

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

Peter D. Smits^{1,*}

1. University of Chicago, Chicago, Illinois 60637.

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

Manuscript elements:

Keywords:

Manuscript type: Article

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

Introduction

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

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

Abstract

6 Place holder text.

A regional species pool is the set of species which form communities in a specific region. Local scale
8 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
10 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?
12 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
14 diversity and diversification.

Functional diversity is frequently broken into or thought of as a set of guilds, which are a set of
16 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
18 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
20 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
22 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
24 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;
26 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.

28 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)
30 (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
32 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;
34 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
36 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
38 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
40 mammal species diversity and diversification.

The principle species trait considered in this study is a species' ecotype, defined here as the unique
42 combination of species dietary category and locomotor category (e.g. arboreal omnivore versus
unguligrade herbivore). These classifications can be considered analogous to guilds or unique
44 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
46 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
48 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
50 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
52 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

- 54 the lens of species interactions. Specifically, this issue arrises when trying to generalize previous
observations from taxonomy-based framework to ecology-based one.
- 56 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
- 58 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
60 expected then that, for this analysis, turnover events or periods of rapid diversification or depletion
should not occur simultaneously for all ecotypes.
- 62 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
64 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
66 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
68 HERE. WHAT DO PEOPLE THINK THE MAMMAL DIVERSITY CURVE REPRESENTS?
CAN BE ANYTHING IF YOU THINK ABOUT IT HARD ENOUGH.
- 70 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
72 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
74 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.,
76 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
78 communities towards the end of the Cenozoic and the latter emergence of “modern” environments
and communities (Janis and Wilhelm, 1993).
- 80 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
82 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
84 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
86 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
88 MORE TO THIS STORY; CHECK THE WORK OF VANVALKENBURG; DANCES WITH
WOLF AVATARS?

90 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;
92 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
94 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).
96 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
98 between life history traits and body size limits processed-based inference because the actual causal
mechanisms underlying an observed pattern are obscured or missing.
100 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
102 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
104 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
106 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
108 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
110 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,
112 it should be possible to distinguish amongst these hypotheses.

Fundamentally, all species respond differently to climate and environmental change (Blois and
114 Hadly, 2009). Macroecological patterns are emergent patterns due to the similarities among species
in how they respond to a similar “stimulus.”

116 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
118 climate-mediated (Alroy, 1996; Alroy et al., 2000; Clyde and Gingerich, 1998; Figueirido et al.,
2012). However, differences in temporal and geographic scale seem to underly the contrast between
120 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,
122 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
124 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
126 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
128 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
130 CITATION. 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 that
132 mechanistic or process based inference is impossible. While emergent patterns might correspond to
taxonomic grouping, this itself is an emergent phenomenon. Instead, by framing hypotheses in
134 terms of species traits and their environmental context, these emergent phenomenon can be
observed rather than assumed.

¹³⁶ The climate history of the Cenozoic is generally characterized by a global cooling trend and the
¹³⁸ development of polar ice-caps during the Neogene; there are, of course, a few notable exceptions to
¹⁴⁰ 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
¹⁴² 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;
Strömberg, 2005).

A lot of the climate and environmental changes observed for North America have been attributed to
¹⁴⁴ 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
¹⁴⁶ 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,
¹⁴⁸ or high crowned teeth, among herbivore groups over the Cenozoic of both North America and
Europe CITATIONS.

¹⁵⁰ 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
¹⁵² 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
¹⁵⁴ 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
¹⁵⁶ 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
¹⁵⁸ 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
¹⁶⁰ 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
¹⁶² 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
164 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)
166 allowing for inference as to how a species ecology can mediate selective pressures do to its
environmental context.

168 Fourth-corner modeling is an approach to explaining the patterns of either species abundance or
presence/absence as a product of species traits, environmental factors, and the interaction between
170 traits and environment (Brown et al., 2014; Jamil et al., 2013; Pollock et al., 2012; Warton et al.,
2015); effectively uniting species distribution modeling (SDMs) with trait-based community
172 assembly (CATS). In modern ecological studies, what is being modeled is species occurrences at
localities distributed across a region (Jamil et al., 2013; Pollock et al., 2012). In this study, what is
174 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
176 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
178 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
180 means that these approaches are complimentary but reveal different patterns of how species are
distributed in time and space.

182 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
184 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
186 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
188 observed (Foote, 2001; Foote and Sepkoski, 1999; Wang et al., 2016; Wang and Marshall, 2016).

Ultimately, the goals of this analysis are to understand when are unique ecotypes enriched or

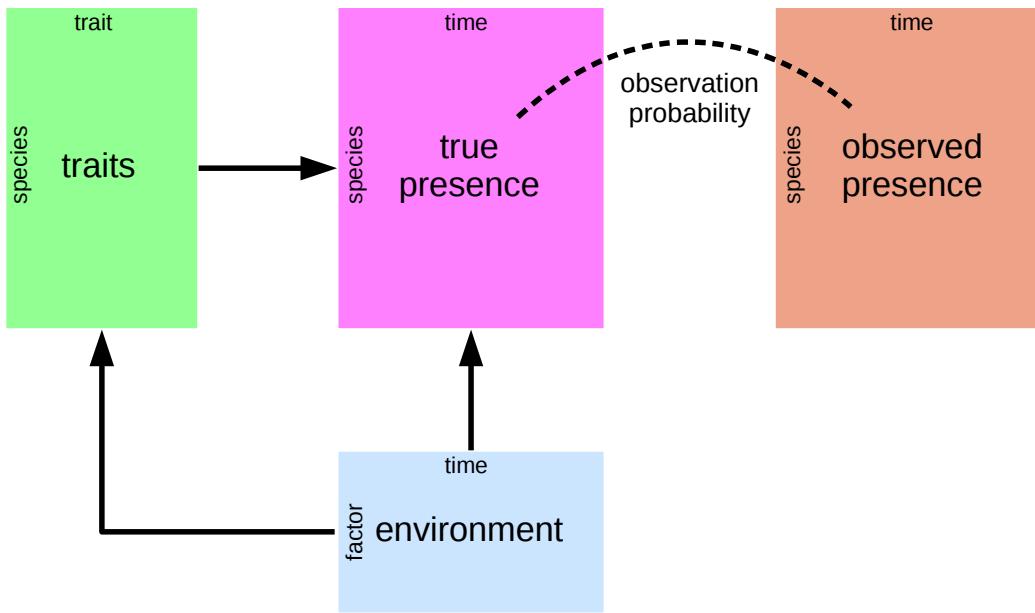


Figure 1: Conceptual diagram of the paleontological fourth corner problem. The observed presence matrix (orange) is the empirical presence/absence pattern for all species for all time points; this matrix is an incomplete observation of the “true” presence/absence pattern (purple). The estimated true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

190 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
 192 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
 194 hierachal Bayesian model combing the forth-corner modeling approach with a model of an
 observation-occurrence or observation-originatation-extinction process. The complexity and nuance
 196 inherent in questions that 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

¹⁹⁸ mitigating complications arising from both this complexity and the plethora of testable hypotheses
(e.g. multiple comparisons, garden of forking paths) CITATIONS.

²⁰⁰ **Materials and Methods**

Taxon occurrences and species-level information

²⁰² 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
²⁰⁴ 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,
²⁰⁶ 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
²⁰⁸ <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.

²¹⁰ After download, the raw occurrence data was then sorted, cleaned, and manipulated
programmatically before analysis. Many species taxonomic assignments as present in the raw
²¹² 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
²¹⁴ 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.
²¹⁶ 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
²¹⁸ pools.

Species ecotype is defined here as the interaction between life habit and diet categories, the goal of
²²⁰ 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
²²² 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
 224 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
 226 associated with their taxonomic group as described in Table 2 (Carrano, 1999). Ecotype categories
 with less than 10 total species of that combination were excluded, yielding a total of 18 observed
 228 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
	Aplodontidae	plantigrade
	Apternodontidae	scansorial

Continued on next page

Table 2 – continued from previous page

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

Continued on next page

Table 2 – continued from previous page

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

Continued on next page

Table 2 – continued from previous page

Order	Family	Stance
	Mustelidae	plantigrade
	Mylagaulidae	fossorial
	Mylodontidae	plantigrade
	Nimravidae	digitigrade
	Nothrotheriidae	plantigrade
Notoungulata		unguligrade
	Oromerycidae	unguligrade
	Oxyaenidae	digitigrade
	Palaeomerycidae	unguligrade
	Palaeoryctidae	plantigrade
	Pampatheriidae	plantigrade
	Pantolambdidae	plantigrade
	Peritychidae	digitigrade
Perissodactyla		unguligrade
	Phenacodontidae	unguligrade
Primates		plantigrade
	Procyonidae	plantigrade
	Proscalopidae	plantigrade
	Protoceratidae	unguligrade
	Reithroparamyidae	plantigrade
	Sciuravidae	plantigrade
	Sciuridae	plantigrade
	Simimyidae	plantigrade
	Soricidae	plantigrade
	Suidae	digitigrade

Continued on next page

Table 2 – continued from previous page

Order	Family	Stance
	Talpidae	fossorial
	Tayassuidae	unguligrade
	Tenrecidae	plantigrade
	Titanoideidae	plantigrade
	Ursidae	plantigrade
	Viverravidae	plantigrade
	Zapodidae	plantigrade

²³⁰ 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
²³² (Brook and Bowman, 2004; Freudenthal and Martín-Suárez, 2013; McKenna, 2011; Raia et al., 2012; Tomiya, 2013; ?); this is similar to what was done in Smits (2015). When species mass was not
²³⁴ 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 mass given
²³⁶ 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). Mass was
²³⁸ 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 covariates are
²⁴⁰ 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)
²⁴² 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).

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	?

244 Environmental and temporal covariates

The environmental covariates used in this study are collectively referred to as group-level covariates
 246 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
 248 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 phases”
 250 defined by Graham CITATION. Global temperature across most of the Cenozoic was calculated
 from Mg/Ca isotope record from deep sea carbonates (Cramer et al., 2011). Mg/Ca based
 252 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
 254 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.
 256 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
 258 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
260 mean and then dividing by twice the standard deviation.

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

270 Modelling species occurrence

Two different models were used in this study: a pure-presence model and a birth-death model. Both
272 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
274 or surviving are considered equal or different (Table 4). Something that is important to realize is
that while there are only two state "codes" in a presence-absence matrix (i.e. 0/1), there are in fact
276 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);
278 this is made obvious in the transition matrices as the probability of an extinct species changing
states is 0 (Table 4). See below for parameter explanations (Tables 6, and 7).

280 Data augmentation

All empirical presence/absence observations are potentially incomplete or observed with error. The
282 hidden Markov model at the core of this analysis allows for observed absences to be used
meaningfully to estimate the number of unobserved species. Of concern in this analysis is the

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

(a) Pure-presence

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

(b) Birth-death

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

- 284 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
286 case of paleontological data, there are obviously many species which were never sampled (Royle and
Dorazio, 2008; Royle et al., 2007).
- 288 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
290 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
292 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,
294 $M = \lfloor 1.25 \times N \rfloor$.

Data imputation is the process of estimating missing data for partially observed covariates given
296 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
298 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
300 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.
302 This approach is an example of imputing covariate information that is missing completely at

Table 5: Observation parameters

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

random (Gelman and Hill, 2007; Royle and Dorazio, 2012). Because log mass values are rescaled as

- 304 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
- 306 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.
- 308 This classification has no biological interpretation.

Observation process

- 310 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
- 312 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
- 314 $\phi = \pi$, 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
316 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 6: Parameters for the model of presence in the pure-presence model

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

318 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 5.

320 Pure-presence process

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

The species-level of the model (Eq. 2) is a logistic regression where the intercept varies by ecotype.
 326 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
 328 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

330 choice of priors for these regression coefficients.

The values of each ecotype's intercept are themselves modeled as regressions using the group-level
332 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
334 covariances between ecotype intercepts, given this group-level regression, are modeled (Gelman and
Hill, 2007).

336 All parameters not modeled elsewhere were given weakly informative priors (Gelman et al., 2013)
CITATION STAN MANUAL STATISTICAL RETHINKING. Weakly informative means that
338 priors do not necessarily encode actual prior information but instead help regularize or weakly
constrain posterior estimates. These priors have a concentrated probability density around and near
340 zero; this has the effect of tempering our estimates and help prevent overfitting the model to the
data (Gelman et al., 2013) CITATION STAN MANUAL STATISTICAL RETHINKING. The
342 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 statement, excluding the imputation of body
344 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

- ³⁴⁶ 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
³⁴⁸ (Eq. 2, Table 4b). The parameters associated with the birth-death presence model are presented in Table 7 and the full sampling statement, including observation (Eq. 1), is described in Equation 3:

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

- ³⁵⁰ 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
³⁵² between mass and both ϕ and π are reflected in the parameterization of the model and choice of priors (Eq. 3).
³⁵⁴ 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 7: Parameters for the model of presence in the pure-presence model

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

356 pure presence model (Eq. 2).

Posterior inference and model adequacy

358 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
360 feature latent discrete parameters in the large matrix z (Tables 5, 6, 7; Eqs. 1, 2, 3). All methods
for posterior inference implemented in Stan are derivative based which causes complications for
362 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
364 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
366 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

368 are calculated, from the end members of the species having existed for the entire study interval and
 370 the species having only existed between the directly observed FAD and LAD to all possible
 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
 372 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),
 374 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
 376 implemented. Instead, an approximate Bayesian approach was used: variational inference. A
 recently developed automatic variational inference algorithm called “automatic differentiation
 378 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

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 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).

Of additional 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 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 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, 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 approach might be more analytically sound than the more ad-hoc approaches that are occasionally used to overcome this hurdle (Royle and Dorazio, 2008). Additionally, because $p_{t=1}$ and $p_{t=T}$ are 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 approach used here may help correct for this reality by drawing these estimates towards the overall means of those parameters.

After fitting both models (Eqs. 2, 3) using ADVI, model adequacy and quality of fit was assessed using a simple posterior predictive check CITATION CITATION. Because all Bayesian models are inherently generative, simulations of new data sets is “free” CITATION. By simulating 100 theoretical data sets using the observed covariate information the congruence 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 $t = 1$; 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 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 simulated

⁴⁰⁸ 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.

- ⁴¹⁰ 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
- ⁴¹² distribution and are instead just an approximation CITATIONS. Because of this, posterior predictive estimates are themselves only approximate checks of model adequacy. The posterior
- ⁴¹⁴ 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.
- ⁴¹⁶ Given parameter estimates, diversity and diversification rates are estimated through posterior predictive simulations. Given the observed presence-absence matrix y , estimates of the true
- ⁴¹⁸ 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
- ⁴²⁰ 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

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

- ⁴²² Given estimates of N^{stand} for all time points, the estimated number of originations O_t are be estimated as

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

- ⁴²⁴ 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

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

426 Results

The results of the analyses described above take one of two forms: direct inspection of parameter
428 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
430 explained below in the comparison of the models' posterior predictive check results.

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

Comparison of the posterior predictive results from the pure-presence and birth-death models
434 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
436 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
438 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
440 parameter estimates from both models gives further insight into the reasons for this difference in
posterior predictive results (Gelman et al., 2013).

442 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
444 probability estimates (Fig. 6). This result supports the idea that changes to the North American

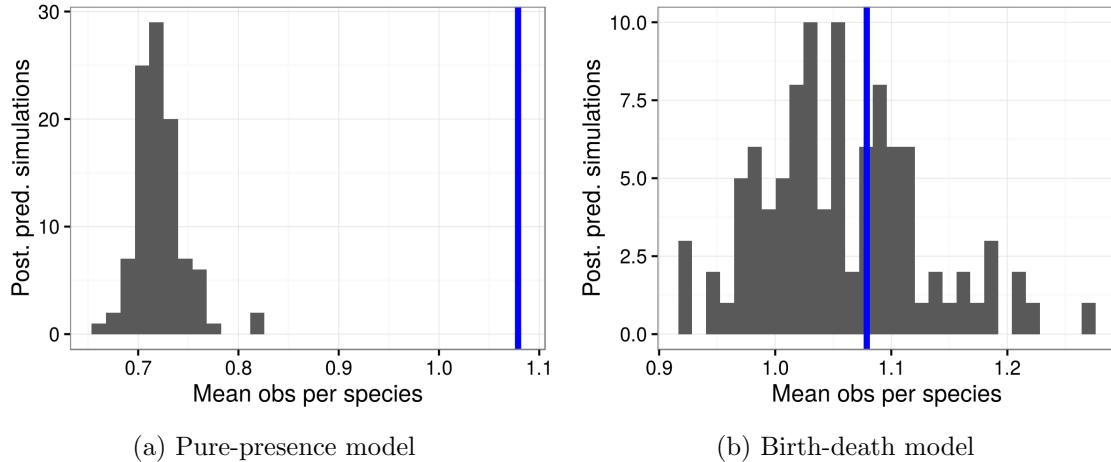


Figure 3: Comparison of the average observed number of occurrences per species (blue line) to the average number of occurrences from 100 posterior predictive datasets using the posterior estimates from the pure-presence and birth-death models.

regional species pool is more likely due to changes in origination than extinction, a result that is

446 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)

448 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

450 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

452 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

454 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

456 no longer originating or originating very rarely.

The dramatic differences in the estimates origination and survival probabilities are indicative of

458 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.

460 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
 462 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
 464 estimates of survival probabilities (Fig. 5, and 6). In logistic regression, high uncertainty in the
 estimates of the underlying log-odds of occurrence, origination, or survival tends to be indicative of
 466 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
 468 effect of which has been mitigated by the hierarchical modeling strategy used here (Gelman et al.,
 2013; Gelman and Hill, 2007) CITATION Statistical Rethinking.

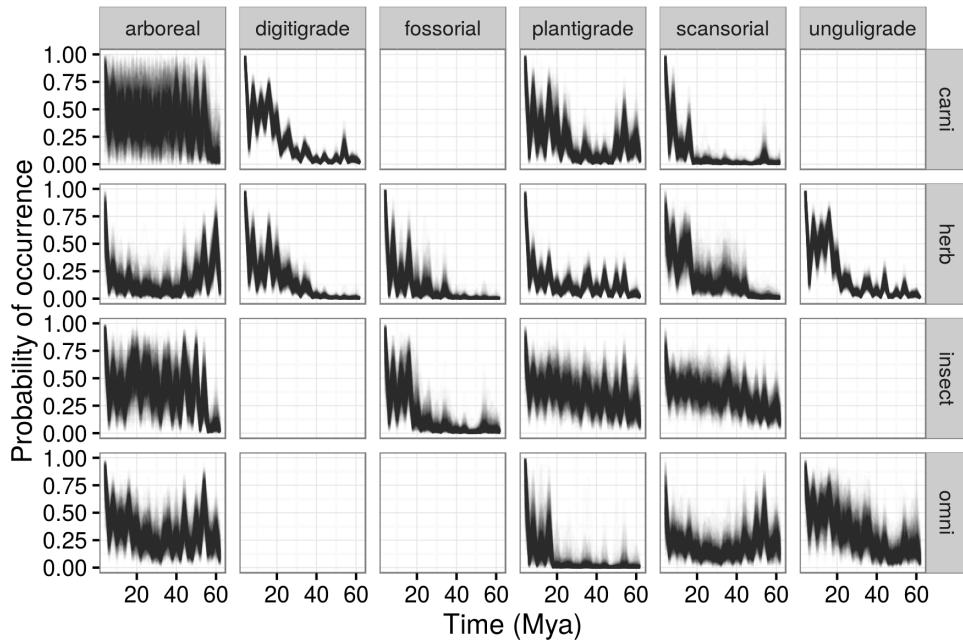


Figure 4: Probability of a mammal ecotype occurring over time as estimated from the pure-presence model. Each panel depicts 100 random samples from the model's posterior. The columns are by locomotor category and rows by dietary category; their intersections are the observed and analyzed ecotypes. Panels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

470 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
 472 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

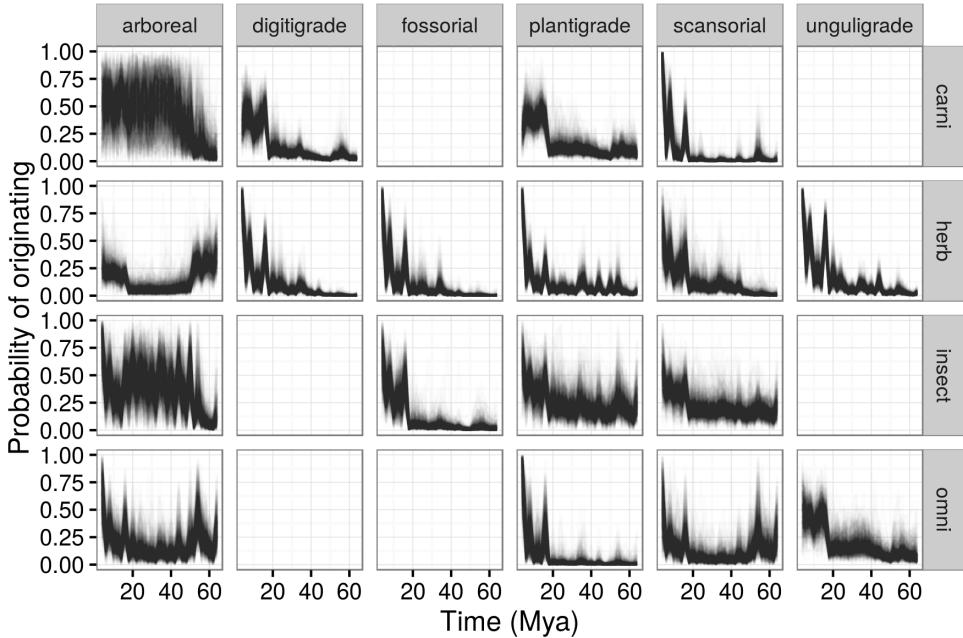


Figure 5: Probability of a mammal ecotype origination probabilities at each time point as estimated from the birth-death model. Each panel depicts 100 random samples from the model’s posterior. The columns are by locomotor category and rows by dietary category; their intersections are the observed and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

- 474 observation such that larger species are less likely to be observed if present than smaller species
 (Fig. 7b).
- 476 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
- 478 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
 480 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
 482 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.
- 484 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,

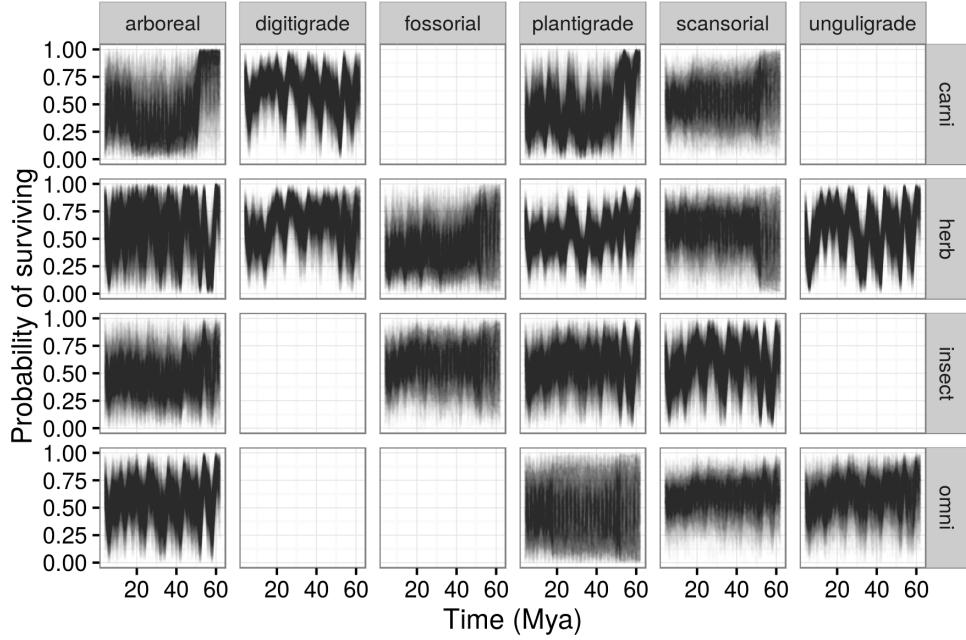


Figure 6: Probability of a mammal ecotype survival probabilities at each time point as estimated from the birth-death model. Each panel depicts 100 random samples from the model’s posterior. The columns are by locomotor category and rows by dietary category; their intersections are the observed and analyzed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

486 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
 488 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
 490 parameter estimates in the model.

The effect of species mass on probability of occurrence as estimated from the pure-presence (Fig. 8)
 492 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
 494 probability of occurrence for species with body sizes closer to the mean than either extremes. This
 result is consistent with the canonically normal distribution of mammal body sizes CITATION; it is
 496 then expected that the most likely to occur species would be those from the middle of the
 distribution, and that species originating will on average be of average mass, especially considering
 498 species shared common ancestry CITATION. Note that all variation in estimates between ecotypes

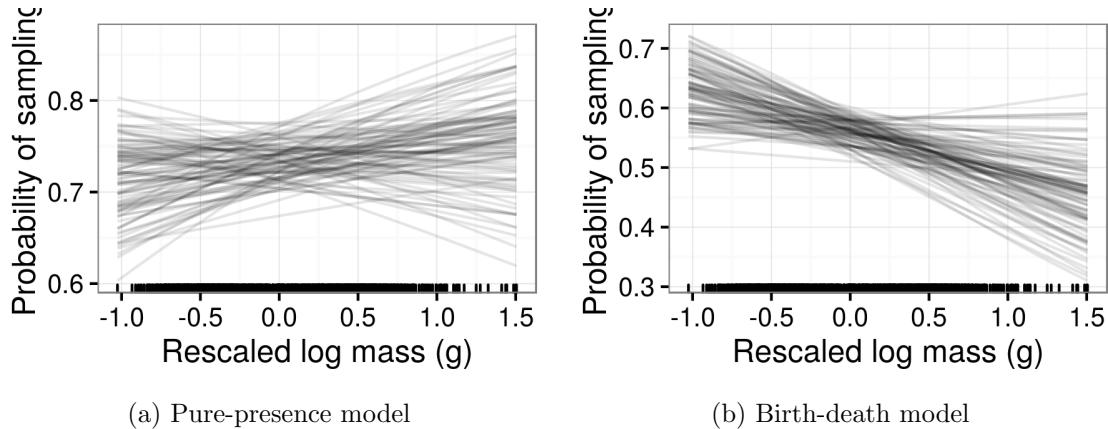


Figure 7: Estimates of the effect of species mass on probability of sampling a present species (p). Mass has been log-transformed, centered, and rescaled; this means that a mass of 0 corresponds to the mean of log-mass of all observed species and that mass is in standard deviation units. Estimates are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

(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.

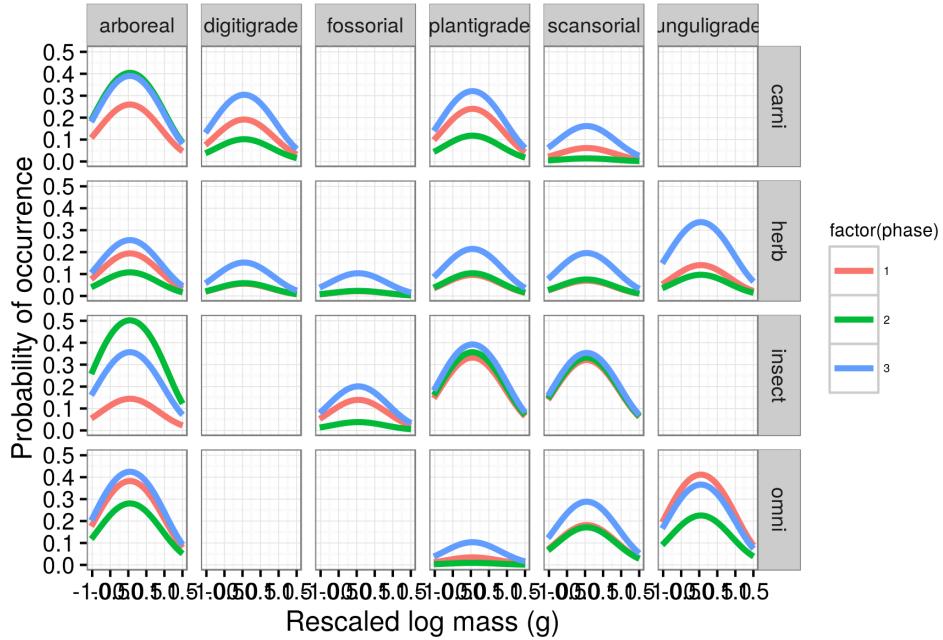


Figure 8: Mean estimate of the effect of species mass on the probability of a species occurrence for each of the three plant phases. The effect of mass is considered constant over time and that the only aspect of the model that changes with plant phase is the intercept of the relationship between mass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

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

516 increases vs decreases? Are their similarities along ecotype axes?

Which ecotypes are associated with a strong effect of either of the temperature covariates? Are

518 these relationships positive (inc temp, inc occ) or negative (dec temp, inc occ)? Are their similarities along ecotype axes?

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

522 Which ecotypes are associated with a strong effect of either of the temperature covariates? Are these relationships positive (inc temp, inc orig) or negative (dec temp, inc orig)? Are their 524 similarities along ecotype axes?

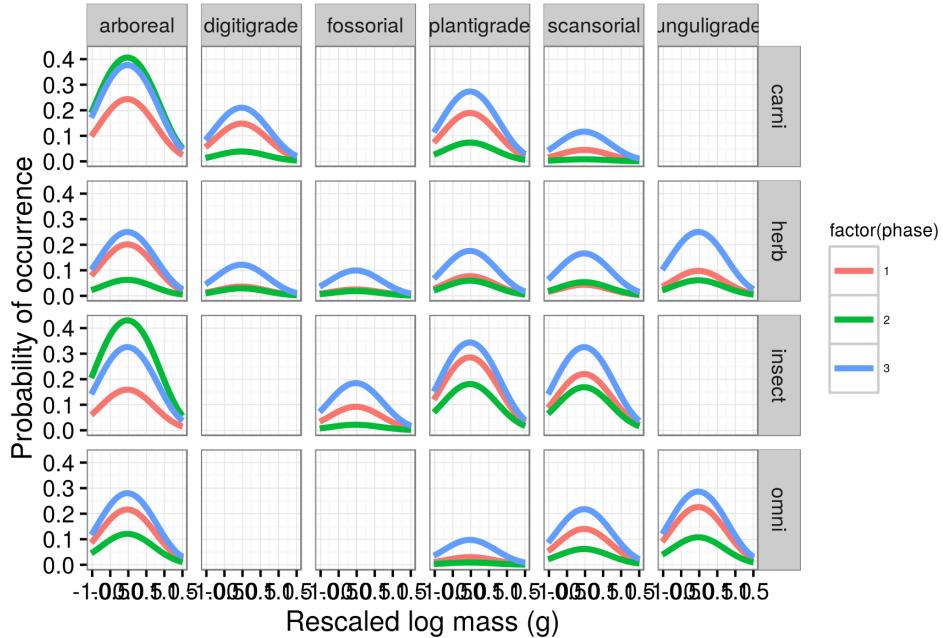


Figure 9: Mean estimate of the effect of species mass on the probability of a species originating for each of the three plant phases. The effect of mass is considered constant over time and that the only aspect of the model that changes with plant phase is the intercept of the relationship between mass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

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

526 increases vs decreases? Are their similarities along ecotype axes?

Which ecotypes are associated with a strong effect of either of the temperature covariates? Are

528 these relationships positive (inc temp, inc surv) or negative (dec temp, inc surv)? Are their similarities along ecotype axes?

530 Analysis of diversity

All of the following analyses of diversification and macroevolutionary rates has been done using

532 only the birth-death model; this is because of the models better posterior predictive check performance (Fig. 3a, and 3b).

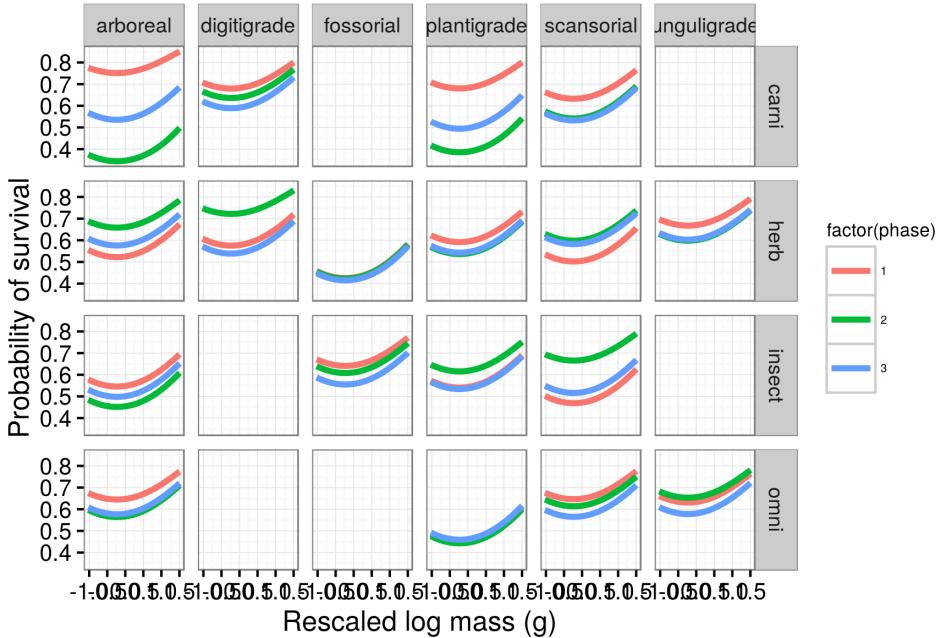


Figure 10: Mean estimate of the effect of species mass on the probability of a species survival for each of the three plant phases. The effect of mass is considered constant over time and that the only aspect of the model that changes with plant phase is the intercept of the relationship between mass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

- 534 The general pattern of total estimated North American mammal diversity for the Cenozoic is
 “stable” in that mean standing diversity does not fluctuate wildly and rapidly over the Cenozoic
 536 (Fig. 14a). In broad strokes, the first 15 or so million years of the Cenozoic are characterized by a
 gradual decline in standing diversity until approximately 45-50 million years ago (early-middle
 538 Eocene). Following this decline, standing diversity is broadly constant from 45 to 18 Mya (early
 Miocene). After this, there is a rapid spike in diversity followed by a slight decline in diversity up to
 540 the Modern. This characterization of the estimated diversity history is knowingly broad strokes and
 diversity time series is not without variation and vagaries.
- 542 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

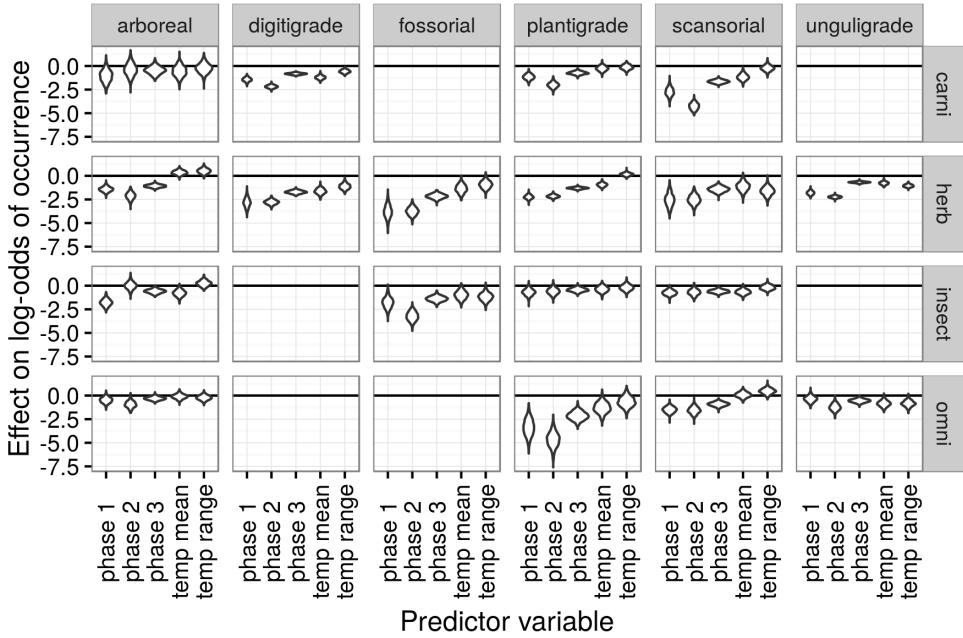


Figure 11: Estimated effects of the group-level covariates describing environmental context on log-odds of species occurrence. These estimates are from the pure-presence model.

- 544 hovers around zero, punctuated by both positive and negative spikes. The largest spike in
diversification rate is at 16 Mya, which is early Oligocene (Fig. 14b). Other notable increases in
546 diversification rate occur at WHEN? Mya (Table 14); other possible increases in diversification rate
are less certain (e.g. 8 Mya). Notable decreases in diversification rate occur at WHEN? (Table 14).
- 548 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 origination
550 and survival probabilities, diversification rate seems most driven by changes in origination rate as
opposed to extinction rate. Extinction rate, on the other hand, demonstrates an almost saw-toothed
552 pattern around a constant mean.

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

- 554 Diversity partitioned by ecotype reveals a lot of the complexity behind the pattern of mammal
diversity for the Cenozoic (Fig. 15).
- 556 Arboreal ecotypes obtain peak diversity early in the Cenozoic and then decline for the rest of the

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

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

time series, becoming increasingly rare or absent as diversity approaches the Modern (Fig. 15).

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

562 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,
 564 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
 566 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
 568 (Fig.15). Between these two peaks digitigrade carnivore diversity dips below average diversity

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

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

following the first peak and then grows slowly until the second peak. Plantigrade carnivores obtain

570 peak diversity in the early Cenozoic and then maintain a relatively stable diversity until another peak at the end of the Cenozoic.

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

580 Many of the estimated ecotype specific diversity histories share a similar increases in diversity to

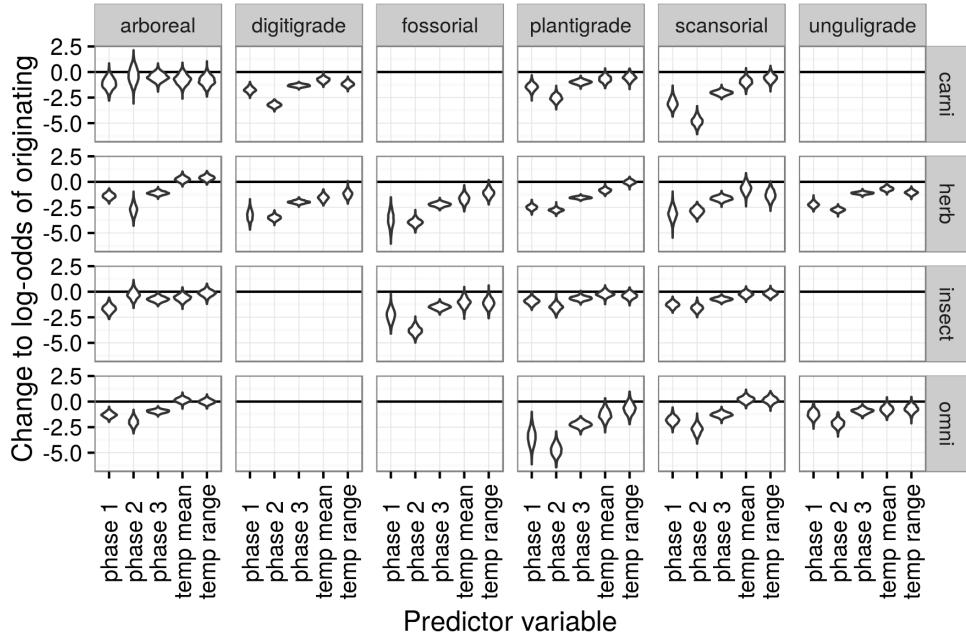


Figure 12: Estimated effects of the group-level covariates describing environmental context on log-odds of species origination. These estimates are from the birth-death model.

one degree or another at the late Cenozoic 16-14 Mya (Fig. 15); these increases are either sustained
 582 or temporary: digitigrade carnivores, plantigrade carnivores, scansorial carnivores, unguilgrade
 herbivores, fossorial insectivores, and plantigrade omnivores.

584 Discussion

Both the composition of a species pool and its environmental context change over time, though not
 586 necessarily at the same rate. Local communities, who's species are drawn from the regional species
 pool, have “roles” in their communities defined by their interactions with a host of biotic and abiotic
 588 interactors (i.e. species niche). For higher level ecological characterizations like ecotypes and guilds,
 these roles are broadly defined and not defined by specific interactions but the genre of interactions
 590 that species within that grouping participate in. The diversity of species within an ecotype or guild
 can be stable over millions of years despite constant species turnover (Jernvall and Fortelius, 2004;
 592 Slater, 2015) CITATIONS. This implies that the size and scope of the role of an ecotype or guild is

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

	P(Phase 1 > Phase 2)	P(Phase 2 > Phase 3)	P(Phase 1 > Phase 3)
arboreal carnivore	0.460	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

preserved even as the individual interactors change. This also implies the structure of regional

594 species pools can be constant over time despite a constantly changing set of “players.”

Comparison of the pure-presence model to the birth-death model support the conclusion that

596 regional species pool dynamics cannot simply be described by a single occurrence probability and is
 instead better modeled as the result of both origination and extinction. Additionally, changes to
 598 ecotypic composition of the North American regional species pool are driven primarily by variation
 in origination rates. These aspects of how regional species pool diversity is shaped is not observable
 600 from studies of the Modern CITATION.

Extinction rate for the entire regional species pool through time is highly variable and demonstrates

602 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
 604 extinction rate is actually variable may not have been equally appreciated. What is most consistent

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

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

with previous observations (Alroy, 1996; Alroy et al., 2000), however, is that diversity seems to be

most structured by origination than extinction.

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

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

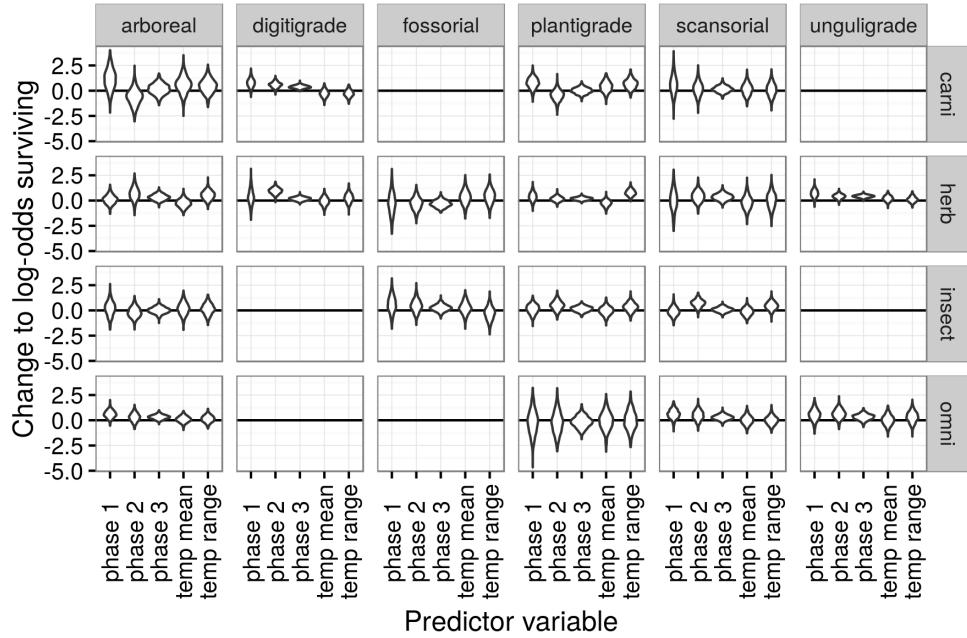


Figure 13: Estimated effects of the group-level covariates describing environmental context on log-odds of species survival. These estimates are from the birth-death model.

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

624 What these results support is a gradual change to the ecotypic diversity of the regional species pool
 for the Cenozoic. The rapidity of Cenozoic environmental change is worth discussing. If change is
 626 rapid, ecotypic composition of species pool does not seem to track environmental change. If change
 is gradual then there is the possibility that changes to ecotypic composition may be tracking
 628 environmental change.

If plant phase is associated with differences in ecotype occurrence this is interpreted to mean that
 630 ecotype enrichment or depletion is due to associations between that ecotype and whatever plants

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

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

are dominate at that time.

- 632 Temperature affects very few of the occurrence, origination, or survival probabilities of the mammal
 633 ecotypes except for a negative relationship between temperature and the origination probabilities of
 634 digitigrade carnivores, and both digitigrade and unguligrade herbivores. The origination
 635 probabilities and diversity of these three groups all increase over the Cenozoic as average global
 636 temperature decreased. This result coupled with the lack of relationship between temperature and
 637 the other ecotypes may be responsible for the continued confusion surrounding the impact of
 638 temperature on mammal diversity and diversification (Alroy, 1996; Alroy et al., 2000; Blois and
 639 Hadly, 2009; Figueirido et al., 2012; Janis and Wilhelm, 1993).
 640 What is the comparative size of the effects of plant phase and temperature? Both seem of “equal”
 641 importance in the sense that they have similar effect sizes on the ecotypes. Perhaps focusing on
 642 temperature and not considering other measures of environmental context has been a mistake and

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

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

perhaps led to increasing confusion in discussions of how “environment” effects mammal diversity

644 and diversification. The environment or climate is not just global or regional temperature, it is the
 set of all possible biotic and abiotic interactions. By including more descriptors of species’
 646 environmental context a more complete “picture” of the diversification process is inferred.

The effect of species mass on either occurrence or origination and extinction was not allowed to

648 vary by ecotype or environmental context even though it is not known if this is the case or not
 CITATION. The primary reason for this modeling choice was that this study focuses on ecotypic
 650 based differences in either occurrence, or origination and extinction. Allowing the effect of body size
 to vary by ecotype, time, and environmental factors would increase the overall complexity of the
 652 model, something that I felt was not necessary because the overall scope of the study. Instead, body
 size was included in order to control for its possible underlying effects CITATION. A control means
 654 that if there is variation due to body mass, having a term to “absorb” that effect is better than

ignoring it which may affect other parameter estimates. Additionally, the effect of body size was
656 allowed to have a second-order polynomial form and no higher order polynomials were considered;
this was done because it is hard to conceive of a more complex third- or higher-order relationship
658 between body size and the other parameters. Additionally, nonlinearity is rarely if ever considered
in the first place, so the simple act of estimating a potential second-order relationship is an
660 opportunity to test more complex hypotheses of the effects of body size on macroevolutionary and
macroecological processes.

662 The only covariate allowed to affect sampling probability was mass and only as a linear predictor.
Other covariates, such as the environmental factors considered here, could have affected the
664 underlying preservation process that limits sampling probability. Their exclusion as covariates of
sampling/observation was the product of a few key decisions: model complexity, model
666 interpretability, the scope of this study, and a lack of good hypotheses related to these covariates to
warrant their inclusion. It should be noted that in other similar studies that use a hidden
668 birth-death model to handle simultaneous estimation of sampling, origination, and extinction have
not considered the possible effects of covariates, both species traits and environmental factors, on
670 sampling CITATION.

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

The potential effects of common ancestry (i.e. phylogeny) on origination and extinction are not
678 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
680 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;

682 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
684 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
686 histories and data augmentation already included in both models.

Phylogenetic comparative community ecology and phylogenetic comparative biogeography also
688 discusses how the macroevolutionary processes helps structure an observed community, though it is
not necessarily phrased that way. However, that community did not form in isolation but it the
690 result of many factors interacting over time including incumbency, competition, limiting similarity,
etc.

692 Acknowledgements

I would like to thank K. Angielczyk, M. Foote, P. D. Polly, and R. Ree for helpful discussion and
694 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
696 M. Uhen for curating most of the mammal occurrences recorded in the PBDB. This is Paleobiology
Database publication XXX.

698 References

- Allen, L. J. S. 2011. An introduction to stochastic processes with applications to biology. 2nd ed.
700 Chapman and Hall/CRC, Boca Raton, FL.
- Alroy, J. 1996. Constant extinction, constrained diversification, and uncoordinated stasis in North
702 American mammals. *Palaeogeography, Palaeoclimatology, Palaeoecology* 127:285–311.
- . 2009. Speciation and extinction in the fossil record of North American mammals. Pages

- 704 302–323 *in* R. K. Butlin, J. R. Bridle, and D. Schluter, eds. *Speciation and patterns of diversity*.
Cambridge University Press, Cambridge.
- 706 Alroy, J., P. L. Koch, and J. C. Zachos. 2000. Global climate change and North American
mammalian evolution. *Paleobiology* 26:259–288.
- 708 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*
710 39:373–399.
- 712 Badgley, C., T. M. Smiley, R. Terry, E. B. Davis, L. R. G. Desantis, D. L. Fox, S. S. B. Hopkins,
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:
714 Modern and Geohistorical Perspectives. *Trends in Ecology & Evolution* pages 1–16.
- 716 Bambach, R. K. 1977. Species richness in marine benthic habitats through the Phanerozoic.
Paleobiology 3:152–167.
- 718 Bambach, R. K., A. M. Bush, and D. H. Erwin. 2007. Autecology and the filling of ecospace: Key
metazoan radiations. *Palaeontology* 50:1–22.
- 720 Blois, J. L., and E. A. Hadly. 2009. Mammalian Response to Cenozoic Climatic Change. *Annual
Review of Earth and Planetary Sciences* 37:181–208.
- 722 Brook, B. W., and D. M. J. S. Bowman. 2004. The uncertain blitzkrieg of Pleistocene megafauna.
Journal of Biogeography 31:517–523.
- 724 Brown, A. M., D. I. Warton, N. R. Andrew, M. Binns, G. Cassis, and H. Gibb. 2014. The
fourth-corner solution - using predictive models to understand how species traits interact with
the environment. *Methods in Ecology and Evolution* 5:344–352.
- 726 Brown, J. H., and B. A. Maurer. 1989. Macroecology: the division of food and space among species
on continents. *Science* 243:1145–1150.
- 728 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
730 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
732 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
734 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
736 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
738 thermal maximum: an isotaphonomic study in the northern Bighorn Basin, Wyoming. *Geology*
26:1011–1014.
- 740 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
742 geochemistry ($\delta^{18}\text{O}$ and Mg/Ca) with sea level history. *Journal of Geophysical Research: Oceans*
116:1–23.
- 744 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
746 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
748 species richness and reveal the mode of macroevolutionary competition. *Ecology Letters*
19:899–906.
- 750 Figueirido, B., C. M. Janis, J. A. Pérez-Claros, M. De Renzi, and P. Palmqvist. 2012. Cenozoic
climate change influences mammalian evolutionary dynamics. *Proceedings of the National
752 Academy of Sciences* 109:722–727.

- Foote, M. 2001. Inferring temporal patterns of preservation, origination, and extinction from
754 taxonomic survivorship analysis. *Paleobiology* 27:602–630.
- Foote, M., and J. J. Sepkoski. 1999. Absolute measures of the completeness of the fossil record.
756 *Nature* 398:415–7.
- Fraser, D., R. Gorelick, and N. Rybczynski. 2015. Macroevolution and climate change influence
758 phylogenetic community assembly of North American hoofed mammals. *Biological Journal of the Linnean Society* 114:485–494.
- 760 Freudenthal, M., and E. Martín-Suárez. 2013. Estimating body mass of fossil rodents. *Scripta Geologica* 145:1–130.
- 762 Gelman, A. 2008. Scaling regression inputs by dividing by two standard deviations. *Statistics in Medicine* pages 2865–2873.
- 764 Gelman, A., J. B. Carlin, H. S. Stern, D. B. Dunson, A. Vehtari, and D. B. Rubin. 2013. Bayesian data analysis. 3rd ed. Chapman and Hall, Boca Raton, FL.
- 766 Gelman, A., and J. Hill. 2007. Data Analysis using Regression and Multilevel/Hierarchical Models. Cambridge University Press, New York, NY.
- 768 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.
- 770 Jamil, T., W. A. Ozinga, M. Kleyer, and C. J. F. Ter Braak. 2013. Selecting traits that explain species-environment relationships: A generalized linear mixed model approach. *Journal of Vegetation Science* 24:988–1000.
772
- Janis, C., J. Damuth, and J. M. Theodor. 2004. The species richness of Miocene browsers, and implications for habitat type and primary productivity in the North American grassland biome. *Palaeogeography, Palaeoclimatology, Palaeoecology* 207:371–398.
- 776 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.

- 778 ———. 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.
- 780 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*
782 97:7899–904.
- 784 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.
- 786 Janis, C. M., K. M. Scott, and L. L. Jacobs. 1998. Evolution of Tertiary mammals of North
America. Vol. 1. Terrestrial carnivores, ungulates, and ungulate-like mammals. Cambridge
788 University Press, Cambridge.
- 790 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.
- 792 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.
- 794 Legendre, S. 1986. Analysis of mammalian communities from the Late Eocene and Oligocene of
Southern France. *Paleovertebrata* 16:191–212.
- 796 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.
- 798 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.
- 800 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.
- 802 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*
804 40:237–254.
- McKenna, R. T. 2011. Potential for Speciation in Mammals Following Vast , Late Miocene Volcanic
806 Interruptions in the Pacific Northwest. Masters. Portland State University.
- Novack-Gottshall, P. M. 2007. Using a theoretical ecospace to quantify the ecological diversity of
808 Paleozoic and modern marine biotas Using a theoretical ecospace to quantify the ecological
diversity of Paleozoic and modern marine biotas. *Paleobiology* 33:273–294.
- 810 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*
812 282:20151952.
- Pollock, L. J., W. K. Morris, and P. A. Vesk. 2012. The role of functional traits in species
814 distributions revealed through a hierarchical model. *Ecography* 35:716–725.
- Quental, T. B., and C. R. Marshall. 2013. How the Red Queen Drives Terrestrial Mammals to
816 Extinction. *Science* 341:290–292.
- Raia, P., F. Carotenuto, F. Passaro, D. Fulgione, and M. Fortelius. 2012. Ecological specialization
818 in fossil mammals explains Cope’s rule. *The American Naturalist* 179:328–37.
- Royle, J. A., and R. M. Dorazio. 2008. Hierarchical modeling and inference in ecology: the analysis
820 of data from populations, metapopulations and communities. Elsevier, London.
- . 2012. Parameter-expanded data augmentation for Bayesian analysis of capture-recapture
822 models. *Journal of Ornithology* 152:521–537.
- Royle, J. A., R. M. Dorazio, and W. a. Link. 2007. Analysis of Multinomial Models With Unknown
824 Index Using Data Augmentation. *Journal of Computational and Graphical Statistics* 16:67–85.
- Royle, J. A., J. D. Nichols, M. Kéry, E. Ranta, and M. Kery. 2014. detection is of species when
826 Modelling occurrence and abundance imperfect 110:353–359.

- Rubin, D. B. 1996. Multiple imputation after 18+ years. *Journal of the American Statistical Association* 91:473–489.
- 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.
- 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.
- Simberloff, D., and T. Dayan. 1991. The Guild Concept and the Structure of Ecological Communities. *Annual Review of Ecology and Systematics* 22:115–143.
- 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.
- 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.
- 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.
- Tomiya, S. 2013. Body Size and Extinction Risk in Terrestrial Mammals Above the Species Level. *The American Naturalist* 182:196–214.
- Valentine, J. W. 1969. Patterns of taxonomic and ecological structure of the shelf benthos during Phanerozoic time. *Paleontology* 12:684–709.
- 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 letters* 14:561–8.

- Wang, S. C., P. J. Everson, H. J. Zhou, D. Park, and D. J. Chudzicki. 2016. Adaptive credible
852 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*
854 *Letters* 12:20150989.
- Warton, D. I., B. Shipley, and T. Hastie. 2015. CATS regression - a model-based approach to
856 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.
- 858 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.
- 860 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.

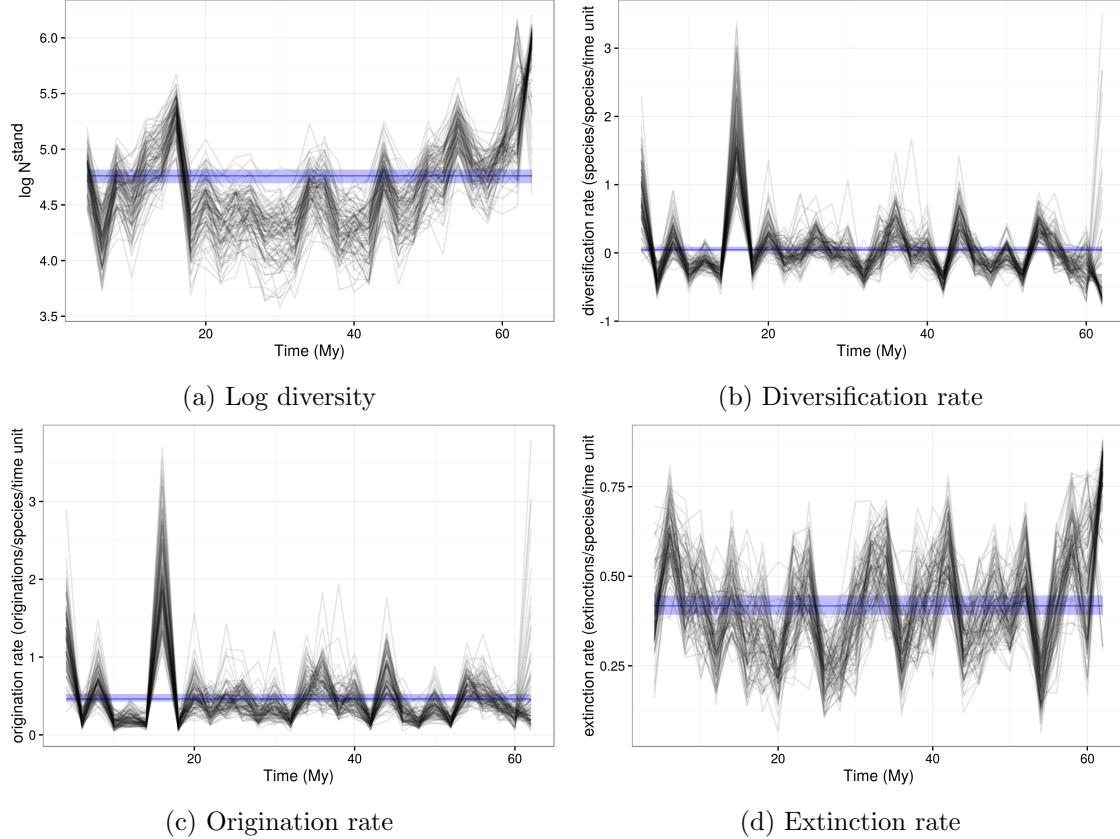


Figure 14: Posterior estimates of the time series of Cenozoic North American mammal diversity and its characteristic macroevolutionary rates; all estimates are from the birth-death model and 100 posterior draws are plotted to indicate the uncertainty in these estimates. The blue horizontal strip corresponds 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 14: Posterior probabilities of diversity N_t^{stand} or diversification rate D_t^{rate} being greater than average standing diversity \bar{N}^{stand} or average diversification rate \bar{D}^{rate} for the whole Cenozoic. The “Time” column corresponds to the top of each of the temporal bins. Diversification rate can not be estimated for the last time point because it is unknown how many more species originated or went extinct following this temporal bin. The estimates are from the birth-death model.

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

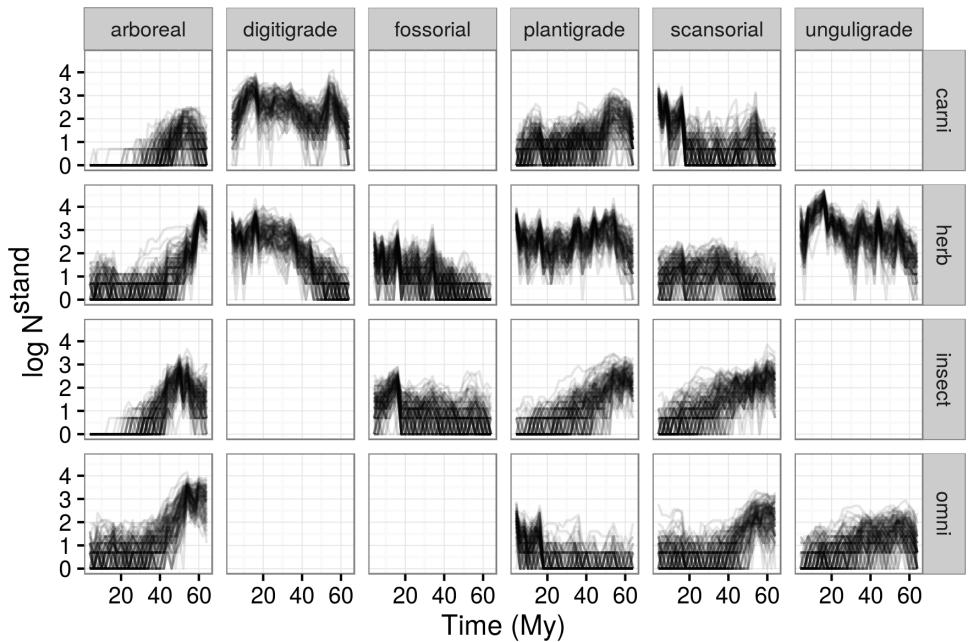


Figure 15: Posterior of standing log-diversity of North American mammals by ecotype for the Cenozoic as estimated from the birth-death model; 100 posterior draws are plotted to indicate the uncertainty in these estimates and what is technically plotted is log of diversity plus 1.