

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

Peter D. Smits<sup>1,\*</sup>

1. University of Chicago, Chicago, Illinois 60637.

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

*Manuscript elements:*

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

*Manuscript type:* Article

Prepared using the suggested L<sup>A</sup>T<sub>E</sub>X template for *Am. Nat.*

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

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

4 **Abstract**

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

## Introduction

30 A regional species pool is the set of species which form communities in a specific region (Harrison  
and Cornell, 2008; Mittelbach and Schemske, 2015). Local scale processes like resource competition  
32 only affect the regional species pool if all communities are affected. The taxonomic and functional  
composition of a regional species pool changes over time due to speciation, migration, and  
34 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? How does global and regional  
36 environmental context affect the set of species ecotypes (e.g. guilds) in a regional species pool? I  
address these questions through an analysis of the ecological and functional diversity of North  
38 American mammals for most of the Cenozoic, or last 65 million years.

Functional diversity is frequently thought of as a set of guilds, which themselves are a set of species  
40 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 guild  
42 are expected to have more similar macroecological dynamics than species in a different guilds.

Building on the concept of guilds and a macroecological niche, Bush et al. (2007) presented a  
44 three-dimensional construct, or ecocube, for describing the macroecological role of a marine  
invertebrate species by their physical position (i.e. tiering), motility, and trophic role. Unique  
46 combinations along the three ecological trait axes indicate which among the possible ecotypes are  
observed. This approach has proven quite popular as it attempts to operationalize the guild concept  
48 in terms of shared characteristics that are indicative of the type of interactions experience by  
species of that macroecology (Bambach et al., 2007; Bush and Bambach, 2011; Bush et al., 2007;  
50 Bush and Novack-Gottshall, 2012; Novack-Gottshall, 2007; Villéger et al., 2011), but the utility of  
this approach is limited due to its condition as just a data type.

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

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

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

Translating previous work into hypotheses applicable to this analysis is difficult for a variety of  
72 reasons. Taxonomic groupings such as order or family are frequently invoked as an important factor  
in many proposed hypotheses for how mammal diversity is structured (Janis, 2008; Janis and  
74 Wilhelm, 1993; Pires et al., 2015; Quental and Marshall, 2013; Slater, 2015). Because taxonomic  
grouping conflates both species macroecology with shared evolutionary history, there are few clear  
76 ways to translate and operationalize these hypotheses in terms of macroecological change viewed  
through the lens of species interactions. Specifically, this issue arises when trying to generalize  
78 previous observations from taxonomy-based framework to ecology-based one.

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

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

Jernvall and Fortelius (2004) found that for the Neogene of Europe the relative abundance of  
86 mammal guilds was stable over time even in the face of high turnover rates, though they only  
considered large bodied taxa from a small set of mammal orders. Similar results have been observed  
88 for some taxonomic groups in North America (Valkenburgh, 1999). These results imply that there  
the types of interactions happening in local communities observed over a region are constant over  
90 time even if the interactors are constantly changing. A diversity-dependent or somehow limited  
diversification history is also theorized for North American mammal diversity as a whole (Alroy,  
92 2009; Rabosky and Hurlbert, 2015), though the evidence for this is limited (Harmon and Harrison,  
2015).

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

104 Within the canid guild of North America (e.g. plantigrade and digitigrade carnivores) there is  
evidence that their diversity is self-regulating or somehow limited (Valkenburgh, 1999). Specifically,  
106 it has been proposed 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  
108 constant 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

110 constant species turnover is consistent with limited possibility of increased diversity, even though  
111 there was no evidence of diversity-dependence in trait (e.g. body size) evolution (Slater, 2015).  
112 There is some uncertainty and a lack of consensus as to the effect of species body size on mammal  
113 diversity and aspects of the diversification processes, specifically extinction (Liow et al., 2008, 2009;  
114 Smits, 2015; Tomiya, 2013). Species body size is frequently framed as an important biological  
115 descriptor because of its correlation with other important and relevant ecological traits such as  
116 metabolic rate and home range size (Brown, 1995). It is also relatively easy to estimate for extinct  
117 species using proxy measures and regression equations, as was done in this study (see below).  
118 However, body size is normally analyzed without simultaneous reference to other species traits  
119 (Huang et al., 2017; Liow et al., 2008; Raia et al., 2012; Smith et al., 2004), but see (Smits, 2015);  
120 this combined with the high amount of correlation between life history traits and body size limits  
121 process-based inference because the actual causal mechanisms underlying an observed pattern are  
122 obscured or missing.

123 Smits (2015) found that the individual traits which form this study's ecotypes have strong effects  
124 on mammal extinction risk. Omnivorous taxa were found to have, on average, a greater duration  
125 than other dietary categories, while arboreal taxa were found to have a shorter duration than other  
126 locomotor categories (Smits, 2015). Two possible scenarios that could yield this pattern were  
127 proposed: the extinction risk faced by arboreal is constant and high or the Paleogene and Neogene  
128 represent different regimes and extinction risk increased in the Neogene, thus driving up the  
129 Cenozoic average extinction risk. These two possible explanations have clear and testable  
130 predictions with respect to the diversity history of arboreal taxa: 1) the extinction risk arboreal  
131 taxa increased in the Neogene compared to the Paleogene, driving the average extinction risk of  
132 arboreal mammals up and leading to the loss of arboreal taxa from the species pool, or 2) if  
133 arboreal taxa have just a generally higher extinction risk than other ecotypes but have maintained  
134 a constant diversity for the Cenozoic. By inspecting the inferred diversity histories of the ecotypes,  
135 it should be possible to distinguish amongst these hypotheses.  
136 Fundamentally, all species respond differently to climate and environmental change (Blois and

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

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

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

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

<sup>164</sup> grasslands environments of the Neogene (Blois and Hadly, 2009; Janis, 1993; Janis et al., 2000;  
Strömberg, 2005).

<sup>166</sup> A lot of the climate and environmental changes observed for North America have been attributed to  
tectonic activity or uplift (Badgley and Finarelli, 2013; Blois and Hadly, 2009; Eronen et al., 2015;  
<sup>168</sup> Janis, 2008). Tectonic uplift changes weather patterns (e.g. rain shadow) and mobilizes grit into the  
environment (Jardine et al., 2012). Increased grit in the environment combined with decreased rain  
<sup>170</sup> fall is considered the primary reason behind the trend of increased hypsodonty, or high crowned  
teeth, among herbivore groups over the Cenozoic of both North America and Europe (Damuth and  
<sup>172</sup> Janis, 2011; Jardine et al., 2012; Jernvall and Fortelius, 2002).

The Eocene-Oligocene transition has been observed to be associated with extinction of many  
<sup>174</sup> ungulate taxa (Janis, 2008). This boundary also marks the transition from the Paleogene to the  
Neogene and from herbivores being browsing dominated to grazing dominated, though not  
<sup>176</sup> concurrently (Janis, 1993; Strömberg, 2005). Additionally, the Paleogene-Neogene boundary marks  
the approximate start of Antarctic ice sheets, which were previously absent (Zachos et al., 2008).  
<sup>178</sup> There is an observed stability in 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  
<sup>180</sup> Mid-Miocene climatic optimum is 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  
<sup>182</sup> biological interactions or 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  
<sup>184</sup> Miocene would be a period of marked by an absence of major changes to diversity or the  
diversification process.

<sup>186</sup> The environmental factors included in this study are estimates of global temperature and the  
changing floral groups present in North America across the Cenozoic (Cramer et al., 2011; Graham,  
<sup>188</sup> 2011). These covariates were chosen because they provide high level characterizations of the  
environmental context of the entire North American regional species pool for most of the Cenozoic.  
<sup>190</sup> Importantly, the effects of a species ecotype on diversity are themselves modeled as functions of

environmental factors (Fig. 1) allowing for inference as to how a species ecology can mediate  
192 selective pressures do to its environmental context.

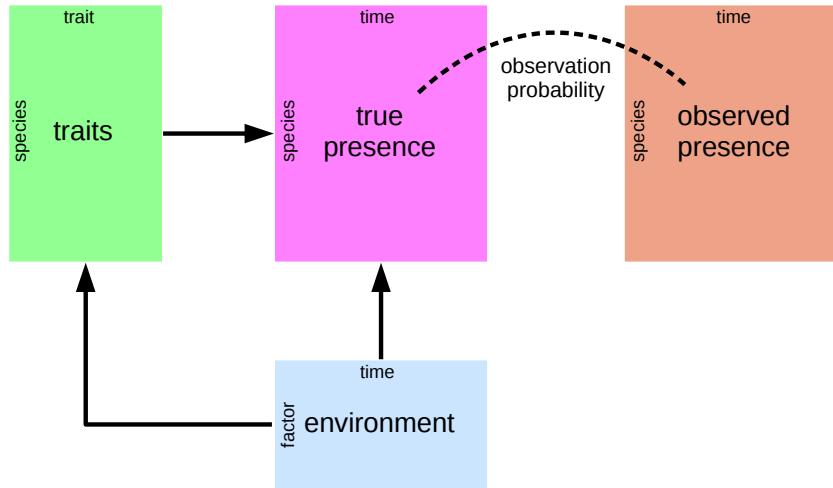


Figure 1: Conceptual diagram of the paleontological fourth corner problem. The observed presence matrix (orange) is the empirical presence/absence pattern for all species for all time points; this matrix is an incomplete observation of the “true” presence/absence pattern (purple). The estimated true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

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

202 Modern) is assembled over time. These two approaches, modern and paleontological, are different  
views of the same three-dimensional pattern: species at localities over time. The temporal  
204 limitations of modern ecological studies and difficulties with uneven spatial occurrences of fossils in  
paleontological studies means that these approaches are complimentary but reveal different patterns  
206 of how species are distributed in time and space.

All observations, paleontological or modern, are made with uncertainty. With presence/absence  
208 data this uncertainty comes from now knowing if an absence is a “true” absence or just a failure to  
observe (Foote, 2001; Foote and Sepkoski, 1999; Lloyd et al., 2011; Royle and Dorazio, 2008; Royle  
210 et al., 2005; Wang and Marshall, 2016). For paleontological data, the incomplete preservation of  
whatever species were present into fossil form combined with incomplete sampling of what  
212 organisms were actually fossilized means that the true times of origination or extinction may not be  
observed (Foote, 2001; Foote and Sepkoski, 1999; Wang et al., 2016; Wang and Marshall, 2016).

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

## Materials and Methods

### **226 Taxon occurrences and species-level information**

All fossil occurrence information used in this analysis was downloaded from the Paleobiology

228 Database (PBDB). The initial download restricted all occurrences to all Mammalia observed in  
North America between the Maastrichtian (72-66 Mya) and Gelasian (2.58-1.8 Mya) stages (Cohen

230 et al., 2015). Occurrences were then further limited to those occurring between 64 and 2 million  
years ago (Mya). Taxonomic, stratigraphic, and ecological metadata for each occurrence and species

232 was also downloaded. A new download for a raw, unfiltered PBDB datafile following the same  
criterion used here is available at <http://goo.gl/2s1geU>. The raw datafiles used as a part of this

234 study, along with all code for filtering and manipulating this download is available at  
<http://github.com/psmits/copings>.

236 After being downloaded, the raw occurrence data was then sorted, cleaned, and manipulated  
programmatically before analysis. Many species taxonomic assignments as present in the raw

238 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

240 were based on taxonomies present in the Encyclopedia of Life (<http://eol.org>) and Janis et al.  
(2008, 1998). All taxa who's life habit was classified as either volant (i.e. Chiroptera) or aquatic (e.g.

242 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

244 pools.

Species ecotype is defined here as the interaction between locomotor category and diet categories,  
246 the goal of which is to classify species based on the manner with which they interact with their

248 environment. Most mammal species records in the PBDB have life habit (i.e. locomotor category)  
and dietary category assignments. In order to simplify interpretation, analysis, and per ecotype

sample size these classifications were coarsened in a similar manner to (Smits, 2015) following Table  
250 1. Additionally for this study, the life history category was further broken up to better reflect the

diversity of mammal locomotor modes. Ground dwelling species locomotor categories were  
252 reassigned based on their ankle posture associated with their taxonomic group as described in Table  
2 (Carrano, 1999). Ecotype categories with less than 10 total species of that combination were  
254 excluded, yielding a total of 18 observed 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
	Aplopontidae	plantigrade
	Apternodontidae	scansorial
	Arctocyonidae	unguligrade
	Barbourofelidae	digitigrade

Continued on next page

**Table 2 – continued from previous page**

Order	Family	Stance
	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
	Geolabididae	plantigrade
	Glyptodontidae	plantigrade

Continued on next page

**Table 2 – continued from previous page**

Order	Family	Stance
	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
	Mustelidae	plantigrade
	Mylagaulidae	fossorial

Continued on next page

**Table 2 – continued from previous page**

Order	Family	Stance
	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
	Talpidae	fossorial
	Tayassuidae	unguligrade

Continued on next page

**Table 2 – continued from previous page**

Order	Family	Stance
	Tenrecidae	plantigrade
	Titanoideidae	plantigrade
	Ursidae	plantigrade
	Viverravidae	plantigrade
	Zapodidae	plantigrade

256 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  
258 (Brook and Bowman, 2004; Freudenthal and Martín-Suárez, 2013; McKenna, 2011; Raia et al., 2012;  
Smith et al., 2004; Tomiya, 2013); this is similar to what was done in Smits (2015). When species  
260 mass was not available, proxy measures were used and then transformed into estimates of mass. For  
example, given a measurement of a mammal tooth size, it is possible and routine to estimate its  
262 mass given some regression equation. The PBDB has one or more body part measures for many  
species. These were used as body size proxies for many species, as was the case in Smits (2015).  
264 Mass was log-transformed and then mean-centered and rescaled by dividing by two-times its  
standard deviation; this insures that the magnitude of effects for both continuous and discrete  
266 covariates are directly comparable (Gelman, 2008; Gelman and Hill, 2007).

All fossil occurrences from 64 to 2 million years long (Mya) were binned into 31 2 million year (My)  
268 bins. This temporal length was chosen because it is approximately the resolution of the North  
American mammal fossil record (Alroy, 1996, 2009; Alroy et al., 2000; Marcot, 2014).

270 **Environmental and temporal covariates**

The environmental covariates used in this study are collectively referred to as group-level covariates  
272 because they predict the response of a “group” of individual-level observations (i.e. species

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	Foster (2009)
General	$\log(m) = 3.68x - 3.83$	skull length	Luo et al. (2001)
Carnivores	$\log(m) = 2.97x + 1.681$	lower m1 length	Van Valkenburgh (1990)
Insectivores	$\log(m) = 1.628x + 1.726$	lower m1 area	Bloch et al. (1998)
Insectivores	$\log(m) = 1.714x + 0.886$	upper M1 area	Bloch et al. (1998)
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	Gordon (2003)
Marsupials	$\log(m) = 1.733x + 1.571$	upper M1 area	Gordon (2003)
Rodentia	$\log(m) = 1.767x + 2.172$	lower m1 area	Legendre (1986)
Ungulates	$\log(m) = 1.516x + 3.757$	lower m1 area	Mendoza et al. (2006)
Ungulates	$\log(m) = 3.076x + 2.366$	lower m2 length	Mendoza et al. (2006)
Ungulates	$\log(m) = 1.518x + 2.792$	lower m2 area	Mendoza et al. (2006)
Ungulates	$\log(m) = 3.113x - 1.374$	lower toothrow length	Mendoza et al. (2006)

occurrences of an ecotype). Additionally, these covariates are defined for temporal bins and not the

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” defined by Graham (2011). Global temperature across most of the Cenozoic was calculated from Mg/Ca isotope record from deep sea carbonates (Cramer et al., 2011). Mg/Ca based temperature estimates are preferable to the frequently used  $\delta^{18}\text{O}$  temperature proxy (Alroy et al., 2000; Figueirido et al., 2012; Zachos et al., 2008, 2001) because Mg/Ca estimates do not conflate temperature with ice sheet volume and depth/stratification changes. The former is particularly important to this analysis as the current polar ice-caps appeared and grew over the Cenozoic. 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 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 mean and then dividing by twice the standard deviation.

The second set of environmental factors included in this study are the Cenozoic plant phases defined

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

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

in Graham (2011). 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 modernity of those plant groups with younger phases representing increasingly modern taxa (Graham, 2011). Graham (2011) defines four intervals from the Cretaceous to the Pliocene, though only three of these intervals are included in this analysis. Graham's plant phases was included as a series of "dummy variables" encoding the three phases included in this analysis (Gelman and Hill, 2007); this means that the first phase is synonymous with the intercept and subsequent phases are defined by their differences from the first phase. The temporal boundries of these plant phases are defined in Table 4.

## Modelling species occurrence

Two different models were used in this study: a pure-presence model and a birth-death model. Both 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 or surviving are considered equal or different (Table 5). Something that is important to realize is that while there are only two state "codes" in a presence-absence matrix (i.e. 0/1), there are in fact 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); this is made obvious in the transition matrices as the probability of an extinct species changing states is 0 (Table 5). See below for parameter explainiations (Tables 6, 7, and 8).

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

(a) Pure-presence

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

(b) Birth-death

Table 5: Transition matrices for the pure-presence (5a) and birth-death (5b) models. Both of these models share the core machinery of discrete-time birth-death processes but make distinct assumptions about the equality of originating and surviving (Eq. 2, and 3). Note also that while there are only two state “codes” (0, 1), there are in fact three states: never having originated  $0_{never}$ , present 1, extinct  $0_{extinct}$  (Allen, 2011).

### 306 Data augmentation

All empirical presence/absence observations are potentially incomplete or observed with error. The 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 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 case of paleontological data, there are obviously many species which were never sampled (Royle and Dorazio, 2008; Royle et al., 2007).

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 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  $\psi$  is the probability that any augmented species should actually be “present.” Because  $M$  is user defined, this approach effectively gives  $\psi$  a uniform prior over  $N$  to  $M$  (Royle and Dorazio, 2008). For this study,  $M = \lfloor 1.25 \times N \rfloor$ .

Data imputation is the process of estimating missing data for partially observed covariates given 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 species are fully imputed species and thus have no known mass so a mass estimate must be imputed

Table 6: Paramaeters associated with the observation process part of the hidden Markov model.

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
$\alpha_0$	1	average log-odds of $p$
$\alpha_1$	1	change in average log-odds of $p$ per change mass
$r$	$T$	difference from $\alpha_0$ associated with time $t$
$\sigma$	1	standard deviation of $r$

for each possible species (Royle and Dorazio, 2012). Assuming that mass values for augmented  
326 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.  
328 This approach is an example of imputing covariate information that is missing completely at  
random (Gelman and Hill, 2007; Royle and Dorazio, 2012). Because log mass values are rescaled as  
330 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  
332 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.  
334 This classification has no biological interpretation.

### Observation process

336 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  $\phi$  of a species surviving from  
338 one time to another, and probability  $\pi$  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  
340  $\phi = \pi$ , 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  
342 time-varying intercept and species mass as a covariate. The effect of species mass on  $p$  was assumed  
linear and constant over time and given a prior reflecting a possible positive relationship; these

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

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

<sup>344</sup> assumptions are reflected in the structure of the model Equation 1, presented here:

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

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

### <sup>346</sup> Pure-presence process

For the pure-presence model there is only a single probability dealing with the presence of a species <sup>348</sup>  $\theta$  (Table 5a). This probability was modeled as multi-level logistic regression with both species-level and group-level covariates (Gelman et al., 2013; Gelman and Hill, 2007). The parameters associated <sup>350</sup> with pure-presence model are presented in Table 7 and the full sampling statement in Equation 2.

The species-level of the model (Eq. 2) is a logistic regression where the intercept varies by ecotype. <sup>352</sup> 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 <sup>354</sup> 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

356 choice of priors for these regression coefficients.

The values of each ecotype's intercept are themselves modeled as regressions using the group-level  
358 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  
360 covariances between ecotype intercepts, given this group-level regression, are modeled (Gelman and  
Hill, 2007).

362 All parameters not modeled elsewhere were given weakly informative priors (Gelman et al., 2013;  
McElreath, 2016; Stan Development Team, 2016). Weakly informative means that priors do not  
364 necessarily encode actual prior information but instead help regularize or weakly constrain posterior  
estimates. These priors have a concentrated probability density around and near zero; this has the  
366 effect of tempering our estimates and help prevent overfitting the model to the data (Gelman et al.,  
2013; McElreath, 2016; Stan Development Team, 2016). The general line of thinking behind this  
368 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 mass associated with the augment  
370 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

- <sup>372</sup> 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  $\theta$  is modeled in the pure-presence model  
<sup>374</sup> (Eq. 2, Table 5b). The parameters associated with the birth-death presence model are presented in  
 Table 8 and the full sampling statement, including observation (Eq. 1), is described in Equation 3:

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

- <sup>376</sup> Similar to the pure-presence model, both  $\phi$  and  $\pi$  are modeled as logistic regressions with  
 varying-intercept and one covariate associated with two parameters. The possible relationships  
<sup>378</sup> between mass and both  $\phi$  and  $\pi$  are reflected in the parameterization of the model and choice of  
 priors (Eq. 3).  
<sup>380</sup> The intercepts of  $\phi$  and  $\pi$  both vary by species ecotype and those values are themselves the product  
 of group-level regression using environmental factors as covariates (Eq. 3); this is identical to the

Table 8: Parameters for the model of presence in the pure-presence model

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

382 pure presence model (Eq. 2).

## Posterior inference and model adequacy

384 Computer programs that implement joint posterior inference for the above models (Eqs. 2, 3) were  
written in the probabilistic programming language Stan (Stan Development Team, 2016). The  
386 models used here both feature latent discrete parameters in the large matrix  $z$  (Tables 6, 7, 8; Eqs.  
1, 2, 3). All methods for posterior inference implemented in Stan are derivative based which causes  
388 complications for actually implementing the above models because integers do not have derivatives.  
Instead of implementing a latent discrete parameterization, the log posterior probabilities of all  
390 possible states 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  
392 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

394 are calculated, from the end members of the species having existed for the entire study interval and  
 the species having only existed between the directly observed FAD and LAD to all possible  
 396 intermediaries (Fig 2) (Stan Development Team, 2016). This process is identical, language-wise, to  
 assuming range-through and then estimating the possibility of range extension due to incomplete  
 398 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),  
 400 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  
 402 implemented. Instead, an approximate Bayesian approach was used: variational inference. A  
 recently developed automatic variational inference algorithm called “automatic differentiation  
 404 variational inference” (ADVI) is implemented in Stan and was used here (Kucukelbir et al., 2015;  
 Stan Development Team, 2016). ADVI assumes that the posterior is Gaussian but still yields a true

406 Bayesian posterior; this assumption is similar to quadratic approximation of the likelihood function  
commonly used in maximum likelihood based inference (McElreath, 2016). The principal limitation  
408 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.  
410 standard deviation).

Of additionaly concern for posterior inference is the partial identifiability of observation parameters  
412  $p_{t=1}$  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  
414 “gaps” in species occurrence histories that are guarenteed to be filled. Instead, the values of the first  
and final columns of the “true” presence-absence matrix  $z$  for thos observations that do not already  
416 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  
418 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-hob approaches that are occationally  
420 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  $\theta$  and origination  $\phi$  at  $t = 1$  and estimates of  $\theta$ ,  $\phi$   
422 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  
424 means of those parameters.

After fitting both models (Eqs. 2, 3) using ADVI, model adequacy and quality of fit was assessed  
426 using a posterior predictive check (Gelman et al., 2013). Because all Bayesian models are inherently  
generative, simulations of new data sets is “free” (Gelman et al., 2013; McElreath, 2016). By  
428 simulating 100 theoretical data sets from the posterior estimates of the model parameters and the  
observed covariate information the congruence between predictions made by the model and the  
430 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  
432 stochastic processes with covariate information given from the empirical observations. Importantly,  
this is fundamentally different from observing the posterior estimates of the “true”

- 434 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 averages; if the  
 436 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.
- 438 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  
 440 distribution and are instead just an approximation (Gelman et al., 2013). Because of this, posterior  
 predictive estimates are themselves only approximate checks of model adequacy. The posterior  
 442 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.
- 444 Given parameter estimates, diversity and diversification rates are estimated through posterior  
 predictive simulations. Given the observed presence-absence matrix  $y$ , estimates of the true  
 446 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  
 448 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

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

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

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

- 452 and number of extinctions  $E_t$  estimated as

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

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

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

## 454 Results

The results of the analyses described above take one of two forms: direct inspection of parameter  
 456 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  
 458 explained below in the comparison of the models' posterior predictive check results.

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

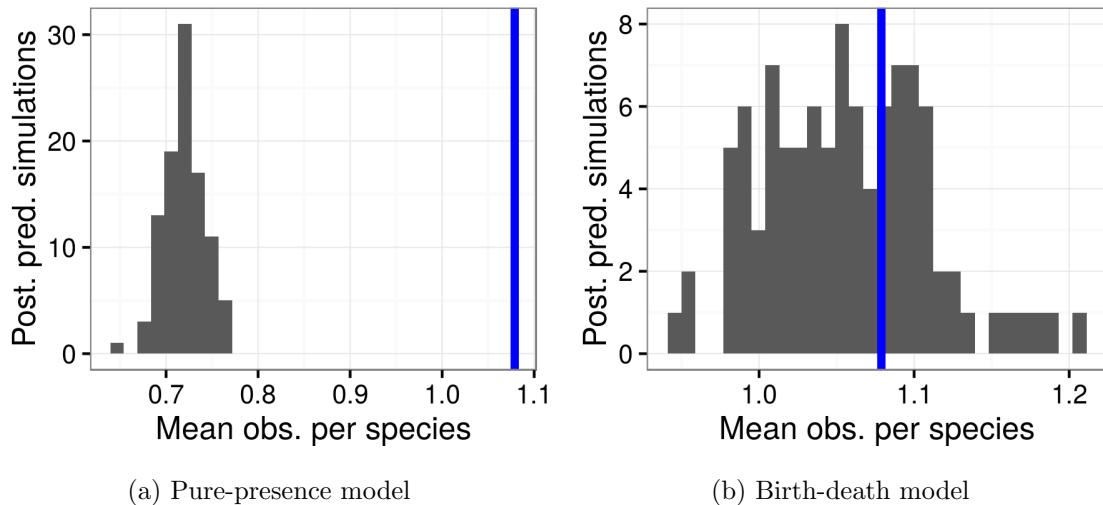


Figure 3: Comparison of the average observed number of occurrences per species (blue line) to the average number of occurrences from 100 posterior predictive datasets using the posterior estimates from the pure-presence and birth-death models.

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

462 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  
464 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  
466 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  
468 parameter estimates from both models gives further insight into the reasons for this difference in  
posterior predictive results (Gelman et al., 2013).

470 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  
472 probability estimates (Fig. 6). This result supports the idea that changes to the North American  
regional species pool is more likely due to changes in origination than extinction, a result that is  
474 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)  
476 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  
478 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  
480 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  
482 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  
484 no longer originating or originating very rarely.

The dramatic differences in the estimates origination and survival probabilities are indicative of  
486 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.  
488 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  
 490 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  
 492 estimates of survival probabilities (Fig. 5, and 6). In logistic regression, high uncertainty in the  
 494 estimates of the underlying log-odds of occurrence, origination, or survival tends to be indicative of  
 which occurs when there is no uncertainty in the effect of a covariate on presence/absence, the  
 496 effect of which has been mitigated by the hierarchical modeling strategy used here (Gelman et al.,  
 2013; Gelman and Hill, 2007; McElreath, 2016).

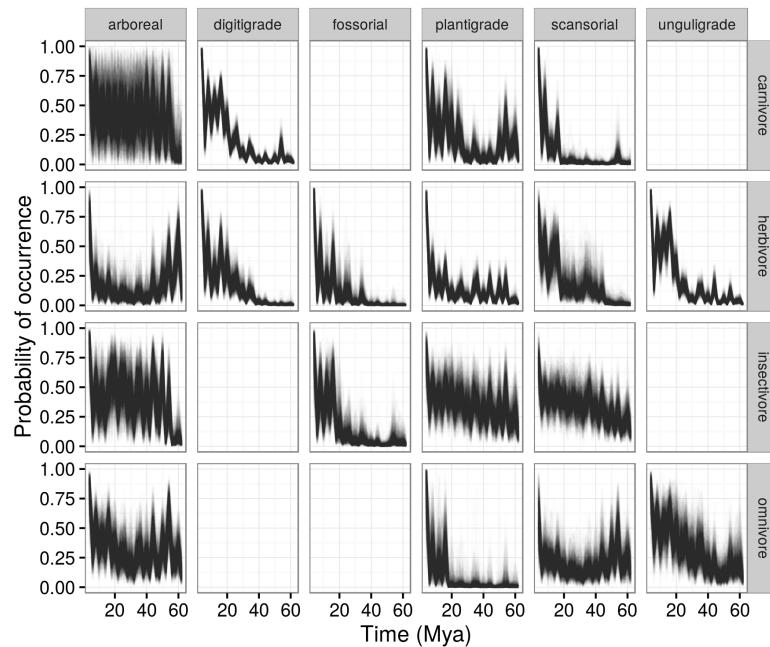


Figure 4: Probability of a mammal ecotype occurring over time as estimated from the pure-presence model. Each panel depicts 100 random samples from the model’s posterior. The columns are by locomotor category and rows by dietary category; their intersections are the observed and analyzed ecotypes. Panels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

498 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  
 500 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

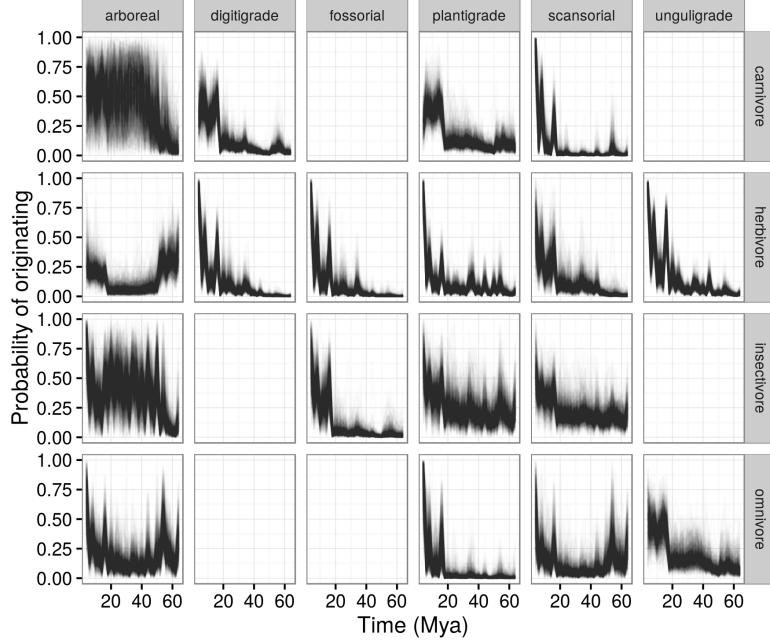


Figure 5: Probability of a mammal ecotype origination probabilities at each time point as estimated from the birth-death model. Each panel depicts 100 random samples from the model’s posterior. The columns are by locomotor category and rows by dietary category; their intersections are the observed and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

observation such that larger species are less likely to be observed if present than smaller species (Fig. 7b). The result from the birth-death model is partially expected based on previous research (Damuth, 1982). What this result means is that the record of large 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 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.

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, 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 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

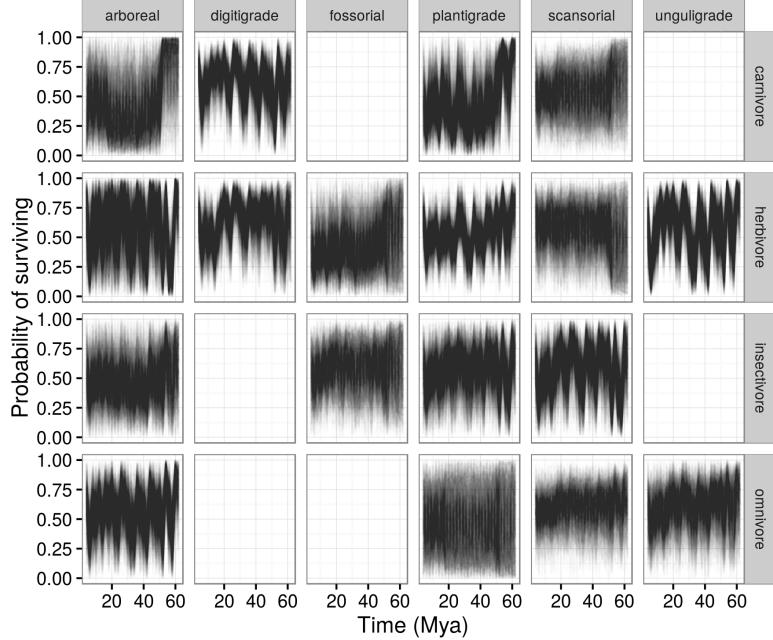


Figure 6: Probability of a mammal ecotype survival probabilities at each time point as estimated from the birth-death model. Each panel depicts 100 random samples from the model’s posterior. The columns are by locomotor category and rows by dietary category; their intersections are the observed and analyzed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

parameter estimates in the model.

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

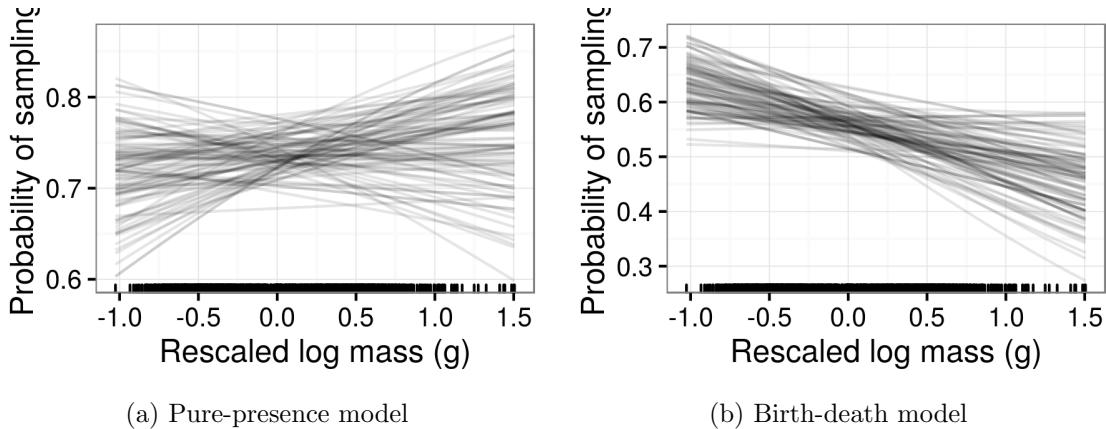


Figure 7: Estimates of the effect of species mass on probability of sampling a present species ( $p$ ). Mass has been log-transformed, centered, and rescaled; this means that a mass of 0 corresponds to the mean of log-mass of all observed species and that mass is in standard deviation units. Estimates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

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.

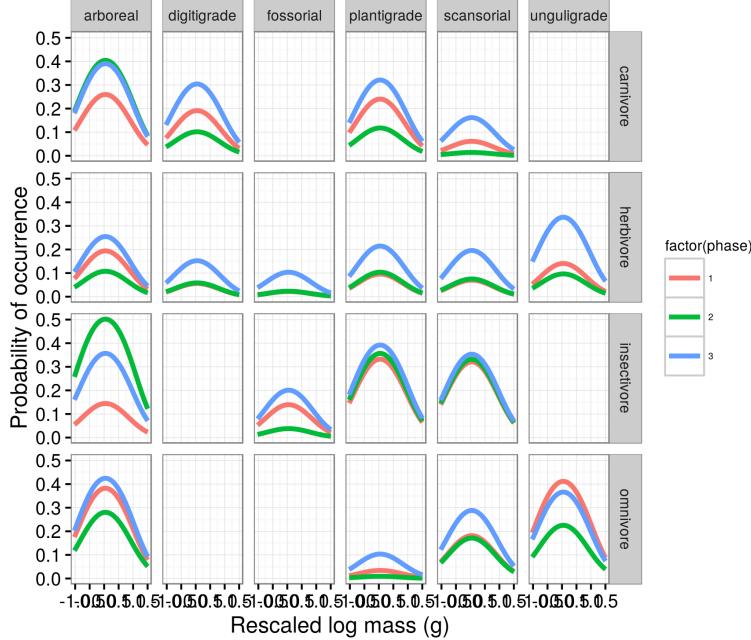


Figure 8: Mean estimate of the effect of species mass on the probability of a species occurrence for each of the three plant phases. The effect of mass is considered constant over time and that the only aspect of the model that changes with plant phase is the intercept of the relationship between mass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

544 for an ecotype may be indicative of an increasing number of possible and available mammal-plant  
 545 interactions (e.g. ecological opportunity; Losos, 2010; Losos and Mahler, 2010; Yoder et al., 2010);  
 546 while adverse conditions may translate to a decreasing set of interactions or loss of appropriate  
 547 environmental context. Remember that favorable versus adverse condition of a plant phase is  
 548 definitionally relative to the other two plant phases.

Plant phases are associated with large differences in log-odds for occurrence and origination  
 550 probabilities (Tables 9, 10), though there is little evidence of plant phase being an important  
 551 distinguishing factor in species survival as only a few ecotypes demonstrate strong affinities with  
 552 some plant phases (Table 11). As with previous comparisons between parameter estimates  
 553 associated with species occurrence and species origination, parameters associated with probability

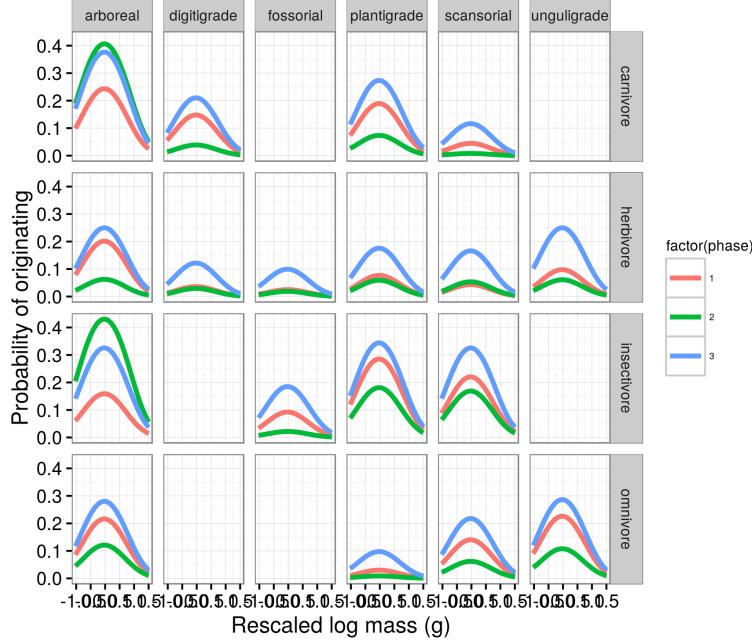


Figure 9: Mean estimate of the effect of species mass on the probability of a species originating for each of the three plant phases. The effect of mass is considered constant over time and that the only aspect of the model that changes with plant phase is the intercept of the relationship between mass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

- 554 of newly originating appear as a more “tempered” version of those associated with probability occurrence.
- 556 The almost universal pattern of the effect of plant phase on ecotype occurrence or origination is that the during first and last plant phases ecotypes have a greater log-odds of occurrence or
- 558 origination than the second plant phase (Fig. 4, 5). The three ecotypes that do not follow this pattern are fossorial herbivores, scansorial herbivores, and arboreal insectivores.
- 560 The difference between the third plant phase and the other two plant phases, for all ecotypes except arboreal carnivores, is obvious upon inspection the occurrence and origination time series as there
- 562 is large up-tick in probability of occurring or originating towards the modern (Fig. 4, 5). The differences in mean probability of occurring or originating attributable to the plant phases are

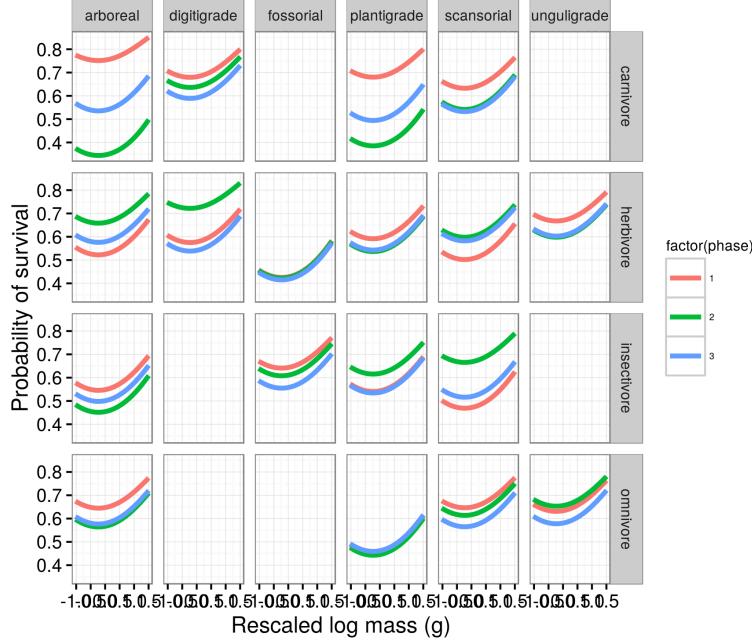


Figure 10: Mean estimate of the effect of species mass on the probability of a species survival for each of the three plant phases. The effect of mass is considered constant over time and that the only aspect of the model that changes with plant phase is the intercept of the relationship between mass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

observable as shifts along the time series corresponding to the phase barriers (Table 4). For example, scansorial herbivore occurrence and origination probabilities demonstrate clear shifts at 50 Mya and 16 Mya (Fig. 4, 5).

Both aspects of global temperature analyzed here are estimated to have strong effects on species occurrence and origination for most mammal ecotypes (Tables 12, 13). Similarly, temperature is only expected to have a strong effect on species extinction for very few ecotypes (Table 14). For the occurrence and origination probabilities of many ecotypes, both temperature covariates have negative estimates which means that as temperature decreases, occurrence or origination are expected to increase. The only strongly positive estimate (e.g. temperature decrease, origination decrease) is for the effect of temperature range on arboreal herbivores. Contrastingly, the only

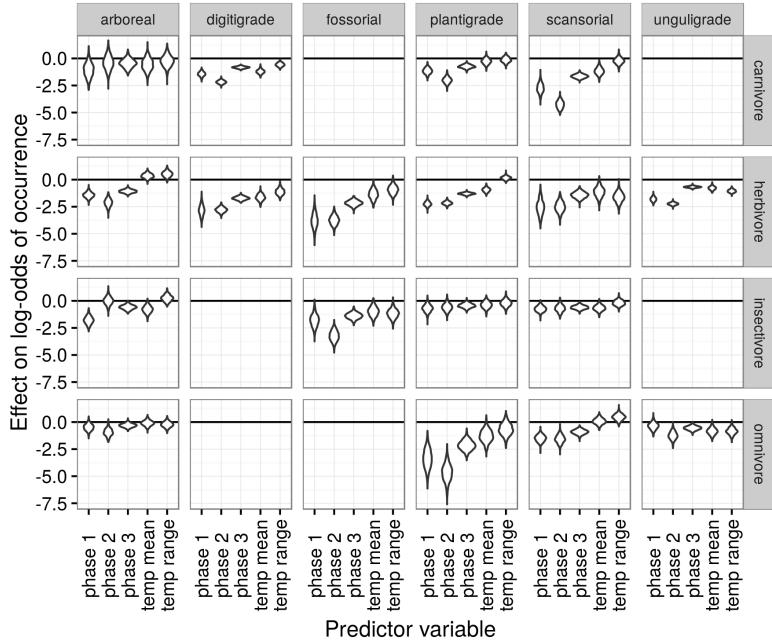


Figure 11: Estimated effects of the group-level covariates describing environmental context on log-odds of species occurrence. These estimates are from the pure-presence model. The effect of plant phase graphed here is calculated as Phase 1 =  $\gamma_{phase\ 1}$ , Phase 2 =  $\gamma_{phase\ 1} + \gamma_{phase\ 2}$ , and so on.

574 strong ecotype associations for either of the temperature covariates are with plantigrade carnivores,  
 575 plantigrade herbivores, and to a less certain extent arboreal herbivores and scansorial insectivores  
 576 (Tab. 14). The effects of the temperature covariates on these ecotypes are all estimated to be  
 positive (e.g. temperature range increase, increase in survival).

## 578 Analysis of diversity

All of the analyses of diversification and macroevolutionary rates has been done using only the  
 580 birth-death model because of the model's better posterior predictive check performance (Fig. 3).

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

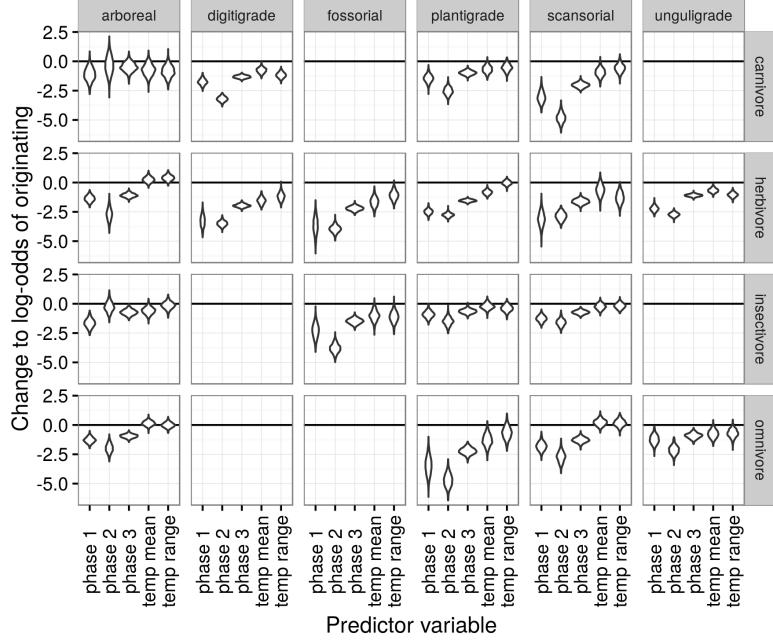


Figure 12: Estimated effects of the group-level covariates describing environmental context on log-odds of species origination. These estimates are from the birth-death model. The effect of plant phase graphed here is calculated as Phase 1 =  $\gamma_{phase\ 1}$ , Phase 2 =  $\gamma_{phase\ 1} + \gamma_{phase\ 2}$ , and so on.

586 Miocene). After this, there is a rapid spike in diversity followed by a slight decline in diversity up to  
 587 the Modern. This characterization of the estimated diversity history is knowingly broad strokes and  
 588 diversity time series is not without variation and vagaries.

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

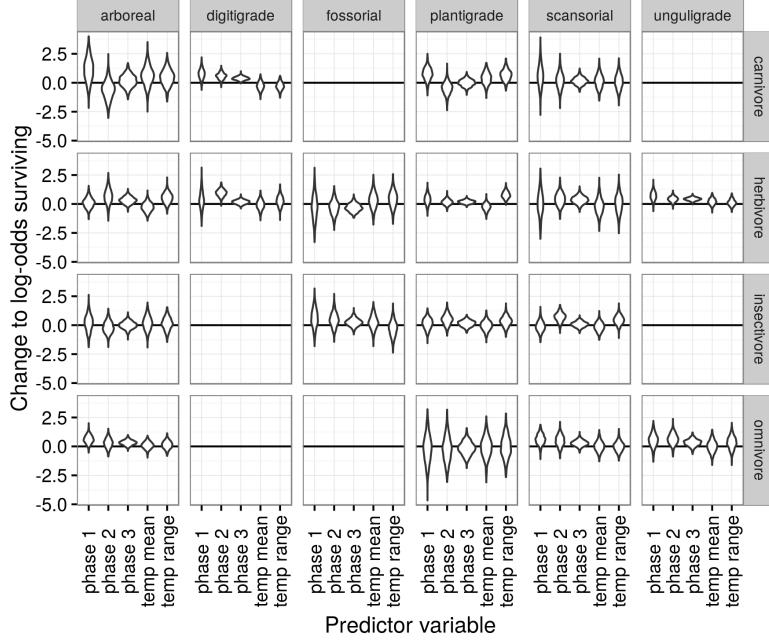


Figure 13: Estimated effects of the group-level covariates describing environmental context on log-odds of species survlval. These estimates are from the birth-death model. The effect of plant phase graphed here is calculated as Phase 1=  $\gamma_{phase\ 1}$ , Phase 2=  $\gamma_{phase\ 1} + \gamma_{phase\ 2}$ , and so on.

The comparison between per capita origination and extinction rate estimates reveals how  
 600 diversification rate is formed (Fig. 14c, 14d). As expected given previous inspection of the ecotype  
 specific estimates of origination and survival probabilities from the birth-death model,  
 602 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 pattern around a constant  
 604 mean (Fig. 14d).

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

Arboreal ecotypes obtain peak diversity early in the Cenozoic and then decline for the rest of the  
 608 time series, becoming increasingly rare or absent as diversity approaches the Modern (Fig. 15).  
 Arboreal herbivores and omnivores obtain peak diversity at the beginning of the Cenozoic then go  
 610 into decline while remaining a small part of the species pool, while arboreal carnivores and  
 insectivores obtain peak diversity 52-50 Mya and then quickly decline and become extremely rare or

Table 9: Posterior probability of the differences in the log-odds of an ecotype occurring based on plant phase. These probabilities are calculated as  $P(\text{Phase 1} > 2) = (\sum \gamma_{\text{phase1}} > \gamma_{\text{phase1}} + \gamma_{\text{phase2}})/100$  and similarly for the other comparisons. These estimates are from the pure-presence model.

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

612 entirely absent from the species pool.

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

614 In contrast, plantigrade herbivore diversity does not have a single, broad-strokes pattern; instead, diversity increases, decreases, and may have then increased till the Modern. Contrastingly, fossorial 616 and scansorial herbivores demonstrate a much flatter history of diversity, with a slight increase in diversity that over time is more pronounced among fossorial taxa than scansorial taxa.

618 Digitigrade carnivores have a multi-modal diversity history, with peaks 54-52 and 12-10 Mya (Fig.15). Between these two peaks digitigrade carnivore diversity dips below average diversity 620 following the first peak and then grows slowly until the second peak. Plantigrade carnivores obtain peak diversity in the early Cenozoic and then maintain a relatively stable diversity until another 622 peak at the end of the Cenozoic.

There are some broad similarities in diversity histories of insectivorous and omnivorous taxa. The

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

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

624 diversity histories of arboreal, plantigrade, and scansorial insectivorous taxa all demonstrate a  
 625 decreasing pattern with time, while fossorial insectivores have a flat diversity history with a peak  
 626 approximately 10 Mya (Fig. 15). Arboreal and scansorial omnivores decrease in diversity from their  
 627 initial peaks early in the Cenozoic, and plantigrade omnivores have a generally flat diversity history  
 628 with a sudden peak in diversity late in the Cenozoic (Fig. 15). Unguligrade omnivores also  
 629 demonstrate a possible decrease in diversity over the Cenozoic, but not as clearly as arboreal and  
 630 scansorial omnivores.

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

When ecotype diversity is decomposed into the number of origination events per time bin (Fig. 16)

Table 11: Posterior probability of the differences in the log-odds of an ecotype surviving based on plant phase. These probabilities are calculated as  $P(\text{Phase 1} > 2) = (\sum \gamma_{\text{phase1}} > \gamma_{\text{phase1}} + \gamma_{\text{phase2}})/100$  and similarly for the other comparisons. These estimates are from the birth-death model.

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

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

## 642 Discussion

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

Table 12: Posterior probability the the effects of the two temperature covariates on the the log-odds of an ecotype occurring are greater than 0. What is estimated is the probability that these estimates are greater than 0; high or low probabilities indicate the “strength” of the covariate in that direction (positive and negative, respectively). These estimates are from the pure-presence model.

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

ecotypes and guilds, these roles are broad and not defined by specific interactions but by the genre

648 of interactions 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  
 650 Fortelius, 2004; Slater, 2015; Valkenburgh, 1999). This implies that the size and scope of the role of  
 an ecotype or guild in local communities, and the regional species pool as a whole, is preserved even  
 652 as the individual interactors change. This also implies the structure of regional species pools can be  
 constant over time despite a constantly changing set of “players.”

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

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

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

regional species pool diversity is shaped is not directly observable in studies of the Modern where

time scales are short and macroevolutionary dynamics are inferable solely from phylogeny (Fritz et al., 2013).

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

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

Table 14: Posterior probability that the effects of the two temperature covariates on the log-odds of an ecotype survival are greater than 0. What is estimated is the probability that these estimates are greater than 0; high or low probabilities indicate the “strength” of the covariate in that direction (positive and negative, respectively). These estimates are from the birth-death model.

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

The closest examples to a sudden increase or decrease of a specific ecotype is the jump in standing

672 diversity of scansorial carnivorans and, to a lesser extent, fossorial insectivores at 16 Mya (i.e. the  
start of the third plant phase). This result may, however, not reflect the dynamics of individual  
674 local communities as this is an analysis of the entire North American mammal regional species pool.

Arboreal taxa disappear from the regional species pool over the Cenozoic, with massive

676 disappearance by the Paleogene-Neogene transition ~22 Mya. This is consistent with one of the two  
possible patterns presented here and in Smits (2015) that would result in arboreal taxa having a  
678 greater extinction risk than other ecotypes: the Paleogene and Neogene were different selective  
regimes and while the earliest Cenozoic may have been neutral wrt arboreal taxa, they disappeared  
680 quickly over the Cenozoic which may account for their higher extinction risk. In addition to all  
arboreal taxa, the diversity of plantigrade and scansorial insectivores decreases with time (Fig. 15).

682 Digitigrade carnivores have a relatively stable diversity history through the Cenozoic and can be

characterized as varying around a constant mean diversity. This ecotype has a large amount of  
684 overlap with the carnivore guild which has been the focus of much research (Janis and Wilhelm,  
1993; Pires et al., 2015; Slater, 2015; Valkenburgh, 1999). This result is consistent with some form  
686 of “control” on the ecotype, such as environmental stability, diversity-dependence, or similar  
(Silvestro et al., 2015; Slater, 2015; Valkenburgh, 1999).

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

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

newly available mammal-plant interactions due to the increased modernization of their floral  
712 context (Graham, 2011).

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

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

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

738 more descriptors of species' environmental context a more complete "picture" of the diversification  
process is inferred.

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

The only covariate allowed to affect sampling probability was mass and only as a linear predictor.  
756 Other covariates, such as the environmental factors considered here, could have affected the  
underlying preservation process that limits sampling probability; their exclusion as covariates of  
758 sampling/observation was the product of a few key decisions: model complexity, model  
interpretability, the scope of this study, and a lack of good hypotheses related to these covariates to  
760 warrant their inclusion.

The potential effects of common ancestry (i.e. phylogeny) on origination and extinction are not  
762 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  
764 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; 766 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 768 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 770 histories and data augmentation already included in both models.

These results support the conclusion that the relative ecotypic diversity of the North American 772 mammal species pool has changed gradually over time. While there is constant species turnover for the entire Cenozoic, there is little evidence of major cross-ecotype upheaval and sudden 774 reorganization of the functional composition of the regional species pool. The results of this study also support the conclusion that mammal diversification over the Cenozoic is driven primarily by 776 changes to origination rate and not extinction rate. There are a number of interesting estimated ecotype diversity patterns. While arboreal ecotypes are diverse in the Paleogene, by the Neogene all 778 arboreal ecotypes dramatically decreased in diversity and became either rare or absent from the regional species pool. The other ecotypes that decrease in diversity over the Cenozoic are 780 plantigrade and scansorial insectivores and scansorial omnivores. Contrastingly, the only ecotypes that demonstrate a sustained pattern of increasing diversity are digitigrade and unguligrade 782 herbivores. Interestingly, when the environmental covariates analyzed here are inferred to affect the diversification of an ecotype, this effect is virtually always for origination and not survival. This 784 analysis provides a much more complete picture of North American mammal diversity and diversification, specifically the dynamics of the ecotypic composition of that diversity. By increasing 786 the complexity of analysis while precisely translating research questions into a statistical model, the context of the results is much better understood. Future studies of diversity and diversification 788 should incorporate as much information as possible into their analyses in order to better understand or at least contextualize the complex processes underlying that diversity.

<sup>790</sup> **Acknowledgements**

I would like to thank K. Angielczyk, M. Foote, P. D. Polly, R. Ree, and G. Slater for helpful  
<sup>792</sup> discussion and advice. This entire study would not have been possible without the Herculean  
effort of the many contributors to the Paleobiology Database. In particular, I would like to thank J.  
<sup>794</sup> Alroy and M. Uhen for curating most of the mammal occurrences recorded in the PBDB. This is  
Paleobiology Database publication XXX.

<sup>796</sup> **References**

- Allen, L. J. S. 2011. An introduction to stochastic processes with applications to biology. 2nd ed.  
<sup>798</sup> Chapman and Hall/CRC, Boca Raton, FL.
- Alroy, J. 1996. Constant extinction, constrained diversification, and uncoordinated stasis in North  
<sup>800</sup> American mammals. *Palaeogeography, Palaeoclimatology, Palaeoecology* 127:285–311.
- . 2009. Speciation and extinction in the fossil record of North American mammals. Pages  
<sup>802</sup> 302–323 *in* R. K. Butlin, J. R. Bridle, and D. Schlüter, eds. *Speciation and patterns of diversity*.  
Cambridge University Press, Cambridge.
- <sup>804</sup> Alroy, J., P. L. Koch, and J. C. Zachos. 2000. Global climate change and North American  
mammalian evolution. *Paleobiology* 26:259–288.
- <sup>806</sup> 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*  
<sup>808</sup> 39:373–399.
- Badgley, C., T. M. Smiley, R. Terry, E. B. Davis, L. R. G. Desantis, D. L. Fox, S. S. B. Hopkins,  
<sup>810</sup> 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:  
<sup>812</sup> Modern and Geohistorical Perspectives. *Trends in Ecology & Evolution* pages 1–16.

- Bambach, R. K. 1977. Species richness in marine benthic habitats through the Phanerozoic.  
814      Paleobiology 3:152–167.
- Bambach, R. K., A. M. Bush, and D. H. Erwin. 2007. Autecology and the filling of ecospace: Key  
816      metazoan radiations. Palaeontology 50:1–22.
- Bloch, J. I., K. D. Rose, and P. D. Gingerich. 1998. New species of Batodonoides (Lipotyphla,  
818      Geolabididae) from the Early Eocene of Wyoming: smallest known mammal? Journal of  
Mammalogy 79:804–827.
- Blois, J. L., and E. A. Hadly. 2009. Mammalian Response to Cenozoic Climatic Change. Annual  
820      Review of Earth and Planetary Sciences 37:181–208.
- Brook, B. W., and D. M. J. S. Bowman. 2004. The uncertain blitzkrieg of Pleistocene megafauna.  
Journal of Biogeography 31:517–523.  
822
- 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  
824      the environment. Methods in Ecology and Evolution 5:344–352.
- Brown, J. H., and B. A. Maurer. 1989. Macroecology: the division of food and space among species  
on continents. Science 243:1145–1150.  
828
- Brown, J. J. 1995. Macroecology. University of Chicago Press, Chicago.
- Bush, A. M., and R. K. Bambach. 2011. Paleoecologic Megatrends in Marine Metazoa. Annual  
Review of Earth and Planetary Sciences 39:241–269.  
830
- Bush, A. M., R. K. Bambach, and G. M. Daley. 2007. Changes in theoretical ecospace utilization in  
marine fossil assemblages between the mid-Paleozoic and late Cenozoic. Paleobiology 33:76–97.  
832
- Bush, A. M., and P. M. Novack-Gottshall. 2012. Modelling the ecological-functional diversification  
of marine Metazoa on geological time scales. Biology Letters 8:151–155.  
834
- Cantalapiedra, J. L., J. L. Prado, and M. T. Alberdi. 2017. Decoupled ecomorphological evolution  
and diversification in Neogene-Quaternary horses. Science 355:627–630.  
836

- 838 Carrano, M. T. 1999. What, if anything, is a cursor? Categories versus continua for determining  
locomotor habit in mammals and dinosaurs. *Journal of Zoology* 247:29–42.
- 840 Clyde, W. C., and P. D. Gingerich. 1998. Mammalian community response to the latest Paleocene  
thermal maximum: an isotaphonomic study in the northern Bighorn Basin, Wyoming. *Geology*  
842 26:1011–1014.
- 844 Cohen, K. M., S. C. Finney, P. L. Gibbard, and J.-X. Fan. 2015. The ICS International  
Chronostratigraphic Chart.
- 846 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  
geochemistry ( $\delta^{18}\text{O}$  and Mg/Ca) with sea level history. *Journal of Geophysical Research: Oceans*  
848 116:1–23.
- 850 Damuth, J. 1982. Analysis of the preservation of community structure in assemblages of fossil  
mammals. *Paleobiology* 8:434–446.
- 852 Damuth, J., and C. M. Janis. 2011. On the relationship between hypsodonty and feeding ecology in  
ungulate mammals, and its utility in palaeoecology. *Biological Reviews* 86:733–758.
- 854 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  
America and Europe. *Proceedings of the Royal Society B: Biological Sciences* 282:20150136.
- 856 Ezard, T. H. G., A. Purvis, and H. Morlon. 2016. Environmental changes define ecological limits to  
species richness and reveal the mode of macroevolutionary competition. *Ecology Letters*  
858 19:899–906.
- Felsenstein, J. 1985. Phylogenies and the comparative method. *The American Naturalist* 125:1–15.
- 860 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  
862 Academy of Sciences* 109:722–727.

- Foote, M. 2001. Inferring temporal patterns of preservation, origination, and extinction from  
864 taxonomic survivorship analysis. *Paleobiology* 27:602–630.
- Foote, M., and J. J. Sepkoski. 1999. Absolute measures of the completeness of the fossil record.  
866 *Nature* 398:415–7.
- Foster, J. R. 2009. Preliminary body mass estimates for mammalian genera of the Morrison  
868 Formation (Upper Jurassic, North America). *PaleoBios* 28:114–122.
- Fraser, D., R. Gorelick, and N. Rybczynski. 2015. Macroevolution and climate change influence  
870 phylogenetic community assembly of North American hoofed mammals. *Biological Journal of the Linnean Society* 114:485–494.
- Freudenthal, M., and E. Martín-Suárez. 2013. Estimating body mass of fossil rodents. *Scripta Geologica* 145:1–130.
- Fritz, S. A., J. Schnitzler, J. T. Eronen, C. Hof, K. Böhning-Gaese, and C. H. Graham. 2013. Diversity in time and space: wanted dead and alive. *Trends in Ecology & Evolution* 28:509–16.
- Gelman, A. 2008. Scaling regression inputs by dividing by two standard deviations. *Statistics in Medicine* pages 2865–2873.
- Gelman, A., J. B. Carlin, H. S. Stern, D. B. Dunson, A. Vehtari, and D. B. Rubin. 2013. Bayesian data analysis. 3rd ed. Chapman and Hall, Boca Raton, FL.
- Gelman, A., and J. Hill. 2007. Data Analysis using Regression and Multilevel/Hierarchical Models. Cambridge University Press, New York, NY.
- Gelman, A., J. Hill, and M. Yajima. 2012. Why We (Usually) Don't Have to Worry About Multiple Comparisons. *Journal of Research on Educational Effectiveness* 5:189–211.
- Gelman, A., and E. Loken. 2014. The statistical crisis in science: data-dependent analysis – a “garden of forking paths” – explains why many statistically significant comparisons don't hold up.  
886 *American Scientist* 102:460–465.

- Gordon, C. L. 2003. A First Look at Estimating Body Size in Dentally Conservative Marsupials.  
888 Journal of Mammalian Evolution page 21.
- Graham, A. 2011. A natural history of the New World: the ecology and evolution of plants in the  
890 Americas. University of Chicago Press, Chicago.
- Harmon, L. J., and S. Harrison. 2015. Species Diversity Is Dynamic and Unbounded at Local and  
892 Continental Scales. *The American Naturalist* 185:000–000.
- Harnik, P. G., P. C. Fitzgerald, J. L. Payne, and S. J. Carlson. 2014. Phylogenetic signal in  
894 extinction selectivity in Devonian terebratulide brachiopods. *Paleobiology* 40:675–692.
- Harrison, S., and H. Cornell. 2008. Toward a better understanding of the regional causes of local  
896 community richness. *Ecology Letters* 11:969–979.
- Huang, S., J. T. Eronen, C. M. Janis, J. J. Saarinen, D. Silvestro, and S. A. Fritz. 2017. Mammal  
898 body size evolution in North America and Europe over 20 Myr: similar trends generated by  
different processes. *Proceedings of the Royal Society B: Biological Sciences* 284:20162361.
- Jamil, T., W. A. Ozinga, M. Kleyer, and C. J. F. Ter Braak. 2013. Selecting traits that explain  
900 species-environment relationships: A generalized linear mixed model approach. *Journal of  
Vegetation Science* 24:988–1000.
- Janis, C., J. Damuth, and J. M. Theodor. 2004. The species richness of Miocene browsers, and  
904 implications for habitat type and primary productivity in the North American grassland biome.  
*Palaeogeography, Palaeoclimatology, Palaeoecology* 207:371–398.
- Janis, C. M. 1993. Tertiary mammal evolution in the context of changing climates, vegetation, and  
906 tectonic events. *Annual Review of Ecology and Systematics* 24:467–500.
- . 2008. An evolutionary history of browsing and grazing ungulates. Pages 21–45 *in* I. J.  
Gordon and H. H. T. Prins, eds. *The Ecology of Browsing and Grazing*. Springer-Verlag.
- 910 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*  
912 97:7899–904.
- Janis, C. M., G. F. Gunnell, and M. D. Uhen. 2008. Evolution of Tertiary mammals of North  
914 America. Vol. 2. Small mammals, xenarthrans, and marine mammals. Cambridge University  
Press, Cambridge.
- 916 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  
918 University Press, Cambridge.
- 920 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.
- 922 Jardine, P. E., C. M. Janis, S. Sahney, and M. J. Benton. 2012. Grit not grass: concordant patterns  
of early origin of hypodonty in Great Plains ungulates and Glires. *Palaeogeography,  
Palaeoclimatology, Palaeoecology* 365–366:1–10.
- 924 Jernvall, J., and M. Fortelius. 2002. Common mammals drive the evolutionary increase of  
hypodonty in the Neogene. *Nature* 417:538–40.
- 926 ———. 2004. Maintenance of trophic structure in fossil mammal communities: site occupancy and  
taxon resilience. *The American Naturalist* 164:614–624.
- 928 Kucukelbir, A., R. Ranganath, A. Gelman, and D. M. Blei. 2015. Automatic Variational Inference  
in Stan. Pages 568–576 in *NIPS*. Vol. 28.
- 930 Legendre, S. 1986. Analysis of mammalian communities from the Late Eocene and Oligocene of  
Southern France. *Paleovertebrata* 16:191–212.
- 932 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  
934 Academy of Sciences* 105:6097–6102.

- Liow, L. H., M. Fortelius, K. Lintulaakso, H. Mannila, and N. C. Stenseth. 2009. Lower Extinction  
936 Risk in SleeporHide Mammals. *The American Naturalist* 173:264–272.
- Lloyd, G. T., J. R. Young, and A. B. Smith. 2011. Taxonomic Structure of the Fossil Record is  
938 Shaped by Sampling Bias. *Systematic Biology* 61:80–89.
- Losos, J. B. 2010. Adaptive radiation, ecological opportunity, and evolutionary determinism. *The*  
940 *American naturalist* 175:623–39.
- Losos, J. B., and D. L. Mahler. 2010. Adaptive radiation: the interaction of ecological opportunity,  
942 adaptation, and speciation. Chap. 15, pages 381–420 *in* M. A. Bell, D. J. Futuyma, W. F. Eanes,  
and J. S. Levinton, eds. *Evolution since Darwin: the first 150 years*. Sinauer Associates,  
944 Sunderland, MA.
- Luo, Z.-X., A. W. Crompton, and A.-L. Sun. 2001. A New Mammaliaform from the Early Jurassic  
946 and Evolution of Mammalian Characteristics. *Science* 292:1535–1540.
- Marcot, J. D. 2014. The fossil record and macroevolutionary history of North American ungulate  
948 ungulate mammals: standardizing variation in intensity and geography of sampling. *Paleobiology*  
40:237–254.
- 950 McElreath, R. 2016. *Statistical rethinking: a Bayesian course with examples in R and Stan*. CRC  
Press, Boca Raton, FL.
- 952 McKenna, R. T. 2011. Potential for Speciation in Mammals Following Vast , Late Miocene Volcanic  
Interruptions in the Pacific Northwest. Masters. Portland State University.
- 954 Mendoza, M., C. M. Janis, and P. Palmqvist. 2006. Estimating the body mass of extinct ungulates:  
a study on the use of multiple regression. *Journal of Zoology* 270:90–101.
- 956 Mittelbach, G. G., and D. W. Schemske. 2015. Ecological and evolutionary perspectives on  
community assembly. *Trends in Ecology and Evolution* 30:241–247.
- 958 Novack-Gottshall, P. M. 2007. Using a theoretical ecospace to quantify the ecological diversity of

- Paleozoic and modern marine biotas Using a theoretical ecospace to quantify the ecological  
960 diversity of Paleozoic and modern marine biotas. *Paleobiology* 33:273–294.
- Pires, M. M., D. Silvestro, and T. B. Quental. 2015. Continental faunal exchange and the  
962 asymmetrical radiation of carnivores. *Proceedings of the Royal Society B: Biological Sciences*  
282:20151952.
- 964 Pollock, L. J., W. K. Morris, and P. A. Vesk. 2012. The role of functional traits in species  
distributions revealed through a hierarchical model. *Ecography* 35:716–725.
- 966 Quental, T. B., and C. R. Marshall. 2013. How the Red Queen Drives Terrestrial Mammals to  
Extinction. *Science* 341:290–292.
- 968 Rabosky, D. L., and A. H. Hurlbert. 2015. Species Richness at Continental Scales Is Dominated by  
Ecological Limits. *The American Naturalist* 185:000–000.
- 970 Raia, P., F. Carotenuto, F. Passaro, D. Fulgione, and M. Fortelius. 2012. Ecological specialization  
in fossil mammals explains Cope’s rule. *The American Naturalist* 179:328–37.
- 972 Royle, J. A., and R. M. Dorazio. 2008. Hierarchical modeling and inference in ecology: the analysis  
of data from populations, metapopulations and communities. Elsevier, London.
- 974 ———. 2012. Parameter-expanded data augmentation for Bayesian analysis of capture-recapture  
models. *Journal of Ornithology* 152:521–537.
- 976 Royle, J. A., R. M. Dorazio, and W. a. Link. 2007. Analysis of Multinomial Models With Unknown  
Index Using Data Augmentation. *Journal of Computational and Graphical Statistics* 16:67–85.
- 978 Royle, J. A., J. D. Nichols, and M. Kéry. 2005. Modelling occurrence and abundance of species  
when detection is imperfect. *Oikos* 110:353–359.
- 980 Rubin, D. B. 1996. Multiple imputation after 18+ years. *Journal of the American Statistical  
Assocaition* 91:473–489.
- 982 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.
- Smith, F. A., J. Brown, J. Haskell, and S. Lyons. 2004. Similarity of mammalian body size across the taxonomic hierarchy and across space and time. *The American Naturalist* 163:672–691.
- 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.
- Stan Development Team. 2016. Stan Modeling Language Users Guide and Reference Manual.
- 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.
- Valkenburgh, B. V. 1999. Major patterns in the history of carnivorous mammals. *Annual Review of Earth and Planetary Sciences* 27:463–493.
- Van Valkenburgh, B. 1990. Skeletal and dental predictors of body mass in carnivores. Pages

- 181–205 in J. Damuth and B. J. Macfadden, eds. Body size in mammalian paleobiology:  
1008 estimation and biological implications. Cambridge University Press, Cambridge.
- Villéger, S., P. M. Novack-Gottshall, and D. Mouillot. 2011. The multidimensionality of the niche  
1010 reveals functional diversity changes in benthic marine biotas across geological time. Ecology  
letters 14:561–8.
- 1012 Wang, S. C., P. J. Everson, H. J. Zhou, D. Park, and D. J. Chudzicki. 2016. Adaptive credible  
intervals on stratigraphic ranges when recovery potential is unknown. Paleobiology 42:240–256.
- 1014 Wang, S. C., and C. R. Marshall. 2016. Estimating times of extinction in the fossil record. Biology  
Letters 12:20150989.
- 1016 Warton, D. I., B. Shipley, and T. Hastie. 2015. CATS regression - a model-based approach to  
studying trait-based community assembly. Methods in Ecology and Evolution 6:389–398.
- 1018 Wilson, J. B. 1999. Guilds, functional types and ecological groups. Oikos 86:507–522.
- Yoder, J. B., E. Clancey, S. Des Riches, J. M. Eastman, L. Gentry, W. Godsoe, T. J. Hagey,  
1020 D. Jochimsen, B. P. Oswald, J. Robertson, B. A. J. Sarver, J. J. Schenk, S. F. Spear, and L. J.  
Harmon. 2010. Ecological opportunity and the origin of adaptive radiations. Journal of  
1022 Evolutionary Biology 23:1581–1596.
- Zachos, J. C., G. R. Dickens, and R. E. Zeebe. 2008. An early Cenozoic perspective on greenhouse  
1024 warming and carbon-cycle dynamics. Nature 451:279–283.
- Zachos, J. C., M. Pagani, L. Sloan, E. Thomas, and K. Billups. 2001. Trends, rhythms, and  
1026 aberrations in global climate 65 Ma to present. Science 292:686–693.

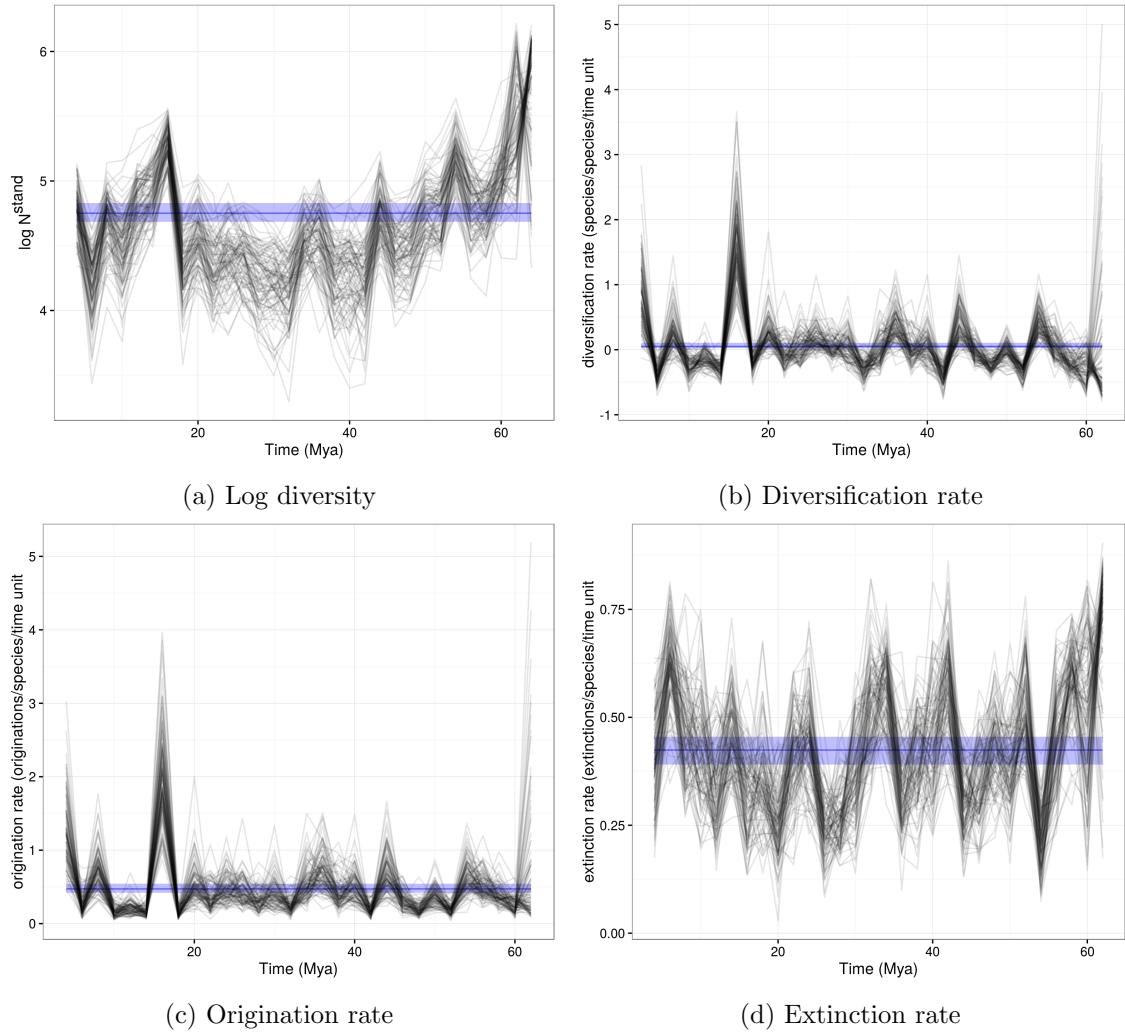


Figure 14: Posterior estimates of the time series of Cenozoic North American mammal diversity and its characteristic macroevolutionary rates; all estimates are from the birth-death model and 100 posterior draws are plotted to indicate the uncertainty in these estimates. The blue horizontal strip corresponds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

Table 15: Posterior probabilities of diversity  $N_t^{stand}$  or diversification rate  $D_t^{rate}$  being greater than average standing diversity  $\bar{N}^{stand}$  or average diversification rate  $\bar{D}^{rate}$  for the whole Cenozoic. The “Time” column corresponds to the top of each of the temporal bins. Diversification rate can not be estimated for the last time point because it is unknown how many more species originated or went extinct following this tempral bin. The estimates are from the birth-death model.

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

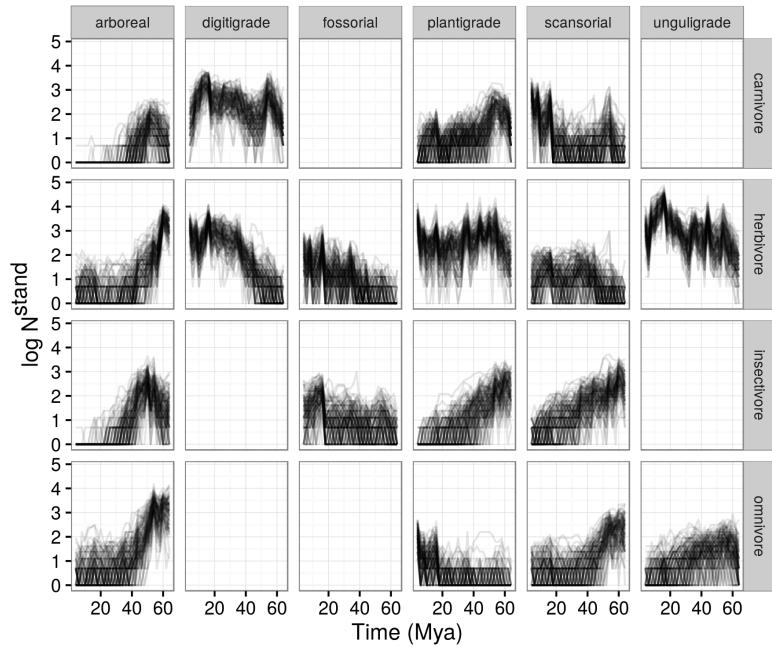


Figure 15: Posterior of standing log-diversity of North American mammals by ecotype for the Cenozoic as estimated from the birth-death model; 100 posterior draws are plotted to indicate the uncertainty in these estimates and what is technically plotted is log of diversity plus 1.

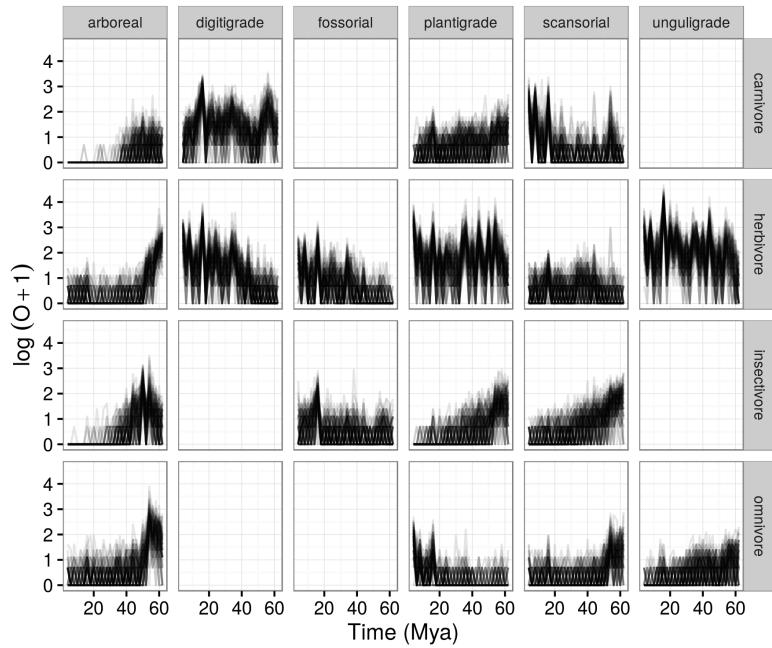


Figure 16: Posterior estimates of the number of origination events from one temporal bin to another, plotted at the bin they originate from. 100 posterior draws are plotted to indicate the uncertainty in these estimates. Also, what is plotted is log of the number of originations plus 1.

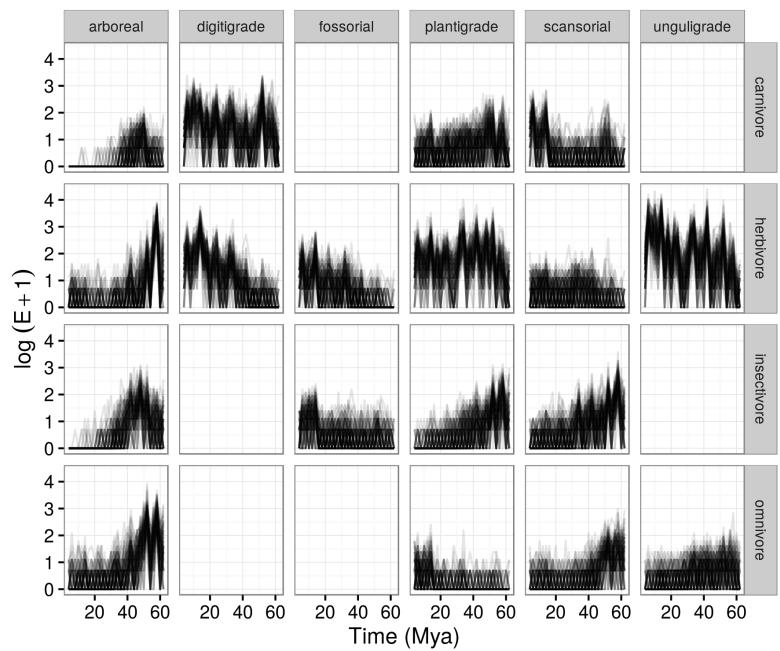


Figure 17: Posterior estimates of the number of extinction events from one temporal bin to another, plotted at the bin they go extinct from. 100 posterior draws are plotted to indicate the uncertainty in these estimates. Also, what is plotted is  $\log$  of the number of extinctions plus 1.