BE ANYTHING IF YOU THINK ABOUT IT HARD ENOUGH.

94 The diversity history of ungulate herbivores has been characterized as more recently originating
taxa having longer legs, higher crowned teeth, and a shift from graze-dominated to
96 browse-dominated diets than their earlier originating counterparts (Cantalapiedra et al., 2017;
Fraser et al., 2015; Janis et al., 2004; Janis, 2008; Janis et al., 2000; Janis and Wilhelm, 1993); all of
98 which have all been attributed to some combination of environmental change itself or tectonic
activity driving environmental change (Badgley et al., 2017; Blois and Hadly, 2009; Eronen et al.,
100 2015; Janis, 2008). Additionally, it has been observed that these cursorial ungulate forms arose
prior to cursorial carnivore forms, an observation attributed to the reorganization of plant
102 communities towards the end of the Cenozoic and the latter emergence of “modern” environments
and communities (Janis and Wilhelm, 1993).

104 Within the canid guild of North America (e.g. plantigrade and digitigrade carnivores) there is
evidence that their diversity is self-regulating or somehow limited. Specifically, it has been proposed
106 that different canid clades have replaced each other as the dominate members of that
macroecological role within the species pool (Silvestro et al., 2015). A pattern of generally constant
108 diversity through time is also observed within the canid carnivore subguilds of hypercarnivore,
hypocarnivore, and mesocarnivores identified by Slater (2015) even in the face of constant species
110 turnover is consistent with limited possibility of increased diversity, even though there was no

evidence of diversity-dependence in trait (e.g. body size) evolution (Slater, 2015). THERE IS
112 MORE TO THIS STORY; CHECK THE WORK OF VANVALKENBURG; DANCES WITH
WOLF AVATARS?

114 There is some uncertainty and a lack of consensus as to the effect of species body size on mammal
diversity and aspects of the diversification processes, specifically extinction (Liow et al., 2008, 2009;
116 Smits, 2015; Tomiya, 2013). Species body size is frequently framed as an important biological
descriptor because of how it is correlated with other important and relevant traits such as
118 metabolic rate and home range size CITATIONS. It is also relatively easy to estimate for extinct
species using proxy measures and regression equations, as was done in this study (see below).
120 However, body size is normally considered without reference to other ecological descriptors of the
species (Liow et al., 2008), but see (Smits, 2015); this combined with the high amount of correlation
122 between life history traits and body size limits process-based inference because the actual causal
mechanisms underlying an observed pattern are obscured or missing.
124 Smits (2015) found that the individual traits which form this study's ecotypes have strong effects
on mammal extinction risk. Omnivorous taxa were found to have, on average, a greater duration
126 than other dietary categories, while arboreal taxa were found to have a shorter duration than other
locomotor categories (Smits, 2015). Two possible scenarios that could yield this pattern were
128 proposed: the extinction risk faced by arboreal is constant and high or the Paleogene and Neogene
represent different regimes and extinction risk increased in the Neogene, thus driving up the
130 Cenozoic average extinction risk. These two possible explanations have clear and testable
predictions with respect to the diversity history of arboreal taxa: 1) the extinction risk arboreal
132 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
134 arboreal taxa have just a generally higher extinction risk than other ecotypes but have maintained
a constant diversity for the Cenozoic. By inspecting the inferred diversity histories of the ecotypes,
136 it should be possible to distinguish amongst these hypotheses.

Fundamentally, all species respond differently to climate and environmental change (Blois and

¹³⁸ Hadly, 2009). Macroecological patterns are emergent patterns due to the similarities among species in how they respond to a similar “stimulus.”

¹⁴⁰ The effect of climate on diversity and the diversification process has been the focus of considerable research with a slight consensus favoring diversification being more biologically-mediated than
¹⁴² climate-mediated (Alroy, 1996; Alroy et al., 2000; Clyde and Gingerich, 1998; Figueirido et al.,
¹⁴⁴ 2012). However, differences in temporal and geographic scale seem to underly the contrast between
¹⁴⁶ these two perspectives. For example when the mammal fossil record analyzed at small temporal and
geographic scales a correlation between diversity and climate are observable (Clyde and Gingerich,
¹⁴⁸ 1998). However, when the record is analyzed at the scale of the continent and most of the Cenozoic
there is no correlation with diversity and climate (Alroy et al., 2000). This results, however, does not
¹⁵⁰ go against the idea that there may be short periods of correlation and that the correlation between
diversity and climate can change or even reverse direction over time; this type result means that
there is no single direction of correlation between diversity and climate (Figueirido et al., 2012).

In the case of a fluctuating correlation between diversity and climate it is hard to make the
¹⁵² argument for an actual causal link between the two without modeling the underlying ecological
differences between species; after all, species respond differently based on their individual ecologies
¹⁵⁴ (Blois and Hadly, 2009). When analysis is based on diversity or taxonomy alone no mechanisms are
possible to infer. Taxonomy, like body size, stands in for many important species traits to the point
¹⁵⁶ that mechanistic or process based inference is impossible. While emergent patterns might
correspond to taxonomic grouping, this itself is an emergent phenomenon. Instead, by framing
¹⁵⁸ hypotheses in terms of species traits and their environmental context, these emergent phenomena
can be observed rather than assumed.

¹⁶⁰ The climate history of the Cenozoic is generally characterized by a global cooling trend and the
development of polar ice-caps during the Neogene; there are, of course, a few notable exceptions to
¹⁶² this broad characterization (Cramer et al., 2011; Zachos et al., 2008, 2001). The environmental
context of North America for the Cenozoic is additionally characterized by an environmental
¹⁶⁴ transition from the closed, partially forested environments of the Paleogene to the savannah and

grasslands environments of the Neogene (Blois and Hadly, 2009; Janis, 1993; Janis et al., 2000;
166 Strömberg, 2005).

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

174 The Eocene-Oligocene transition has been observed to be associated with extinction of many
ungulate taxa (Janis, 2008). This boundary also marks the transition from the Paleogene to the
176 Neogene and from herbivores being browsing dominated to grazing dominated CITATION. WHEN
DO ICE CAPS DEVELOP? WHEN DO GRASSES DOMINATE? There is an observed stability in
178 estimates of global temperature from the E/O transition till the end of the Miocene called the
Mid-Miocene climatic optimum (Zachos et al., 2008, 2001). The Mid-Miocene climatic optimum is
180 bookended by periods of temperature decline. We would then expect that, for the Miocene,
turnover and other diversification events would most likely be due to biological interactions or
182 immigration and not biotic-abiotic interactions because of the constancy of the climate, and that
those groups that are driven primarily by environmental factors, the Miocene would be a period of
184 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
186 changing floral groups present in North America across the Cenozoic CITATIONS. These covariates
were chosen because they provide high level characterizations of the environmental context of the
188 entire North 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)
190 allowing for inference as to how a species ecology can mediate selective pressures do to its
environmental context.

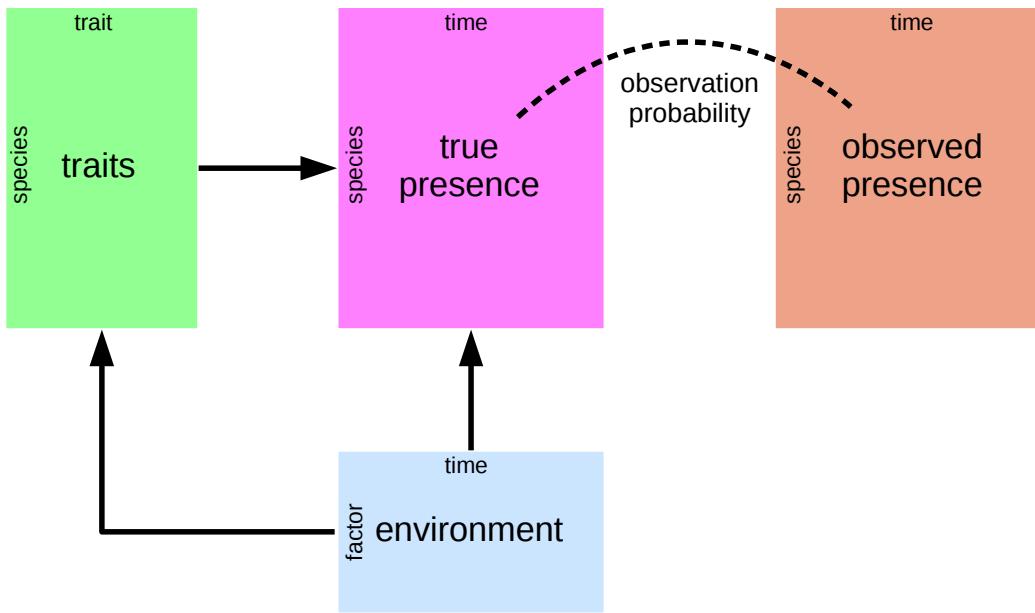


Figure 1: Conceptual diagram of the paleontological fourth corner problem. The observed presence matrix (orange) is the empirical presence/absence pattern for all species for all time points; this matrix is an incomplete observation of the “true” presence/absence pattern (purple). The estimated true presence matrix is modeled as a function of both environmental factors over time (blue) and multiple species traits (green). Additionally, the effect 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).

192 Fourth-corner modeling is an approach to explaining the patterns of either species abundance or
 193 presence/absence as a product of species traits, environmental factors, and the interaction between
 194 traits and environment (Brown et al., 2014; Jamil et al., 2013; Pollock et al., 2012; Warton et al.,
 195 2015); effectively uniting species distribution modeling (SDMs) with trait-based community
 196 assembly (CATS). In modern ecological studies, what is being modeled is species occurrences at
 197 localities distributed across a region (Jamil et al., 2013; Pollock et al., 2012). In this study, what is
 198 being modeled is the pattern of species occurrence over time for most of the Cenozoic in North
 America (Fig. 1). By adding an additional dimension (time) to the fourth-corner framework we can

200 gain better inference of how an instantaneous species pool (i.e. the Modern) is assembled over time.
These two approaches, modern and paleontological, are different views of the same
202 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
204 means that these approaches are complimentary but reveal different patterns of how species are
distributed in time and space.

206 All observations, paleontological or modern, are made with uncertainty. With presence/absence
data this uncertainty comes from now knowing if an absence is a “true” absence or just a failure to
208 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 incomplete preservation of
210 whatever species were present into fossil form combined with incomplete sampling of what
organisms were actually fossilized means that the true times of origination or extinction may not be
212 observed (Foote, 2001; Foote and Sepkoski, 1999; Wang et al., 2016; Wang and Marshall, 2016).

214 Ultimately, the goals of this analysis are to understand when are unique ecotypes enriched or
depleted in the North American mammal regional species pool and how changes in ecotypic
diversity are related to changes in species’ environmental context. In the analyses done here, many
216 covariates which describe both a species’ macroecology and its environmental context are considered.
In order to analyze this complex and highly structured data set, I developed a hierachal Bayesian
218 model combing the forth-corner modeling approach with a model of an observation-occurrence or
observation-origination-extinction process. The complexity and nuance inherent in questions that
220 are focus of this study, it is possible to consider and test a large number of possible hypotheses. The
hierarchical Bayesian modeling approach used here is appropriate for mitigating complications
222 arising from both this complexity and the plethora of testable hypotheses (e.g. multiple
comparisons, garden of forking paths) (Gelman et al., 2013, 2012; Gelman and Loken, 2014).

²²⁴ **Materials and Methods**

Taxon occurrences and species-level information

²²⁶ All fossil occurrence information used in this analysis was downloaded from the Paleobiology Database (PBDB). The initial download restricted all occurrences to all Mammalia observed in ²²⁸ North America between the Maastrichtian (XX Mya) and Gelasian (XX Mya) stages. Occurrences were then further limited to those occurring between 64 and 2 million years ago (Mya). Taxonomic, ²³⁰ stratigraphic, and ecological metadata for each occurrence and species was also downloaded. A new download for a raw, unfiltered PBDB datafile following the same criterion used here is available at ²³² <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.

²³⁴ After download, the raw occurrence data was then sorted, cleaned, and manipulated programmatically before analysis. Many species taxonomic assignments as present in the raw ²³⁶ 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 ²³⁸ were based on taxonomies present in the Encyclopedia of Life WEBSITE and Janis et al. (2008, 1998). All taxa who's life habit was classified as either volant (i.e. Chiroptera) or aquatic (e.g. ²⁴⁰ 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 ²⁴² pools.

²⁴⁴ Species ecotype is defined here as the interaction between life habit and diet categories, the goal of which is to classify species based on the manner with which they interact with their environment. Most mammal species records in the PBDB have life habit and dietary category assignments. In ²⁴⁶ order to simplify 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 ²⁴⁸ life history category was further broken up to better reflect the diversity of mammal locomotor modes. Ground dwelling species locomotor categories were reassigned based on their ankle posture

250 associated with their taxonomic group as described in Table 2 (Carrano, 1999). Ecotype categories
 251 with less than 10 total species of that combination were excluded, yielding a total of 18 observed
 252 ecotypes out of a possible 24.

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
	Aploodontidae	plantigrade
	Apternodontidae	scansorial
	Arctocyonidae	unguligrade
	Barbourofelidae	digitigrade
	Barylambdidae	plantigrade

Continued on next page

Table 2 – continued from previous page

Order	Family	Stance
	Bovidae	unguligrade
	Camelidae	unguligrade
	Canidae	digitigrade
	Cervidae	unguligrade
	Cimolodontidae	scansorial
	Coryphodontidae	plantigrade
	Cricetidae	plantigrade
	Cylindrodontidae	plantigrade
	Cyriacotheriidae	plantigrade
	Dichobunidae	unguligrade
Dinocerata		unguligrade
	Dipodidae	digitigrade
	Elephantidae	digitigrade
	Entelodontidae	unguligrade
	Eomyidae	plantigrade
	Erethizontidae	plantigrade
	Erinaceidae	plantigrade
	Esthonychidae	plantigrade
	Eutypomyidae	plantigrade
	Felidae	digitigrade
	Florentiamyidae	plantigrade
	Gelocidae	unguligrade
	Geolabididae	plantigrade
	Glyptodontidae	plantigrade
	Gomphotheriidae	unguligrade

Continued on next page

Table 2 – continued from previous page

Order	Family	Stance
	Hapalodectidae	plantigrade
	Heteromyidae	digitigrade
	Hyaenidae	digitigrade
	Hyaenodontidae	digitigrade
	Hypertragulidae	unguligrade
	Ischyromyidae	plantigrade
	Jimomyidae	plantigrade
Lagomorpha		digitigrade
	Leptictidae	plantigrade
	Leptochoeridae	unguligrade
	Leptomerycidae	unguligrade
	Mammutidae	unguligrade
	Megalonychidae	plantigrade
	Megatheriidae	plantigrade
	Mephitidae	plantigrade
	Merycoidodontidae	digitigrade
Mesonychia		unguligrade
	Mesonychidae	digitigrade
	Micropternodontidae	plantigrade
	Mixodectidae	plantigrade
	Moschidae	unguligrade
	Muridae	plantigrade
	Mustelidae	plantigrade
	Mylagaulidae	fossorial
	Mylodontidae	plantigrade

Continued on next page

Table 2 – continued from previous page

Order	Family	Stance
	Nimravidae	digitigrade
	Nothrotheriidae	plantigrade
Notoungulata		unguligrade
	Oromerycidae	unguligrade
	Oxyaenidae	digitigrade
	Palaeomerycidae	unguligrade
	Palaeoryctidae	plantigrade
	Pampatheriidae	plantigrade
	Pantolambdidae	plantigrade
	Peritychidae	digitigrade
Perissodactyla		unguligrade
	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

Continued on next page

Table 2 – continued from previous page

Order	Family	Stance
	Titanoideidae	plantigrade
	Ursidae	plantigrade
	Viverravidae	plantigrade
	Zapodidae	plantigrade

254 Estimates of species mass used as a covariate in this study were sourced from multiple databases
and papers, especially those focusing on similar macroevolutionary or macrecological questions
256 (Brook and Bowman, 2004; Freudenthal and Martín-Suárez, 2013; McKenna, 2011; Raia et al., 2012;
Smith et al., 2004; Tomiya, 2013); this is similar to what was done in Smits (2015). When species
258 mass was not available, proxy measures 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
260 mass given some regression equation. 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).
262 Mass was log-transformed and then mean-centered and rescaled by dividing by two-times its
standard deviation; this insures that the magnitude of effects for both continuous and discrete
264 covariates are directly comparable (Gelman, 2008; Gelman and Hill, 2007).

All fossil occurrences from 64 to 2 million years long (Mya) were binned into 31 2 million year (My)
266 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).

268 Environmental and temporal covariates

The environmental covariates used in this study are collectively referred to as group-level covariates
270 because they predict the response of a “group” of individual-level observations (i.e. species
occurrences of an ecotype). Additionally, these covariates are defined for temporal bins and not the

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	?

272 species 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
 274 phases” defined by Graham (2011). 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
 276 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
 278 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.
 280 These properties make Mg/Ca based temperature estimates preferable for macroevolutionary and
 macroecological studies (Ezard et al., 2016). Two aspects of the Mg/Ca-based temperature curve
 282 were included in this analysis: mean and range. Both were calculated as the mean of all respective
 estimates for each 2 My temporal bins. Both mean and range were then rescaled by subtracting the
 284 mean and then dividing by twice the standard deviation.

The second set of environmental factors included in this study are the Cenozoic plant phases defined
 286 in Graham (2011). Graham’s plant phases are holistic descriptors of the taxonomic composition of

Table 4: Definitions of the start and stop times of the three plant phases used this study as defined by Graham (2011).

Plant phase	Phase number	Start	Stop
Paleocene-Eocene	1	66	50
Eocene-Miocene	2	50	16
Miocene-Pleistocene	3	16	2

12 ecosystem types, which plants are present at a given time, and the relative modernity of those
 288 plant groups with younger phases representing increasingly modern taxa (Graham, 2011). Graham
 (2011) defines four intervals from the Cretaceous to the Pliocene, though only three of these
 290 intervals are included in this analysis. Graham’s plant phases was included as a series of “dummy
 variables” encoding the three phases included in this analysis. This means that the first phase is
 292 synonymous with the intercept and subsequent phases are defined by their differences from the first
 phase. The temporal boundries of these plant phases are defined in Table 4.

294 Modelling species occurrence

Two different models were used in this study: a pure-presence model and a birth-death model. Both
 296 models at their core are a hidden Markov model where the latent process has an absorbing state
 (Allen, 2011). The difference between these two models is if the probabilities of a species originating
 298 or surviving are considered equal or different (Table 5). Something that is important to realize is
 that while there are only two state “codes” in a presence-absence matrix (i.e. 0/1), there are in fact
 300 three states in a birth-death model: not having originated (yet), extant, and extinct. The last of
 these is the absorbing state, as once a species has gone extinct it cannot re-originate (Allen, 2011);
 302 this is made obvious in the transition matrices as the probability of an extinct species changing
 states is 0 (Table 5). See below for parameter explainations (Tables 7, and 8).

304 Data augmentation

All empirical presence/absence observations are potentially incomplete or observed with error. The
 306 hidden Markov model at the core of this analysis allows for observed absences to be used

		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 5: Transition matrices for the pure-presence (5a) and birth-death (5b) 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).

meaningfully to estimate the number of unobserved species. Of concern in this analysis is the
 308 unknown “true” size of the dataset; how many species could have actually been observed? While
 many species have been observed, the natural incompleteness of all observations, especially in the
 310 case of paleontological data, there are obviously many species which were never sampled (Royle and
 Dorazio, 2008; Royle et al., 2007).

312 Let N by the total number of observed species, M be the upper limit of possible species that could
 have existed given some model of species occurrence, and N^* is the number of all-zero histories
 314 added to the presence absence matrix y where $N^* = M - N$. 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
 316 that any augmented species should actually be “present.” Because M is user defined, this approach
 effectively gives ψ a uniform prior over N to M (Royle and Dorazio, 2008). For this study,
 318 $M = \lfloor 1.25 \times N \rfloor$.

Data imputation is the process of estimating missing data for partially observed covariates given
 320 the other fully-observed observations and some model (Gelman and Hill, 2007; Rubin, 1996), this is
 simple in a Bayesian context because data are also parameters (Gelman et al., 2013). Augmented
 322 species are fully imputed species and thus have no known mass so a mass estimate must be imputed
 for each possible species (Royle and Dorazio, 2012). Assuming that mass values for augmented
 324 species are from the same distribution as observed species, the distribution of observed mass values
 are estimated as part of the model and new mass values are then generated from this distribution.

Table 6: 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

- 326 This approach is an example of imputing covariate information that is missing completely at
 random (Gelman and Hill, 2007; Royle and Dorazio, 2012). Because log mass values are rescaled as
 328 a part of this study, the body mass distribution is already known ($\mathcal{N}(0, 0.5)$) the body mass of the
 augmented species are generated by simple random draws from this distribution. In addition to
 330 body mass information, the augmented species need an ecotype classification. Because these species
 are completely unknown, they were all classified as “augmented” to indicate their unknown biology.
 332 This classification has no biological interpretation.

Observation process

- 334 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
 336 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
 338 $\phi = \pi$, 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
 340 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

Table 7: 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

342 assumptions are reflected in the structure of the model Equation 1, presented here:

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

The parameters associated with Equation 1 are described in Table 6.

344 Pure-presence process

For the pure-presence model there is only a single probability dealing with the presence of a species
 346 θ (Table 5a). 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
 348 with pure-presence model are presented in Table 7 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.
 350 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
 352 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

354 choice of priors for these regression coefficients.

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

360 All parameters not modeled elsewhere were given weakly informative priors (Gelman et al., 2013;
McElreath, 2016) CITATION STAN MANUAL STATISTICAL RETHINKING. Weakly
362 informative means that priors do not necessarily encode actual prior information but instead help
regularize or weakly constrain posterior estimates. These priors have a concentrated probability
364 density around and near zero; this has the effect of tempering our estimates and help prevent
overfitting the model to the data (Gelman et al., 2013; McElreath, 2016) CITATION STAN
366 MANUAL STATISTICAl RETHINKING. The 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. The sampling
368 statement, excluding the imputation of body mass associated with the augment species, is as follows

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

Birth-death process

- ³⁷⁰ In the birth-death version of the model, $\phi \neq \pi$ and so each of these probabilities are modeled separately but each is handled in a similar manner to how θ is modeled in the pure-presence model
³⁷² (Eq. 2, Table 5b). The parameters associated with the birth-death presence model are presented in Table 8 and the full sampling statement, including observation (Eq. 1), is described in Equation 3:

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

- ³⁷⁴ Similar to the pure-presence model, both ϕ and π are modeled as logistic regressions with varying-intercept and one covariate associated with two parameters. The possible relationships
³⁷⁶ between mass and both ϕ and π are reflected in the parameterization of the model and choice of priors (Eq. 3).
³⁷⁸ The intercepts of ϕ and π both vary by species ecotype and those values are themselves the product of group-level regression using environmental factors as covariates (Eq. 3); this is identical to the

Table 8: 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

380 pure presence model (Eq. 2).

Posterior inference and model adequacy

382 Computer programs that implement joint posterior inference for the above models (Eqs. 2, 3) were
written in the probabilistic programming language Stan CITATION. The models used here both
384 feature latent discrete parameters in the large matrix z (Tables 6, 7, 8; Eqs. 1, 2, 3). All methods
for posterior inference implemented in Stan are derivative based which causes complications for
386 actually implementing the above models because integers do not have derivatives. Instead of
implementing a latent discrete parameterization, the log posterior probabilities of all possible states
388 of the latent parameters z were calculated and summed (i.e. marginalized).

Species durations at minimum range-through from the FAD to the LAD, but the incompleteness of
390 all observations means that the actual time of origination or extinction is unknown. The
marginalization approach used here means that the probabilities all possible histories for a species

392 are calculated, from the end members of the species having existed for the entire study interval and
 the species having only existed between the directly observed FAD and LAD to all possible
 394 intermediaries (Fig 2) CITATION STAN MANUAL. This process is identical, language-wise, to
 assuming range-through and then estimating the possibility of range extension due to incomplete
 396 sampling.

	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	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

The combined size of the dataset and large number of parameters in both models (Eqs. 2, 3),
 398 specifically the total number of latent parameters that are the matrix z , means that stochastic
 approximate posterior inference is computationally very slow even using NUTS based HMC
 400 implemented. Instead, an approximate Bayesian approach was used: variational inference. A
 recently developed automatic variational inference algorithm called “automatic differentiation
 402 variational inference” (ADVI) is 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

404 similar to quadratic approximation of the likelihood function commonly used in maximum
 likelihood based inference CITATION. The principal limitation of assuming the joint posterior is
 406 Gaussian is that the true topology of the log-posterior isn't estimated; this is a particular burden
 for scale parameters which are bound to be positive (e.g. standard deviation).

408 Of additionally concern for posterior inference is the partial identifiability of observation parameters
 $p_{t=1}$ and $p_{t=T}$ (Royle and Dorazio, 2008). This issue means that the estimates of sampling
 410 probabilities at the “edges” of the time series cannot fully be estimated because there are no known
 “gaps” in species occurrence histories that are guaranteed to be filled. Instead, the values of the first
 412 and final columns of the “true” presence-absence matrix z for those observations that do not already
 have presences in the observed presence-absence matrix y cannot be estimated (Royle and Dorazio,
 414 2008). The hierarchical modeling approach used here helps mitigate this problem by pulling the
 values of $p_{t=1}$ and $p_{t=T}$ towards the overall mean of p (Gelman et al., 2013), and in fact this
 416 approach might be more analytically sound than the more ad-hob approaches that are occasionally
 used to overcome this hurdle (Royle and Dorazio, 2008). Additionally, because $p_{t=1}$ and $p_{t=T}$ are
 418 only partially identifiable, estimates of occurrence θ and origination ϕ at $t = 1$ and estimates of θ , ϕ
 and survival pi at $t = T$ may suffer from similar edge effects. Again, the hierarchical modeling
 420 approach used here may help correct for this reality by drawing these estimates towards the overall
 means of those parameters.

422 After fitting both models (Eqs. 2, 3) using ADVI, model adequacy and quality of fit was assessed
 using a simple posterior predictive check (Gelman et al., 2013). Because all Bayesian models are
 424 inherently generative, simulations of new data sets is “free” (Gelman et al., 2013; McElreath, 2016).
 By simulating 100 theoretical data sets using the observed covariate information the congruence
 426 between predictions made by the model and the observed empirical data can be assessed. These
 datasets are simulated by starting with the observed states of the presence-absence matrix at $t = 1$;
 428 from their, the time series roll forward as stochastic processes with covariate information given from
 the empirical observations. Importantly, this is fundamentally different from observing the posterior
 430 estimates of the “true” presence-absence matrix z . The posterior predictive check used in this study
 is to compare the observed average number of observations per species to a distribution of

- 432 simulated averages; if the empirically observed value sits in the middle of the distribution than the
model can be considered adequate in reproducing the observed number of occurrences per species.
- 434 The ADVI assumption of a purely Gaussian posterior limits the utility and accuracy of the
posterior predictive checks because parameter estimates do not reflect the true posterior
- 436 distribution and are instead just an approximation CITATIONS. Because of this, posterior
predictive estimates are themselves only approximate checks of model adequacy. The posterior
- 438 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.
- 440 Given parameter estimates, diversity and diversification rates are estimated through posterior
predictive simulations. Given the observed presence-absence matrix y , estimates of the true
442 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
444 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

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

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

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

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

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

450 Results

The results of the analyses described above take one of two forms: direct inspection of parameter
 452 posterior estimates from both models, and downstream estimates of diversity and diversification
 rates based on posterior predictive simulations from the birth-death model; the reason for which is
 454 explained below in the comparison of the models' posterior predictive check results.

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

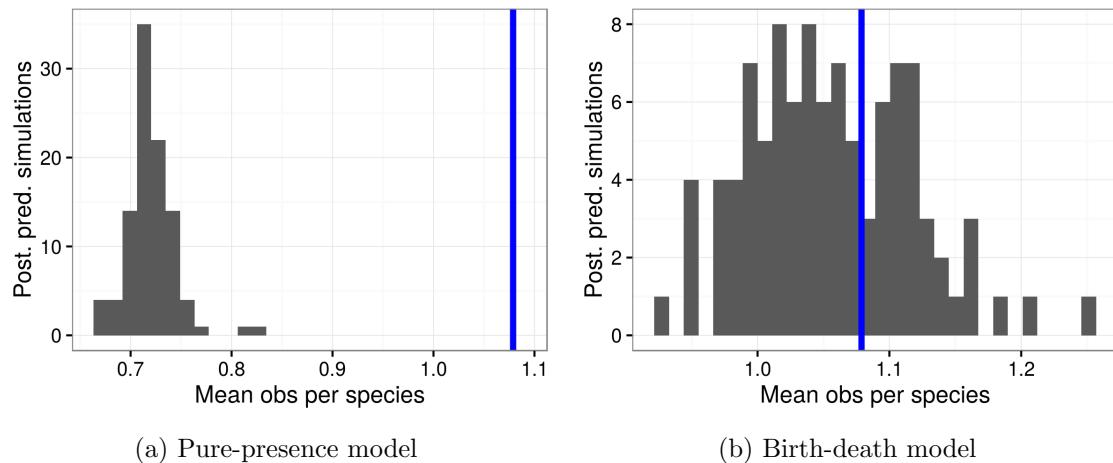


Figure 3: Comparison of the average observed number of occurrences per species (blue line) to the average number of occurrences from 100 posterior predictive datasets using the posterior estimates from the pure-presence and birth-death models.

Comparison of the posterior predictive results from the pure-presence and birth-death models

458 reveals a striking difference in performance of either model to predict the structure of the
underlying data (Fig. 3). The simulated datasets generated from the birth-death model are clearly
460 able to 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
462 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
464 parameter estimates from both models gives further insight into the reasons for this difference in
posterior predictive results (Gelman et al., 2013).

466 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) but not the survival
468 probability estimates (Fig. 6). This result supports the idea that changes to the North American
regional species pool is more likely due to changes in origination than extinction, a result that is
470 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)
472 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
474 had that occurrence by the last time point, so our certainty in a species occurring must increase
with time. Importantly, there are potential issues surrounding the partial identifiability of the
476 observation parameters p which may contribute to edge effects in estimates of occurrence,
origination, and extinction (Royle and Dorazio, 2008). Notably, ecotypes with arboreal components
478 do not appear to follow a similar pattern; instead, occurrence and origination probabilities appear
relatively flat for most of the Cenozoic; this is most likely caused by those species of those ecotypes
480 no longer originating or originating very rarely.

The dramatic differences in the estimates origination and survival probabilities are indicative of
482 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.
484 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
 486 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
 488 estimates of survival probabilities (Fig. 5, and 6). In logistic regression, high uncertainty in the
 490 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
 492 effect of which has been mitigated by the hierarchical modeling strategy used here (Gelman et al.,
 2013; Gelman and Hill, 2007) CITATION Statistical Rethinking.

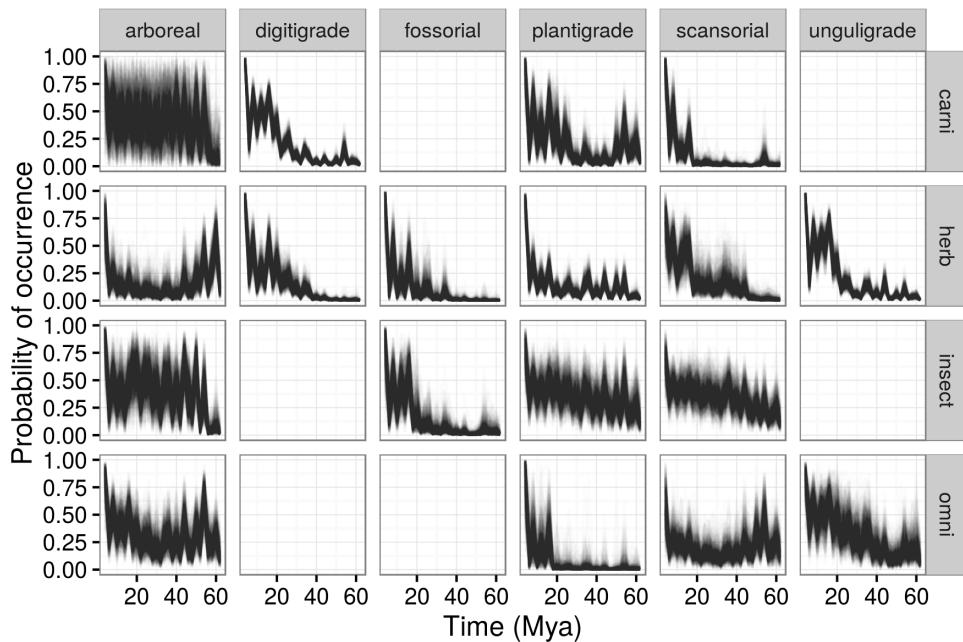


Figure 4: Probability of a mammal ecotype occurring over time as estimated from the pure-presence model. Each panel depicts 100 random samples from the model’s posterior. The columns are by locomotor category and rows by dietary category; their intersections are the observed and analyzed ecotypes. Panels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

494 The pure-presence and birth-death models also differ in the estimated effect of mass on the
 probability of sampling a species that is present (Fig. 7). For the pure-presence model, mass is
 496 estimated to not have a strong 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

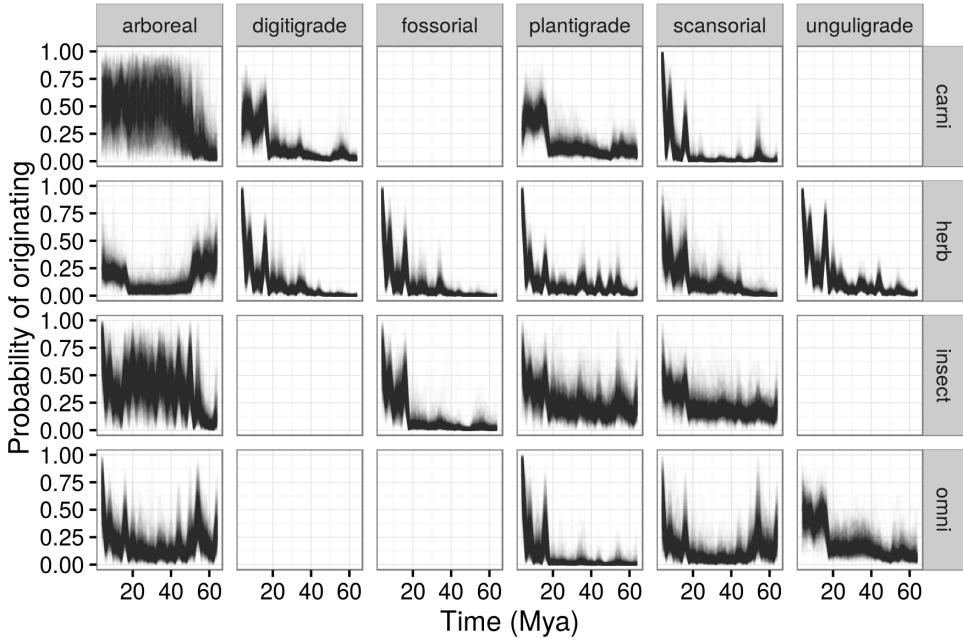


Figure 5: Probability of a mammal ecotype origination probabilities at each time point as estimated from the birth-death model. Each panel depicts 100 random samples from the model’s posterior. The columns are by locomotor category and rows by dietary category; their intersections are the observed and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

- 498 observation such that larger species are less likely to be observed if present than smaller species
 (Fig. 7b).
- 500 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
- 502 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
 504 bodied species is expected on average to have more gaps in sampling and a less consistent record
 from time point to time point than smaller bodied species. Additionally, as this is presence/absence
 506 data 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 multiple time points.
- 508 The average sampling probabilities for both the pure-presence model and birth-death model are
 both at the point where (rescaled log) mass equals 0; visual comparison indicates that, on average,

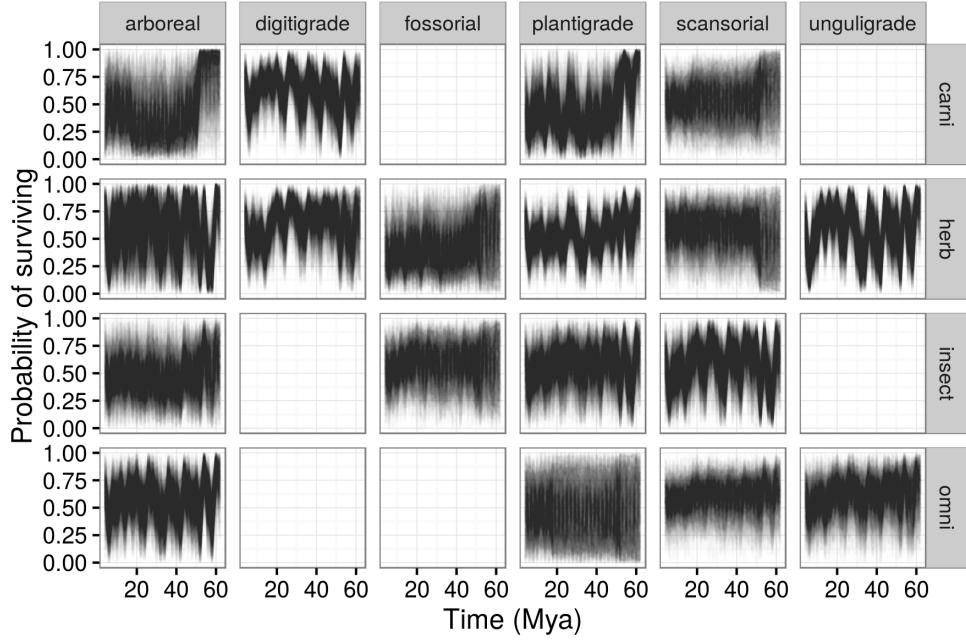


Figure 6: Probability of a mammal ecotype survival probabilities at each time point as estimated from the birth-death model. Each panel depicts 100 random samples from the model’s posterior. The columns are by locomotor category and rows by dietary category; their intersections are the observed and analyzed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

510 sampling probability has greater posterior estimate in the pure-presence model than the birth-death
 model (Fig. 7). The probability that one estimate is different from the other, however, are not
 512 directly calculable as they come from different models; what this tells us is how adding more
 information to the model (i.e. replacing occurrence with origination and extinction) changes
 514 parameter estimates in the model.

The effect of species mass on probability of occurrence as estimated from the pure-presence (Fig. 8)
 516 are most similar to the estimated effect of species mass on probability of origination for the
 birth-death model (Fig. 9). The striking pattern observable in both sets of estimates is the higher
 518 probability of occurrence for species with body sizes closer to the mean than either extremes. This
 result is consistent with the canonically normal distribution of mammal body sizes (Smith et al.,
 520 2004); it is then expected that the most likely to occur species would be those from the middle of
 the distribution, and that species originating will on average be of average mass, especially
 522 considering species shared common ancestry (Felsenstein, 1985). Note that all variation in estimates

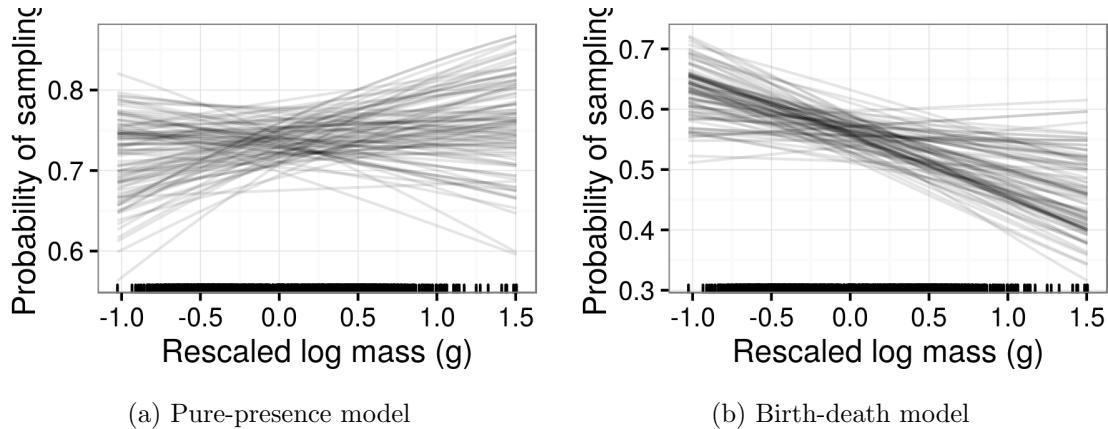


Figure 7: Estimates of the effect of species mass on probability of sampling a present species (p). Mass has been log-transformed, centered, and rescaled; this means that a mass of 0 corresponds to the mean of log-mass of all observed species and that mass is in standard deviation units. Estimates are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

between ecotypes (Fig. 9) 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.

In contrast, the effect of species mass on probability of survival as estimated from the birth-death model (Fig. 10) is consistent with previous findings that there is little effect of mass on extinction for North American mammals for the Cenozoic (Smits, 2015; Tomiya, 2013). Note that all variation between ecotypes depicted in Figure 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 (Eqs. 2, 3).

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 estimates from the pure-presence and birth-death models, the effects of the group-level covariates in the pure-presence model (Fig. 11) are more similar to those estimates of the group-level effects on 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 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.

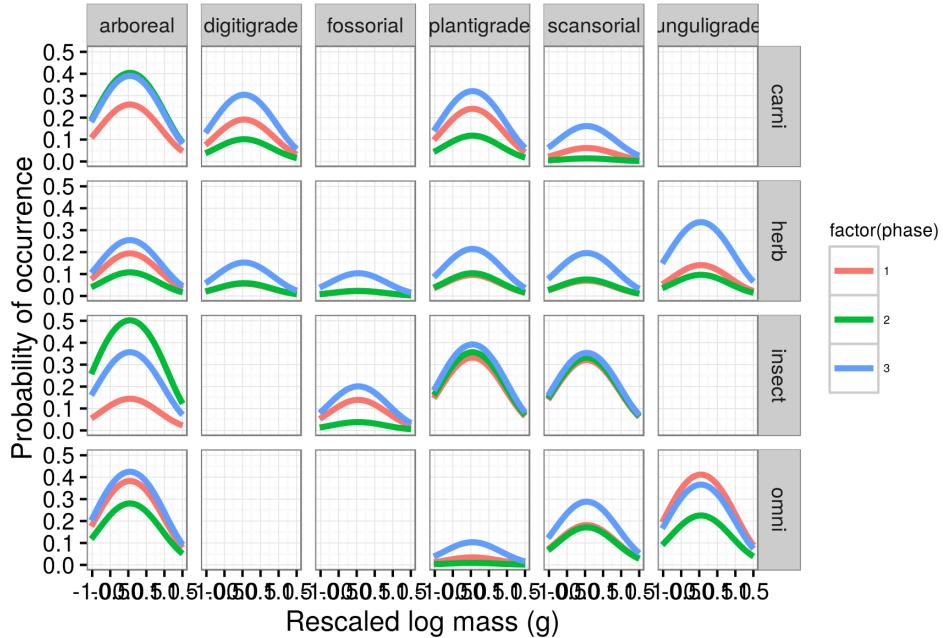


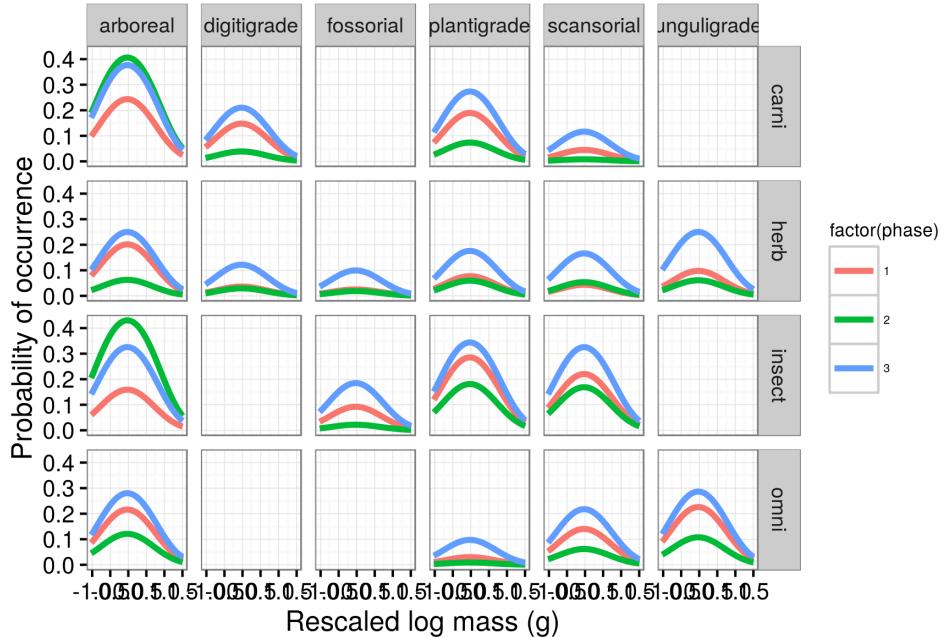
Figure 8: Mean estimate of the effect of species mass on the probability of a species occurrence for each of the three plant phases. The effect of mass is considered constant over time and that the only aspect of the model that changes with plant phase is the intercept of the relationship between mass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

An association between plant phase and differences in the log-odds of occurrence (Fig. 11),

540 origination (Fig. 12), or extinction (Fig. 13) is interpreted to mean that if the set of possible
 541 mammal-plant interactions was either favorable (positive association) or adverse (negative
 542 association) to those ecotypes. In the case of species origination, for example, favorable conditions
 543 for an ecotype may be indicative of an increasing number of possible and available mammal-plant
 544 interactions (e.g. ecological opportunity); while adverse conditions may translate to a decreasing set
 545 of interactions or loss of appropriate environmental context. Note that favorable versus adverse
 546 condition of a plant phase is definitionally relative to the other two plant phases.

Plant phases are associated with large differences in log-odds for occurrence and origination

548 probabilities (Tab. 9), 10), though there is little evidence of plant phase being an important



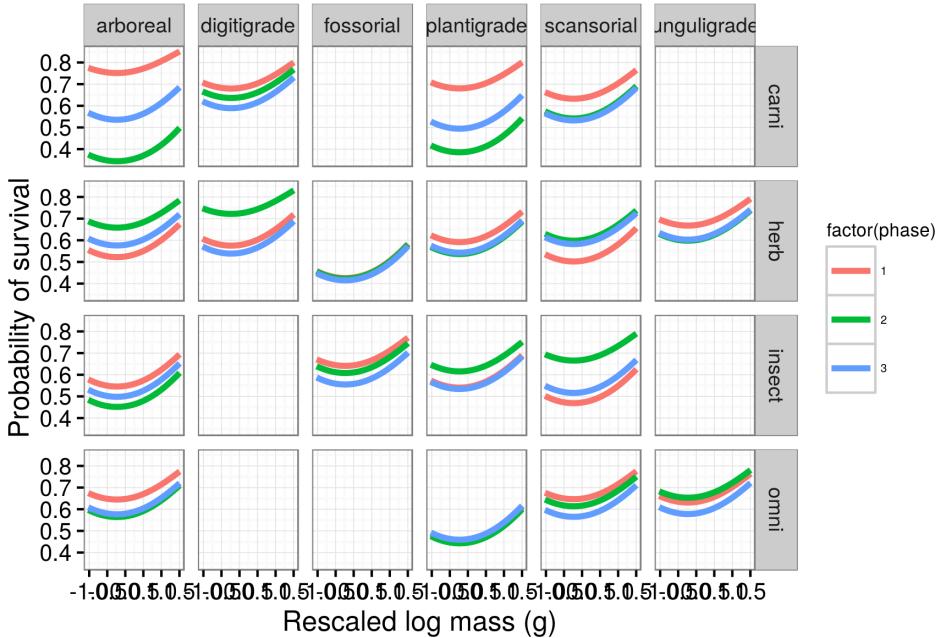


Figure 10: Mean estimate of the effect of species mass on the probability of a species survival for each of the three plant phases. The effect of mass is considered constant over time and that the only aspect of the model that changes with plant phase is the intercept of the relationship between mass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

is large up-tick in probability of occurring or originating towards the modern (Fig. 4, 5). The
560 differences in mean probability of occurring or originating attributable to the plant phases are

observable as shifts along the time series correponding to the phase barriers (Table 4). For example,

562 scansorial herbivore occurrence and origination probabilities demonstrate clear shifts at 50 Mya and
16 Mya (Fig. 4, 5).

564 Both aspects of global temperature analyzed here are estimated to have strong effects on species
occurrence and origination for most mammal ecotypes (Tables 12, 13). Similarity, temperature is
566 only expected to have a strong effect on species extinction for very few ecotypes (Tab. 14). For the
occurrence and origination probabilities of many ecotypes, both temperature covariates have
568 negative estimates which means that as temperature decreases, occurrence or origination are

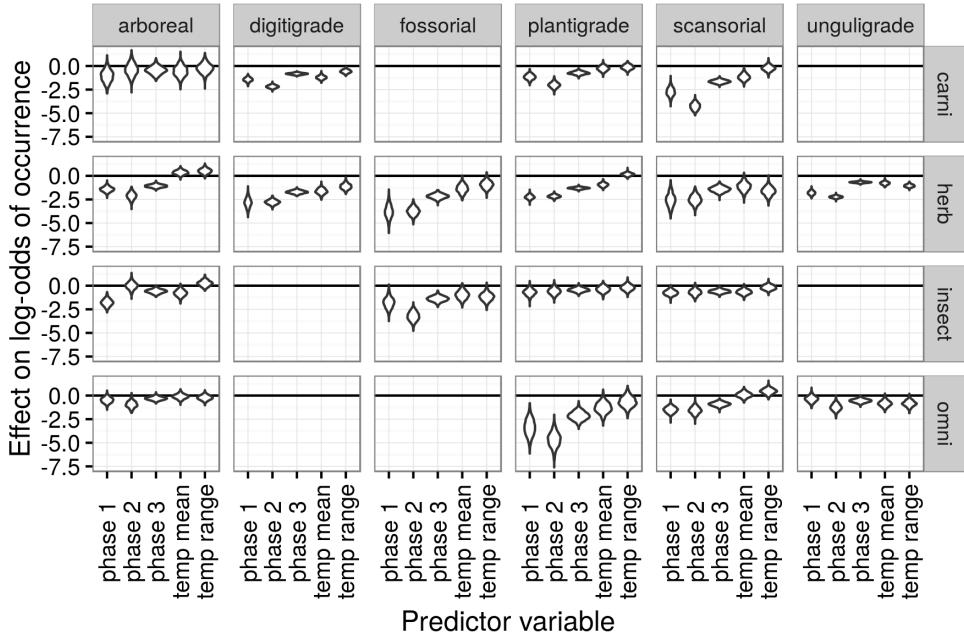


Figure 11: Estimated effects of the group-level covariates describing environmental context on log-odds of species occurrence. These estimates are from the pure-presence model.

expected to increase. The only strongly positive estimate (e.g. temperature decrease, origination
570 decrease) is for the effect of temperature range on arboreal herbivores. Contrastingly, the only
strong ecotype associations for either of the temperature covariates are with plantigrade carnivores,
572 plantigrade herbivores, and to a less certain extent arboreal herbivores and scansorial insectivores
(Tab. 14). The effects of the temperature covariates on these ecotypes are all estimated to be
574 positive (e.g. temperature range increase, increase in survival).

Analysis of diversity

576 All of the analyses of diversification and macroevolutionary rates has been done using only the
birth-death model because of the model's better posterior predictive check performance (Fig. 3).
578 The general pattern of the estimated North American total mammal diversity for the Cenozoic is
“stable” in that mean standing diversity does not fluctuate wildly and rapidly over the Cenozoic
580 (Fig. 14a). In broad strokes, the first 15 or so million years of the Cenozoic are characterized by a

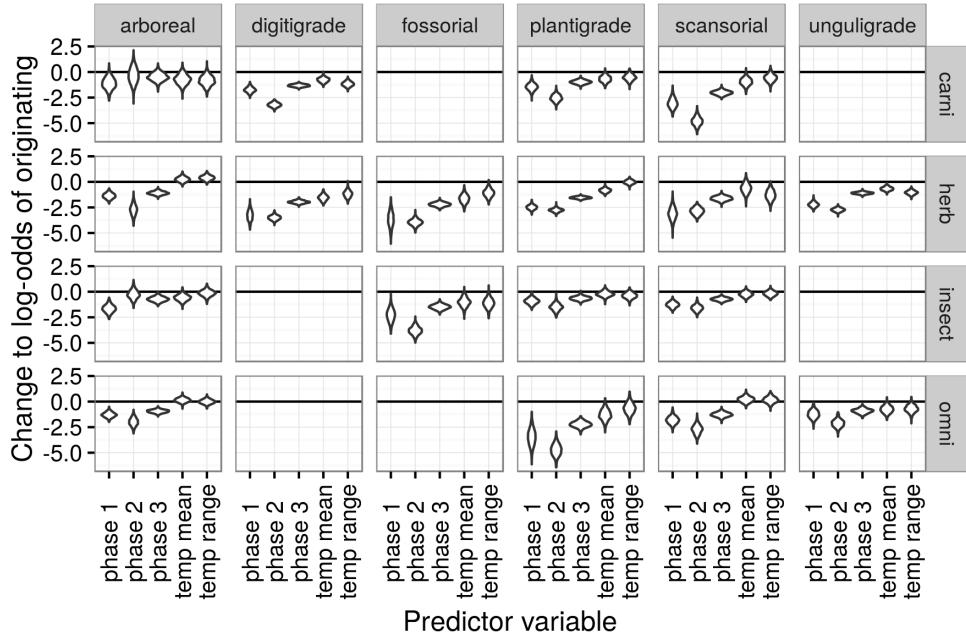


Figure 12: Estimated effects of the group-level covariates describing environmental context on log-odds of species origination. These estimates are from the birth-death model.

gradual decline in standing diversity until approximately 45-50 million years ago (early-middle

582 Eocene). Following this decline, standing diversity is broadly constant from 45 to 18 Mya (early
583 Miocene). After this, there is a rapid spike in diversity followed by a slight decline in diversity up to
584 the Modern. This characterization of the estimated diversity history is knowingly broad strokes and
diversity time series is not without variation and vagaries.

586 When viewed through the lens of diversification rate, some of the structure behind the estimated
diversity history begins to take shape (Fig. 14b). For most of the Cenozoic, the diversification rate
588 hovers around zero, punctuated by both positive and negative spikes. The largest spike in
diversification rate is at 18 Mya, which is early Oligocene (Fig. 14b). Other notable increases in
590 diversification rate occur 56, 46, 38, and 6 Mya (Table 15), though the last of these may be due
edge effects surrounding the partial-identifiability of $p_{t=T}$. Notable decreases in diversification rate
592 occur 60, 54, 50, 44, 34, 20, 16, 12, and 8 Mya (Table 15), meaning that diversification rate has
more major decreases than increases. Given that diversification rate more closely resembles
594 origination rate than extinction rate (Fig. 14b, 14c, 14d), these decreases in diversification rate may

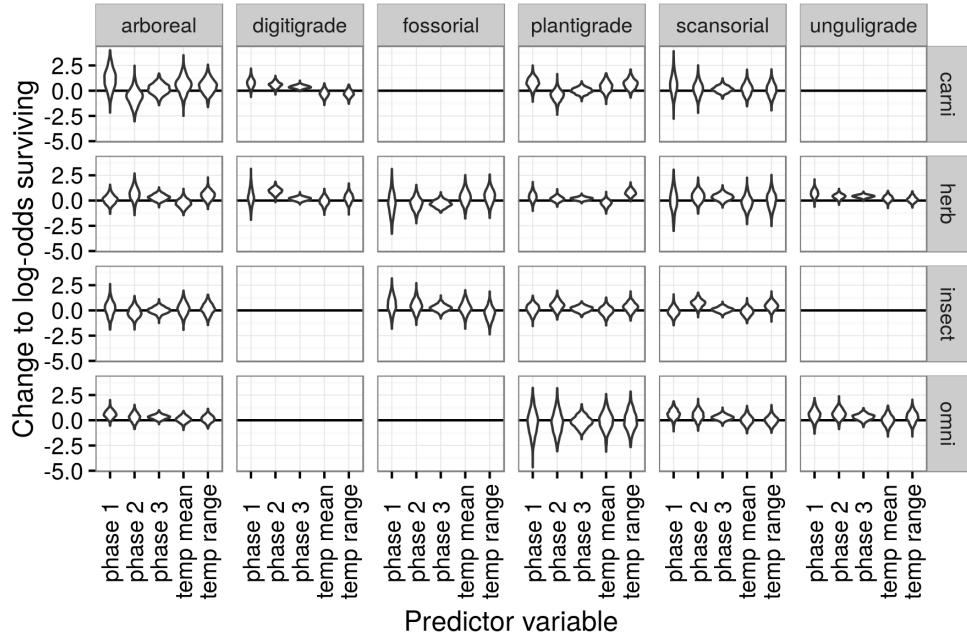


Figure 13: Estimated effects of the group-level covariates describing environmental context on log-odds of species survival. These estimates are from the birth-death model.

be indicative of “depletions” (failure to replace extinct taxa) rather than pulses of extinction.

- 596 The comparison between per capita origination and extinction rate estimates reveals how
diversification rate is formed (Fig. 14c, 14d). As expected given previous inspection of the ecotype
598 specific estimates of origination and survival probabilities from the birth-death model,
diversification rate seems most driven by changes in origination rate as opposed to extinction rate.
600 Extinction rate, on the other hand, demonstrates an almost saw-toothed pattern around a constant
mean (Fig. 14d).

602 Diversity partitioned by ecotype reveals a lot of the complexity behind the pattern of mammal
diversity for the Cenozoic (Fig. 15).

604 Arboreal ecotypes obtain peak diversity early in the Cenozoic and then decline for the rest of the
time series, becoming increasingly rare or absent as diversity approaches the Modern (Fig. 15).
606 Arboreal herbivores and omnivores obtain peak diversity at the beginning of the Cenozoic then go
into decline while remaining a small part of the species pool, while arboreal carnivores and

Table 9: Posterior probability of the differences in the log-odds of an ecotype originating 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. 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

608 insectivores obtain peak diversity 52-50 Mya and then quickly decline and become extremely rare or entirely absent from the species pool.

610 The diversity of both digitigrade and unguligrade herbivores increase over the Cenozoic (Fig. 15). In contrast, plantigrade herbivore diversity does not have a single, broad-strokes pattern; instead, 612 diversity increases, decreases, and may have then increased till the Modern. Contrastingly, fossorial and scansorial herbivores demonstrate a much flatter history of diversity, with a slight increase in 614 diversity that over time is more pronounced among fossorial taxa than scansorial taxa.

Digitigrade carnivores have a multi-modal diversity history, with peaks 54-52 and 12-10 Mya 616 (Fig.15). Between these two peaks digitigrade carnivore diversity dips below average diversity following the first peak and then grows slowly until the second peak. Plantigrade carnivores obtain 618 peak diversity in the early Cenozoic and then maintain a relatively stable diversity until another peak at the end of the Cenozoic.

Table 10: Posterior probability of the 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. 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

- 620 There are some broad similarities in diversity histories of insectivorous and omnivorous taxa. The
 621 diversity histories of arboreal, plantigrade, and scansorial insectivorous taxa all demonstrate a
 622 decreasing pattern with time, while fossorial insectivores have a flat diversity history with a peak
 623 approximately 10 Mya (Fig. 15). Arboreal and scansorial omnivores decrease in diversity from their
 624 initial peaks early in the Cenozoic, and plantigrade omnivores have a generally flat diversity history
 625 with a sudden peak in diversity late in the Cenozoic (Fig. 15). Unguligrade omnivores also
 626 demonstrate a possible decrease in diversity over the Cenozoic, but not as clearly as arboreal and
 627 scansorial omnivores.
- 628 Many of the estimated ecotype specific diversity histories share a similar increases in diversity to
 629 one degree or another at the late Cenozoic 16-14 Mya (Fig. 15); these increases are either sustained
 630 or temporary: digitigrade carnivores, plantigrade carnivores, scansorial carnivores, unguligrade
 631 herbivores, fossorial insectivores, and plantigrade omnivores.
- 632 When ecotype diversity is decomposed into the number of origination events per time bin (Fig. 16)

Table 11: Posterior probability of the differences in the log-odds of an ecotype surviving 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. 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

and the number of extinction events per time bin (Fig. 17) the estimates are clearly similar; there

634 are no obvious major cross-ecotype origination or extinction events, and there is no evidence of a
 sudden turnover as expected peaks in originations proceed peaks in peaks in the number of
 636 extinctions. Also, it is clear that the sustained increases in digitigrade and unguligrade herbivore
 diversity observed above (Fig. 15) is driven by an increase in the average number of originations as
 638 with a relatively constant number of extinctions over time (Fig. 16, 17).

Discussion

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

Table 12: Posterior probability the the effects of the two temperature covariates on the the log-odds of an ecotype occurring are greater than 0. 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

644 ecotypes and guilds, these roles are broad and not defined by specific interactions but by the genre
 of interactions species within that grouping participate in. The diversity of species within an
 646 ecotype or guild 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
 648 ecotype or guild in local communities, and the regional species pool as a whole, is preserved even as
 the individual interactors change. This also implies the structure of regional species pools can be
 650 constant over time despite a constantly changing set of “players.”

Comparison of the posterior predictive check results for the pure-presence and birth-death models
 652 supports the conclusion that regional species pool dynamics cannot simply be described by a single
 occurrence probability and is instead the result of the interplay between origination and extinction.
 654 Additionally, changes to the ecotypic composition and diversification rate for the North American
 regional species pool are driven primarily by variation in origination rates. These aspects of how

Table 13: Posterior probability that the effects of the two temperature covariates on the log-odds of an ecotype origination are greater than 0. 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

656 regional species pool diversity is shaped is not directly observable in studies of the Modern where
 time scales are short and macroevolutionary dynamics are inferable solely from phylogeny (Fritz
 658 et al., 2013).

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

Comparison of the ecotype specific diversity histories supports the conclusion that there was no
 666 large scale or sudden simultaneous reorganization of all local communities; instead these results
 support a more gradual and idiosyncratic shifts in relative ecotypic diversity over time (Fig. 15).

Table 14: Posterior probability that the effects of the two temperature covariates on the log-odds of an ecotype survival are greater than 0. 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

- 668 The closest examples to a sudden increase or decrease of a specific ecotype is the jump in standing
 diversity of scansorial carnivorans and, to a lesser extent, fossorial insectivores at 16 Mya (i.e. the
 670 start of the third plant phase). This result may, however, not reflect the dynamics of individual
 local communities as this is an analysis of the entire North American mammal regional species pool.
 672 Arboreal taxa disappear from the regional species pool over the Cenozoic, with massive
 disappearance by the Paleogene-Neogene transition ~22 Mya. This is consistent with one of the two
 674 possible patterns presented here and in Smits (2015) that would result in arboreal taxa having a
 greater extinction risk than other ecotypes: the Paleogene and Neogene were different selective
 676 regimes and while the earliest Cenozoic may have been neutral wrt arboreal taxa, they disappeared
 quickly over the Cenozoic which may account for their higher extinction risk. In addition to all
 678 arboreal taxa, the diversity of plantigrade and scansorial insectivores decreases with time (Fig. 15).
 Digitigrade carnivorans have a relatively stable diversity history through the Cenozoic and can be

680 characterized as varying around a constant mean diversity. This ecotype has a large amount of
overlap with the carnivore guild which has been the focus of much research CITATIONS. This
682 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).

684 Both digitigrade and unguligrade herbivores increase in diversity over the Cenozoic. The increase of
these cursorial forms is consistent with the gradual opening up of the North American landscape
686 (Blois and Hadly, 2009; Graham, 2011; Strömberg, 2005). These herbivore increase in diversity over
the Cenozoic which may be indicative of a long-term shift in the interactors associated with those
688 ecotypes leading to increased contribution to the regional species pool. This result may be
comparable to the increasing percentage of hypsodont (high-crowned teeth) mammals in the
690 Neogene of Europe being due to an enrichment of hyposodont taxa and not a depletion of
non-hypsodont taxa. Smaller scale increases in fossorial herbivore species, and a lesser extent
692 plantigrade herbivores, suggests that the increase of interactors may be associated mostly with the
herbivore dietary category with locomotor category tempering that relationship.

694 An association between plant phase and differences in ecotype occurrence or origination-extinction
probabilities is interpreted to mean that an ecotype enrichment or depletion is due to to
696 associations between that ecotype and whatever plants are dominate at that time and are thus a
contributing factor to the constancy of an ecotype, or the lack there of. Plant phase clearly
698 structures the occurrence and origination probability time series (Fig. 4, 5). These differences in
occurrence or origination translate opaquely to the estimates of diversity and diversification rate;
700 the largest spike in both diversity, diversification rate, and origination rate all correspond to the
onset of the last plant phase (Fig. 14). The clearest example of the diversity of an ecotype
702 increasing at this particular transition is in scansorial carnivores (Fig. 15); similar shifts in other
ecotypes are much more subtle, as was previously noted for fossorial insectivores. Interestingly, all
704 of the ecotypes with sudden changes to diversity at this transition increase in diversity, even if only
temporarily. My interpretation of this result is that, because plant phase associations are only with
706 occurrence or origination probabilities and not survival, these ecotypes were well suited for the
newly available mammal-plant interactions due to the increased modernization of their floral

708 context (Graham, 2011).

There are fewer, less obvious shifts in diversity surrounding the transition from the first to second,
710 with the following ecotypes having apparent shifts in diversity at 50 My: digitigrade carnivores
(down), plantigrade carnivores (down), plantigrade herbivores (up), arboreal omnivores (down), and
712 scansorial omnivores (down). Because plant phase has been found to structure
occurrence/origination (Fig. 4, 5), but not survival (Fig. 6, my interpretation of these results is that
714 new species were not entering the system because there were fewer available mammal-plant
interactions available for those ecotypes. Instead, these ecotypes were poorly suited for the newly
716 available mammal-plant interactions brought upon by the changing environmental context
(Graham, 2011).

718 The estimated effects of temperature on occurrence and origination-extinction probabilities are
similar to those of the plant phases. The occurrence and origination probabilities of many mammal
720 ecotypes have strong relationships with the two temperature covariates (Tables 12, 13). In most
cases, there is a negative association between temperature and probability of occurring or first
722 originating; this means that if temperature decreases, we would then expect the probability of
occurring or first originating would increase. Contrastingly, only temperature range are estimated
724 to be good predictors of survival in four mammal ecotypes and only marginally for two of those
(Table 14). Additionally, in all four of these cases are expected to have positive relationships,
726 meaning that if temperature decreases it is expected that species survival will also decrease.

What is the comparative size of the effects of plant phase and temperature are approximately equal
728 in importance in the sense that they have similar effect sizes on the ecotypes. The focus in previous
research on temperature and major climatic or geological events without other measures of
730 environmental context may have been a mistake and perhaps led to increasing confusion in
discussions of how the “environment” affects mammal diversity and diversification. The
732 environment or climate is not just global or regional temperature, it is the set of all possible biotic
and abiotic interactions that can be experienced by a member of the species pool. By including
734 more descriptors of species’ environmental context a more complete “picture” of the diversification

process is inferred.

736 The effect of species mass on either occurrence or origination and extinction was not allowed to
vary by ecotype or environmental context even though it is not known if this is the case or not
738 CITATION. The primary reason for this modeling choice was that this study focuses on ecotypic
based differences in either occurrence, or origination and extinction. Allowing the effect of body size
740 to vary by ecotype, time, and environmental factors would increase the overall complexity of the
model, something that I felt was not necessary because the overall scope of the study. Instead, body
742 size was included in order to control for its possible underlying effects (McElreath, 2016). A control
means that if there is variation due to body mass, having a term to “absorb” that effect is better
744 than ignoring it which may affect other parameter estimates. Additionally, the effect of body size
was allowed to have a second-order polynomial form and no higher order polynomials were
746 considered; 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. Finally, parameteric forms of nonlinearity
748 have not previously been considered, so the simple act of estimating a potential second-order
relationship is an opportunity to test more complex hypotheses of the relationship between body
750 size and both macroevolutionary and macroecological processes.

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

752 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
754 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
756 warrant their inclusion.

The potential effects of common ancestry (i.e. phylogeny) on origination and extinction are not
758 directly considered in this analysis. While a birth-death process approximates the
speciation-extinction process underlying the phylogeny (Silvestro et al., 2014) this is not same as
760 considering how the similarity between closely related species may affect the estimates of the effects
of species traits to environmental factors on both origination and extinction (Harnik et al., 2014;

⁷⁶² Smits, 2015). One of the principle barriers to the inclusion of the effect of phylogeny in either the pure-presence or birth-death models is computational; with well over 1000 tips, the calculation of
⁷⁶⁴ the scale parameter defining the phylogenetic effect would be very slow and further increase the already slow computation time necessary for both the marginalization of the discrete occurrence
⁷⁶⁶ histories and data augmentation already included in both models.

HERE GOES A CONCLUSION

⁷⁶⁸ Acknowledgements

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

⁷⁷⁴ References

- Allen, L. J. S. 2011. An introduction to stochastic processes with applications to biology. 2nd ed.
⁷⁷⁶ Chapman and Hall/CRC, Boca Raton, FL.
- Alroy, J. 1996. Constant extinction, constrained diversification, and uncoordinated stasis in North
⁷⁷⁸ American mammals. *Palaeogeography, Palaeoclimatology, Palaeoecology* 127:285–311.
- . 2009. Speciation and extinction in the fossil record of North American mammals. Pages
⁷⁸⁰ 302–323 *in* R. K. Butlin, J. R. Bridle, and D. Schlüter, eds. *Speciation and patterns of diversity*. Cambridge University Press, Cambridge.
- ⁷⁸² Alroy, J., P. L. Koch, and J. C. Zachos. 2000. Global climate change and North American
mammalian evolution. *Paleobiology* 26:259–288.

- 784 Badgley, C., and J. A. Finarelli. 2013. Diversity dynamics of mammals in relation to tectonic and
climatic history: comparison of three Neogene records from North America. *Paleobiology*
786 39:373–399.
- Badgley, C., T. M. Smiley, R. Terry, E. B. Davis, L. R. G. Desantis, D. L. Fox, S. S. B. Hopkins,
788 T. Jezkova, M. D. Matocq, N. Matzke, J. L. McGuire, A. Mulch, B. R. Riddle, V. L. Roth, J. X.
Samuels, C. A. E. Strömberg, and B. J. Yanites. 2017. Biodiversity and Topographic Complexity:
790 Modern and Geohistorical Perspectives. *Trends in Ecology & Evolution* pages 1–16.
- Bambach, R. K. 1977. Species richness in marine benthic habitats through the Phanerozoic.
792 *Paleobiology* 3:152–167.
- Bambach, R. K., A. M. Bush, and D. H. Erwin. 2007. Autecology and the filling of ecospace: Key
794 metazoan radiations. *Palaeontology* 50:1–22.
- Blois, J. L., and E. A. Hadly. 2009. Mammalian Response to Cenozoic Climatic Change. Annual
796 Review of Earth and Planetary Sciences 37:181–208.
- Brook, B. W., and D. M. J. S. Bowman. 2004. The uncertain blitzkrieg of Pleistocene megafauna.
798 *Journal of Biogeography* 31:517–523.
- Brown, A. M., D. I. Warton, N. R. Andrew, M. Binns, G. Cassis, and H. Gibb. 2014. The
800 fourth-corner solution - using predictive models to understand how species traits interact with
the environment. *Methods in Ecology and Evolution* 5:344–352.
- 802 Brown, J. H., and B. A. Maurer. 1989. Macroecology: the division of food and space among species
on continents. *Science* 243:1145–1150.
- 804 Bush, A. M., and R. K. Bambach. 2011. Paleoecologic Megatrends in Marine Metazoa, vol. 39.
- Bush, A. M., R. K. Bambach, and G. M. Daley. 2007. Changes in theoretical ecospace utilization in
806 marine fossil assemblages between the mid-Paleozoic and late Cenozoic. *Paleobiology* 33:76–97.
- Bush, A. M., and P. M. Novack-Gottshall. 2012. Modelling the ecological-functional diversification
808 of marine Metazoa on geological time scales. *Biology Letters* 8:151–155.

- Cantalapiedra, J. L., J. L. Prado, and M. T. Alberdi. 2017. Decoupled ecomorphological evolution
810 and diversification in Neogene-Quaternary horses. *Science* 355:627–630.
- Carrano, M. T. 1999. What, if anything, is a cursor? Categories versus continua for determining
812 locomotor habit in mammals and dinosaurs. *Journal of Zoology* 247:29–42.
- Clyde, W. C., and P. D. Gingerich. 1998. Mammalian community response to the latest Paleocene
814 thermal maximum: an isotaphonomic study in the northern Bighorn Basin, Wyoming. *Geology*
26:1011–1014.
- 816 Cramer, B. S., K. Miller, P. Barrett, and J. Wright. 2011. Late Cretaceous-Neogene trends in deep
ocean temperature and continental ice volume: Reconciling records of benthic foraminiferal
818 geochemistry ($\delta^{18}\text{O}$ and Mg/Ca) with sea level history. *Journal of Geophysical Research: Oceans*
116:1–23.
- 820 Eronen, J. T., C. M. Janis, C. P. Chamberlain, and A. Mulch. 2015. Mountain uplift explains
differences in Palaeogene patterns of mammalian evolution and extinction between North
822 America and Europe. *Proceedings of the Royal Society B: Biological Sciences* 282:20150136.
- Ezard, T. H. G., A. Purvis, and H. Morlon. 2016. Environmental changes define ecological limits to
824 species richness and reveal the mode of macroevolutionary competition. *Ecology Letters*
19:899–906.
- 826 Felsenstein, J. 1985. Phylogenies and the comparative method. *The American Naturalist* 125:1–15.
- Figueirido, B., C. M. Janis, J. A. Pérez-Claros, M. De Renzi, and P. Palmqvist. 2012. Cenozoic
828 climate change influences mammalian evolutionary dynamics. *Proceedings of the National
Academy of Sciences* 109:722–727.
- 830 Foote, M. 2001. Inferring temporal patterns of preservation, origination, and extinction from
taxonomic survivorship analysis. *Paleobiology* 27:602–630.
- 832 Foote, M., and J. J. Sepkoski. 1999. Absolute measures of the completeness of the fossil record.
Nature 398:415–7.

- 834 Fraser, D., R. Gorelick, and N. Rybczynski. 2015. Macroevolution and climate change influence
phylogenetic community assembly of North American hoofed mammals. *Biological Journal of the*
Linnean Society 114:485–494.
- 836
- Freudenthal, M., and E. Martín-Suárez. 2013. Estimating body mass of fossil rodents. *Scripta*
Geologica 145:1–130.
- 838
- Fritz, S. A., J. Schnitzler, J. T. Eronen, C. Hof, K. Böhning-Gaese, and C. H. Graham. 2013.
840 Diversity in time and space: wanted dead and alive. *Trends in Ecology & Evolution* 28:509–16.
- Gelman, A. 2008. Scaling regression inputs by dividing by two standard deviations. *Statistics in*
842 *Medicine* pages 2865–2873.
- Gelman, A., J. B. Carlin, H. S. Stern, D. B. Dunson, A. Vehtari, and D. B. Rubin. 2013. Bayesian
844 data analysis. 3rd ed. Chapman and Hall, Boca Raton, FL.
- Gelman, A., and J. Hill. 2007. Data Analysis using Regression and Multilevel/Hierarchical Models.
846 Cambridge University Press, New York, NY.
- Gelman, A., J. Hill, and M. Yajima. 2012. Why We (Usually) Don't Have to Worry About Multiple
848 Comparisons. *Journal of Research on Educational Effectiveness* 5:189–211.
- Gelman, A., and E. Loken. 2014. The statistical crisis in science: data-dependent analysis – a
850 “garden of forking paths” – explains why many statistically significant comparisons don't hold up.
American Scientist 102:460–465.
- 852 Graham, A. 2011. A natural history of the New World: the ecology and evolution of plants in the
Americas. University of Chicago Press, Chicago.
- 854 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.
- 856 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*
858 *Vegetation Science* 24:988–1000.

- Janis, C., J. Damuth, and J. M. Theodor. 2004. The species richness of Miocene browsers, and
860 implications for habitat type and primary productivity in the North American grassland biome.
Palaeogeography, Palaeoclimatology, Palaeoecology 207:371–398.
- 862 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.
- 864 ———. 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.
- 866 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
868 97:7899–904.
- 869 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.
- 872 Janis, C. M., K. M. Scott, and L. L. Jacobs. 1998. Evolution of Tertiary mammals of North
America. Vol. 1. Terrestrial carnivores, ungulates, and ungulatelike mammals. Cambridge
874 University Press, Cambridge.
- 876 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.
- 878 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.
- 880 Legendre, S. 1986. Analysis of mammalian communities from the Late Eocene and Oligocene of
Southern France. Paleovertebrata 16:191–212.
- 882 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.

- 884 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.
- 886 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.
- 888 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*
890 40:237–254.
- 892 McElreath, R. 2016. Statistical rethinking: a Bayesian course with examples in R and Stan. CRC
Press, Boca Raton, FL.
- 894 McKenna, R. T. 2011. Potential for Speciation in Mammals Following Vast , Late Miocene Volcanic
Interruptions in the Pacific Northwest. Masters. Portland State University.
- 896 Novack-Gottshall, P. M. 2007. Using a theoretical ecospace to quantify the ecological diversity of
Paleozoic and modern marine biotas Using a theoretical ecospace to quantify the ecological
diversity of Paleozoic and modern marine biotas. *Paleobiology* 33:273–294.
- 898 Pires, M. M., D. Silvestro, and T. B. Quental. 2015. Continental faunal exchange and the
asymmetrical radiation of carnivores. *Proceedings of the Royal Society B: Biological Sciences*
900 282:20151952.
- 902 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.
- 904 Quental, T. B., and C. R. Marshall. 2013. How the Red Queen Drives Terrestrial Mammals to
Extinction. *Science* 341:290–292.
- 906 Raia, P., F. Carotenuto, F. Passaro, D. Fulgione, and M. Fortelius. 2012. Ecological specialization
in fossil mammals explains Cope’s rule. *The American Naturalist* 179:328–37.
- 908 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.

- . 2012. Parameter-expanded data augmentation for Bayesian analysis of capture-recapture models. *Journal of Ornithology* 152:521–537.
- 910
- 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.
- 912
- 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.
- 914
- Rubin, D. B. 1996. Multiple imputation after 18+ years. *Journal of the American Statistical Association* 91:473–489.
- 916
- Silvestro, D., A. Antonelli, N. Salamin, and T. B. Quental. 2015. The role of clade competition in the diversification of North American canids. *Proceedings of the National Academy of Sciences of the United States of America* 112:8684–9.
- 918
- Silvestro, D., J. Schnitzler, L. H. Liow, A. Antonelli, and N. Salamin. 2014. Bayesian estimation of speciation and extinction from incomplete fossil occurrence data. *Systematic biology* 63:349–67.
- 920
- Simberloff, D., and T. Dayan. 1991. The Guild Concept and the Structure of Ecological Communities. *Annual Review of Ecology and Systematics* 22:115–143.
- 922
- Slater, G. J. 2015. Iterative adaptive radiations of fossil canids show no evidence for diversity-dependent trait evolution. *Proceedings of the National Academy of Sciences* 112:4897–4902.
- 924
- Smith, F. A., J. Brown, J. Haskell, and S. Lyons. 2004. Similarity of mammalian body size across the taxonomic hierarchy and across space and time. *The American Naturalist* 163:672–691.
- 926
- 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.
- 928
- Strömberg, C. A. E. 2005. Decoupled taxonomic radiation and ecological expansion of open-habitat grasses in the Cenozoic of North America. *Proceedings of the National Academy of Sciences of the United States of America* 102:11980–4.
- 930

- 934 Tomiya, S. 2013. Body Size and Extinction Risk in Terrestrial Mammals Above the Species Level.
The American Naturalist 182:196–214.
- 936 Valentine, J. W. 1969. Patterns of taxonomic and ecological structure of the shelf benthos during
Phanerozoic time. Paleontology 12:684–709.
- 938 Villéger, S., P. M. Novack-Gottshall, and D. Mouillot. 2011. The multidimensionality of the niche
reveals functional diversity changes in benthic marine biotas across geological time. Ecology
940 letters 14:561–8.
- Wang, S. C., P. J. Everson, H. J. Zhou, D. Park, and D. J. Chudzicki. 2016. Adaptive credible
942 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
944 Letters 12:20150989.
- Warton, D. I., B. Shipley, and T. Hastie. 2015. CATS regression - a model-based approach to
946 studying trait-based community assembly. Methods in Ecology and Evolution 6:389–398.
- Wilson, J. B. 1999. Guilds, functional types and ecological groups. Oikos 86:507–522.
- 948 Zachos, J. C., G. R. Dickens, and R. E. Zeebe. 2008. An early Cenozoic perspective on greenhouse
warming and carbon-cycle dynamics. Nature 451:279–283.
- 950 Zachos, J. C., M. Pagani, L. Sloan, E. Thomas, and K. Billups. 2001. Trends, rhythms, and
aberrations in global climate 65 Ma to present. Science 292:686–693.

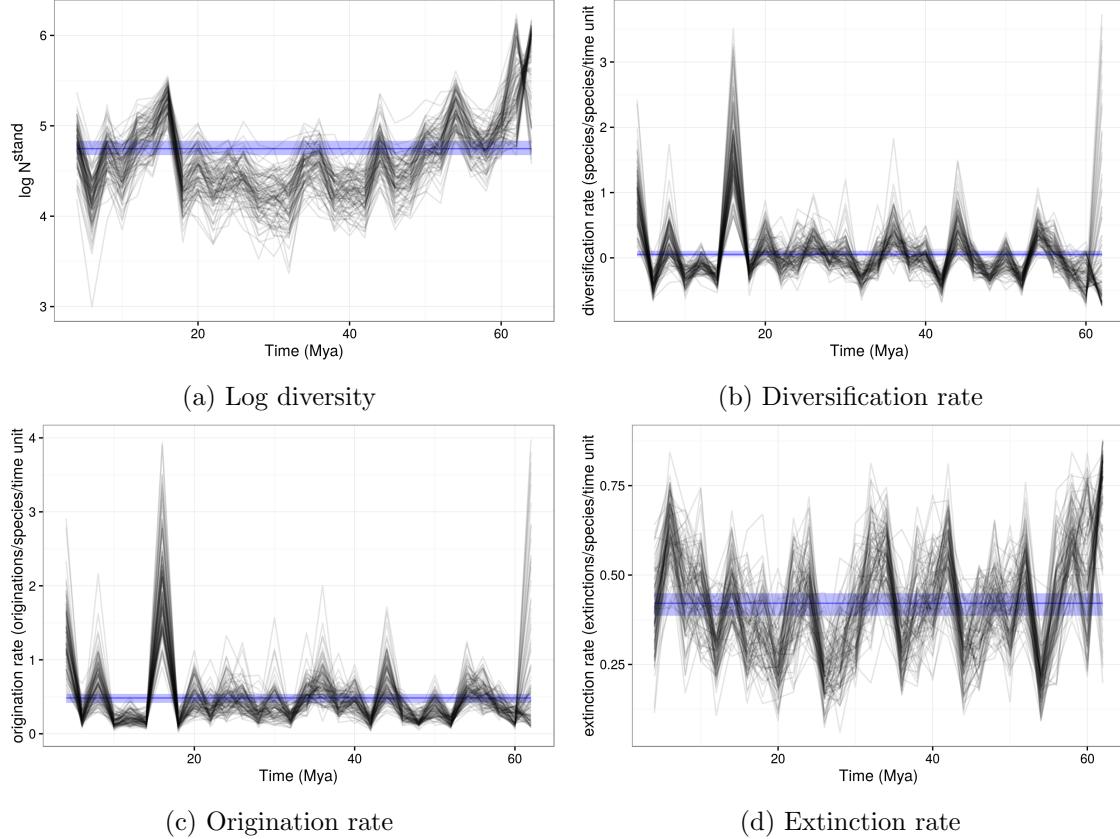


Figure 14: Posterior estimates of the time series of Cenozoic North American mammal diversity and its characteristic macroevolutionary rates; all estimates are from the birth-death model and 100 posterior draws are plotted to indicate the uncertainty in these estimates. The blue horizontal strip corresponds 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 15: 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	

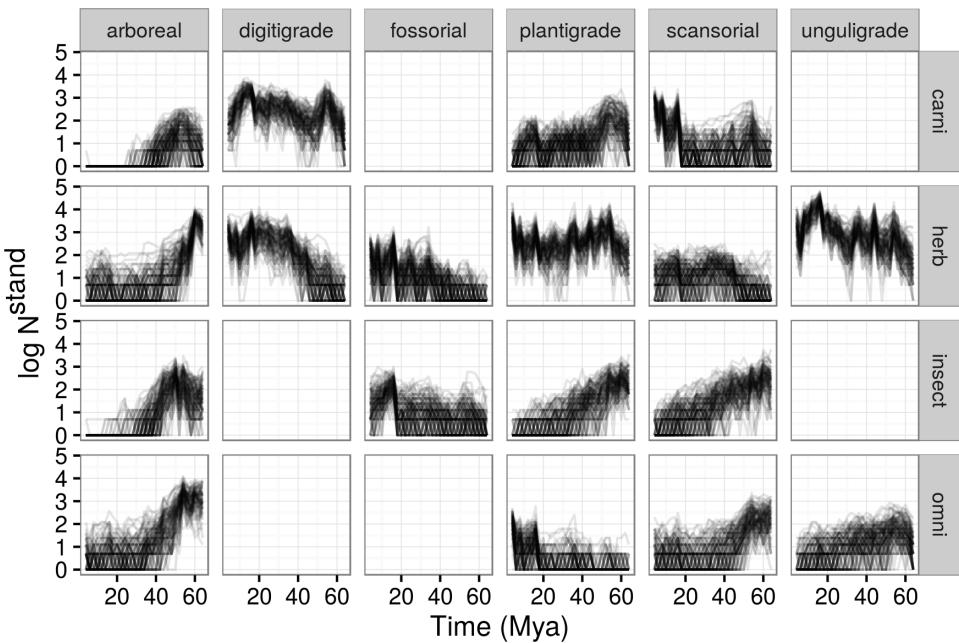


Figure 15: Posterior of standing log-diversity of North American mammals by ecotype for the Cenozoic as estimated from the birth-death model; 100 posterior draws are plotted to indicate the uncertainty in these estimates and what is technically plotted is log of diversity plus 1.

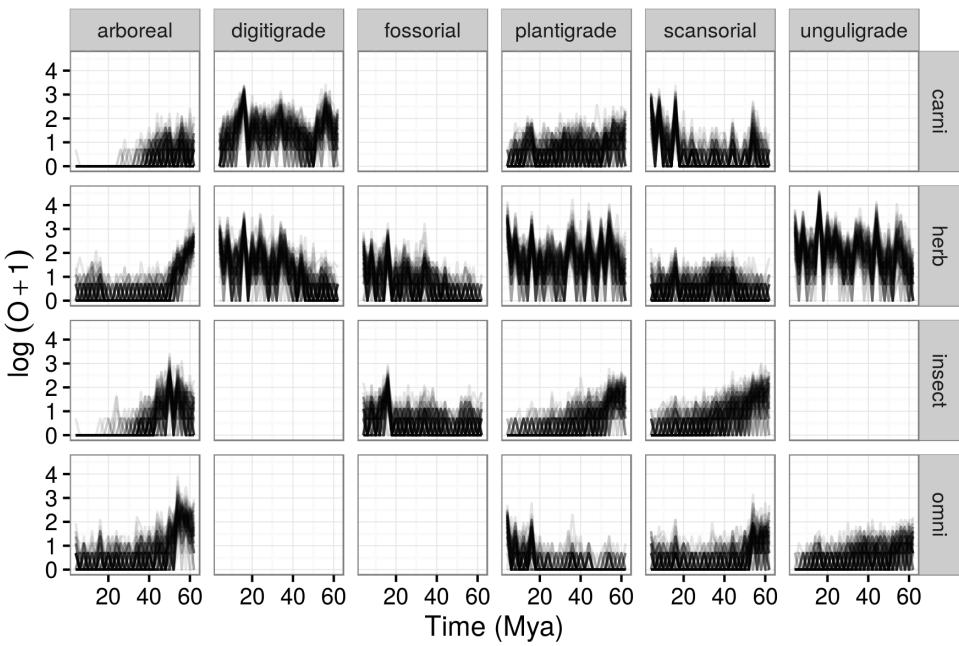


Figure 16: Posterior estimates of the number of origination events from one temporal bin to another, plotted at the bin they originate from. 100 posterior draws are plotted to indicate the uncertainty in these estimates. Also, what is plotted is log of the number of originations plus 1.

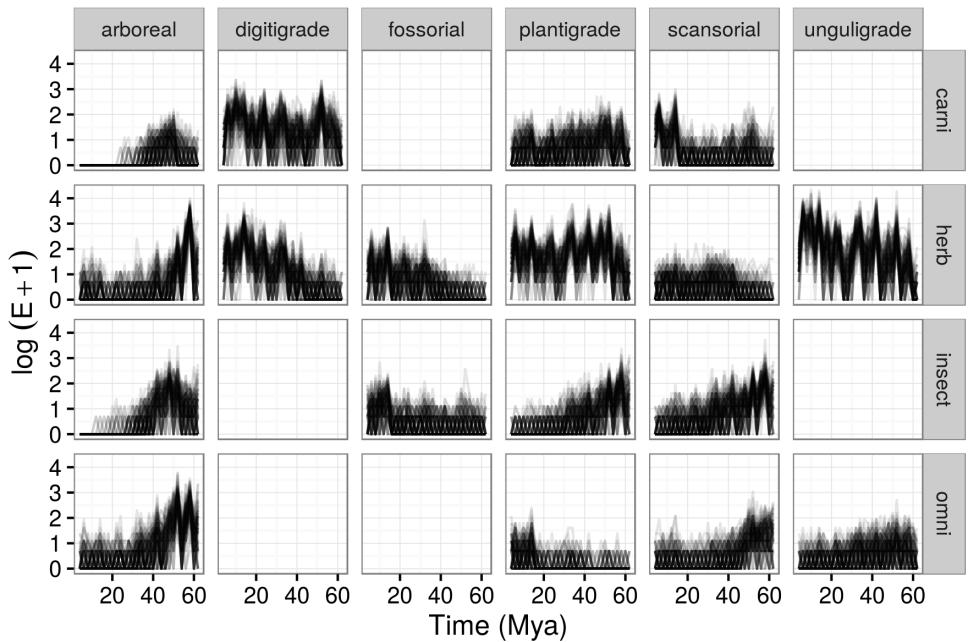


Figure 17: Posterior estimates of the number of extinction events from one temporal bin to another, plotted at the bin they go extinct from. 100 posterior draws are plotted to indicate the uncertainty in these estimates. Also, what is plotted is \log of the number of extinctions plus 1.