

# Species occurrence as a function of 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 through extinction and extirpation. How a regional  
8 species pool changes over time is the product of many processes acting at multiple levels of  
organization. Changes in the functional composition of a regional species pool are changes that  
10 occur across all local communities drawn from that species pool. While a species' presence in a  
local community is due to the availability of the necessary biotic-biotic or biotic-abiotic  
12 interactions that enable coexistence, a species' presence in a regional species pool just requires  
that at least one local community has that set of necessary interactions. The goal of this  
14 analysis is to understand when, and possibly for what reasons, mammal ecotypes are enriched or  
depleted relative to their average diversity. Here, I analyze the diversity history of North  
16 American mammals ecotypes for most of the Cenozoic (the last 65 million years). This analysis  
frames mammal diversity in terms of both their means of interacting with the biotic and abiotic  
18 environment (i.e. functional group or ecotype) as well as their regional and global environmental  
context. Using two hierarchical Bayesian hidden Markov models of diversity, I find that changes  
20 to mammal diversity are driven more by the influx of new species than by selective extinction. I  
also find that the only ecotypes which experience a near constant increase in diversity over time  
22 are digitigrade and unguligrade herbivores, while arboreal 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 origination of some mammal ecotypes but,  
24 in almost all cases, does not affect the extinction of mammal ecotypes.

## Introduction

26 Changes to species diversity are the result of evolutionary and ecological processes acting both in  
concert and continually. Local communities are shaped by dispersal and local ecological processes  
28 such as resource competition and predator-prey relationships. The constituent species of these

community are drawn from a regional species pool, or the set of all species that are present in at least one community within a region (Harrison and Cornell, 2008; Mittelbach and Schemske, 2015; Urban et al., 2008). Species dispersal from the regional species pool to the local communities is a sorting process shaped by biotic and abiotic environmental filters which are mediated by those species traits (Cottenie, 2005; Elith and Leathwick, 2009; Harrison and Cornell, 2008; Loeuille and Leibold, 2008; Shipley et al., 2006; Urban et al., 2008). Regional species pools are shaped by speciation, extinction, migration, and extirpation. The gain or loss regional diversity reflects macroevolutionary dynamics which, in turn, shape downstream macroecological dynamics of the species pool and its constituent local communities (Harrison and Cornell, 2008; Mittelbach and Schemske, 2015; Urban et al., 2008).

Fundamentally, all species respond differently to climate and environmental change (Blois and Hadly, 2009). Similarities in ecological roles of species within a regional species pool can be described as a collection of guilds or functional groups (Bambach, 1977; Brown and Maurer, 1989; Simberloff and Dayan, 1991; Valentine, 1969; Wilson, 1999). Species within the same functional group are expected to have more similar macroecological dynamics than to species of a different functional group. By focusing on the relative diversity of functional groups, changes to diversity are interpretable as changes to the set of ways species within a species pool could interact with the biotic and abiotic environment.

A key question when comparing communities or regional species pools based their functional composition is whether a functional group is enriched or depleted and why; what are the processes that led to a species pool having the functional composition it does (Blois and Hadly, 2009; Brown and Maurer, 1989; McGill et al., 2006; Smith et al., 2008; Weber et al., 2017). Comparisons of contemporaneous regional species pools only determines if a functional group is enriched or depleted relative to other species pools; this does not reveal if that functional group is enriched or depleted relative to its diversity in the regional species pool over time (Blois and Hadly, 2009). While a species pool may be depleted of a functional group relative to other contemporaneous species pools, that same functional group may be actually be enriched in that species pool relative to its historical diversity. Because the processes which shape regional species pool diversity (e.g.

origination, extinction) operate on much longer time scales than is possible for studies of the

58 Modern, paleontological data provides a unique opportunity to observe and estimate the changes to  
functional diversity and how species functional traits and environmental context can shape the

60 enrichment or depletion of functional groups within a regional species pool (Blois and Hadly, 2009;  
Smith et al., 2008).

62 The paleontological record of North American mammals for the Cenozoic (66 million years ago to  
present) provides one of the best opportunities for understanding how regional species pool

64 functional diversity. The North American mammal record is a relatively complete temporal  
sequence for the entire Cenozoic which primarily, but not exclusively, based on fossil localities from

66 the Western Interior of North America (Alroy, 1996, 2009; Alroy et al., 2000). Additionally,  
mammal fossils preserve a lot of important physiological information, such as teeth, so that

68 functional traits like the dietary/trophic category of species are easy to estimate (Eronen et al.,  
2010; Polly et al., 2011, 2015).

70 The goals of this study are to understand when are unique functional groups, called ecotypes,  
enriched or depleted in the North American mammal regional species pool and to estimate the

72 relationship between changes to regional ecotypic diversity and changes to their environmental  
context.

## 74 **Background**

The diversity history of North American mammals for the Cenozoic is relatively well understood as  
76 it has been the focus of considerable study (Alroy, 1996, 2009; Alroy et al., 2000; Badgley and  
Finarelli, 2013; Blois and Hadly, 2009; Figueirido et al., 2012; Fraser et al., 2015; Janis, 1993; Janis  
78 and Wilhelm, 1993; Pires et al., 2015; Quental and Marshall, 2013; Silvestro et al., 2015; Slater,  
2015; Smits, 2015). Previous approaches to understanding mammal diversity, both in North  
80 America and elsewhere, fall into a number of overlapping categories: total diversity (Alroy, 1996;  
Alroy et al., 2000; Figueirido et al., 2012; Liow et al., 2008), with/between guild comparisons (Janis  
82 et al., 2004; Janis, 2008; Janis et al., 2000; Janis and Wilhelm, 1993; Jernvall and Fortelius, 2004;

Pires et al., 2015), within/between clade comparisons (Cantalapiedra et al., 2017; Fraser et al.,  
84 2015; Quental and Marshall, 2013; Silvestro et al., 2015; Slater, 2015), and estimating the impact of  
environmental process on diversity (Alroy et al., 2000; Badgley and Finarelli, 2013; Badgley et al.,  
86 2017; Blois and Hadly, 2009; Eronen et al., 2015; Fraser et al., 2015; Janis, 1993; Janis and  
Wilhelm, 1993). Each of these individual perspectives provide a limited perspective on the  
88 macroevolutionary and macroecological processes shaping diversity and diversification. Integration  
across perspectives is necessary for producing a holistic and internally consistent picture of how the  
90 North American mammal species pool has changed through time. One of the goals of this study is  
to present a framework for approaching hypotheses about diversity and diversification through  
92 multiple lenses simultaneously so that our inferences are better constrained and the relative  
importance of various functional traits and environmental factors may be better elucidated.  
  
94 The narrative of the diversification of North American mammals over the Cenozoic is one of  
gradual change. There is little convincing evidence that there have been any major or sudden  
96 cross-functional group or cross-taxonomic turnover events for the entire Cenozoic (Alroy, 1996, 2009;  
Alroy et al., 2000; Eronen et al., 2015; Janis, 1993). Instead of being concentrated in time, turnover  
98 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 should not occur simultaneously for  
100 all functional groups under study. Additionally, changes to mammal diversification seem to be  
primarily driven by changes to origination rate and not to extinction (Alroy, 1996, 2009; Alroy  
102 et al., 2000). An unresolved aspect of the general history of mammal diversification is whether that  
diversity is limited or self-regulating; namely, to what extent is mammal diversification  
104 diversity-dependent (Alroy, 2009; Harmon and Harrison, 2015; Rabosky, 2013; Rabosky and  
Hurlbert, 2015). Similarly, this question can also be asked of specific functional groups (Jernvall  
106 and Fortelius, 2004; Quental and Marshall, 2013; Silvestro et al., 2015; Valkenburgh, 1999).

Within the overall narrative of mammal diversity, the histories of a selection of taxonomic and  
108 functional groups are better understood. These groups have particularly good fossil records and/or  
have been the focus of previous analyses.

110 The diversity history of ungulate herbivores has been characterized by more recently originating  
111 taxa having longer legs, higher crowned teeth, and a shift from graze-dominated to  
112 browse-dominated diets than their earlier originating counterparts (Cantalapiedra et al., 2017;  
113 Fraser et al., 2015; Janis et al., 2004; Janis, 2008; Janis et al., 2000; Janis and Wilhelm, 1993). The  
114 mechanisms which drive this pattern are theorized to be some combination of tectonic activity  
115 driving environmental change such as the drying of the western interior of North America due  
116 mountain building and global temperature and environmental change such as the formation of polar  
icecaps (Badgley et al., 2017; Blois and Hadly, 2009; Eronen et al., 2015; Janis, 2008).

118 In contrast, the origination of modern cursorial carnivore forms was not until much later in the  
119 Cenozoic; this is not to say that carnivore diversity only grew in the late Cenozoic, but that those  
120 forms were late entrants. Instead, the diversity history of carnivores is reflective of  
121 density-dependence or some other form of self-regulation (Silvestro et al., 2015; Slater, 2015;  
122 Valkenburgh, 1999). Specifically, it has been proposed that different canid clades have replaced each  
123 other as the dominate members of that macroecological role within the species pool (Silvestro et al.,  
124 2015; Valkenburgh, 1999). It is then expected that for this analysis the diversity of digitigrade and  
125 plantigrade carnivores (i.e. the “carnivore” guild of Valkenburgh (1999)) should be relatively  
126 constant for the Cenozoic, but possibly more so the Neogene than the Paleogene.

In a relevant study, Smits (2015) found that functional traits such as a species dietary or locomotor  
128 category structure differences in mammal extinction risk. In particular, arboreal taxa were found to  
have a shorter duration on average than species from other locomotor categories (Smits, 2015). Two  
130 possible scenarios that could yield this pattern were proposed: the extinction risk faced by arboreal  
species is constant and high for the entire Cenozoic or the Paleogene and Neogene represent  
132 different regimes and extinction risk increased in the Neogene, thus driving up the Cenozoic average  
extinction risk. These two possible explanations have clear and testable predictions with respect to  
134 the diversity history of arboreal taxa: 1) if arboreal taxa always have an elevated extinction risk  
when compared to other taxa, then the diversity history of arboreal taxa is expected to be constant  
136 with time, albeit possibly at low diversity; and 2) if the Paleogene and Neogene represent difference  
selective regimes with the former being associated with lower extinction risk than the latter, then

138 the diversity history of arboreal taxa are expected to be present in the Paleogene but depleted or absent from the species pool during the Neogene.

140 The climate history of the Cenozoic can be broadly described as a gradual cooling trend, with polar ice-caps forming in the Neogene (Cramer et al., 2011; Zachos et al., 2008, 2001). There are of course  
142 exceptions to this pattern such as the Eocene climatic optimum, the mid-Miocene climatic optimum, and the sudden drop in temperature at the Eocene/Oligocene boundary (Zachos et al.,  
144 2008, 2001). In terms of the biotic environment of North America, the Cenozoic is additionally characterized by major transition from having closed, partially forested biomes being common in  
146 the Paleogene to the landscape being dominated by savannah and grasslands biomes by the Neogene (Blois and Hadly, 2009; Janis, 1993; Janis et al., 2000; Strömberg, 2005); though this  
148 transition was not itself coincident with the transition within herbivores from browse-dominated to graze-dominated diets (Strömberg, 2005). Additionally, the fundamental landscape of North  
150 America changed substantially over the Cenozoic with mountain uplift and other tectonic actives in Western North America (Badgley and Finarelli, 2013; Blois and Hadly, 2009; Eronen et al., 2015;  
152 Janis, 2008). This type of geological activity affects both local climates as well as continental weather patterns while also mobilizing increased grit into the environment, something which may  
154 be responsible for increasing trend of hyposodony (high crowned teeth) among herbivores (Damuth and Janis, 2011; Jardine et al., 2012; Jernvall and Fortelius, 2002).

156 The effect of climate on mammal diversity and its accompanying diversification process has been the focus of considerable research with a slight consensus favoring mammal diversification being  
158 more biologically-mediated than climate-mediated (Alroy, 1996; Alroy et al., 2000; Clyde and Gingerich, 1998; Figueirido et al., 2012). However, differences in temporal and geographic scale  
160 seem to underly the contrast between these two perspectives. For example when the mammal fossil record analyzed at small temporal and geographic scales a correlation between diversity and climate  
162 is observable (Clyde and Gingerich, 1998). However, when the record is analyzed at the scale of the continent and most of the Cenozoic there is no correlation with diversity and climate (Alroy et al.,  
164 2000). This result, however, does not go against the idea that there may be short periods of correlation and that the correlation between diversity and climate can change or even reverse

<sup>166</sup> direction over time; this type result means that there is no single direction of correlation between diversity and climate (Figueirido et al., 2012).

<sup>168</sup> 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

<sup>170</sup> 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

<sup>172</sup> 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

<sup>174</sup> 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

<sup>176</sup> can be observed rather than assumed.

## Foreground

<sup>178</sup> Fourth-corner modeling is an approach to explaining the patterns of either species abundance or presence/absence as a product of species traits, environmental factors, and the interaction between

<sup>180</sup> traits and environment (Brown et al., 2014; Jamil et al., 2013; Pollock et al., 2012; Warton et al., 2015); effectively uniting species distribution modeling (SDMs) with trait-based community

<sup>182</sup> assembly (CATS, MaxEnt). In modern ecological studies, what is being modeled is species occurrences at localities distributed across a region (Jamil et al., 2013; Pollock et al., 2012). In this

<sup>184</sup> study, what is being modeled is the pattern of species occurrence over time for most of the Cenozoic in North America (Fig. 1). By incorporating an additional dimension (time) to the

<sup>186</sup> fourth-corner framework we can gain better inference of how an instantaneous species pool (i.e. the Modern) is assembled over time. These two approaches, modern and paleontological, are different

<sup>188</sup> views of the same three-dimensional pattern: species at localities over time. The temporal limitations of modern ecological studies and difficulties with uneven spatial occurrences of fossils in

<sup>190</sup> paleontological studies means that these approaches are complementary but reveal different patterns of how species are distributed in time and space.

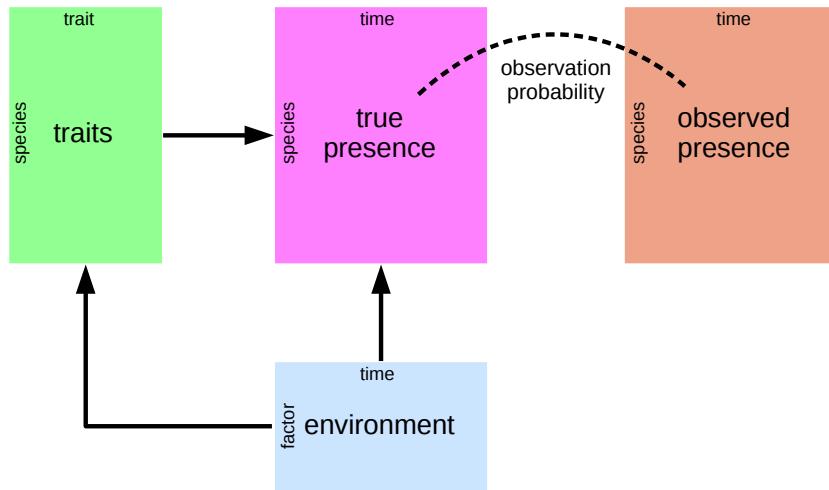


Figure 1: Conceptual diagram of the paleontological fourth corner problem. The observed presence matrix (orange) is the empirical presence/absence pattern for all species for all time points; this matrix is an incomplete observation of the “true” presence/absence pattern (purple). The estimated true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

- 192 My approach to delimiting and assigning mammal functional groups is inspired on the ecocube  
 heuristic used to classify marine invertebrate species by three functional traits (Bambach et al.,  
 194 2007; Bush and Bambach, 2011; Bush et al., 2007; Bush and Novack-Gottshall, 2012;  
 Novack-Gottshall, 2007; Villéger et al., 2011). Unique combinations of traits represent ecotypes,  
 196 which are equivalent to functional groups defined by functional traits instead of a holistic  
 understanding how a taxon interacts with its environment. In this study, the two functional traits  
 198 used to define a species’ ecotype are dietary (e.g. herbivore, carnivore, etc.) and locomotor category  
 (e.g. arboreal, unguligrade, etc.). Species body mass was also included as a species trait in this  
 200 analysis, but not as a functional trait for defining ecotypes; instead, its inclusion is principally to  
 control for differences in species dynamics that driven by mass and not ecotype.
- 202 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,  
204). 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.  
206 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  
208 selective pressures due to its environmental context.

All observations, paleontological or modern, are made with uncertainty. With presence/absence  
210 data this uncertainty comes from not 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  
212 et al., 2005; Wang and Marshall, 2016). For paleontological data, the incomplete preservation and  
sampling of species means that the true times of origination or extinction may not be observed  
214 (Foote, 2001; Foote and Sepkoski, 1999; Wang et al., 2016; Wang and Marshall, 2016). The model(s)  
I propose below represent an attempt to translate the verbal/visual model described here (Fig. 1)  
216 into a statistical model for estimating the relative diversity of mammal ecotypes over time and how  
those ecotypes respond to changes to environmental context while taking into account the  
218 fundamental incompleteness of the fossil record.

Ultimately, the goals of this analysis are to understand when unique ecotypes enriched or  
220 depleted in the North American mammal regional species pool and how these changes in ecotypic  
diversity are related to changes in species’ environmental context. In the analyses done here, many  
222 covariates which describe a species’ macroecology and its environmental context are considered. In  
order to analyze this complex and highly structured data set, I developed a hierachal Bayesian  
224 model combing the fourth-corner modeling approach with a model of an observation-occurrence or  
observation-origination-extinction process.

<sup>226</sup> **Materials and Methods**

**Taxon occurrences and species-level information**

<sup>228</sup> All fossil occurrence information used in this analysis was downloaded from the Paleobiology Database (PBDB). The initial download restricted occurrences to Mammalia observed in North America between the Maastrichtian (72-66 Mya) and Gelasian (2.58-1.8 Mya) stages (Cohen et al., 2015). Occurrences were then further limited to those occurring between 64 and 2 million years ago (Mya); this age restriction was to insure that observation time series lines up with the temperature time series (Cramer et al., 2011). Taxonomic, stratigraphic, and ecological metadata for each occurrence and species was also downloaded. A new download for a raw, unfiltered PBDB datafile following the same criterion used here is available at <http://goo.gl/2s1geU>. The raw datafile used as a part of this study, along with all code for filtering and manipulating this download is available at <http://github.com/psmits/copings>.

<sup>238</sup> 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 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 were based on taxonomies present in the Encyclopedia of Life (<http://eol.org>) and Janis et al. (2008, 1998). All taxa whose life habit was classified as either volant (i.e. Chiroptera) or aquatic (e.g. Cetacea) were excluded from this analysis because of their lack of direct applicability to the study of terrestrial species pools.

<sup>246</sup> Species ecotype is defined based on a combination of locomotor and diet categories; the goal is to classify species based on the manner with which they interact with their environment. Most mammal species records in the PBDB have life habit (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 1. Additionally for this study, the life history category was further broken up to better reflect the diversity of

252 mammal locomotor modes. Ground dwelling species locomotor categories were reassigned based on  
 the ankle posture associated with their taxonomic group, as described in Table 2 (Carrano, 1999).  
 254 Ankle posture was assumed uniform for all species within a taxonomic group except for those  
 species assigned a non-ground dwelling locomotor category by the PBDB. All species for which it  
 256 was possible to assign a locomotor category had one assigned, including species for which  
 post-crana are unknown but for which a taxonomic grouping is known. Ground dwelling species  
 258 which were unable to be reassigned based on posture were excluded from analysis. Finally, ecotype  
 categories with less than 10 total species were excluded, yielding a total of 18 observed ecotypes out  
 260 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.

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.

Order	Family	Stance
	Ailuridae	plantigrade
	Allomyidae	plantigrade
	Amphicyonidae	plantigrade
	Amphilemuridae	plantigrade
	Anthracotheriidae	digitigrade
	Antilocapridae	unguligrade
	Apheliscidae	plantigrade

Continued on next page

**Table 2 – continued from previous page**

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

Continued on next page

**Table 2 – continued from previous page**

Order	Family	Stance
	<i>Florentiamyidae</i>	plantigrade
	<i>Gelocidae</i>	unguligrade
	<i>Geolabididae</i>	plantigrade
	<i>Glyptodontidae</i>	plantigrade
	<i>Gomphotheriidae</i>	unguligrade
	<i>Hapalodectidae</i>	plantigrade
	<i>Heteromyidae</i>	digitigrade
	<i>Hyaenidae</i>	digitigrade
	<i>Hyaenodontidae</i>	digitigrade
	<i>Hypertragulidae</i>	unguligrade
	<i>Ischyromyidae</i>	plantigrade
	<i>Jimomyidae</i>	plantigrade
Lagomorpha		digitigrade
	<i>Leptictidae</i>	plantigrade
	<i>Leptochoeridae</i>	unguligrade
	<i>Leptomerycidae</i>	unguligrade
	<i>Mammutidae</i>	unguligrade
	<i>Megalonychidae</i>	plantigrade
	<i>Megatheriidae</i>	plantigrade
	<i>Mephitidae</i>	plantigrade
	<i>Merycoidodontidae</i>	digitigrade
Mesonychia		unguligrade
	<i>Mesonychidae</i>	digitigrade
	<i>Micropternodontidae</i>	plantigrade
	<i>Mixodectidae</i>	plantigrade

Continued on next page

**Table 2 – continued from previous page**

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

Continued on next page

**Table 2 – continued from previous page**

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

- <sup>262</sup> 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
- <sup>264</sup> (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
- <sup>266</sup> 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
- <sup>268</sup> 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).
- <sup>270</sup> Mass was log-transformed and then rescaled by first subtracting mean log-mass from all mass estimates, then dividing by two-times its standard deviation; this insures that the magnitude of
- <sup>272</sup> effects for both continuous and discrete covariates are directly comparable (Gelman, 2008; Gelman and Hill, 2007).
- <sup>274</sup> All fossil occurrences from 64 to 2 million years ago (Mya) were binned into 31 2 million year (My) bins. This temporal length was chosen because it is approximately the resolution of the North

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)

276 American mammal fossil record (Alroy, 1996, 2009; Alroy et al., 2000; Marcot, 2014).

## Environmental and temporal covariates

278 The environmental covariates used in this study are collectively referred to as group-level covariates  
because they predict the response of a “group” of individual-level observations (i.e. species  
280 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  
282 group-level covariates in this study are two global temperature estimates and the Cenozoic “plant  
phases” defined by Graham (2011).  
284 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  
286 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  
288 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

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

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 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 take place during the time frame being analyzed. 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 boundaries 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 lies in whether the probabilities of a species originating or surviving are considered equal or different (Table 5). 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:

		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).

Table 6: Parameters 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 $r$

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). Thus, in the transition matrices the probability of an extinct species changing states is 0 (Table 5). See below for parameter explanations (Tables 6, 7, and 8).

## Observation process

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 one time to another, and probability  $\pi$  of a species first appearing (Royle and Dorazio, 2008). In this formulation, the probability of a species becoming extinct is  $1 - \phi$ . For the pure-presence model  $\phi = \pi$ , while for the birth-death model  $\phi \neq \pi$ .

The probability  $p$  of observing a species that is present is modeled as a logistic regression with a

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 $a_d$

322 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  
 324 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.

### 326 Pure-presence process

For the pure-presence model there is only a single probability dealing with the presence of a species  
 328  $\theta$  (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  
 330 with the pure-presence model are presented in Table 7, and the full sampling statement in Equation  
 2.

332 Presence is a logistic regression where the intercept varies by ecotype (Eq. 2). Species mass was  
 included as a covariate associated with two regression coefficients allowing a quadratic relationship

<sup>334</sup> with log-odds of occurrence. Because the distribution of mammal species body mass is unimodal  
and approximately normal (Smith et al., 2004), I assume that species of intermediate body size will  
<sup>336</sup> be more common than species of very large or very small mass. These assumptions are also  
reflected in the choice of priors for  $b_1$  and  $b_2$  where the latter is given a weakly informative prior  
<sup>338</sup> with most of the density is below 0 (Eq. 2).

The values of each ectype's intercept are themselves modeled as regressions using the group-level  
<sup>340</sup> covariates associated with environmental context. Each of these regressions has an associated  
variance of possible values of each ectype's intercept (Gelman and Hill, 2007). In addition, the  
<sup>342</sup> covariances between ectype intercepts, given this group-level regression, are modeled (Gelman and  
Hill, 2007).

<sup>344</sup> 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  
<sup>346</sup> 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  
<sup>348</sup> 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  
<sup>350</sup> approach is that a result of 0 or “no effect” is more preferable to a wrong or extremely weak result.

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

## Birth-death process

- 352 In the birth-death version of the model,  $\phi \neq \pi$  and so each of these probabilities is modeled  
 separately but each is handled in a similar manner to how  $\theta$  is modeled in the pure-presence model  
 354 (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:

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

- 356 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 between  
 358 mass and both  $\phi$  and  $\pi$  are reflected in the parameterization of the model and choice of priors (Eq.  
 3).  
 360 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 $a_d^\pi$

<sup>362</sup> pure presence model (Eq. 2).

## Posterior inference and model adequacy

<sup>364</sup> 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  
<sup>366</sup> models used here 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; this causes  
<sup>368</sup> 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  
<sup>370</sup> possible states of the latent parameters  $z$  were calculated and summed (i.e. marginalized).

Species durations at minimum range through from their first appearance to their last appearance in  
<sup>372</sup> the fossil record, but the incompleteness of all observations means that the actual times of  
origination and extinction are unknown. The marginalization approach used here means that the

374 probabilities of all possible histories for a species are calculated, from the end members of the  
 species having existed for the entire study interval and the species having only existed between the  
 376 directly observed FAD and LAD to all possible intermediaries (Fig 2) (Stan Development Team,  
 2016). This process is identical, language-wise, to assuming range-through and then estimating the  
 378 possibility of range extension due to incomplete 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. By marginalizing over all possible occurrence histories, the probability of each potential history is estimated. The process of parameter marginalization is described in the text.

The combined size of the dataset and large number of parameters in both models (Eqs. 2, 3),  
 380 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  
 382 implemented. Instead, an approximate Bayesian approach was used: variational inference. A  
 recently developed automatic variational inference algorithm called “automatic differentiation  
 384 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

386 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  
388 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.  
390 standard deviation).

Of additional concern for posterior inference is the partial identifiability of observation parameters  
392  $p_{t=1}$  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  
394 “gaps” in species occurrence histories that are guaranteed to be filled. Instead, the values of the first  
and final columns of the “true” presence-absence matrix  $z$  for those observations that do not already  
396 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  
398 values of  $p_{t=1}$  and  $p_{t=T}$  towards the overall mean of  $p$  (Gelman et al., 2013), and in fact this  
approach might be more analytically sound than the more ad-hoc approaches that are occasionally  
400 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  $\theta$ ,  $\phi$   
402 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  
404 means of those parameters.

After fitting both models (Eqs. 2, 3) using ADVI, model adequacy and quality of fit were assessed  
406 using a posterior predictive check (Gelman et al., 2013). By simulating 100 theoretical data sets  
from the posterior estimates of the model parameters and the observed covariate information the  
408 congruence between predictions made by the model and the observed empirical data can be  
assessed. These datasets are simulated by starting with the observed states of the presence-absence  
410 matrix at  $t = 1$ ; from there, the time series roll forward as stochastic processes with covariate  
information given from the empirical observations. Importantly, this is fundamentally different from  
412 observing the posterior estimates of the “true” presence-absence matrix  $z$ . The posterior predictive  
check used in this study is to compare the observed average number of observations per species to a

<sup>414</sup> distribution of simulated averages; if the empirically observed value sits in the middle of the  
 distribution then the model can be considered adequate in reproducing the observed number of  
<sup>416</sup> occurrences per species.

The ADVI assumption of a purely Gaussian posterior limits the utility and accuracy of the  
<sup>418</sup> posterior predictive checks because parameter estimates do not reflect the true posterior  
 distribution and are instead just an approximation (Gelman et al., 2013). Because of this, posterior  
<sup>420</sup> predictive estimates are themselves only approximate checks of model adequacy. The posterior  
 predictive check that is used in this study focuses on mean occurrence and not to any scale  
<sup>422</sup> parameters that might be most affected by the ADVI assumptions.

Given parameter estimates, diversity and diversification rates are estimated through posterior  
<sup>424</sup> predictive simulations. Given the observed presence-absence matrix  $y$ , estimates of the true  
 presence-absence matrix  $z$  can be simulated and the distribution of possible occurrence histories  
<sup>426</sup> can be analyzed. This is conceptually similar to marginalization where the probability of each  
 possible occurrence history is estimated (Fig. 2).

<sup>428</sup> The posterior distribution of  $z$  gives the estimate of standing diversity  $N_t^{stand}$  for all time points as

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

Given estimates of  $N^{stand}$  for all time points, the estimated number of originations  $O_t$  is estimated  
<sup>430</sup> as

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

and number of extinctions  $E_t$  estimated as

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

<sup>432</sup> Per-capita growth  $D^{rate}$ , origination  $O^{rate}$  and extinction  $E^{rate}$  rates are then calculated as

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

## Results

<sup>434</sup> The results of the analyses described above take one of two forms: direct inspection of parameter posterior estimates from both models, and downstream estimates of diversity and diversification <sup>436</sup> rates based on posterior predictive simulations from the birth-death model because this model has a better fit to the observed occurrence information.

<sup>438</sup> **Comparing parameter estimates from the pure-presence and birth-death models**

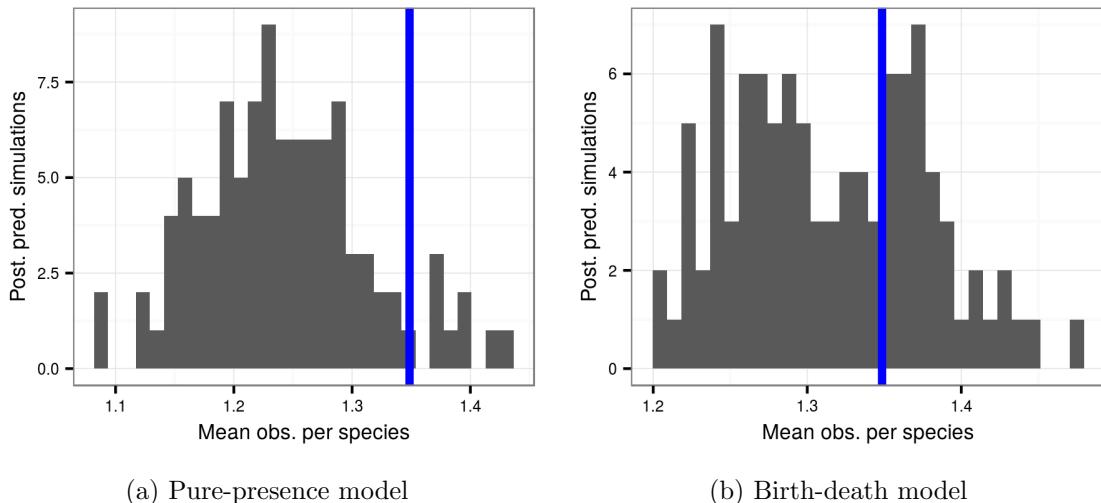


Figure 3: Comparison of the average observed number of occurrences per species (blue line) to the average number of occurrences from 100 posterior predictive datasets using the posterior estimates from the pure-presence and birth-death models.

<sup>440</sup> Comparison of the posterior predictive results from the pure-presence and birth-death models

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

Occurrence probabilities estimated from the pure-presence model (Fig. 4) are broadly similar to the  
450 estimates of origination probability from the birth-death model (Fig. 5) but not the survival  
probability estimates (Fig. 6). This result supports the idea that changes to the North American  
452 regional species pool is more likely due to changes in origination than extinction, a result to which I  
will return to later in the discussion of per-capita diversification, origination, and extinction rates.

454 For most ecotypes, both estimated occurrence probabilities from the pure-presence model (Fig. 4)  
and origination probabilities estimated from the birth-death model (Fig. 5) increase with time. This  
456 makes sense given that, over time, all species that have at least one observed occurrence must have  
had that occurrence by the last time point, so our certainty in a species occurring must increase  
458 with time. Importantly, there are potential issues surrounding the partial identifiability of the  
observation parameters  $p$  which may contribute to edge effects in estimates of occurrence,  
460 origination, and extinction (Royle and Dorazio, 2008). Notably, ecotypes with arboreal components  
do not appear to follow a similar pattern; instead, occurrence and origination probabilities appear  
462 relatively flat with high posterior variance for most of the Cenozoic.

The dramatic differences in the estimates of origination and survival probabilities are indicative of  
464 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.  
466 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

<sup>468</sup> probabilities have relatively stable means for the entire Cenozoic (Fig. 5, and 6).

For most ecotypes, the estimates of origination probabilities involve less uncertainty than similar  
<sup>470</sup> estimates of survival probabilities (Fig. 5, and 6). In logistic regression, high uncertainty in the  
estimates of the underlying log-odds of occurrence, origination, or survival tends to indicate  
<sup>472</sup> extreme rarity or complete absence of the specific ecotype. The latter is called complete separation  
and occurs when there is no uncertainty in the effect of a covariate on presence/absence. The  
<sup>474</sup> problem of complete separation is mitigated by the hierarchical modeling strategy used here  
(Gelman et al., 2013; Gelman and Hill, 2007; McElreath, 2016).

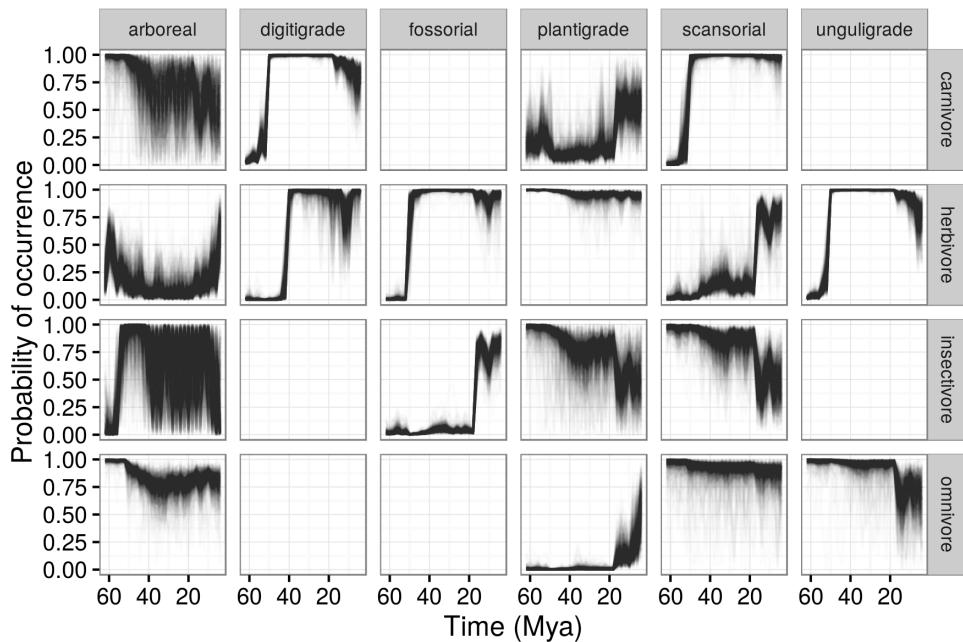


Figure 4: Probability of a mammal ecotype occurring over time as estimated from the pure-presence model. Each panel depicts 100 random samples from the model's posterior. The columns are by locomotor category and rows by dietary category; their intersections are the observed and analyzed ecotypes. Panels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

<sup>476</sup> 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  
<sup>478</sup> estimated to not have a strong effect on the probability of sampling a species that is presence (Fig.  
7a). In contrast, for the birth-death model mass is found to have a negative relationship with  
<sup>480</sup> observation such that larger species are less likely to be observed if present than smaller species

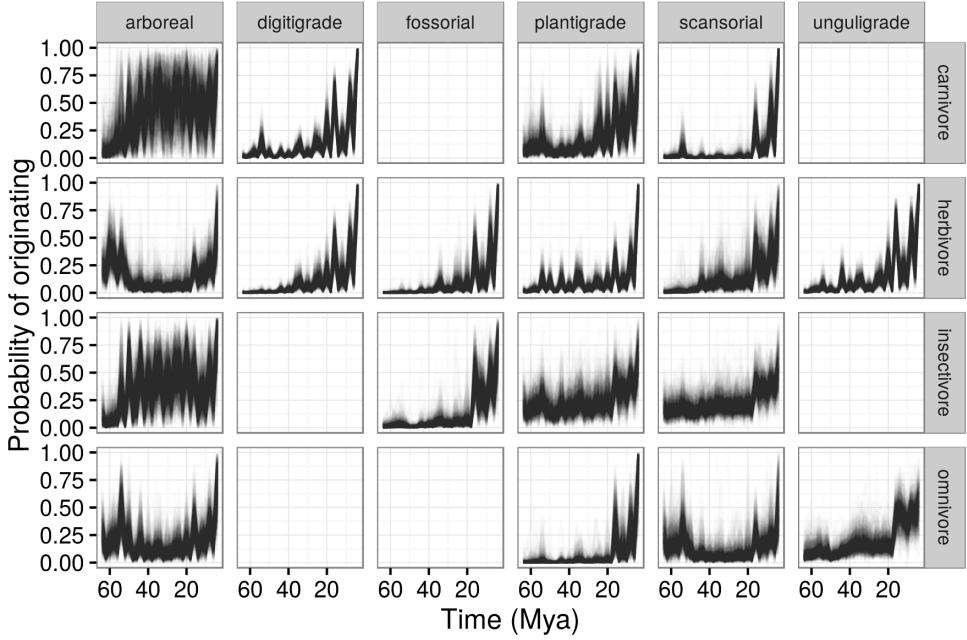


Figure 5: Probability of a mammal ecotype origination probabilities at each time point as estimated from the birth-death model. Each panel depicts 100 random samples from the model’s posterior. The columns are by locomotor category and rows by dietary category; their intersections are the observed and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

(Fig. 7b). The result from the birth-death model is partially expected based on previous research

482 (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  
 484 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  
 486 necessarily translate to high preservation over multiple time points.

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

488 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  
 490 model (Fig.7). The probability that one estimate is different from the other, however, is not directly  
 calculable as they come from different models; what this tells us is how adding more information to  
 492 the model (i.e. replacing occurrence with origination and extinction) changes parameter estimates  
 in the model.

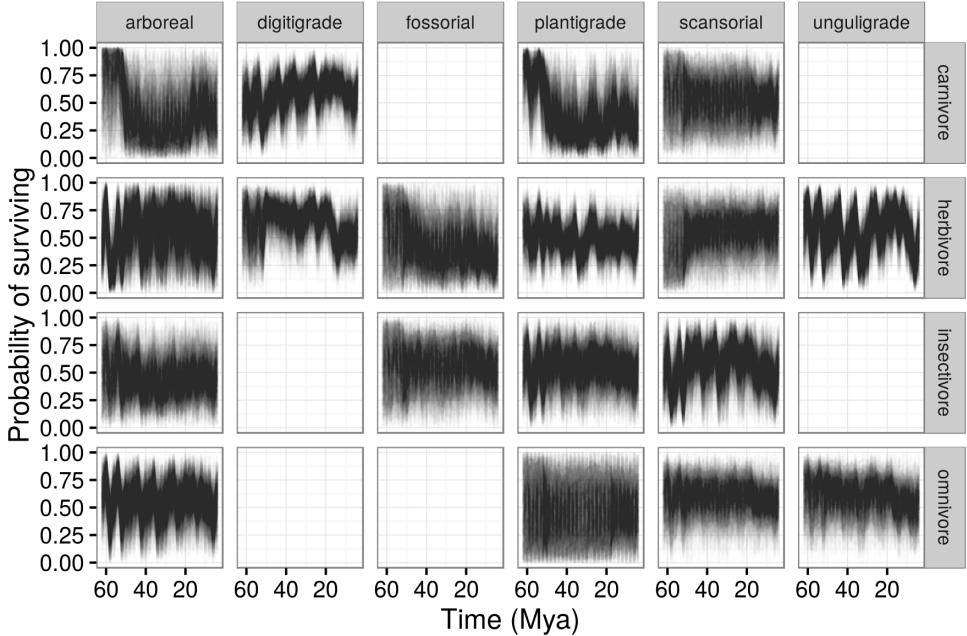


Figure 6: Probability of a mammal ecotype survival probabilities at each time point as estimated from the birth-death model. Each panel depicts 100 random samples from the model’s posterior. The columns are by locomotor category and rows by dietary category; their intersections are the observed and analyzed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

- 494 The effect of species mass on probability of occurrence as estimated from the pure-presence model  
 (Fig. 8) is most similar to the estimated effect of species mass on probability of origination for the  
 496 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  
 498 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  
 500 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  
 502 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.
- 504 In contrast, the effect of species mass on probability of survival as estimated from the birth-death  
 model (Fig. 10) is consistent with previous findings that there is little effect of mass on extinction  
 506 for North American mammals for the Cenozoic (Smits, 2015; Tomiya, 2013). Note that all variation

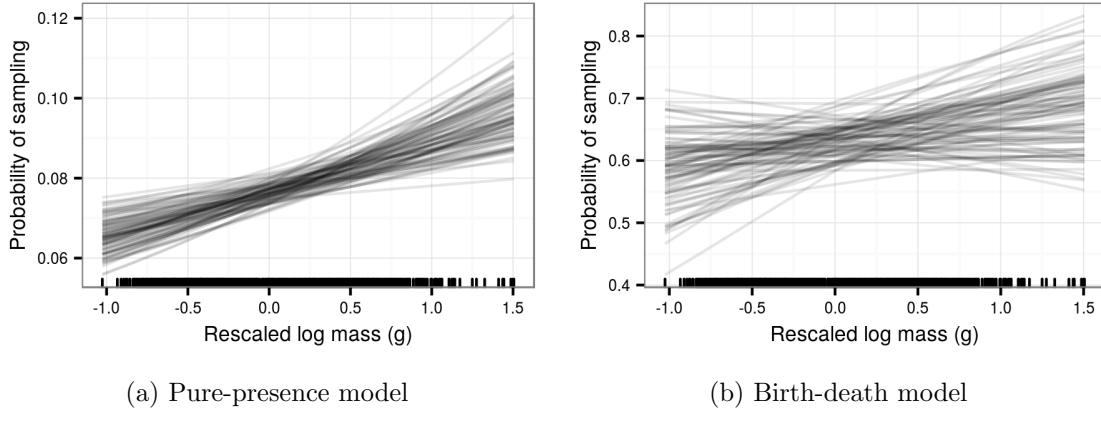


Figure 7: Estimates of the effect of species mass on probability of sampling a present species ( $p$ ). Mass has been log-transformed, centered, and rescaled; this means that a mass of 0 corresponds to the mean of log-mass of all observed species and that mass is in standard deviation units. Estimates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

between ecotypes depicted in Figure 10 is due to differences in ecotype-specific survival probability  
508 and the associated effects of plant phase; the effect of mass was considered constant for all ecotypes  
(Eqs. 2, 3).

510 Similarities in parameter estimates between ecotypes may be due to a similar response to  
environmental factors (Fig. 11, 12, and 13). As with previous comparisons between posterior  
512 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  
514 origination (Fig. 12) as opposed to survival (Fig. 13). As demonstrated in the comparisons of the  
effect of body mass on occurrence from the pure-presence model (Fig. 8) with the effect of body  
516 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.

518 An association between plant phase and differences in the log-odds of occurrence (Fig. 11),  
origination (Fig. 12), or extinction (Fig. 13) is interpreted to mean that the set of possible  
520 mammal-plant interactions was relatively more favorable (positive association) or less so (negative  
association) to those ecotypes. In the case of species origination, for example, more favorable  
522 conditions for an ecotype may indicate an increasing number of possible and available  
mammal-plant interactions (e.g. ecological opportunity; Losos, 2010; Losos and Mahler, 2010; Yoder

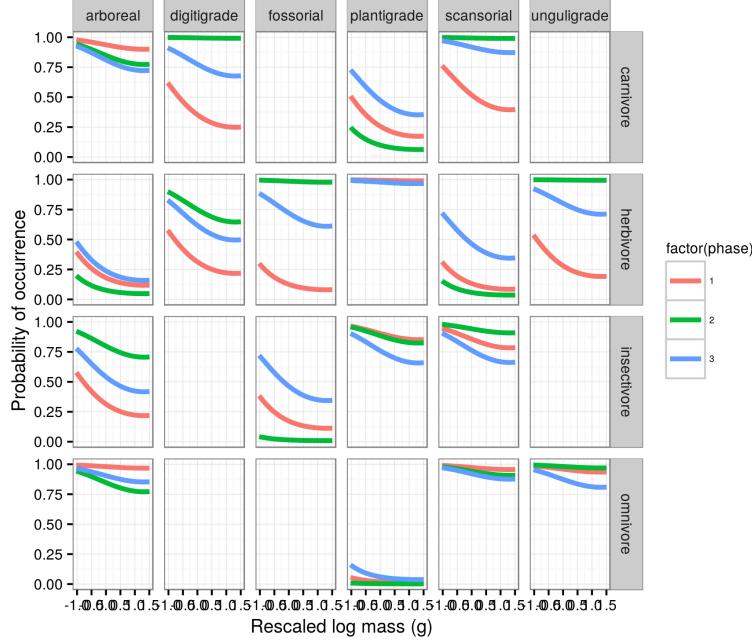


Figure 8: Mean estimate of the effect of species mass on the probability of a species occurrence for each of the three plant phases. The effect of mass is considered constant over time and that the only aspect of the model that changes with plant phase is the intercept of the relationship between mass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

et al., 2010); while adverse conditions may translate to a decreasing set of interactions or loss of appropriate environmental context. Remember that favorable versus adverse condition of a plant phase is definitionally relative to the other two plant phases.

One of the limitations to this interpretation is the almost deterministic increase in probability of occurrence and origination for most ecotypes (Fig. 4, 5). This “pull of the Recent” means that interpreting the biological meaning of differences between the final plant phase and the two previous phases is difficult as the guaranteed occurrence of the later taxa increases the average probability for that phase, which in turn affects the other time bins in that phase.

Plant phases are associated with large differences in log-odds for occurrence and origination probabilities (Tables 9, 10), though there is little evidence of plant phase being an important distinguishing factor in species survival, as only a few ecotypes demonstrate strong affinities with

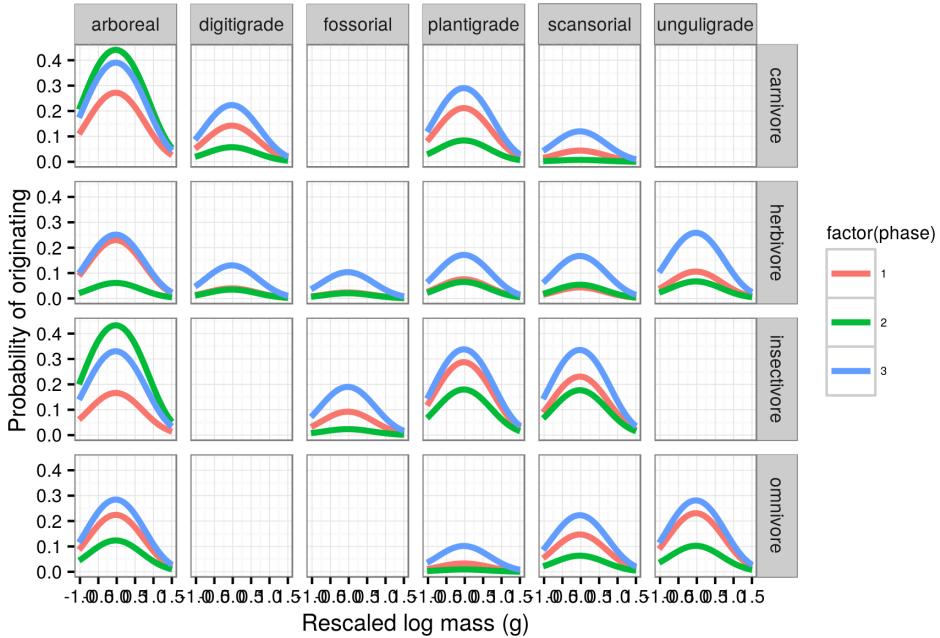


Figure 9: Mean estimate of the effect of species mass on the probability of a species originating for each of the three plant phases. The effect of mass is considered constant over time and that the only aspect of the model that changes with plant phase is the intercept of the relationship between mass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

some plant phases (Table 11). As with previous comparisons between parameter estimates

536 associated with species occurrence and species origination, parameters associated with probability  
537 of newly originating appear as a more “tempered” version of those associated with probability  
538 occurrence.

The almost universal pattern of the effect of plant phase on ecotype occurrence or origination is

540 that the during first and last plant phases ecotypes have a greater log-odds of occurrence or  
541 origination than the second plant phase (Fig. 4, 5). The three ecotypes that do not follow this  
542 pattern are fossorial herbivores, scansorial herbivores, and arboreal insectivores.

The difference between the third plant phase and the other two plant phases, for all ecotypes except  
544 arboreal carnivores, is obvious upon inspection of the occurrence and origination time series as  
545 there is large up-tick in probability of occurring or originating towards the modern (Fig. 4, 5). The

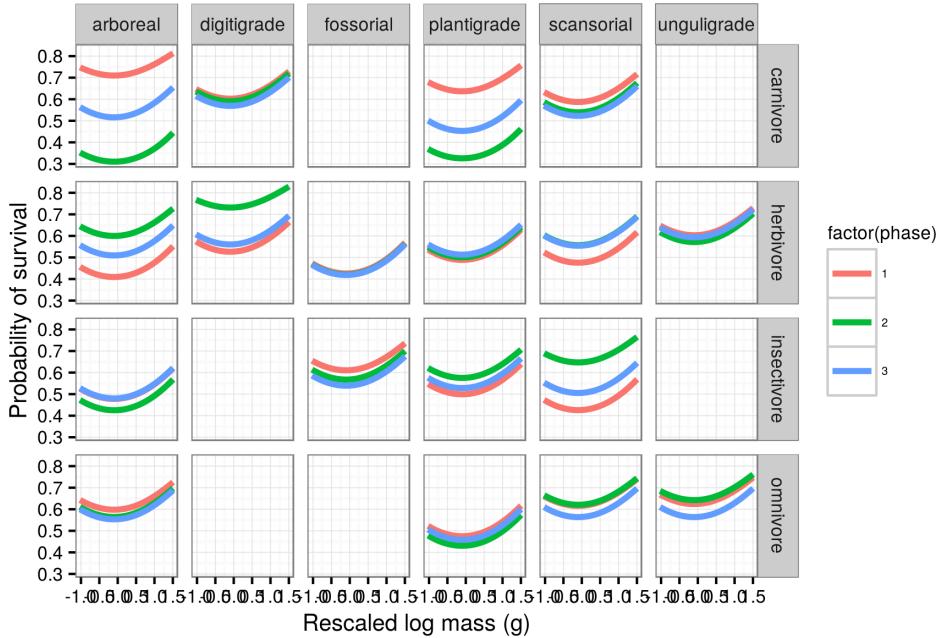


Figure 10: Mean estimate of the effect of species mass on the probability of a species survival for each of the three plant phases. The effect of mass is considered constant over time and that the only aspect of the model that changes with plant phase is the intercept of the relationship between mass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

546 differences in mean probability of occurring or originating attributable to the plant phases are  
 observable as shifts along the time series corresponding to the phase barriers (Table 4). For  
 548 example, scansorial herbivore occurrence and origination probabilities demonstrate clear shifts at 50  
 Mya and 16 Mya (Fig. 4, 5).  
 550 Both aspects of global temperature analyzed here are estimated to have strong effects on species  
 occurrence and origination for most mammal ecotypes (Tables 12, 13). Similarity, temperature is  
 552 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  
 554 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  
 556 decrease) is for the effect of temperature range on arboreal herbivores. In the case of survival, the

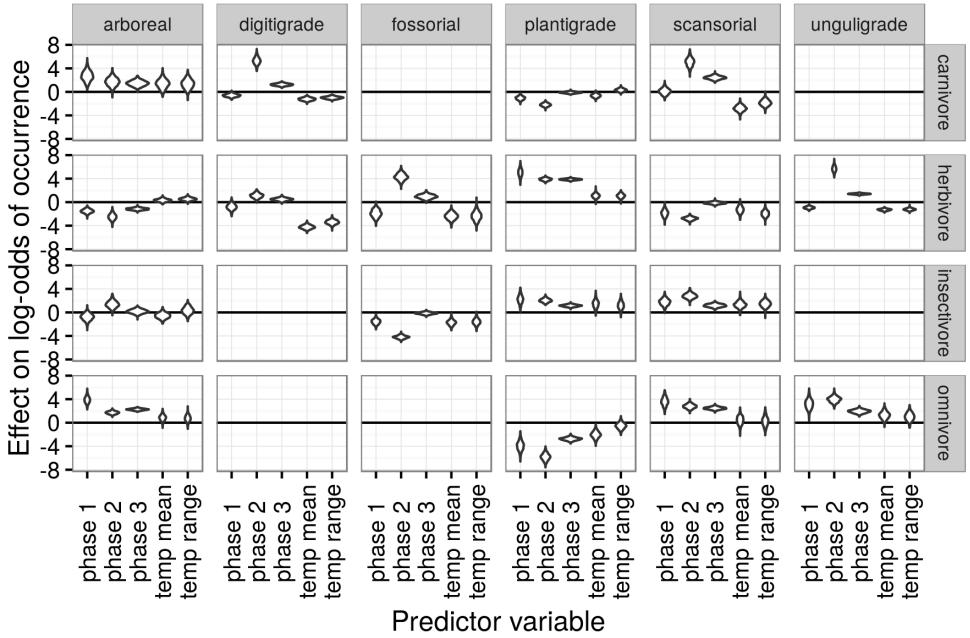


Figure 11: Estimated effects of the group-level covariates describing environmental context on log-odds of species occurrence. These estimates are from the pure-presence model. The effect of plant phase graphed here is calculated as Phase 1 =  $\gamma_{phase\ 1}$ , Phase 2 =  $\gamma_{phase\ 1} + \gamma_{phase\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

only strong ecotype associations for either of the temperature covariates are with plantigrade

558 carnivores, plantigrade herbivores, and to a less certain extent arboreal herbivores and scansorial  
insectivores (Tab. 14). The effects of the temperature covariates on these ecotypes are all estimated  
560 to be positive (e.g. temperature range increase, increase in survival).

## Analysis of diversity

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

564 The general pattern of the estimated North American total mammal diversity for the Cenozoic is  
“stable” in that diversity fluctuates around a constant mean standing diversity, does not fluctuate  
566 wildly and rapidly over the Cenozoic, and demonstrates no sustained directional trends (Fig. 14a).

In broad strokes, the first 15 or so million years of the Cenozoic are characterized by a gradual  
568 decline in standing diversity until approximately 45-50 million years ago (early-middle Eocene).

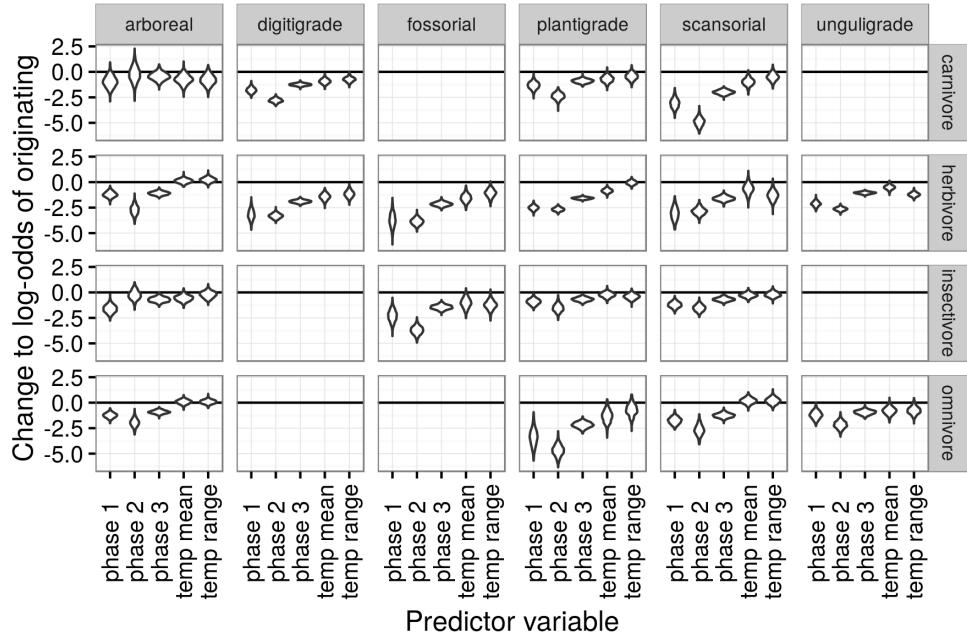


Figure 12: Estimated effects of the group-level covariates describing environmental context on log-odds of species origination. These estimates are from the birth-death model. The effect of plant phase graphed here is calculated as Phase 1 =  $\gamma_{phase\ 1}$ , Phase 2 =  $\gamma_{phase\ 1} + \gamma_{phase\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

Following this decline, standing diversity is broadly constant from 45 to 18 Mya (early Miocene).

- 570 After this, there is a rapid spike in diversity followed by a slight decline in diversity up to the Modern.
- 572 The pattern exhibited by the diversity history estimated in this study (Fig. 14a) has some major similarities with previous mammal diversity curves (Alroy, 2009): both curves begin with an
- 574 increase in diversity most of the major increases in diversity are retained including the large diversity spike during the Miocene. Note that my time series begins at a different point than that
- 576 Alroy (2009) and I'm able to interpolate over unsampled/poorly sampled time periods. Additionally, the time intervals used by Alroy (2009) are slightly shorter than those used here, so this may cause
- 578 some of the minor differences between the curves. Also, please note that the diversity values are plotted at the “ceiling” of each temporal interval and not at the midpoint (Fig. 14a).
- 580 When viewed through the lens of diversification rate, some of the structure behind the estimated diversity history begins to take shape (Fig. 14b). For most of the Cenozoic, the diversification rate

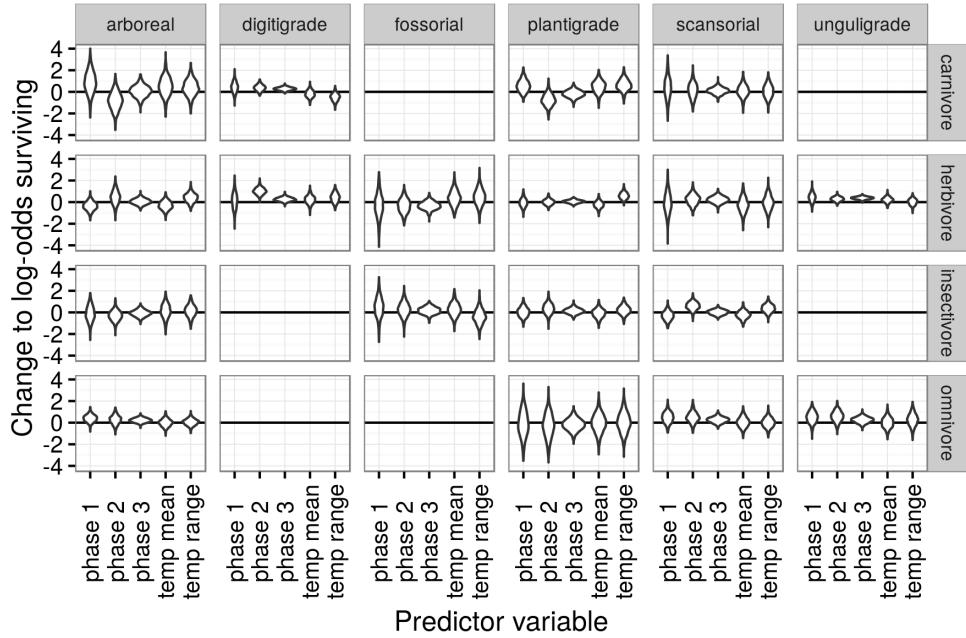


Figure 13: Estimated effects of the group-level covariates describing environmental context on log-odds of species survlval. These estimates are from the birth-death model. The effect of plant phase graphed here is calculated as Phase 1=  $\gamma_{phase\ 1}$ , Phase 2=  $\gamma_{phase\ 1} + \gamma_{phase\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

- 582 hovers around zero, punctuated by both positive and negative spikes. The largest spike in  
 diversification rate is at 18 Mya, which is early Oligocene (Fig. 14b). Other notable increases in  
 584 diversification rate occur 56, 46, 38, and 6 Mya (Table 15), though the last of these may be due to  
 edge effects surrounding the partial-identifiability of  $p_{t=T}$ . Notable decreases in diversification rate  
 586 occur at 60, 54, 50, 44, 34, 20, 16, 12, and 8 Mya (Table 15), meaning that diversification rate has  
 more major decreases than increases. Given that diversification rate more closely resembles  
 588 origination rate than extinction rate (Fig. 14b, 14c, 14d), these decreases in diversification rate may  
 be indicative of “depletions” (failure to replace extinct taxa) rather than pulses of extinction.
- 590 The comparison between per capita origination and extinction rate estimates reveals how  
 diversification rate is formed (Fig. 14c, 14d). As expected given previous inspection of the ecotype  
 592 specific estimates of origination and survival probabilities from the birth-death model,  
 diversification rate seems most driven by changes in origination rate as opposed to extinction rate.  
 594 Extinction rate, on the other hand, demonstrates an almost saw-toothed pattern around a constant

Table 9: Posterior probability of the differences in the log-odds of an ecotype occurring based on plant phase. These probabilities are calculated as  $P(\text{Phase 1} > \text{2}) = (\sum \gamma_{\text{phase1}} > \gamma_{\text{phase1}} + \gamma_{\text{phase2}})/100$  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	0.315	0.043	0.000
digitigrade carnivore	0.000	1.000	0.000
plantigrade carnivore	1.000	0.000	0.793
scansorial carnivore	0.000	0.618	0.000
arboreal herbivore	0.997	0.367	1.000
digitigrade herbivore	0.023	0.654	0.024
fossorial herbivore	0.000	1.000	0.001
plantigrade herbivore	0.475	0.000	0.000
scansorial herbivore	1.000	0.000	0.778
unguligrade herbivore	0.000	1.000	0.000
arboreal insectivore	0.007	0.961	0.309
fossorial insectivore	1.000	0.000	0.796
plantigrade insectivore	0.001	0.189	0.000
scansorial insectivore	0.000	0.845	0.000
arboreal omnivore	0.999	0.000	0.000
plantigrade omnivore	1.000	0.289	1.000
scansorial omnivore	0.141	0.000	0.000
unguligrade omnivore	0.000	0.592	0.000

mean (Fig. 14d). These results are broadly consistent with those from previous analyses of North

596 American mammals diversity and diversification (Alroy, 1996, 2009; Alroy et al., 2000).

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

598 diversity for the Cenozoic (Fig. 15).

Arboreal ecotypes obtain peak diversity early in the Cenozoic and then decline for the rest of the

600 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

602 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

604 entirely absent from the species pool. This is consistent with increasing extinction risk in the

Neogene compared to the Paleogene as proposed by Smits (2015).

606 The diversity of digitigrade and unguligrade herbivores increases over the Cenozoic (Fig. 15). In

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	0.373	0.810	0.873
digitigrade carnivore	1.000	0.066	1.000
plantigrade carnivore	1.000	0.036	1.000
scansorial carnivore	1.000	0.019	1.000
arboreal herbivore	1.000	0.134	1.000
digitigrade herbivore	1.000	0.985	1.000
fossorial herbivore	1.000	0.919	1.000
plantigrade herbivore	1.000	0.996	1.000
scansorial herbivore	1.000	0.843	1.000
unguligrade herbivore	1.000	0.001	1.000
arboreal insectivore	0.096	0.996	1.000
fossorial insectivore	1.000	0.019	1.000
plantigrade insectivore	0.993	0.331	1.000
scansorial insectivore	1.000	0.293	1.000
arboreal omnivore	0.998	0.378	1.000
plantigrade omnivore	1.000	0.277	1.000
scansorial omnivore	0.999	0.353	1.000
unguligrade omnivore	1.000	0.224	1.000

contrast, plantigrade herbivore diversity does not have a single, broad-strokes pattern; instead,  
 608 diversity increases, decreases, and may have then increased till the Recent. In contrast, fossorial and  
 scansorial herbivores demonstrate a much flatter history of diversity, with a slight increase in  
 610 diversity that over time is more pronounced among fossorial taxa than scansorial taxa. The  
 expansion of digitigrade and unguligrade herbivores over the Cenozoic is consistent with the  
 612 gradual expansion of grasslands which these ecotypes are better adapted to than closed  
 environments (Blois and Hadly, 2009; Strömberg, 2005).

614 Digitigrade carnivores have a multi-modal diversity history, with peaks at 54-52 and 12-10 Mya  
 (Fig.15). Between these two peaks digitigrade carnivore diversity dips below average diversity  
 616 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  
 618 peak at the end of the Cenozoic. The generally flat diversity history digitigrade carnivores lacks any

Table 11: Posterior probability of the differences in the log-odds of an ecotype surviving based on plant phase. These probabilities are calculated as  $P(\text{Phase 1} > \text{2}) = (\sum \gamma_{\text{phase1}} > \gamma_{\text{phase1}} + \gamma_{\text{phase2}})/100$  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	0.919	0.128	0.439
digitigrade carnivore	0.358	0.238	0.029
plantigrade carnivore	0.898	0.244	0.702
scansorial carnivore	0.462	0.480	0.373
arboreal herbivore	0.244	0.707	0.436
digitigrade herbivore	0.002	0.939	0.080
fossorial herbivore	0.491	0.696	0.822
plantigrade herbivore	0.593	0.331	0.343
scansorial herbivore	0.489	0.335	0.228
unguligrade herbivore	0.699	0.016	0.003
arboreal insectivore	0.735	0.370	0.610
fossorial insectivore	0.417	0.472	0.324
plantigrade insectivore	0.295	0.560	0.303
scansorial insectivore	0.034	0.935	0.464
arboreal omnivore	0.435	0.317	0.133
plantigrade omnivore	0.541	0.514	0.623
scansorial omnivore	0.286	0.493	0.176
unguligrade omnivore	0.212	0.534	0.189

sustained temporal trends and seems to reflect previous findings of limited diversity in spite of

620 constant turnover and morphological evolution (Silvestro et al., 2015; Slater, 2015; Valkenburgh,  
1999)

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

630 Many of the estimated ecotype-specific diversity histories share a similar increase in diversity in the  
late Cenozoic, 16-14 Mya (Fig. 15). These increases are either sustained or temporary and are seen

Table 12: Posterior probabilities that the effects of the two temperature covariates on 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	0.954	0.955
digitigrade carnivore	0.000	0.000
plantigrade carnivore	0.012	0.883
scansorial carnivore	0.000	0.001
arboreal herbivore	0.907	0.984
digitigrade herbivore	0.000	0.000
fossorial herbivore	0.000	0.003
plantigrade herbivore	0.998	0.996
scansorial herbivore	0.010	0.000
unguligrade herbivore	0.000	0.000
arboreal insectivore	0.117	0.684
fossorial insectivore	0.000	0.002
plantigrade insectivore	0.986	0.975
scansorial insectivore	0.989	0.997
arboreal omnivore	0.955	0.911
plantigrade omnivore	0.000	0.143
scansorial omnivore	0.708	0.607
unguligrade omnivore	0.978	0.944

- 632 in digitigrade carnivores, plantigrade carnivores, scansorial carnivores, unculigrade herbivores,  
fossorial insectivores, and plantigrade omnivores.
- 634 When ecotype diversity is decomposed into the number of origination events per time bin (Fig. 16)  
and the number of extinction events per time bin (Fig. 17) the estimates are clearly similar; there  
636 are no obvious major cross-ecotype origination or extinction events, and there is no evidence of a  
sudden turnover as expected peaks in originations preceed peaks in the number of extinctions. Also,  
638 it is clear that the sustained increases in digitigrade and unculigrade herbivore diversity observed  
above (Fig. 15) is driven by an increase in the average number of originations as with a relatively  
640 constant number of extinctions over time (Fig. 16, 17).

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	0.060	0.062
digitigrade carnivore	0.000	0.001
plantigrade carnivore	0.014	0.098
scansorial carnivore	0.003	0.101
arboreal herbivore	0.709	0.833
digitigrade herbivore	0.000	0.000
fossorial herbivore	0.000	0.002
plantigrade herbivore	0.000	0.393
scansorial herbivore	0.100	0.003
unguligrade herbivore	0.006	0.000
arboreal insectivore	0.030	0.260
fossorial insectivore	0.014	0.003
plantigrade insectivore	0.173	0.055
scansorial insectivore	0.107	0.207
arboreal omnivore	0.659	0.649
plantigrade omnivore	0.015	0.104
scansorial omnivore	0.743	0.720
unguligrade omnivore	0.014	0.034

## Discussion

- 642 Both the composition of a species pool and its environmental context change over time, though not  
 necessarily at the same rate or concurrently. Local communities, whose species are drawn from the  
 644 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  
 646 ecotypes and guilds, these roles are broad and not defined by specific interactions but by the genre  
 of interactions species within that grouping participate in. The diversity of species within an  
 648 ecotype or guild can be stable over millions of years despite constant species turnover (Jernvall and  
 Fortelius, 2004; Slater, 2015; Valkenburgh, 1999). This implies that the size and scope of the role of  
 650 an ecotype or guild in local communities, and the regional species pool as a whole, is preserved even  
 as the individual interactors change. This also implies that the structure of regional species pools  
 652 can be constant over time despite a constantly changing set of “players.” This result supports the

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	0.716	0.704
digitigrade carnivore	0.289	0.071
plantigrade carnivore	0.800	0.902
scansorial carnivore	0.557	0.529
arboreal herbivore	0.241	0.831
digitigrade herbivore	0.725	0.828
fossorial herbivore	0.704	0.733
plantigrade herbivore	0.223	0.979
scansorial herbivore	0.328	0.458
unguligrade herbivore	0.809	0.545
arboreal insectivore	0.592	0.616
fossorial insectivore	0.675	0.276
plantigrade insectivore	0.437	0.697
scansorial insectivore	0.262	0.836
arboreal omnivore	0.490	0.576
plantigrade omnivore	0.487	0.507
scansorial omnivore	0.527	0.503
unguligrade omnivore	0.500	0.682

hypothesis that ecotypes or guilds are at least partially self-organizing and truly emergent (Scheffer

654 and van Nes, 2006).

Comparison of the results from the posterior predictive checks for the pure-presence and

656 birth-death models supports the conclusion that regional species pool dynamics cannot simply be  
658 described by a single occurrence probability and are instead the result of the interplay between  
658 origination and extinction. Additionally, changes to the ecotypic composition and diversification  
660 rate for the North American regional species pool are driven primarily by variation in origination  
660 rates. These aspects of how regional species pool diversity is shaped are not directly observable in  
662 studies of the Recent where time scales are short and macroevolutionary dynamics are inferable  
662 solely from phylogeny (Fritz et al., 2013).

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

664 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  
666 extinction rate is actually variable may not have been equally appreciated. What is most consistent  
with previous observations, however, is that diversity seems to be most structured by changes to  
668 origination rather than changes to extinction (Alroy, 1996; Alroy et al., 2000).

Comparison of the ecotype specific diversity histories supports the conclusion that there were no  
670 major, simultaneous changes in diversity between the functional groups of the regional speies pool;  
instead these results support a more gradual and idiosyncratic shifts in relative ecotypic diversity  
672 over time (Fig. 15). The closest examples to a sudden increase or decrease of a specific ecotype is  
the jump in standing diversity of scansorial carnivores and, to a lesser extent, fossorial insectivores  
674 at 16 Mya (i.e. the start of the third plant phase). However, this result may not reflect the  
dynamics of individual local communities, as this is an analysis of the entire North American  
676 mammal regional species pool.

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

Digitigrade carnivores have a relatively stable diversity history through the Cenozoic and can be  
686 characterized as varying around a constant mean diversity. This ecotype has a large amount of  
overlap with the carnivore guild which has been the focus of much research (Janis and Wilhelm,  
688 1993; Pires et al., 2015; Slater, 2015; Valkenburgh, 1999). This result is consistent with some form  
of “control” on the diversity of this ecotype, such as diversity-dependent diversification (Silvestro  
690 et al., 2015; Slater, 2015; Valkenburgh, 1999).

Both digitigrade and unguligrade herbivores increase in diversity over the Cenozoic. The increase of

692 these cursorial forms is consistent with the gradual opening up of the North American landscape  
693 (Blois and Hadly, 2009; Graham, 2011; Strömberg, 2005) and may indicate a long-term shift in the  
694 interactors associated with those ecotypes leading to increased contribution to the regional species  
695 pool. This result may be comparable to the increasing percentage of hypsodont (high-crowned  
696 teeth) mammals in the Neogene of Europe being due to an enrichment of hyposodont taxa and not  
697 a depletion of non-hypsodont taxa. Smaller scale increases in fossorial herbivore species, and a lesser  
698 extent plantigrade herbivores, suggests that the increase of interactors may be associated mostly  
699 with the herbivore dietary category with locomotor category tempering that relationship. These  
700 results support the conclusion that the increase in digitigrade and unguligrade herbivores is the  
701 result of an enrichment of these ecotypes as opposed to being caused by the depletion of other  
702 herbivorous ecotypes; this is further supported by the lack of major changes to the number of  
703 extinctions of all herbivore ecotypes (Fig. 17).

704 An association between plant phase and differences in ecotype occurrence or origination-extinction  
705 probabilities is interpreted to mean that an ecotype enrichment or depletion is due to associations  
706 between that ecotype and whatever plants are dominant at that time. Plant phase clearly  
707 structures the occurrence and origination probability time series (Fig. 4, 5). These differences in  
708 occurrence or origination translate to the estimates of diversity and diversification rate; the largest  
709 spike in diversity, diversification rate, and origination rate all correspond to the onset of the last  
710 plant phase (Fig. 14). The clearest example of the diversity of an ecotype increasing at this  
711 particular transition is in scansorial carnivores (Fig. 15); similar shifts in other ecotypes are much  
712 more subtle, as was previously noted for fossorial insectivores.

713 Interestingly, for all of the ecotypes with sudden changes in diversity at this transition the change is  
714 an increase, even if only temporarily. There are two interpretations of these results. A biological  
715 interpretation of this result is that, because plant phase associations are only with occurrence or  
716 origination probabilities and not survival, these ecotypes were well suited for the newly available  
717 mammal-plant interactions due to the increased modernization of their floral context (Graham,  
718 2011). Alternatively, the increase in diversity associated with the third plant phase may be caused  
719 by the edge effect in origination probability that is artificially inflating the number of origination

720 events (Fig. 5). However, the estimated number of origination events does not have a tremendous  
spike at this transition, nor is a major increase in the number of origination events sustained (Fig.  
722 16).

There are fewer, less obvious shifts in diversity surrounding the transition from the first to second  
724 plant phase, with the following ecotypes having apparent shifts in diversity at 50 My: digitigrade  
carnivores (down), plantigrade carnivores (down), plantigrade herbivores (up), arboreal omnivores  
726 (down), and 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  
728 that 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  
730 available mammal-plant interactions brought upon by the changing environmental context  
(Graham, 2011).

732 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  
734 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  
736 originating; this means that if temperature decreases, we would then expect an increase in the  
probability of occurring or first originating. In contrast, temperature range is estimated to be a  
738 good predictor of survival in only four mammal ecotypes and only marginally for two of those  
(Table 14). Additionally, all four of these cases have positive relationships, meaning that if  
740 temperature decreases it is expected that species survival will also decrease.

The result that temperature does not affect the extinction probability of most ecotypes is consistent  
742 with previous analysis of mammal diversity (Alroy et al., 2000). The result that temperature affects  
origination probability, on the other hand, is in strong contrast to the results Alroy et al. (2000).  
744 An important difference between the analyses presented here and that of Alroy et al. (2000) is I am  
considering the effect of temperature on the probability of a species originating, assuming it hasn't  
746 originated yet while Alroy et al. (2000) analyzes the correlation between the first differences of the

origination and extinction rates with an oxygen isotope curve (Zachos et al., 2001). Origination or  
748 extinction rates have very different properties than the origination probabilities estimated here  
brought upon by the difference both in definition and units. Origination probability is the expected  
750 probability that a species that has never been present and is not present at time  $t$  will be present at  
time  $t + 1$ ; origination probability is defined for a single species. In contrast, per capita rates are  
752 defined (for origination) as the expected number of new species to have originated between time  $t$   
and  $t + 1$  given the total number of species present at time  $t$ ; per capita rates are defined for the  
754 standing diversity. It is also important to note that even though the occurrence and origination  
probabilities increase over time because of the increasingly deterministic occurrence of those species  
756 which had not yet originated (Fig. 4, 5, the corresponding rates and population level birth/death  
dynamics do not share that pattern (Fig. 14, 16, 17). In effect, the finding that temperature has an  
758 effect on origination simply because as time approaches the present the number of species which  
have originated increases and not because of climatic forcing of origination.

760 Analysis of relationship between temperature and origination rate is probably better suited for a  
continuous-time birth-death model instead of a discrete-time model because the former estimates  
762 rates while the latter estimates probabilities (Allen, 2011). The PyRate model(s) are based on a  
continuous-time birth-death process (Silvestro et al., 2015, 2014). Unfortunately, a continuous-time  
764 model may be unsuited for most paleontological data as the fossil record is naturally discrete; fossils  
are assigned to temporal units, such as stages, which have age ranges. Fossils are not assigned  
766 individual numeric ages. This reality was in fact my one of motivations for using discrete-time  
birth-death model instead of one in continuous-time.

768 The effects of plant phase and temperature on ecotypes are approximately equal in importance. The  
focus in previous research on temperature and major climatic or geological events without other  
770 measures of environmental context may have led to confusion in discussions of how the  
“environment” affects mammal diversity and diversification. The environment or climate is not just  
772 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 more descriptors of species’  
774 environmental context a more complete “picture” of the diversification process is inferred.

The effect of species mass on either occurrence or origination and extinction was not allowed to  
776 vary by ecotype or environmental context. The primary reason for this modeling choice was that  
this study focuses on ecotypic based differences in either occurrence, or origination and extinction.  
778 Allowing the effect of body size to vary by ecotype, time, and environmental factors would increase  
the overall complexity of the model beyond the scope of the study. Instead, body size was included  
780 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 than ignoring it.  
782 The only covariate allowed to affect sampling probability was mass and only as a linear predictor.  
Other covariates, such as the environmental factors considered here, could have affected the  
784 underlying preservation process that limits sampling probability; their exclusion as covariates of  
sampling/observation was the product of a few key decisions: model complexity, model  
786 interpretability, the scope of this study, and a lack of good hypotheses related to these covariates to  
warrant their inclusion.  
788 The potential effects of common ancestry (i.e. phylogeny) on origination and extinction are not  
directly considered in this analysis. While a birth-death process approximates the  
790 speciation-extinction process underlying the phylogeny (Silvestro et al., 2014) this is not same as  
considering how the similarity between closely related species may affect the estimates of the effects  
792 of species traits to environmental factors on both origination and extinction (Harnik et al., 2014;  
Smits, 2015). The inclusion of phylogeny can help disentangle if the functional composition of  
794 species diversity is shaped either by many closely related species occurring at the same time or if  
closely related species are more evenly distributed in time; this is analogous to if species within  
796 local communities are clumped or dispersed relative to their relatedness (Cavender-Bares et al.,  
2009; Kraft et al., 2007; Webb et al., 2002). One of the principal barriers to the inclusion of the  
798 effect of phylogeny in either the pure-presence or birth-death models is computational; with well  
over 1000 tips, the calculation of the scale parameter defining the phylogenetic effect would be very  
800 slow and further increase the already slow computation time necessary for the marginalization of all  
possible discrete occurrence histories for  $z$ .

802 **Conclusions**

These results support the conclusion that the ecotypic diversity of the North American mammal  
804 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 reorganization of the  
806 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 changes to  
808 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  
810 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  
812 plantigrade and scansorial insectivores and scansorial omnivores. The only ecotypes that  
demonstrate a sustained pattern of increasing diversity are digitigrade and unguligrade herbivores.  
814 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 analysis provides a  
816 much more complete picture of North American mammal diversity and diversification, specifically  
the dynamics of the ecotypic composition of that diversity. By increasing the complexity of analysis  
818 while precisely translating research questions into a statistical model, the context of the results is  
much better understood. Future studies of diversity and diversification should incorporate as much  
820 information as possible into their analyses in order to better understand or at least contextualize  
the complex processes underlying that diversity.

822 **Acknowledgements**

I would like to thank K. Angielczyk, M. Foote, P. D. Polly, R. Ree, and G. Slater for helpful  
824 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.  
826 Alroy and M. Uhen for curating most of the mammal occurrences recorded in the PBDB. This is

828 **References**

- Allen, L. J. S. 2011. An introduction to stochastic processes with applications to biology. 2nd ed.  
830 Chapman and Hall/CRC, Boca Raton, FL.
- Alroy, J. 1996. Constant extinction, constrained diversification, and uncoordinated stasis in North  
832 American mammals. *Palaeogeography, Palaeoclimatology, Palaeoecology* 127:285–311.
- . 2009. Speciation and extinction in the fossil record of North American mammals. Pages  
834 302–323 *in* R. K. Butlin, J. R. Bridle, and D. Schluter, eds. *Speciation and patterns of diversity*.  
Cambridge University Press, Cambridge.
- 836 Alroy, J., P. L. Koch, and J. C. Zachos. 2000. Global climate change and North American  
mammalian evolution. *Paleobiology* 26:259–288.
- 838 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*  
840 39:373–399.
- Badgley, C., T. M. Smiley, R. Terry, E. B. Davis, L. R. G. Desantis, D. L. Fox, S. S. B. Hopkins,  
842 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:  
844 Modern and Geohistorical Perspectives. *Trends in Ecology & Evolution* pages 1–16.
- Bambach, R. K. 1977. Species richness in marine benthic habitats through the Phanerozoic.  
846 *Paleobiology* 3:152–167.
- Bambach, R. K., A. M. Bush, and D. H. Erwin. 2007. Autecology and the filling of ecospace: Key  
848 metazoan radiations. *Palaeontology* 50:1–22.
- Bloch, J. I., K. D. Rose, and P. D. Gingerich. 1998. New species of Batodonoides (Lipotyphla,

- 850      Geolabididae) from the Early Eocene of Wyoming: smallest known mammal? *Journal of Mammalogy* 79:804–827.
- 852      Blois, J. L., and E. A. Hadly. 2009. Mammalian Response to Cenozoic Climatic Change. *Annual Review of Earth and Planetary Sciences* 37:181–208.
- 854      Brook, B. W., and D. M. J. S. Bowman. 2004. The uncertain blitzkrieg of Pleistocene megafauna. *Journal of Biogeography* 31:517–523.
- 856      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  
858      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  
860      on continents. *Science* 243:1145–1150.
- Bush, A. M., and R. K. Bambach. 2011. Paleoecologic Megatrends in Marine Metazoa. *Annual  
862      Review of Earth and Planetary Sciences* 39:241–269.
- Bush, A. M., R. K. Bambach, and G. M. Daley. 2007. Changes in theoretical ecospace utilization in  
864      marine fossil assemblages between the mid-Paleozoic and late Cenozoic. *Paleobiology* 33:76–97.
- Bush, A. M., and P. M. Novack-Gottshall. 2012. Modelling the ecological-functional diversification  
866      of marine Metazoa on geological time scales. *Biology Letters* 8:151–155.
- Cantalapiedra, J. L., J. L. Prado, and M. T. Alberdi. 2017. Decoupled ecomorphological evolution  
868      and diversification in Neogene-Quaternary horses. *Science* 355:627–630.
- Carrano, M. T. 1999. What, if anything, is a cursor? Categories versus continua for determining  
870      locomotor habit in mammals and dinosaurs. *Journal of Zoology* 247:29–42.
- Cavender-Bares, J., K. H. Kozak, P. V. a. Fine, and S. W. Kembel. 2009. The merging of  
872      community ecology and phylogenetic biology. *Ecology letters* 12:693–715.
- Clyde, W. C., and P. D. Gingerich. 1998. Mammalian community response to the latest Paleocene

- 874 thermal maximum: an isotaphonomic study in the northern Bighorn Basin, Wyoming. *Geology*  
26:1011–1014.
- 876 Cohen, K. M., S. C. Finney, P. L. Gibbard, and J.-X. Fan. 2015. The ICS International  
Chronostratigraphic Chart.
- 878 Cottenie, K. 2005. Integrating environmental and spatial processes in ecological community  
dynamics. *Ecology Letters* 8:1175–1182.
- 880 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  
882 geochemistry ( $\delta^{18}\text{O}$  and Mg/Ca) with sea level history. *Journal of Geophysical Research: Oceans*  
116:1–23.
- 884 Damuth, J. 1982. Analysis of the preservation of community structure in assemblages of fossil  
mammals. *Paleobiology* 8:434–446.
- 886 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.
- 888 Elith, J., and J. R. Leathwick. 2009. Species distribution models: ecological explanation and  
prediction across space and time. *Annual Review of Ecology, Evolution, and Systematics*  
40:677–697.
- 892 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.
- 894 Eronen, J. T., P. D. Polly, M. FRED, J. Damuth, D. C. FRANK, V. Mosbrugger,  
C. SCHEIDECKER, N. C. Stenseth, and M. Fortelius. 2010. Ecometrics: The traits that bind  
896 the past and present together. *Integrative Zoology* 5:88–101.
- Ezard, T. H. G., A. Purvis, and H. Morlon. 2016. Environmental changes define ecological limits to

- 898 species richness and reveal the mode of macroevolutionary competition. *Ecology Letters*  
19:899–906.
- 900 Felsenstein, J. 1985. Phylogenies and the comparative method. *The American Naturalist* 125:1–15.
- 902 Figueirido, B., C. M. Janis, J. A. Pérez-Claros, M. De Renzi, and P. Palmqvist. 2012. Cenozoic  
climate change influences mammalian evolutionary dynamics. *Proceedings of the National  
Academy of Sciences* 109:722–727.
- 904 Foote, M. 2001. Inferring temporal patterns of preservation, origination, and extinction from  
taxonomic survivorship analysis. *Paleobiology* 27:602–630.
- 906 Foote, M., and J. J. Sepkoski. 1999. Absolute measures of the completeness of the fossil record.  
*Nature* 398:415–7.
- 908 Foster, J. R. 2009. Preliminary body mass estimates for mammalian genera of the Morrison  
Formation (Upper Jurassic, North America). *PaleoBios* 28:114–122.
- 910 Fraser, D., R. Gorelick, and N. Rybczynski. 2015. Macroevolution and climate change influence  
phylogenetic community assembly of North American hoofed mammals. *Biological Journal of the  
Linnean Society* 114:485–494.
- 914 Freudenthal, M., and E. Martín-Suárez. 2013. Estimating body mass of fossil rodents. *Scripta  
Geologica* 145:1–130.
- 916 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.
- 918 Gelman, A. 2008. Scaling regression inputs by dividing by two standard deviations. *Statistics in  
Medicine* pages 2865–2873.
- 920 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.
- 922 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. *American Scientist* 102:460–465.
- Gordon, C. L. 2003. A First Look at Estimating Body Size in Dentally Conservative Marsupials. *Journal of Mammalian Evolution* page 21.
- Graham, A. 2011. A natural history of the New World: the ecology and evolution of plants in the Americas. University of Chicago Press, Chicago.
- Harmon, L. J., and S. Harrison. 2015. Species Diversity Is Dynamic and Unbounded at Local and 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 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 community richness. *Ecology Letters* 11:969–979.
- Jamil, T., W. A. Ozinga, M. Kleyer, and C. J. F. Ter Braak. 2013. Selecting traits that explain species-environment relationships: A generalized linear mixed model approach. *Journal of Vegetation Science* 24:988–1000.
- Janis, C., J. Damuth, and J. M. Theodor. 2004. The species richness of Miocene browsers, and implications for habitat type and primary productivity in the North American grassland biome. *Palaeogeography, Palaeoclimatology, Palaeoecology* 207:371–398.
- Janis, C. M. 1993. Tertiary mammal evolution in the context of changing climates, vegetation, and tectonic events. *Annual Review of Ecology and Systematics* 24:467–500.
- . 2008. An evolutionary history of browsing and grazing ungulates. Pages 21–45 in I. J. Gordon and H. H. T. Prins, eds. *The Ecology of Browsing and Grazing*. Springer-Verlag.

- 948 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*  
950 97:7899–904.
- 952 Janis, C. M., G. F. Gunnell, and M. D. Uhen. 2008. Evolution of Tertiary mammals of North  
America. Vol. 2. Small mammals, xenarthrans, and marine mammals. Cambridge University  
Press, Cambridge.
- 954 Janis, C. M., K. M. Scott, and L. L. Jacobs. 1998. Evolution of Tertiary mammals of North  
America. Vol. 1. Terrestrial carnivores, ungulates, and ungulatelike mammals. Cambridge  
956 University Press, Cambridge.
- 958 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.
- 960 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.
- 962 Jernvall, J., and M. Fortelius. 2002. Common mammals drive the evolutionary increase of  
hypodonty in the Neogene. *Nature* 417:538–40.
- 964 ———. 2004. Maintenance of trophic structure in fossil mammal communities: site occupancy and  
taxon resilience. *The American Naturalist* 164:614–624.
- 966 Kraft, N. J. B., W. K. Cornwell, C. O. Webb, and D. D. Ackerly. 2007. Trait evolution, community  
assembly, and the phylogenetic structure of ecological communities. *The American Naturalist*  
968 170:271–283.
- 970 Kucukelbir, A., R. Ranganath, A. Gelman, and D. M. Blei. 2015. Automatic Variational Inference  
in Stan. Pages 568–576 *in* NIPS. Vol. 28.
- 972 Legendre, S. 1986. Analysis of mammalian communities from the Late Eocene and Oligocene of  
Southern France. *Paleovertebrata* 16:191–212.

- Liow, L. H., M. Fortelius, E. Bingham, K. Lintulaakso, H. Mannila, L. Flynn, and N. C. Stenseth.  
974 2008. Higher origination and extinction rates in larger mammals. *Proceedings of the National  
Academy of Sciences* 105:6097–6102.
- 976 Lloyd, G. T., J. R. Young, and A. B. Smith. 2011. Taxonomic Structure of the Fossil Record is  
Shaped by Sampling Bias. *Systematic Biology* 61:80–89.
- 978 Loeuille, N., and M. a. Leibold. 2008. Evolution in metacommunities: on the relative importance of  
species sorting and monopolization in structuring communities. *The American naturalist*  
980 171:788–99.
- 982 Losos, J. B. 2010. Adaptive radiation, ecological opportunity, and evolutionary determinism. *The  
American naturalist* 175:623–39.
- 984 Losos, J. B., and D. L. Mahler. 2010. Adaptive radiation: the interaction of ecological opportunity,  
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,  
986 Sunderland, MA.
- 988 Luo, Z.-X., A. W. Crompton, and A.-L. Sun. 2001. A New Mammaliaform from the Early Jurassic  
and Evolution of Mammalian Characteristics. *Science* 292:1535–1540.
- 990 Marcot, J. D. 2014. The fossil record and macroevolutionary history of North American ungulate  
mammals: standardizing variation in intensity and geography of sampling. *Paleobiology*  
40:237–254.
- 992 McElreath, R. 2016. *Statistical rethinking: a Bayesian course with examples in R and Stan*. CRC  
Press, Boca Raton, FL.
- 994 McGill, B. J., B. J. Enquist, E. Weiher, and M. Westoby. 2006. Rebuilding community ecology  
from functional traits. *TRENDS in Ecology and Evolution* 21:178–185.
- 996 McKenna, R. T. 2011. Potential for Speciation in Mammals Following Vast , Late Miocene Volcanic  
Interruptions in the Pacific Northwest. Masters. Portland State University.

- 998 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.
- 1000 Mittelbach, G. G., and D. W. Schemske. 2015. Ecological and evolutionary perspectives on community assembly. *Trends in Ecology and Evolution* 30:241–247.
- 1002 Novack-Gottshall, P. M. 2007. Using a theoretical ecospace to quantify the ecological diversity of Paleozoic and modern marine biotas Using a theoretical ecospace to quantify the ecological diversity of Paleozoic and modern marine biotas. *Paleobiology* 33:273–294.
- 1004 Pires, M. M., D. Silvestro, and T. B. Quental. 2015. Continental faunal exchange and the asymmetrical radiation of carnivores. *Proceedings of the Royal Society B: Biological Sciences* 282:20151952.
- 1008 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.
- 1010 Polly, P., J. Eronen, M. Fred, G. P. Dietl, V. Mosbrugger, C. Scheidegger, D. C. Frank, J. Damuth, N. C. Stenseth, and M. Fortelius. 2011. History matters: ecometrics and integrative climate change biology. *Proceedings of the Royal Society B: Biological Sciences* 278:1131–1140.
- 1012 Polly, P. D., A. M. Lawing, J. T. Eronen, and J. Schnitzler. 2015. Processes of ecometric patterning: modelling functional traits, environments, and clade dynamics in deep time. *Biological Journal of the Linnean Society* pages n/a–n/a.
- 1016 Quental, T. B., and C. R. Marshall. 2013. How the Red Queen Drives Terrestrial Mammals to Extinction. *Science* 341:290–292.
- 1018 Rabosky, D. L. 2013. Diversity-Dependence, Ecological Speciation, and the Role of Competition in Macroevolution. *Annual Review of Ecology, Evolution, and Systematics* 44:1–22.
- 1020 Rabosky, D. L., and A. H. Hurlbert. 2015. Species Richness at Continental Scales Is Dominated by Ecological Limits. *The American Naturalist* 185:000–000.

- 1022 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.
- 1024 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.
- 1026 Royle, J. A., J. D. Nichols, and M. Kéry. 2005. Modelling occurrence and abundance of species  
when detection is imperfect. *Oikos* 110:353–359.
- 1028 Scheffer, M., and E. H. van Nes. 2006. Self-organized similarity, the evolutionary emergence of  
groups of similar species. *Proceedings of the National Academy of Sciences* 103:6230–6235.
- 1030 Shipley, B., D. Vile, and E. Garnier. 2006. From plant traits to plant communities: a statistical  
mechanistic approach to biodiversity. *Science* 314:812–814.
- 1032 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*  
1034 *the United States of America* 112:8684–9.
- 1036 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.
- 1038 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  
1040 diversity-dependent trait evolution. *Proceedings of the National Academy of Sciences*  
112:4897–4902.
- 1042 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.
- 1044 Smith, F. A., S. K. Lyons, S. Morgan Ernest, and J. H. Brown. 2008. Macroecology: more than the  
division of food and space among species on continents. *Progress in Physical Geography*  
1046 32:115–138.

- Smits, P. D. 2015. Expected time-invariant effects of biological traits on mammal species duration.  
1048      Proceedings of the National Academy of Sciences 112:13015–13020.
- Stan Development Team. 2016. Stan Modeling Language Users Guide and Reference Manual.
- 1050      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  
1052      the United States of America 102:11980–4.
- Tomiya, S. 2013. Body Size and Extinction Risk in Terrestrial Mammals Above the Species Level.  
1054      The American Naturalist 182:196–214.
- Urban, M. C., M. A. Leibold, P. Amarasekare, L. De Meester, R. Gomulkiewicz, M. E. Hochberg,  
1056      C. A. Klausmeier, N. Loeuille, C. de Mazancourt, J. Norberg, J. H. Pantel, S. Y. Strauss,  
M. Vellend, and M. J. Wade. 2008. The evolutionary ecology of metacommunities. Trends in  
1058      Ecology and Evolution 23:311–317.
- Valentine, J. W. 1969. Patterns of taxonomic and ecological structure of the shelf benthos during  
1060      Phanerozoic time. Paleontology 12:684–709.
- Valkenburgh, B. V. 1999. Major patterns in the history of carnivorous mammals. Annual Review of  
1062      Earth and Planetary Sciences 27:463–493.
- Van Valkenburgh, B. 1990. Skeletal and dental predictors of body mass in carnivores. Pages  
1064      181–205 *in* J. Damuth and B. J. Macfadden, eds. Body size in mammalian paleobiology:  
estimation and biological implications. Cambridge University Press, Cambridge.
- 1066      Villéger, S., P. M. Novack-Gottshall, and D. Mouillot. 2011. The multidimensionality of the niche  
reveals functional diversity changes in benthic marine biotas across geological time. Ecology  
1068      letters 14:561–8.
- Wang, S. C., P. J. Everson, H. J. Zhou, D. Park, and D. J. Chudzicki. 2016. Adaptive credible  
1070      intervals on stratigraphic ranges when recovery potential is unknown. Paleobiology 42:240–256.

- Wang, S. C., and C. R. Marshall. 2016. Estimating times of extinction in the fossil record. *Biology Letters* 12:20150989.
- 1072
- Warton, D. I., B. Shipley, and T. Hastie. 2015. CATS regression - a model-based approach to  
1074 studying trait-based community assembly. *Methods in Ecology and Evolution* 6:389–398.
- Webb, C. O., D. D. Ackerly, M. a. McPeek, and M. J. Donoghue. 2002. Phylogenies and  
1076 Community Ecology. *Annual Review of Ecology and Systematics* 33:475–505.
- Weber, M. G., C. E. Wagner, R. J. Best, L. J. Harmon, and B. Matthews. 2017. Evolution in a  
1078 Community Context: On Integrating Ecological Interactions and Macroevolution. *Trends in  
Ecology & Evolution* xx:1–14.
- 1080 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,  
1082 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  
1084 Evolutionary Biology* 23:1581–1596.
- Zachos, J. C., G. R. Dickens, and R. E. Zeebe. 2008. An early Cenozoic perspective on greenhouse  
1086 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  
1088 aberrations in global climate 65 Ma to present. *Science* 292:686–693.

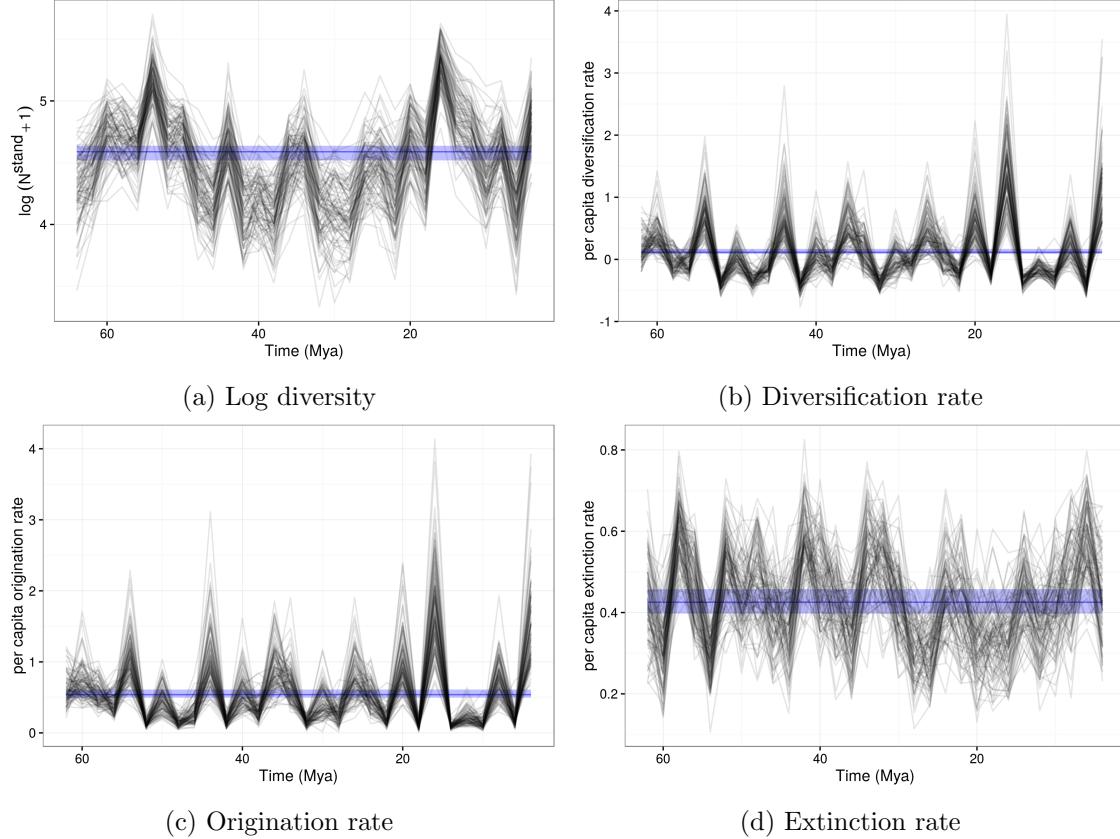


Figure 14: Posterior estimates of the time series of Cenozoic North American mammal diversity and its characteristic macroevolutionary rates; all estimates are from the birth-death model and 100 posterior draws are plotted to indicate the uncertainty in these estimates. The blue horizontal strip corresponds to the 80% credible interval of estimated mean standing diversity, diversification rate, origination rate, and extinction rate respectively; the median estimate is also indicated. What is also plotted is the The dramatic differences between diversity estimates at the first and second time points and the penultimate and last time points in this series are caused by well known edge effects in discrete-time birth-death models caused by  $p_{-,t=1}$  and  $p_{-,t=T}$  being partially unidentifiable (Royle and Dorazio, 2008); the hierarchical modeling strategy used here helps mitigate these effects but they are still present (Gelman et al., 2013; Royle and Dorazio, 2008). Diversification rate is in units of species gained per species present per time unit (2 My), origination rate is in units of species originating per species present per time unit, and extinction rate is in units of species becoming extinct per species present per time unit.

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

Time (Mya)	$P(N_t^{stand} > \bar{N}^{stand})$	$P(D_t^{rate} > \bar{D}^{rate})$
64.00	0.07	0.63
62.00	0.28	0.94
60.00	0.86	0.13
58.00	0.68	0.18
56.00	0.62	0.99
54.00	1.00	0.00
52.00	0.68	0.41
50.00	0.80	0.00
48.00	0.12	0.04
46.00	0.01	0.98
44.00	0.64	0.00
42.00	0.02	0.47
40.00	0.03	0.08
38.00	0.00	0.89
36.00	0.40	0.46
34.00	0.52	0.00
32.00	0.02	0.27
30.00	0.06	0.09
28.00	0.02	0.88
26.00	0.22	0.39
24.00	0.38	0.03
22.00	0.09	0.96
20.00	0.81	0.00
18.00	0.29	1.00
16.00	1.00	0.00
14.00	0.95	0.02
12.00	0.80	0.01
10.00	0.13	0.83
8.00	0.67	0.00
6.00	0.02	1.00
4.00	0.91	

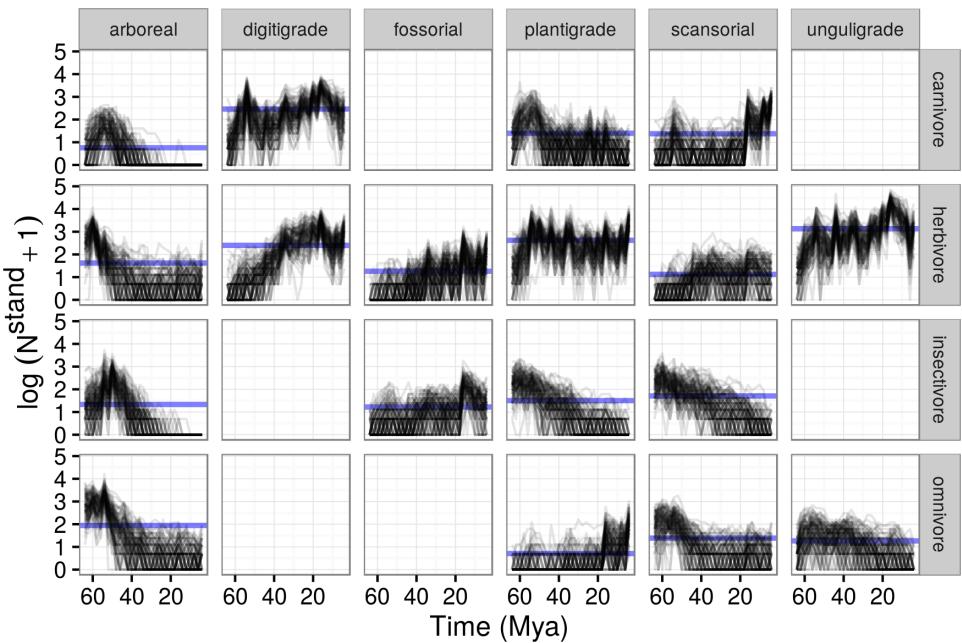


Figure 15: Posterior of standing log-diversity of North American mammals by ecotype for the Cenozoic as estimated from the birth-death model; 100 posterior draws are plotted to indicate the uncertainty in these estimates and what is technically plotted is log of diversity plus 1.

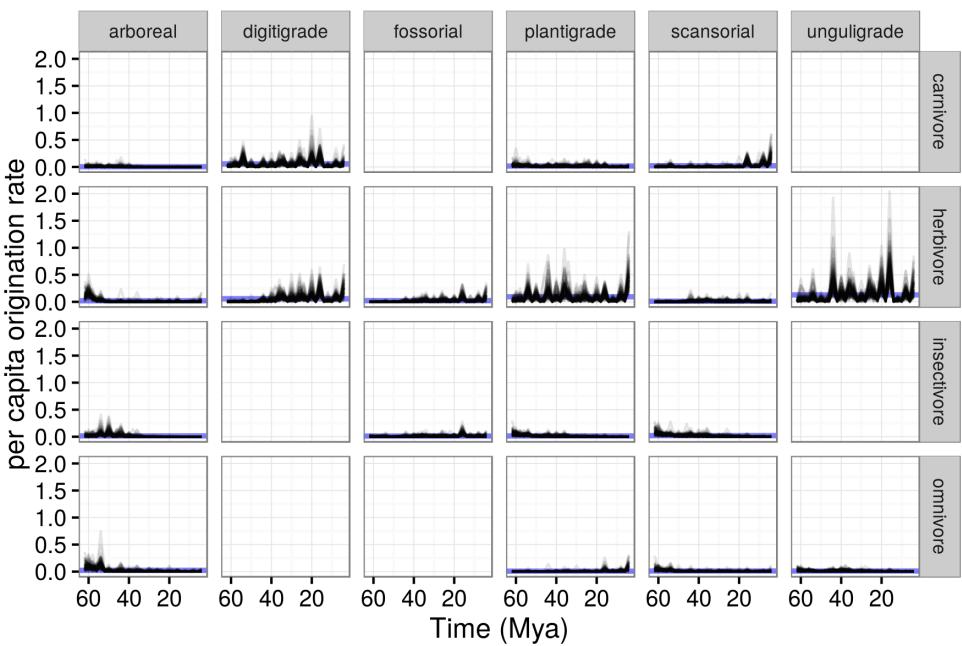


Figure 16: Posterior estimates of the number of origination events from one temporal bin to another, plotted at the bin they originate from. 100 posterior draws are plotted to indicate the uncertainty in these estimates. Also, what is plotted is log of the number of originations plus 1.

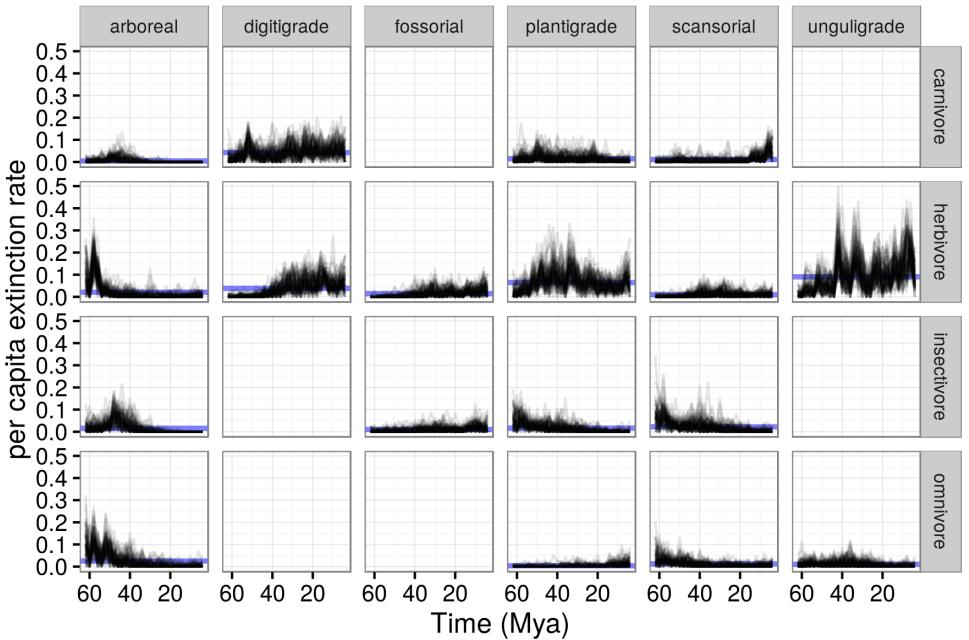


Figure 17: Posterior estimates of the number of extinction events from one temporal bin to another, plotted at the bin they go extinct from. 100 posterior draws are plotted to indicate the uncertainty in these estimates. Also, what is plotted is log of the number of extinctions plus 1.