

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

Peter D. Smits^{1,*}

1. University of Chicago, Chicago, Illinois 60637.

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

Manuscript elements:

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

Manuscript type: Article

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

Abstract

The set of species in a region changes over time as new species enter through speciation or immigration and as species leave the system through extinction and extirpation. How a regional 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 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 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 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 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 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 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 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, in almost all cases, does not affect the extinction of mammal ecotypes.

Introduction

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 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
30 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
32 speciation, extinction, migration, and extirpation. The gain or loss of regional diversity is the result
of macroevolutionary dynamics which, in turn, shape the downstream macroecological dynamics of
34 the regional species pool and its constituent local communities (Harrison and Cornell, 2008;
Mittelbach and Schemske, 2015; Urban et al., 2008).

36 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
38 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
40 group are expected to have more similar macroecological dynamics to each other than to species of
a different functional group. By focusing on the relative diversity of functional groups, changes to
42 diversity are interpretable as changes to the set of ways species within a species pool could interact
with the biotic and abiotic environment.

44 A key question when comparing communities or regional species pools based on their functional
composition is whether specific functional groups are enriched or depleted and why; what are the
46 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)?
48 Comparisons of contemporaneous regional species pools only determines if a functional group is
enriched or depleted in one species pool relative to other species pools. This type of comparison
50 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
52 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
54 which shape regional species pool diversity (e.g. origination, extinction) operate on much longer
time scales than is possible for studies of the Recent, paleontological data provides a unique
56 opportunity to observe and estimate the changes to functional diversity and how species functional

traits and environmental context can shape the enrichment or depletion of functional groups within
58 a regional species pool (Blois and Hadly, 2009; Smith et al., 2008). Being able to identify which if
the diversity of any functional groups are depleted relative to their long term average diversity in
60 the species pool is particularly useful in conservation settings; species in depleted groups are most
likely more at risk of extinction than species in enriched groups, even if those enriched groups are
62 relatively rare when compared to the functional composition of other contemporaneous species
pools.

64 The paleontological record of North American mammals for the Cenozoic (\sim 66 million years ago to
present) provides one of the best opportunities for understanding how regional species pool
66 functional diversity changes over time. The North American mammal record is a relatively complete
temporal sequence for the entire Cenozoic which primarily, but not exclusively, based on fossil
68 localities from 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
70 that functional traits like the dietary/trophic category of species are easy to estimate (Eronen et al.,
2010; Polly et al., 2011, 2015).

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

76 **Background**

The diversity history of North American mammals for the Cenozoic is relatively well understood as
78 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
80 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
82 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 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., 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., 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 macroevolutionary and macroecological processes shaping diversity and diversification. Integration across perspectives is necessary for producing a holistic and internally consistent picture of how the 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 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.

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 cross-functional group or cross-taxonomic turnover events in mammal diversity at any point in the Cenozoic record of North America (Alroy, 1996, 2009; Alroy et al., 2000; Eronen et al., 2015; Janis, 1993). Instead of being concentrated at specific time intervals, species turnover has been found to be distributed through time. It is then expected then that, for this analysis, turnover events or periods of rapid diversification or depletion should not occur simultaneously for 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 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 diversity-dependent (Alroy, 2009; Harmon and Harrison, 2015; Rabosky, 2013; Rabosky and Hurlbert, 2015). Similarity, this question can also be asked of specific functional groups (Jernvall and Fortelius, 2004; Quental and Marshall, 2013; Silvestro et al., 2015; Van Valkenburgh, 1999).

Within the overall narrative of mammal diversity, the histories of a selection of taxonomic and

functional groups are better understood. These groups have particularly good fossil records and/or
112 have been the focus of previous analyses.

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

In contrast, the origination of modern cursorial carnivore forms was not until later in the Cenozoic;
122 this is not to say that carnivore diversity only grew in the late Cenozoic, but that those forms were
late entrants (Janis and Wilhelm, 1993). Instead, the diversity history of carnivores is reflective of
124 density-dependence or some other form of self-regulation (Silvestro et al., 2015; Slater, 2015; Van
Valkenburgh, 1999). Specifically, it has been proposed that different canid clades have replaced each
126 other as the dominate members of their functional group within the species pool (Silvestro et al.,
2015; Van Valkenburgh, 1999). It is then expected that, for this analysis, the diversity of digitigrade
128 and plantigrade carnivores (i.e. the “carnivore” guild of Van Valkenburgh (1999)) should be
relatively constant for the Cenozoic or at least have plateaued by the Neogene.

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

138 when compared to other taxa, then the diversity history of arboreal taxa is expected to be constant
with time, albeit possibly at low diversity; and 2) if the Paleogene and Neogene represent difference
140 selective regimes with the former being associated with lower extinction risk than the latter, then
the diversity history of arboreal taxa are expected to be present in the Paleogene but depleted or
142 absent from the species pool during the Neogene.

The climate history of the Cenozoic can be broadly described as a gradual cooling trend, with polar
144 ice-caps forming in the Neogene (Cramer et al., 2011; Zachos et al., 2008, 2001). There are of course
exceptions to this pattern such as the Eocene climatic optimum, the mid-Miocene climatic optimum,
146 and the sudden drop in temperature at the Eocene/Oligocene boundary (Zachos et al., 2008, 2001).
In terms of the North American biotic environment, the Cenozoic is additionally characterized by
148 major transition from having closed, partially forested biomes being common in the Paleogene to
the landscape being dominated by savannah and grasslands biomes by the Neogene (Blois and
150 Hadly, 2009; Janis, 1993; Janis et al., 2000; Strömberg, 2005). Additionally, the landscape structure
and topology of North America changed substantially over the Cenozoic with mountain uplift and
152 other tectonic actives in Western North America (Badgley and Finarelli, 2013; Blois and Hadly,
2009; Eronen et al., 2015; Janis, 2008). This type of geological activity affects both local climates as
154 well as continental weather patterns while also mobilizing increased grit into the environment,
something which may be responsible for increasing trend of hyposodony (high crowned teeth)
156 among herbivores (Damuth and Janis, 2011; Jardine et al., 2012; Jernvall and Fortelius, 2002).

The effect of climate on mammal diversity and its accompanying diversification process has been the
158 focus of considerable research with a slight consensus favoring mammal diversification being more
biologically-mediated than climate-mediated (Alroy, 1996; Alroy et al., 2000; Clyde and Gingerich,
160 1998; Figueirido et al., 2012). However, differences in temporal and geographic scale seem to underly
the contrast between these two perspectives. For example when the mammal fossil record analyzed
162 at small temporal and geographic scales a correlation between diversity and climate is observable
(Clyde and Gingerich, 1998). However, when the record is analyzed at the scale of the continent and
164 most of the Cenozoic this correlation disappears (Alroy et al., 2000). This result, however, does not
go against the idea that there may be short periods of correlation between diversity and climate and

¹⁶⁶ that this relationship can change or even reverse direction over time; this type result means that
there is no single direction of correlation between diversity and climate (Figueirido et al., 2012).

¹⁶⁸ 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
¹⁷⁰ 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
¹⁷² 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
¹⁷⁴ 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
¹⁷⁶ can be observed and analyzed rather than assumed.

Foreground

¹⁷⁸ Fourth-corner modeling is an approach to explaining the patterns of either species abundance or
presence/absence in a community as a product of species traits, environmental factors, and the
¹⁸⁰ interaction between traits and environment (Brown et al., 2014; Jamil et al., 2013; Pollock et al.,
2012; Warton et al., 2015); effectively uniting climate-based species distribution modeling (SDMs)
¹⁸² with trait-based community assembly models (CATS, MaxEnt). In modern ecological studies, what
is being modeled is species occurrences at localities distributed across a region (Jamil et al., 2013;
¹⁸⁴ Pollock et al., 2012). In this study, what is being modeled is the pattern of species occurrence over
time for most of the Cenozoic record for North America (Fig. 1). By analyzing assemblages over
¹⁸⁶ time instead of space in fourth-corner framework we can gain better inference of how an
instantaneous species pool (i.e. the Recent) is assembled over time. These two approaches, modern
¹⁸⁸ and paleontological, are different views of the same three-dimensional pattern: species at localities
over time. The temporal limitations of modern ecological studies and difficulties with uneven spatial
¹⁹⁰ occurrences of fossils in paleontological studies means that these approaches are complementary and
reveal different patterns of how species are distributed in time and space.

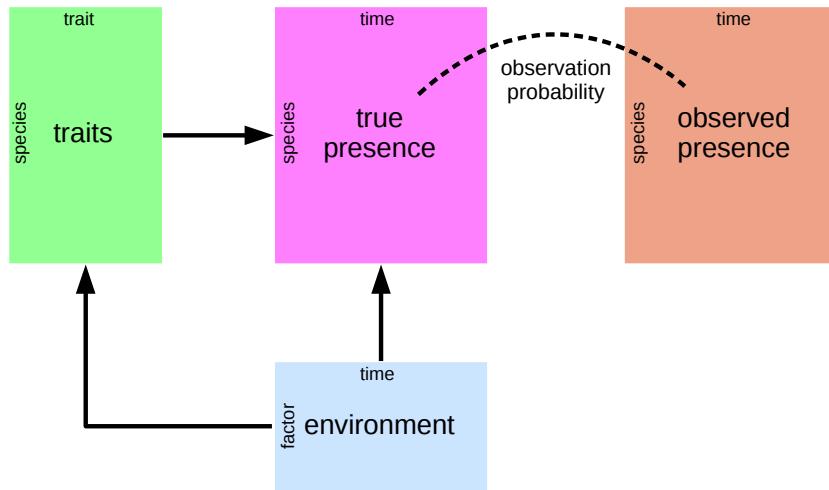


Figure 1: Conceptual diagram of the paleontological fourth corner problem. The observed presence matrix (orange) is the empirical presence/absence pattern for all species for all time points; this matrix is an incomplete observation of the “true” presence/absence pattern (purple). The estimated true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

²²⁶ **Materials and Methods**

Taxon occurrences and species-level information

²²⁸ 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>.

²³⁸ 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.

²⁴⁶ 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. The life history category was then further broken up to better reflect the diversity of 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). 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 was possible to assign a locomotor category had one assigned, including species for which post-crania are unknown but for which a taxonomic grouping is known. Ground dwelling species which were unable to be reassigned based on ankle posture were excluded from analysis. Finally, ecotype categories with less than 10 total species were excluded, yielding a total of 18 observed ecotypes out of a possible 24.

Table 1: Species trait assignments in this study are a coarser version of the information available in the PBDB. Information was coarsened to improve per category sample size.

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
	Aplopontidae	plantigrade

Continued on next page

Table 2 – continued from previous page

Order	Family	Stance
	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
	Florentiamyidae	plantigrade

Continued on next page

Table 2 – continued from previous page

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

Continued on next page

Table 2 – continued from previous page

Order	Family	Stance
	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
	Soricidae	plantigrade

Continued on next page

Table 2 – continued from previous page

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

260

- Estimates of species mass used in this study were sourced from multiple databases and papers,
262 especially those focusing on similar macroevolutionary or macroecological questions (Brook and
Bowman, 2004; Freudenthal and Martín-Suárez, 2013; McKenna, 2011; Raia et al., 2012; Smith
264 et al., 2004; Tomiya, 2013); this is similar to what was done in Smits (2015). When species mass was
not available, proxy measures were used and then transformed into estimates of mass. For example,
266 given a measurement of a mammal tooth size, it is possible and routine to estimate its mass given
some regression equation. The PBDB has one or more body part measures for many species. These
268 were used as body size proxies for many species, as was the case in Smits (2015). Mass was
log-transformed and then rescaled by first subtracting mean log-mass from all mass estimates, then
270 dividing by two-times its standard deviation; this insures that the magnitude of effects for both
continuous and discrete covariates are directly comparable (Gelman, 2008; Gelman and Hill, 2007).
- 272 In total, 1400 mammal species occurrence histories were included in this study after applying all of
the restrictions above.
- 274 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 during the second half of the Cenozoic. These properties make

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

290 Mg/Ca based temperature estimates preferable for macroevolutionary and macroecological studies
 (Ezard et al., 2016). Two aspects of the Mg/Ca-based temperature curve were included in this
 292 analysis: mean and range. Both were calculated as the mean of all respective estimates for each 2
 My temporal bins. The distributions of the temperature mean and range estimates were then
 294 rescaled by subtracting their respective means from all values and then dividing by twice their
 respective standard deviations.

296 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
 298 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
 300 (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
 302 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
 304 and subsequent phases are defined by their differences from the first phase. The temporal
 boundaries of these plant phases are defined in Table 4.

306 Modelling species occurrence

Two different models were used in this study: a pure-presence model and a birth-death model. Both
 308 models at their core are hidden Markov models where the latent process has an absorbing state
 (Allen, 2011). The difference between these two models lies in whether the probabilities of a species
 310 originating or surviving are considered equal or different (Table 5). While there are only two state

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

(a) Pure-presence

		State at $t + 1$		
		0_{never}	1	$0_{extinct}$
State at t	0_{never}	$1 - \pi$	π	0
	1	0	ϕ	$1 - \phi$
	$0_{extinct}$	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
α_0	1	average log-odds of p
α_1	1	change in average log-odds of p per change mass
r	T	difference from α_0 associated with time t
σ	1	standard deviation of r

“codes” in a presence-absence matrix (i.e. 0/1), there are in fact three states in a birth-death model:

- 312 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
 314 probability of an extinct species changing states is 0 (Table 5). See below for parameter explanations (Tables 6, 7, and 8).

316 Observation process

The type of hidden Markov model used in this study has three characteristic probabilities:

- 318 probability p of observing a species given that it is present, probability ϕ of a species surviving from one time to another, and probability π of a species first appearing (Royle and Dorazio, 2008). In
 320 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 $\phi \neq \pi$.

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

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

322 The probability p of observing a species that is present is modeled as a logistic regression with a
 time-varying intercept and species mass as a covariate. The effect of species mass on p was assumed
 324 linear and constant over time. These assumptions are reflected in the structure of this part of the
 model being presented here:

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

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

Pure-presence process

328 For the pure-presence model there is only a single probability dealing with the presence of a species
 θ (Table 5a). This probability was modeled as multi-level logistic regression with both species-level
 330 and group-level covariates (Gelman et al., 2013; Gelman and Hill, 2007). The parameters associated
 with the pure-presence model are presented in Table 7, and the full sampling statement in Equation
 332 2.

Species mass was included as a covariate with two regression coefficients allowing for a quadratic
 334 relationship 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

³³⁶ body size will 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
³³⁸ prior with most of its density below 0 (Eq. 2).

The values of each ectype's intercept are themselves modeled as regressions using the group-level
³⁴⁰ 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
³⁴² covariances between ectype intercepts, given this group-level regression, are modeled (Gelman and
Hill, 2007). The prior choice for the covariance matrix separates it into a vector of scales τ and a
³⁴⁴ correlation matrix Ω . The elements of the former are given weakly informative, independent
half-Normal priors while the latter was given a weakly informative LKJ prior as recommended in
³⁴⁶ the Stan manual (Stan Development Team, 2016).

All parameters not modeled elsewhere were given weakly informative priors (Gelman et al., 2013;
³⁴⁸ McElreath, 2016; Stan Development Team, 2016). Weakly informative means that priors do not
necessarily encode actual prior information but instead help regularize or weakly constrain posterior
³⁵⁰ estimates. These priors have a concentrated probability density around and near zero; this has the
effect of tempering our estimates and help prevent overfitting the model to the data (Gelman et al.,
³⁵² 2013; McElreath, 2016; Stan Development Team, 2016). The general line of thinking behind this

approach is that a result of 0 or “no effect” is more preferable to a wrong or extremely weak result.

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

³⁵⁴ Birth-death process

In the birth-death version of the model, $\phi \neq \pi$ and so each of these probabilities is modeled ³⁵⁶ separately but each is handled in a similar manner to how θ is modeled in the pure-presence model (Eq. 2, Table 5b). The parameters associated with the birth-death presence model are presented in

³⁵⁸ Table 8 and the full sampling statement, including observation (Eq. 1), is described in Equation 3:

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

Similar to the pure-presence model, both ϕ and π are modeled as logistic regressions with varying
³⁶⁰ intercept and one covariate associated with two parameters. The possible relationships between
mass and both ϕ and π are reflected in the parameterization of the model and choice of priors (Eq.
³⁶² 3).

The intercepts of ϕ and π both vary by species ecotype and those values are themselves the product
³⁶⁴ of group-level regression using environmental factors as covariates (Eq. 3); this is identical to the
pure presence model (Eq. 2).

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

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

366 Posterior inference and model adequacy

Computer programs that implement joint posterior inference for the above models (Eqs. 2, 3) were
 368 written in the probabilistic programming language Stan (Stan Development Team, 2016). Both
 models feature a large matrix of latent discrete parameters z (Tables 6, 7, 8; Eqs. 1, 2, 3). All
 370 methods for posterior inference implemented in Stan are derivative-based; this causes complications
 for actually implementing the above models, because integers do not have derivatives. Instead of
 372 implementing a latent discrete parameterization, the log posterior probabilities of all possible states
 of the latent parameters z were calculated and summed (i.e. marginalized).

374 Species durations at minimum range through from a species first appearance to their last
 appearance in the fossil record, but the incompleteness of all observations means that the actual
 376 times of origination and extinction are unknown. The marginalization approach used here means
 that the probabilities of all possible histories for a species are calculated, from the end members of
 378 the species having existed for the entire study interval and the species having only existed between

the directly observed first and last appearances to all possible intermediaries (Fig 2) (Stan
 380 Development Team, 2016). This process is identical, language-wise, to assuming range-through and
 then estimating the possibility of all possible 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	0	1	1	1	1	1	1	0
Potential	1	1	1	1	1	1	1	0
Potential	0	0	0	1	1	1	1	1
Potential	0	0	1	1	1	1	1	1
Potential	0	1	1	1	1	1	1	1
Potential	1	1	1	1	1	1	1	1

Figure 2: Conceptual figure of all possible occurrence histories for an observed species. The first row represents the observed presence/absence pattern for a single species at eight time points. The second row corresponds to the known aspects of the “true” occurrence history of that species. The remaining rows correspond to all possible occurrence histories that are consistent with the observed data. 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.

382 The combined size of the dataset and large number of parameters in both models (Eqs. 2, 3),
 specifically the total number of latent parameters that are the matrix z , means that stochastic
 384 approximate posterior inference is computationally very slow even using NUTS based HMC as
 implemented in Stan (Stan Development Team, 2016). Instead, an approximate Bayesian approach
 386 was used: variational inference. A recently developed automatic variational inference algorithm
 called “automatic differentiation 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 Bayesian posterior; this assumption is similar to quadratic
 390 approximation of the likelihood function commonly used in maximum likelihood based inference

(McElreath, 2016). The principal limitation of assuming the joint posterior is Gaussian is that the
392 true topology of the log-posterior isn't estimated; this is a particular burden for scale parameters
which are bounded to be positive (e.g. standard deviation).

394 Of additional concern for posterior inference is the partial identifiability of observation parameters
 $p_{t=1}$ and $p_{t=T}$ (Royle and Dorazio, 2008). This issue means that the estimates of sampling
396 probabilities at the “edges” of the time series cannot fully be estimated because there are no known
“gaps” in species occurrence histories that are guaranteed to be filled. Instead, the values of the first
398 and final columns of the “true” presence-absence matrix z for those observations that do not already
have presences in the observed presence-absence matrix y cannot be estimated (Royle and Dorazio,
400 2008). The hierarchical modeling approach used here helps mitigate this problem by pulling the
values of $p_{t=1}$ and $p_{t=T}$ towards the overall mean of p (Gelman et al., 2013), and in fact this
402 approach might be more analytically sound than the more ad-hoc approaches that are occasionally
used to overcome this hurdle (Royle and Dorazio, 2008). Additionally, because $p_{t=1}$ and $p_{t=T}$ are
404 only partially identifiable, estimates of occurrence θ and origination ϕ at $t = 1$ and estimates of θ , ϕ
and survival π at $t = T$ may suffer from similar edge effects. Again, the hierarchical modeling
406 approach used here may help correct for this reality by drawing these estimates towards the overall
means of those parameters.

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

occurrences per species.

- 420 The ADVI assumption of a purely Gaussian posterior limits the utility and accuracy of the
 posterior predictive checks because parameter estimates do not reflect the true posterior
- 422 distribution and are instead just an approximation (Gelman et al., 2013). Because of this, posterior
 predictive estimates are themselves only approximate checks of model adequacy. The posterior
- 424 predictive check that is used in this study focuses on mean occurrence and not to any scale
 parameters that might be most affected by the ADVI assumptions.
- 426 Given parameter estimates, diversity and diversification rates are estimated through posterior
 predictive simulations. Given the observed presence-absence matrix y , estimates of the true
- 428 presence-absence matrix z can be simulated and the distribution of possible occurrence histories
 can be analyzed. This is conceptually similar to marginalization where the probability of each
- 430 possible occurrence history is estimated (Fig. 2), but now these occurrence histories are generated
 relative to their estimated probabilities.
- 432 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
 434 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

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

⁴³⁶ 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

⁴³⁸ 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 ⁴⁴⁰ rates based on posterior predictive simulations from the birth-death model because this model has a better fit to the observed occurrence information.

⁴⁴² **Comparing parameter estimates from the pure-presence and birth-death models**

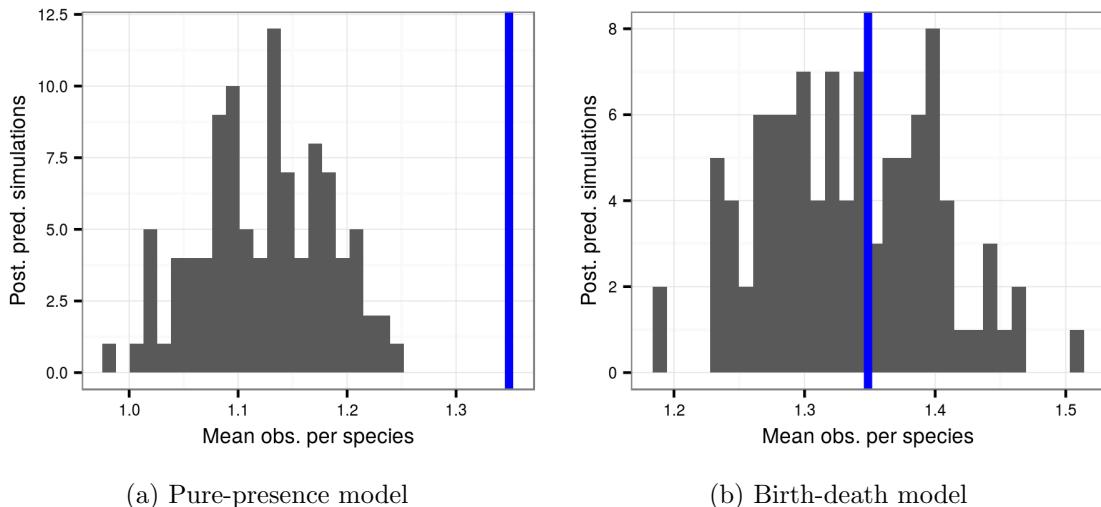


Figure 3: Comparison of the average observed number of occurrences per species (blue line) to the average number of occurrences from 100 posterior predictive datasets using the posterior estimates from the pure-presence and birth-death models.

⁴⁴⁴ 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
446 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
448 which 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
450 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
452 posterior predictive results (Gelman et al., 2013).

Increases in the occurrence probability of an ecotype is interpreted as an increase in the
454 commonness of that ecotype in the species pool. In turn, decreases in the occurrence probability of
an ecotype are interpreted to a decrease in the commonness of that ecotype in the species pool.
456 Additionally, when the uncertainty surrounding a probability estimate is very high, as with arboreal
insectivores, this is interpreted as complete separation which means that that ecotype has most
458 likely all but disappeared from the species pool (Gelman and Hill, 2007). In logistic regression, high
uncertainty in the estimates of the underlying log-odds of occurrence, origination, or survival tends
460 to indicate 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.
462 The problem of complete separation is mitigated by the hierarchical modeling strategy used here
(Gelman et al., 2013; Gelman and Hill, 2007; McElreath, 2016).

464 There are a few shared patterns of ecotype occurrence probabilities as estimated from the
pure-presence model (Fig. 4). The first pattern I identify is that shared by digitigrade carnivores,
466 scansorial carnivores, digitigrade herbivores, fossorial herbivores, and unguligrade herbivores. All of
these ecotypes begin with the Cenozoic with very low occurrence probability and then quickly
468 jumping up to very high probability of presence at approximately 50 Mya except for digitigrade
herbivores whose occurrence probability jumps closer to 40 Mya. The occurrence probabilities of
470 these ecotypes stay high until the start of the third plant phase 18 Mya where it decreases slightly.
The second pattern I identify is where occurrence probability begins very high and then slowly
472 decreases over the Cenozoic. Example ecotypes that fit into this category are arboreal carnivores,

plantigrade insectivores, scansorial insectivores, arboreal omnivores, scansorial omnivores, and
474 unculigrade omnivores. The third pattern I identify is that shared by scansorial herbivores, fossorial
insectivores, and plantigrade omnivores. The occurrence probabilities of all these ecotypes begin
476 very low and jump up at the start of the third plant phase 18 Mya. The final shared pattern is a
“washtub” pattern with occurrence being higher in the first plant phase than the second, which in
478 turn is lower than the occurrence probabilities for the third plant phase. The only ecotype that
doesn’t share its pattern of occurrence probability with any other ecotype are arboreal insectivores.
480 Instead, the occurrence probability time series for this ecotype is with such uncertainty that it is
not possible to distinguish a simple pattern.

482 Origination probabilities estimated from the birth-death model are very different from the
occurrence probability estimates from the pure-presence model (Fig. 5); this is also the case for the
484 survival probability estimates (Fig. 6). For most ecotypes, origination probabilities estimated from
the birth-death model (Fig. 5) increase with time. This makes sense given that, over time, all
486 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 with time. Notably, ecotypes with
488 arboreal components do not appear to follow the same pattern as most other ecotypes; instead,
origination probabilities appear relatively flat with high posterior variance for most of the Cenozoic.
490 For most ecotypes, the origination probability is estimated with less uncertainty than its estimate
of survival probability (Fig. 5, and 6).

492 The dramatic differences in the estimates origination and survival probabilities are indicative of
how differently these processes affect the diversification process and may also be responsible for the
494 better posterior predictive performance of the birth-death model over the pure-presence model (Fig.
3a, and 3b). While the estimates at all points along both time series have high variance, what is
496 striking is how mean origination probability changes over time while most ecotype survival
probabilities have relatively stable means for the entire Cenozoic (Fig. 5, and 6).

498 The pure-presence and birth-death models have similar estimates of effect of mass on the
probability of sampling a species that is present (Fig. 7). For both models this relationship is

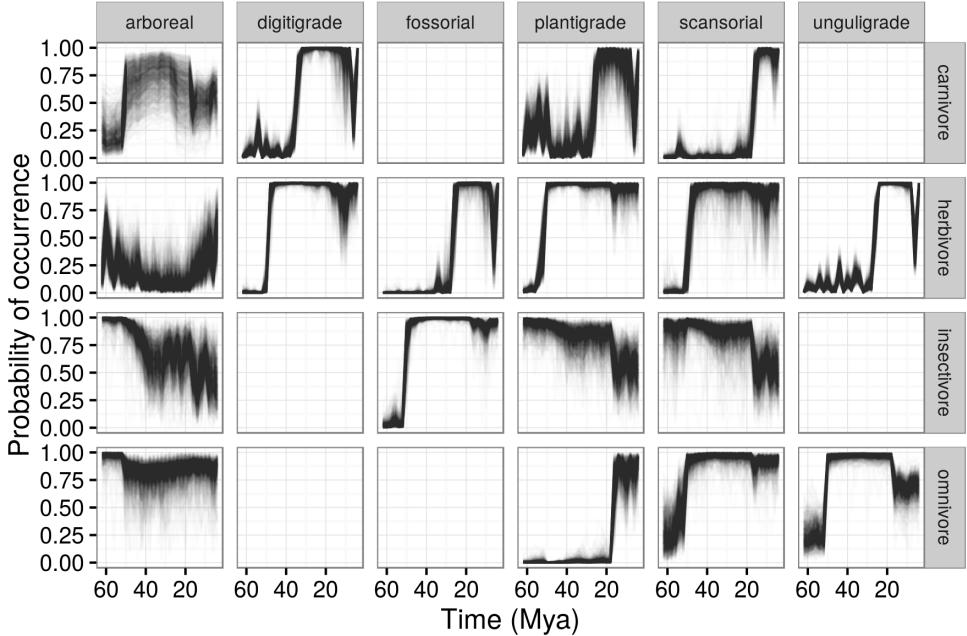


Figure 4: Probability of a mammal ecotype occurring over time as estimated from the pure-presence model. Each panel depicts 100 random samples from the model’s posterior. The columns are by locomotor category and rows by dietary category; their intersections are the observed and analyzed ecotypes. Panels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

positive, which means that as species body mass increases it is expected that they are more likely to be sampled if present. Additionally, the estimated relationship from the pure-presence model is with less uncertainty than that from the birth-death model (Fig. 7). This result is consistent with intuition as larger fossils are more visible to the eye and thus easier to sample. In turn, this means that observed occurrence histories small bodied species are more likely to have gaps, where $y = 0$ for that species the true state z is 1.

For the pure-presence model, species mass is found to have either no relationship with occurrence or a negative one (Fig. 8). A negative relationship between body size and occurrence is interpreted to mean that large bodied species are likely to occur less frequently than smaller bodied species. Note that all variation in estimates between ecotypes (Fig. 8) is due to differences in ecotype-specific occurrence probability and the associated effects of plant phase; the effect of mass was considered constant for all ecotypes.

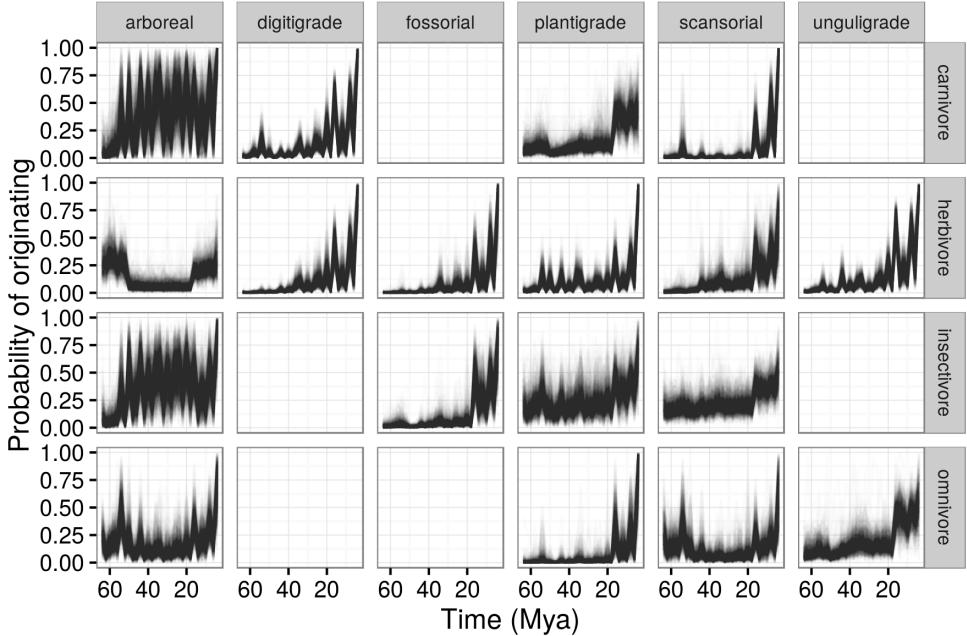


Figure 5: Probability of a mammal ecotype origination probabilities at each time point as estimated from the birth-death model. Each panel depicts 100 random samples from the model’s posterior. The columns are by locomotor category and rows by dietary category; their intersections are the observed and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

- 512 is most similar to the estimated effect of species mass on probability of origination for the
 birth-death model (Fig. 9). The striking pattern is higher probability of origination for species with
 514 body sizes closer to the mean than either extremes. This result is consistent with the canonically
 normal distribution of mammal body sizes (Smith et al., 2004); it is then expected that the most
 516 likely to occur species would be those from the middle of the distribution, and that species
 originating will on average be of average mass, especially considering species shared common
 518 ancestry (Felsenstein, 1985). As with the results from the pure-presence model (Fig. 8) all variation
 in estimates between ecotypes (Fig. 9) is due to differences in ecotype-specific origination
 520 probabilities and the associated effects of plant phase; the effect of mass was considered constant for
 all ecotypes.
- 522 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
 524 for North American mammals for the Cenozoic (Smits, 2015; Tomiya, 2013). Note that all variation

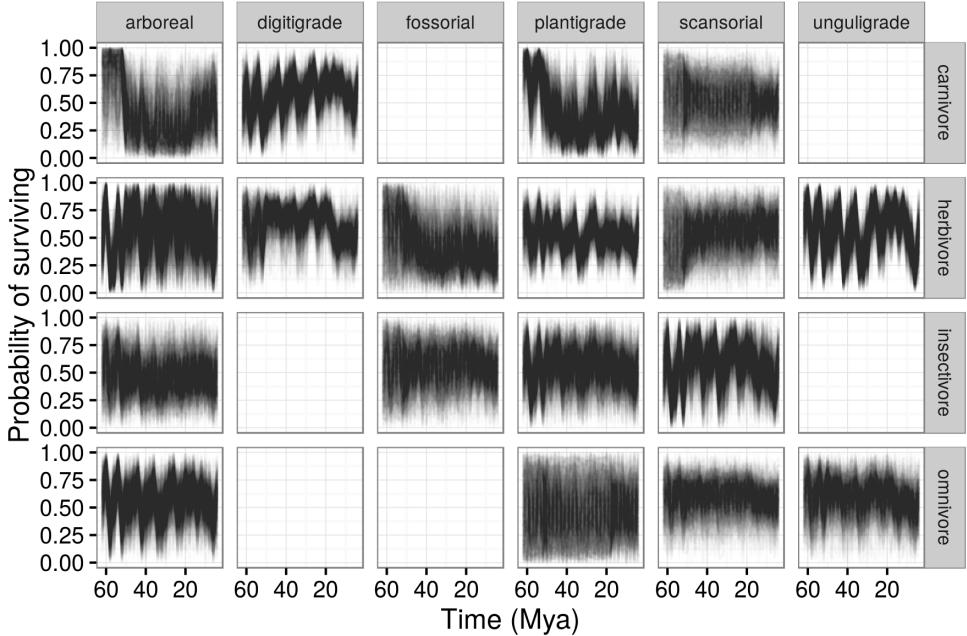


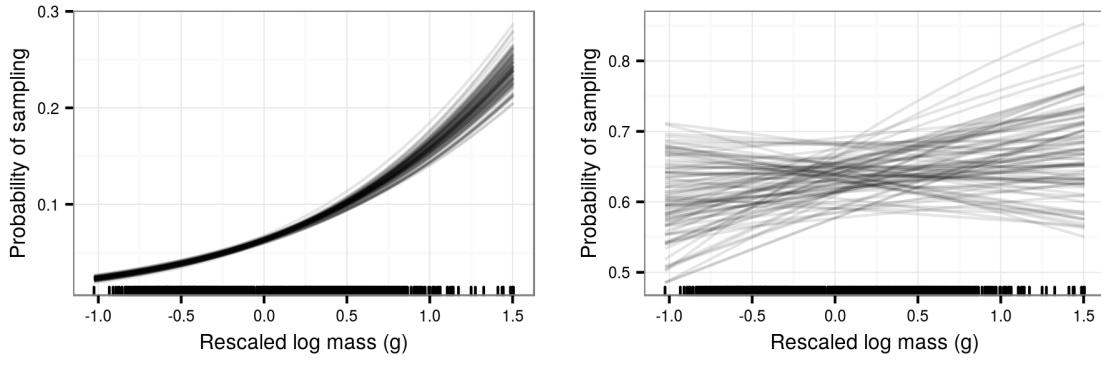
Figure 6: Probability of a mammal ecotype survival probabilities at each time point as estimated from the birth-death model. Each panel depicts 100 random samples from the model’s posterior. The columns are by locomotor category and rows by dietary category; their intersections are the observed and analyzed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

between ecotypes depicted in Figure 10 is due to differences in ecotype-specific survival probability

526 and the associated effects of plant phase; the effect of mass was considered constant for all ecotypes
(Eqs. 2, 3).

528 Similarities in parameter estimates between ecotypes may be due to a similar response to environmental factors (Fig. 11, 12, and 13). The estimated group-level effects on ecotype occurrence,
530 origination, or survival are all very different from each other. At best, the effects of temperature on occurrence and origination can be considered congruent (Fig. 11, 12). As demonstrated in the
532 comparisons of the effect of body mass on occurrence from the pure-presence model (Fig. 8) with the effect of body mass on origination and survival from the birth-death model (Fig. 9, and 10),
534 there is considerable variation in the effect of plant phases on ecotype-specific estimates.

An association between plant phase and differences in the log-odds of occurrence (Fig. 11),
536 origination (Fig. 12), or extinction (Fig. 13) is interpreted to mean that the set of possible



(a) Pure-presence model

(b) Birth-death model

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

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 conditions for an ecotype may indicate an increasing number of possible and available

538
540
542 mammal-plant interactions (e.g. ecological opportunity; Losos, 2010; Losos and Mahler, 2010; Yoder 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.

544 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 546 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 548 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 550 probabilities (Tables 9, 10), though there is little evidence of plant phase being an important 552 distinguishing factor in species survival, as only a few ecotypes demonstrate strong affinities with some plant phases (Table 11). Curiously, the effects of plant phase on occurrence probability (Fig. 11 are almost the opposite of the effects of plant phase on origination probability (Fig. 12). This

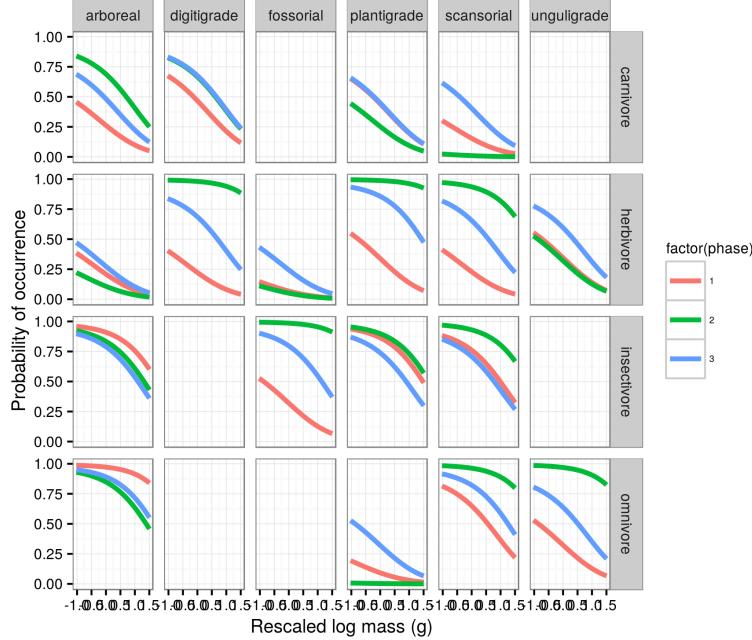


Figure 8: Mean estimate of the effect of species mass on the probability of a species occurrence for each of the three plant phases. The effect of mass is considered constant over time and that the only aspect of the model that changes with plant phase is the intercept of the relationship between mass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

discrepancy is another example of how modeling species occurrence as two separate processes reveals how different origination and extinction are from each other and how they shape diversity in different ways.

The most common pattern of the effects of plant phase on ecotype occurrence is that during first and last plant phases ecotypes have a lower log-odds of occurrence than the second plant phase (Fig. 4). This is the case for all ecotypes identified as the first major pattern in occurrence probabilities (Fig. 4). Even within those ecotypes, the slightly different occurrence probability estimates for digitgrade herbivores is reflected in the less extreme increase in log-odds of occurrence for the second plant phase (Fig. 11). In contrast, the almost universal pattern of the effect of plant phase on ecotype origination is that during first and last plant phases ecotypes have a greater log-odds of occurrence or origination than the second plant phase (Fig. 4, 5). The three ecotypes that do not

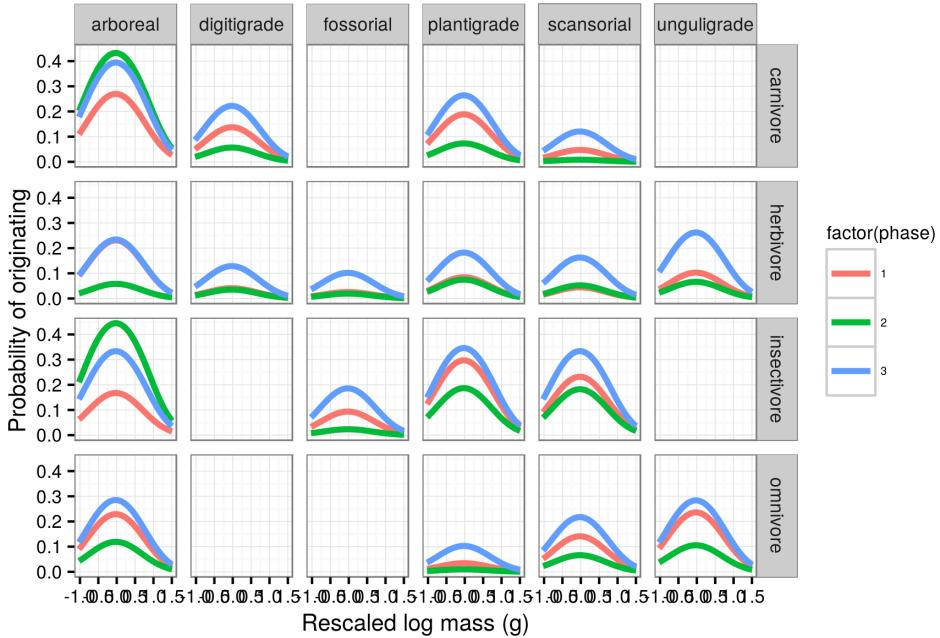


Figure 9: Mean estimate of the effect of species mass on the probability of a species originating for each of the three plant phases. The effect of mass is considered constant over time and that the only aspect of the model that changes with plant phase is the intercept of the relationship between mass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

follow this pattern are fossorial herbivores, scansorial herbivores, and arboreal insectivores.

- 566 Both aspects of global temperature analyzed here are estimated to have strong effects on species
 567 occurrence and origination for most mammal ecotypes (Tables 12, 13). Similarly, the probability
 568 that temperature has a large effect on species extinction is very low for all ecotypes (Table 14). The
 569 effects of the temperature covariates on ecotype occurrence and origination tend to be negative,
 570 which means that as temperature decreases, occurrence or origination are expected to
 571 increase. However, there is a positive relationship between mean temperature and the occurrence
 572 probabilities for about six ecotypes: arboreal carnivores, plantigrade herbivores, plantigrade
 573 insectivores, scansorial insectivores, arboreal omnivores, and unguiligrade omnivores (Table 12). A
 574 positive relationship is interpreted to mean that as temperature increases, the occurrence of those
 575 ecotypes is expected to increase. In the case of survival, the only strong ecotype association for

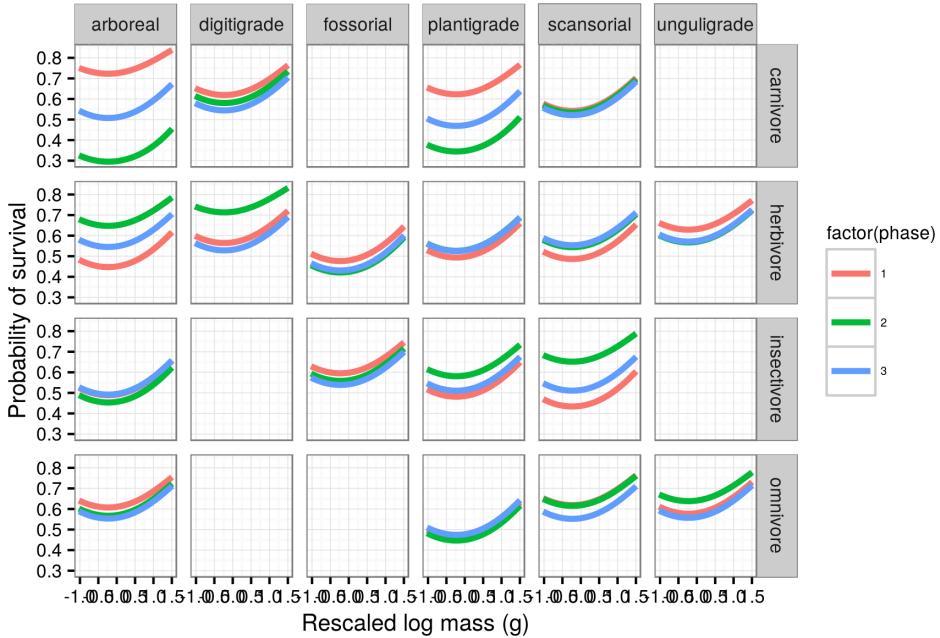


Figure 10: Mean estimate of the effect of species mass on the probability of a species survival for each of the three plant phases. The effect of mass is considered constant over time and that the only aspect of the model that changes with plant phase is the intercept of the relationship between mass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

576 either of the temperature covariates is a positive relationship between temperature range and
 occurrence probabilities of with plantigrade herbivores, and to a less certain extent plantigrade
 578 herbivores (Tab. 14).

Analysis of diversity

580 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).
 582 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
 584 wildly and rapidly over the Cenozoic, and demonstrates no sustained directional trends (Fig. 14a).

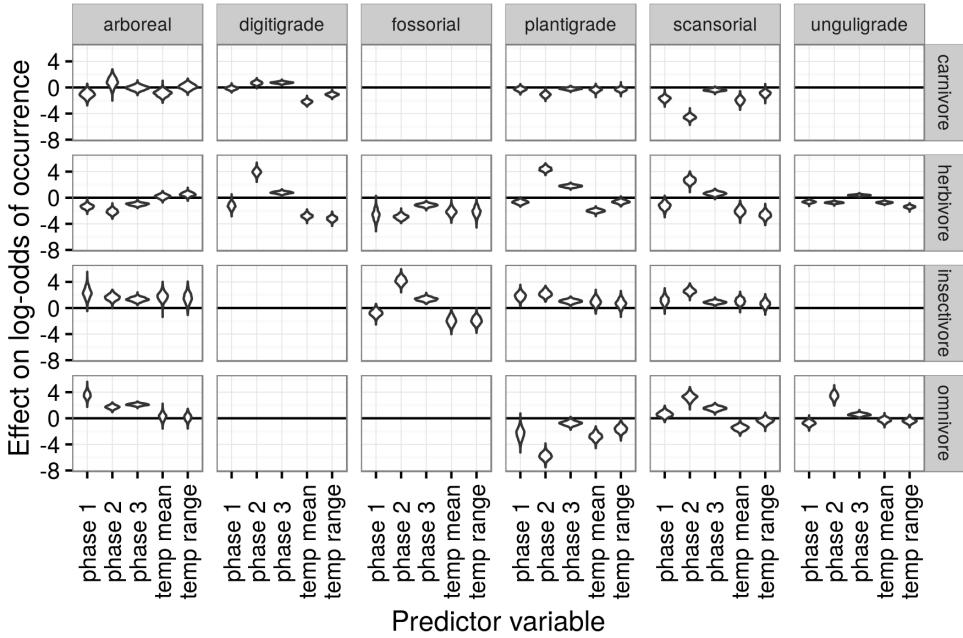


Figure 11: Estimated effects of the group-level covariates describing environmental context on log-odds of species occurrence. These estimates are from the pure-presence model. What is plotted is a violin of the distribution of 1000 samples from the approximate posterior. The effect of plant phase graphed here is calculated as Phase 1 = $\gamma_{phase\ 1}$, Phase 2 = $\gamma_{phase\ 1} + \gamma_{phase\ 2}$, and so on.

In broad strokes, the first 15 or so million years of the Cenozoic are characterized by first an

586 increase and then a decline in standing diversity at approximately 45-50 Mya (early-middle Eocene).
Following this decline, standing diversity is broadly constant from 45 to 18 Mya (early Miocene).
588 After this, there is a rapid spike in diversity followed by a slight decline in diversity up to the
Recent.

590 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
592 increase in diversity most of the major increases in diversity are retained including the large
diversity spike during the Miocene. Unlike subsampling based approaches to estimating diversity
594 (Alroy, 2010), I'm able to interpolate over unsampled/poorly sampled time periods because of how
the hierarchical model can share information across the different units Gelman et al. (2013); for
596 cases like unsampled temporal bins, this may lead to estimates with high uncertainty, but that is
preferable to no estimate at all. Finally, the Bayesian framework here gives a distribution of

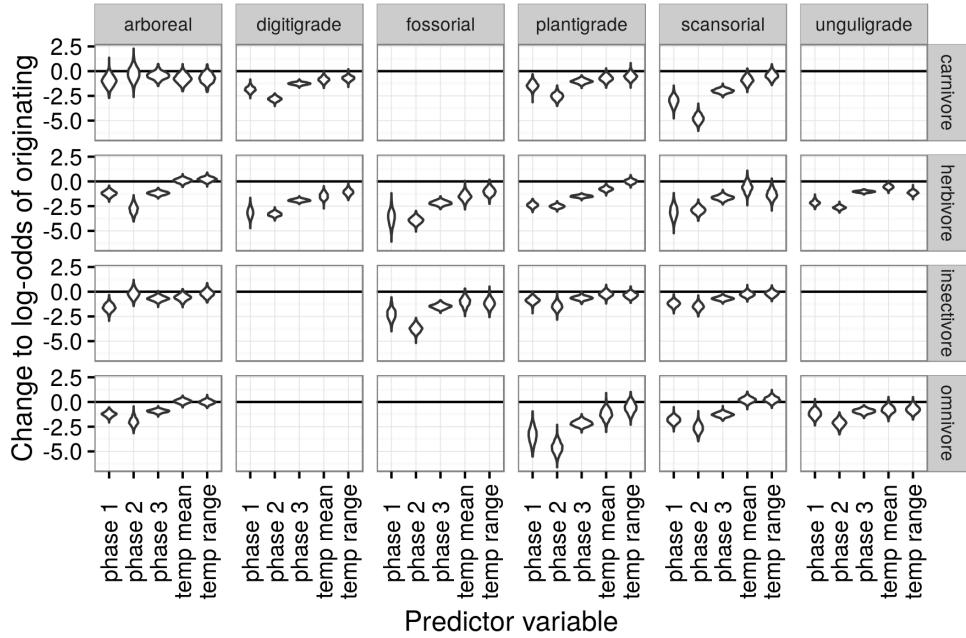


Figure 12: Estimated effects of the group-level covariates describing environmental context on log-odds of species origination. These estimates are from the birth-death model. What is plotted is a violin of the distribution of 1000 samples from the approximate posterior. The effect of plant phase graphed here is calculated as Phase 1 = $\gamma_{\text{phase 1}}$, Phase 2 = $\gamma_{\text{phase 1}} + \gamma_{\text{phase 2}}$, and so on.

possible estimates of diversity allowing for direct inspection of the uncertainty of our inferences, something that is preferable to both traditional and resampling based confidence interval estimates (Gelman et al., 2013). Note that my time series of estimated diversity begins at a slightly different point than that of Alroy (2009) and that the time intervals used by Alroy (2009) are slightly shorter than those used here, so this may cause 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).

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 hovers around zero, punctuated by both positive and negative spikes. The largest spike in diversification rate is at 16 Mya, which is early Oligocene (Fig. 14b). Other notable increases in diversification rate occur 56, 46, 22, 18, 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

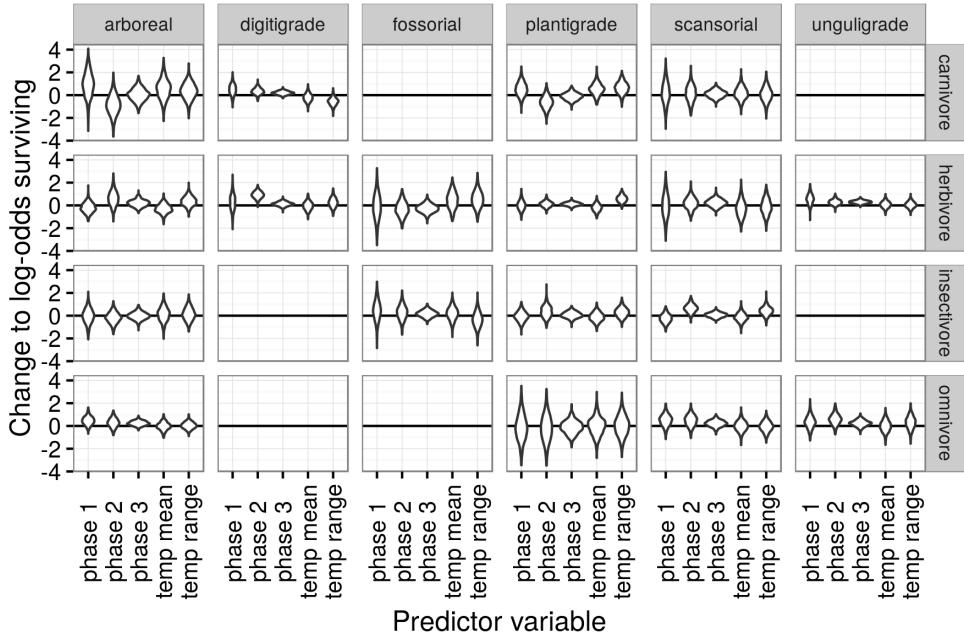


Figure 13: Estimated effects of the group-level covariates describing environmental context on log-odds of species survival. These estimates are from the birth-death model. What is plotted is a violin of the distribution of 1000 samples from the approximate posterior. The effect of plant phase graphed here is calculated as Phase 1 = $\gamma_{phase\ 1}$, Phase 2 = $\gamma_{phase\ 1} + \gamma_{phase\ 2}$, and so on.

rate occur at 54, 50, 48, 44, 40, 34, 30, 24, 20, 16, 12, and 8 Mya (Table 15), meaning that

612 diversification rate has more major decreases than increases. While diversification rates significantly
 lower than average are more common than diversification rates greater than average, when
 614 diversification rate does increase it is with a greater magnitude than most decreases (Fig. 14b).
 Given that diversification rate more closely resembles origination rate than extinction rate (Fig.
 616 14b, 14c, 14d), these decreases in diversification rate may be indicative of “depletions” (failure to
 replace extinct taxa) rather than pulses of extinction.

618 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
 620 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.
 622 Extinction rate, on the other hand, demonstrates an almost saw-toothed pattern around a constant
 mean (Fig. 14d). These results are broadly consistent with those from previous analyses of North

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

624 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
626 diversity for the Cenozoic (Fig. 15).

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

Arboreal herbivores and omnivores obtain peak diversity at the beginning of the Cenozoic then go
630 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
632 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).

634 The diversity of digitigrade and unguligrade herbivores increases over the Cenozoic (Fig. 15). In
contrast, plantigrade herbivore diversity does not have a single, broad-strokes pattern; instead,

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

- 636 diversity increases, decreases, and may have then increased till the Recent. In contrast, fossorial and
 637 scansorial herbivores demonstrate a much flatter history of diversity, with a slight increase in
 638 diversity that over time is more pronounced among fossorial taxa than scansorial taxa. The
 639 expansion of digitigrade and unguligrade herbivores over the Cenozoic is consistent with the
 640 gradual expansion of grasslands which these ecotypes are better adapted to than closed
 641 environments (Blois and Hadly, 2009; Strömberg, 2005).
- 642 Digitigrade carnivores have a multi-modal diversity history, with peaks at 54-52 and 12-10 Mya
 643 (Fig.15). Between these two peaks digitigrade carnivore diversity dips below average diversity
 644 following the first peak and then grows slowly until the second peak. Plantigrade carnivores obtain
 645 peak diversity in the early Cenozoic and then maintain a relatively stable diversity until another
 646 peak at the end of the Cenozoic. The generally flat diversity history digitigrade carnivores lacks any
 647 sustained temporal trends and seems to reflect previous findings of limited diversity in spite of

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

- 648 constant turnover and morphological evolution (Silvestro et al., 2015; Slater, 2015; Van
 Valkenburgh, 1999)
- 650 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
 652 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
 654 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
 656 demonstrate a possible decrease in diversity over the Cenozoic, but not as clearly as arboreal and
 scansorial omnivores.
- 658 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
 660 in digitigrade carnivores, plantigrade carnivores, scansorial carnivores, unguiligrade herbivores,

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

fossorial insectivores, and plantigrade omnivores.

- 662 When ecotype diversity is decomposed into per capita origination (Fig. 16) and per capita extinction
 rates (Fig. 17) the way in which their diversity developed can be exemplified. For ecotype-specific
 664 origination and extinction rates, the number of origination or extinction events for each ecotype was
 calculated and that number was divided by the total standing diversity of all mammals at the time.
 666 As should be expected, origination rates have a much greater range of values with a few very large
 spikes that line up with the spikes in over all diversification rate (Fig. 14b). Importantly, the source
 668 of the massive increase in diversification rate at 16 Mya can be attributed almost solely to the
 origination of unguiligrade herbivores (Fig. 16). Additionally, by decomposing origination rate by
 670 ecotype, it is possible to identify a few possible cross-ecotype increases in origination rate. For
 example, digitigrade carnivores, digitigrade herbivores, and plantigrade herbivores share a lot of
 672 increases in origination rate with unguiligrade herbivores; these are all ecotypes that demonstrate an

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

obvious increase in diversity during the Paleogene and then maintain relatively high diversity

674 through out the Neogene (Fig. 15).

In contrast to ecotype-specific per capita origination rates which demonstrate distinct peaks, the

676 estimates of ecotype-specific per capita are more of a smear (Fig. 17). There are few increases in extinction rate that are shared across ecotypes. The per capita extinction rates of digitigrade,

678 plantigrade, and unguligrade herbivores are lower in Paleogene than the Neogene. This result is interpreted to mean that as the diversity of these three ecotypes was increasing, the number of

680 extinction events was also increasing. Also, the per capita extinction rate of arboreal taxa is higher in the Paleogene than the Neogene. While this result may seem odd considering the observed

682 diversity pattern for these ecotypes (Fig. 15), I argue that this result is actually extremely intuitive: if there are no species of that ecotype originating or present, than there can be extinctions. This

684 result highlights the distinction between extinction risk and extinction rate; an ecotype can have a

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

high extinction risk, but if that ecotype is not present in the species pool in the first place than it

686 has no associated extinction rate.

The apparent similarities in origination rate of digitgrade carnivores, digitigrade herbivores,

688 plantigrade herbivores, and unculigrade herbivores (Fig. 5, 16) can be tested by inspecting the

estimates of the two correlation matrices Ω^ϕ and Ω^π . The elements of these matrices are estimates

690 of the correlation in origination and survival probabilities, respectively, between each of the 18

observed ecotypes. However, because ADVI-based inference assumes that the joint posterior is

692 Gaussian, estimates of scale and correlation parameters are very approximate as these parameters

tend to have decidedly non-Gaussian true posterior distributions (Gelman et al., 2013). Because of

694 this fundamental limitation, inference based on these correlation matrices are very approximate and

subject to change.

696 Consistent with visual inspection of the ecotype origination probability time series, there are some

strong positive correlations between a few ecotypes (Fig. 18). Given the posterior distribution of
698 correlation estimates, the probability of these correlation estimates being greater than 0 can be
estimated. Again, because of the assumed Gaussian posterior, these probability estimates are at
700 best approximate and are subject to change given full posterior inference. To visualize these results,
I've plotted an association graph of the correlations between ecotypes that are estimated to have a
702 greater than 95% posterior probability of being greater than 0 (Fig. 19); in total there are 35
correlations that fit this criterion. The ecotypes correlated with the most number (9) of other
704 ecotypes are digitigrade carnivores, plantigrade herbivores, and unguligrade herbivores. Scansorial
carnivores, digitigrade herbivores, fossorial herbivores are correlated with seven other ecotypes,
706 each. These results support the conclusion that the origination probabilities of many ecotypes are
correlated which I interpret to mean that changes to the species pool's environment context can
708 lead to the simultaneous enrichment of multiple ecotypes instead of each ecotype respondingly
independently. This result is most obvious for digitigrade carnivores, scansorial carnivores,
710 digitigrade herbivores, fossorial herbivores, plantigrade herbivores, and unguligrade herbivores;
these ecotypes are heavily cross-correlated with each other.

712 In contrast to the visually obvious correlations in ecotype origination probability, visual inspection
of the ecotype-specific survival probabilities (Fig. 6) and per capita extinction rates (Fig. 17) does
714 not indicate that many strong correlations between ecotype survival probabilities. This conclusion
is supported by the estimated correlations in ecotype survival probability (Fig. 20). There are very
716 few large magnitude estimates of correlation between any of the ecotypes. This result is further
supported by the fact that only a single correlation, survival probability of digitigrade carnivores
718 and unguligrade herbivores, has a greater than 95% posterior probability of being positive (Fig. 21).
This single correlation, however, adds more nuance to the interpretations from the origination
720 probability correlations. In addition to correlation in enrichment, these ecotypes are correlated in
their depletions. This result supports the conclusion that the diversity history of digitigrade
722 carnivores and unguligrade herbivores are strongly related to each other both in terms of origination
and survival, which stands in contrast to those ecotypes which are only correlated for origination.

⁷²⁴ **Discussion**

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 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 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 ecotype or guild can be stable over millions of years despite constant species turnover (Jernvall and Fortelius, 2004; Slater, 2015; Van Valkenburgh, 1999). This implies that the size and scope of the role of an ecotype or guild in local communities, and the regional species pool as a whole, is preserved even as the individual interactors change. This also implies that the structure of regional species pools can be constant over time despite a constantly changing set of “players.” This result supports the hypothesis that ecotypes or guilds are at least partially self-organizing and truly emergent (Scheffer and van Nes, 2006).

⁷³⁸ Comparison of the results from the posterior predictive checks for the pure-presence and birth-death models supports the conclusion that regional species pool dynamics cannot simply be described by a single occurrence probability and are instead the result of the interplay between origination and extinction. Additionally, changes to the ecotypic composition and diversification rate for the North American regional species pool are driven primarily by variation in origination rates. These aspects of how regional species pool diversity is shaped are not directly observable in studies of the Recent where time scales are short and macroevolutionary dynamics are inferable solely from phylogeny (Fritz et al., 2013).

⁷⁴⁶ Extinction rate for the entire regional species pool through time is highly variable and demonstrates a saw-toothed pattern around an apparently constant mean. While a constant mean extinction rate is consistent with previous observation (Alroy, 1996; Alroy et al., 2000), the degree to which extinction rate is actually variable may not have been equally appreciated. What is most consistent with previous observations, however, is that diversity seems to be most structured by changes to

origination rather than changes to extinction (Alroy, 1996; Alroy et al., 2000).

- 752 Comparison of the ecotype specific diversity histories supports the conclusion that there were no
major, simultaneous changes in diversity between the functional groups of the regional species pool;
754 instead these results support a more gradual and idiosyncratic shifts in relative ecotypic diversity
over time (Fig. 15). The closest examples to a sudden increase or decrease of a specific ecotype is
756 the jump in standing diversity of scansorial carnivores and, to a lesser extent, fossorial insectivores
at 16 Mya (i.e. the start of the third plant phase). However, this result may not reflect the
758 dynamics of individual local communities, as this is an analysis of the entire North American
mammal regional species pool.
- 760 Arboreal taxa disappear from the regional species pool over the Cenozoic, with long term decline
over the Paleogene leading to the disappearance by start of Neogene ~22 Mya. This is consistent
762 with one of the two possible patterns presented here and in Smits (2015) that would result in
arboreal taxa having a greater extinction risk than other ecotypes: the Paleogene and Neogene were
764 different selective regimes and, while the earliest Cenozoic may have been neutral with respect to
arboreal taxa, they disappeared quickly over the Cenozoic which may account for their higher
766 extinction risk. In addition to all arboreal taxa, the diversity of plantigrade and scansorial
insectivores decreases with time (Fig. 15).
- 768 Digitigrade carnivores have a relatively stable diversity history through the Cenozoic and can be
characterized as varying around a constant mean diversity. This ecotype has a large amount of
770 overlap with the carnivore guild which has been the focus of much research (Janis and Wilhelm,
1993; Pires et al., 2015; Slater, 2015; Van Valkenburgh, 1999). This result is consistent with some
772 form of “control” on the diversity of this ecotype, such as diversity-dependent diversification
(Silvestro et al., 2015; Slater, 2015; Van Valkenburgh, 1999).
- 774 Both digitigrade and unguligrade herbivores increase in diversity over the Cenozoic. The increase of
these cursorial forms is consistent with the gradual opening up of the North American landscape
776 (Blois and Hadly, 2009; Graham, 2011; Strömberg, 2005) and may indicate a long-term shift in the
interactors associated with those ecotypes leading to increased contribution to the regional species

778 pool. This result may be comparable to the increasing percentage of hypsodont (high-crowned
teeth) mammals in the Neogene of Europe being due to an enrichment of hypsodont taxa and not
780 a depletion of non-hypsodont taxa. Smaller scale increases in fossorial herbivore species, and a lesser
extent plantigrade herbivores, suggests that the increase of interactors may be associated mostly
782 with the herbivore dietary category with locomotor category tempering that relationship. These
results support the conclusion that the increase in digitigrade and unguligrade herbivores is the
784 result of an enrichment of these ecotypes as opposed to being caused by the depletion of other
herbivorous ecotypes; this is further supported by the lack of major changes to the number of
786 extinctions of all herbivore ecotypes (Fig. 17).

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

796 Interestingly, for all of the ecotypes with sudden changes in diversity at this transition the change is
an increase, even if only temporarily. There are two interpretations of these results. A biological
798 interpretation of this result is that, because plant phase associations are only with occurrence or
origination probabilities and not survival, these ecotypes were well suited for the newly available
800 mammal-plant interactions due to the increased modernization of their floral context (Graham,
2011). Alternatively, the increase in diversity associated with the third plant phase may be caused
802 by the edge effect in origination probability that is artificially inflating the number of origination
events (Fig. 5). However, the estimated number of origination events does not have a tremendous
804 spike at this transition, nor is a major increase in the number of origination events sustained (Fig.
16).

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

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

824 The result that temperature does not affect the extinction probability of most ecotypes is consistent
825 with previous analysis of mammal diversity (Alroy et al., 2000). The result that temperature affects
826 origination probability, on the other hand, is in strong contrast to the results Alroy et al. (2000).
827 An important difference between the analyses presented here and that of Alroy et al. (2000) is I am
828 considering the effect of temperature on the probability of a species originating, assuming it hasn't
829 originated yet while Alroy et al. (2000) analyzes the correlation between the first differences of the
830 origination and extinction rates with an oxygen isotope curve (Zachos et al., 2001). Origination or
831 extinction rates have very different properties than the origination probabilities estimated here
832 brought upon by the difference both in definition and units. Origination probability is the expected
833 probability that a species that has never been present and is not present at time t will be present at

834 time $t + 1$; origination probability is defined for a single species. In contrast, per capita rates are
defined (for origination) as the expected number of new species to have originated between time t
836 and $t + 1$ given the total number of species present at time t ; per capita rates are defined for the
standing diversity. It is also important to note that even though the occurrence and origination
838 probabilities increase over time because of the increasingly deterministic occurrence of those species
which had not yet originated (Fig. 4, 5, the corresponding rates and population level birth/death
840 dynamics do not share that pattern (Fig. 14, 16, 17). In effect, the finding that temperature has an
effect on origination simply because as time approaches the present the number of species which
842 have originated increases and not because of climatic forcing of origination.

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

The effects of plant phase and temperature on ecotypes are approximately equal in importance. The
852 focus in previous research on temperature and major climatic or geological events without other
measures of environmental context may have led to confusion in discussions of how the
854 “environment” affects mammal diversity and diversification. The environment or climate is not just
global or regional temperature, it is the set of all possible biotic and abiotic interactions that can be
856 experienced by a member of the species pool. By including more descriptors of species’
environmental context a more complete “picture” of the diversification process is inferred.

858 The effect of species mass on either occurrence or origination and extinction was not allowed to
vary by ecotype or environmental context. The primary reason for this modeling choice was that
860 this study focuses on ecotypic based differences in either occurrence, or origination and extinction.

Allowing the effect of body size to vary by ecotype, time, and environmental factors would increase
862 the overall complexity of the model beyond the scope of the study. Instead, body size was included
in order to control for its possible underlying effects (McElreath, 2016). A control means that if
864 there is variation due to body mass, having a term to “absorb” that effect is better than ignoring it.

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

The potential effects of common ancestry (i.e. phylogeny) on origination and extinction are not
872 directly considered in this analysis. While a birth-death process approximates the
speciation-extinction process underlying the phylogeny (Silvestro et al., 2014) this is not same as
874 considering how the similarity between closely related species may affect the estimates of the effects
of species traits to environmental factors on both origination and extinction (Harnik et al., 2014;
876 Smits, 2015). The inclusion of phylogeny can help disentangle if the functional composition of
species diversity is shaped either by many closely related species occurring at the same time or if
878 closely related species are more evenly distributed in time; this is analogous to if species within
local communities are clumped or dispersed relative to their relatedness (Cavender-Bares et al.,
880 2009; Kraft et al., 2007; Webb et al., 2002). One of the principal barriers to the inclusion of the
effect of phylogeny in either the pure-presence or birth-death models is computational; with well
882 over 1000 tips, the calculation of the scale parameter defining the phylogenetic effect would be very
slow and further increase the already slow computation time necessary for the marginalization of all
884 possible discrete occurrence histories for z .

Conclusions

These results support the conclusion that the ecotypic diversity of the North American mammal species pool has changed gradually over time. While there is constant species turnover for the entire Cenozoic, there is little evidence of major cross-ecotype upheaval and sudden reorganization of the functional composition of the regional species pool. The results of this study also support the conclusion that mammal diversification over the Cenozoic is driven primarily by changes to origination rate and not extinction rate. There are a number of interesting estimated ecotype diversity patterns. While arboreal ecotypes are diverse in the Paleogene, by the Neogene all 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 plantigrade and scansorial insectivores and scansorial omnivores. The only ecotypes that demonstrate a sustained pattern of increasing diversity are digitigrade and unguligrade herbivores. 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 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 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 information as possible into their analyses in order to better understand or at least contextualize the complex processes underlying that diversity.

Acknowledgements

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

910 Paleobiology Database publication XXX.

References

- 912 Allen, L. J. S. 2011. An introduction to stochastic processes with applications to biology. 2nd ed. Chapman and Hall/CRC, Boca Raton, FL.
- 914 Alroy, J. 1996. Constant extinction, constrained diversification, and uncoordinated stasis in North American mammals. *Palaeogeography, Palaeoclimatology, Palaeoecology* 127:285–311.
- 916 ———. 2009. Speciation and extinction in the fossil record of North American mammals. Pages 302–323 *in* R. K. Butlin, J. R. Bridle, and D. Schluter, eds. *Speciation and patterns of diversity*.
918 Cambridge University Press, Cambridge.
- . 2010. Fair sampling of taxonomic richness and unbiased estimation of origination and extinction rates. Pages 55–80 *in* J. Alroy and G. Hunt, eds. *Quantitative Methods in Paleobiology*. The Paleontological Society.
- 922 Alroy, J., P. L. Koch, and J. C. Zachos. 2000. Global climate change and North American mammalian evolution. *Paleobiology* 26:259–288.
- 924 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*
926 39:373–399.
- Badgley, C., T. M. Smiley, R. Terry, E. B. Davis, L. R. G. Desantis, D. L. Fox, S. S. B. Hopkins,
928 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:
930 Modern and Geohistorical Perspectives. *Trends in Ecology & Evolution* pages 1–16.
- Bambach, R. K. 1977. Species richness in marine benthic habitats through the Phanerozoic.
932 *Paleobiology* 3:152–167.

- Bambach, R. K., A. M. Bush, and D. H. Erwin. 2007. Autecology and the filling of ecospace: Key
934 metazoan radiations. *Palaeontology* 50:1–22.
- Bloch, J. I., K. D. Rose, and P. D. Gingerich. 1998. New species of Batodonoides (Lipotyphla,
936 Geolabididae) from the Early Eocene of Wyoming: smallest known mammal? *Journal of
Mammalogy* 79:804–827.
- Blois, J. L., and E. A. Hadly. 2009. Mammalian Response to Cenozoic Climatic Change. *Annual
Review of Earth and Planetary Sciences* 37:181–208.
- 940 Brook, B. W., and D. M. J. S. Bowman. 2004. The uncertain blitzkrieg of Pleistocene megafauna.
Journal of Biogeography 31:517–523.
- 942 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
944 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
946 on continents. *Science* 243:1145–1150.
- Bush, A. M., and R. K. Bambach. 2011. Paleoecologic Megatrends in Marine Metazoa. *Annual
Review of Earth and Planetary Sciences* 39:241–269.
- 948 Bush, A. M., R. K. Bambach, and G. M. Daley. 2007. Changes in theoretical ecospace utilization in
marine fossil assemblages between the mid-Paleozoic and late Cenozoic. *Paleobiology* 33:76–97.
- Bush, A. M., and P. M. Novack-Gottshall. 2012. Modelling the ecological-functional diversification
952 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
954 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
956 locomotor habit in mammals and dinosaurs. *Journal of Zoology* 247:29–42.

- 958 Cavender-Bares, J., K. H. Kozak, P. V. a. Fine, and S. W. Kembel. 2009. The merging of
community ecology and phylogenetic biology. *Ecology letters* 12:693–715.
- 960 Clyde, W. C., and P. D. Gingerich. 1998. Mammalian community response to the latest Paleocene
thermal maximum: an isotaphonomic study in the northern Bighorn Basin, Wyoming. *Geology*
26:1011–1014.
- 962 Cohen, K. M., S. C. Finney, P. L. Gibbard, and J.-X. Fan. 2015. The ICS International
Chronostratigraphic Chart.
- 964 Cottennie, K. 2005. Integrating environmental and spatial processes in ecological community
dynamics. *Ecology Letters* 8:1175–1182.
- 966 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
968 geochemistry ($\delta^{18}\text{O}$ and Mg/Ca) with sea level history. *Journal of Geophysical Research: Oceans*
116:1–23.
- 970 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.
- 972 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*
974 40:677–697.
- 976 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.
- 978 Eronen, J. T., P. D. Polly, M. FRED, J. Damuth, D. C. FRANK, V. Mosbrugger,
C. SCHEIDEGGER, N. C. Stenseth, and M. Fortelius. 2010. Ecometrics: The traits that bind
980 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

- 982 species richness and reveal the mode of macroevolutionary competition. *Ecology Letters*
19:899–906.
- 984 Felsenstein, J. 1985. Phylogenies and the comparative method. *The American Naturalist* 125:1–15.
- 986 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.
- 988 Foote, M. 2001. Inferring temporal patterns of preservation, origination, and extinction from
taxonomic survivorship analysis. *Paleobiology* 27:602–630.
- 990 Foote, M., and J. J. Sepkoski. 1999. Absolute measures of the completeness of the fossil record.
Nature 398:415–7.
- 992 Foster, J. R. 2009. Preliminary body mass estimates for mammalian genera of the Morrison
Formation (Upper Jurassic, North America). *PaleoBios* 28:114–122.
- 994 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.
- 998 Freudenthal, M., and E. Martín-Suárez. 2013. Estimating body mass of fossil rodents. *Scripta
Geologica* 145:1–130.
- 1000 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.
- 1002 Gelman, A. 2008. Scaling regression inputs by dividing by two standard deviations. *Statistics in
Medicine* pages 2865–2873.
- 1004 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.
- 1006 Gelman, A., and J. Hill. 2007. Data Analysis using Regression and Multilevel/Hierarchical Models.
Cambridge University Press, New York, NY.

- Gordon, C. L. 2003. A First Look at Estimating Body Size in Dentally Conservative Marsupials.
1008 Journal of Mammalian Evolution page 21.
- Graham, A. 2011. A natural history of the New World: the ecology and evolution of plants in the
1010 Americas. University of Chicago Press, Chicago.
- Harmon, L. J., and S. Harrison. 2015. Species Diversity Is Dynamic and Unbounded at Local and
1012 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
1014 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
1016 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
1018 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
1020 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
1024 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.
1026 Gordon and H. H. T. Prins, eds. *The Ecology of Browsing and Grazing*. Springer-Verlag.
- Janis, C. M., J. Damuth, and J. M. Theodor. 2000. Miocene ungulates and terrestrial primary
1028 productivity: where have all the browsers gone? *Proceedings of the National Academy of Sciences*
97:7899–904.
- Janis, C. M., G. F. Gunnell, and M. D. Uhen. 2008. Evolution of Tertiary mammals of North
1030 America. *Annual Review of Ecology and Systematics* 39:191–229.

- America. Vol. 2. Small mammals, xenarthrans, and marine mammals. Cambridge University
1032 Press, Cambridge.
- Janis, C. M., K. M. Scott, and L. L. Jacobs. 1998. Evolution of Tertiary mammals of North
1034 America. Vol. 1. Terrestrial carnivores, ungulates, and ungulatelike mammals. Cambridge
University Press, Cambridge.
- 1036 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.
- 1038 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,
1040 Palaeoclimatology, Palaeoecology* 365–366:1–10.
- Jernvall, J., and M. Fortelius. 2002. Common mammals drive the evolutionary increase of
1042 hypodonty in the Neogene. *Nature* 417:538–40.
- . 2004. Maintenance of trophic structure in fossil mammal communities: site occupancy and
1044 taxon resilience. *The American Naturalist* 164:614–624.
- Kraft, N. J. B., W. K. Cornwell, C. O. Webb, and D. D. Ackerly. 2007. Trait evolution, community
1046 assembly, and the phylogenetic structure of ecological communities. *The American Naturalist*
170:271–283.
- 1048 Kucukelbir, A., R. Ranganath, A. Gelman, and D. M. Blei. 2015. Automatic Variational Inference
in Stan. Pages 568–576 in *NIPS*. Vol. 28.
- 1050 Legendre, S. 1986. Analysis of mammalian communities from the Late Eocene and Oligocene of
Southern France. *Paleovertebrata* 16:191–212.
- 1052 Liow, L. H., M. Fortelius, E. Bingham, K. Lintulaakso, H. Mannila, L. Flynn, and N. C. Stenseth.
2008. Higher origination and extinction rates in larger mammals. *Proceedings of the National
1054 Academy of Sciences* 105:6097–6102.

- Lloyd, G. T., J. R. Young, and A. B. Smith. 2011. Taxonomic Structure of the Fossil Record is
1056 Shaped by Sampling Bias. *Systematic Biology* 61:80–89.
- Loeuille, N., and M. a. Leibold. 2008. Evolution in metacommunities: on the relative importance of
1058 species sorting and monopolization in structuring communities. *The American naturalist*
171:788–99.
- 1060 Losos, J. B. 2010. Adaptive radiation, ecological opportunity, and evolutionary determinism. *The American naturalist* 175:623–39.
- 1062 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,
1064 and J. S. Levinton, eds. *Evolution since Darwin: the first 150 years*. Sinauer Associates,
Sunderland, MA.
- 1066 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.
- 1068 Marcot, J. D. 2014. The fossil record and macroevolutionary history of North American ungulate
mammals: standardizing variation in intensity and geography of sampling. *Paleobiology*
1070 40:237–254.
- McElreath, R. 2016. *Statistical rethinking: a Bayesian course with examples in R and Stan*. CRC
1072 Press, Boca Raton, FL.
- McGill, B. J., B. J. Enquist, E. Weiher, and M. Westoby. 2006. Rebuilding community ecology
1074 from functional traits. *TRENDS in Ecology and Evolution* 21:178–185.
- McKenna, R. T. 2011. Potential for Speciation in Mammals Following Vast , Late Miocene Volcanic
1076 Interruptions in the Pacific Northwest. Masters. Portland State University.
- Mendoza, M., C. M. Janis, and P. Palmqvist. 2006. Estimating the body mass of extinct ungulates:
1078 a study on the use of multiple regression. *Journal of Zoology* 270:90–101.

- Mittelbach, G. G., and D. W. Schemske. 2015. Ecological and evolutionary perspectives on
1080 community assembly. *Trends in Ecology and Evolution* 30:241–247.
- Novack-Gottshall, P. M. 2007. Using a theoretical ecospace to quantify the ecological diversity of
1082 Paleozoic and modern marine biotas Using a theoretical ecospace to quantify the ecological
diversity of Paleozoic and modern marine biotas. *Paleobiology* 33:273–294.
- 1084 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*
1086 282:20151952.
- Pollock, L. J., W. K. Morris, and P. A. Vesk. 2012. The role of functional traits in species
1088 distributions revealed through a hierarchical model. *Ecography* 35:716–725.
- Polly, P., J. Eronen, M. Fred, G. P. Dietl, V. Mosbrugger, C. Scheidegger, D. C. Frank, J. Damuth,
1090 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.
- 1092 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
1094 the Linnean Society* pages n/a–n/a.
- Quental, T. B., and C. R. Marshall. 2013. How the Red Queen Drives Terrestrial Mammals to
1096 Extinction. *Science* 341:290–292.
- Rabosky, D. L. 2013. Diversity-Dependence, Ecological Speciation, and the Role of Competition in
1098 Macroevolution. *Annual Review of Ecology, Evolution, and Systematics* 44:1–22.
- Rabosky, D. L., and A. H. Hurlbert. 2015. Species Richness at Continental Scales Is Dominated by
1100 Ecological Limits. *The American Naturalist* 185:000–000.
- Raia, P., F. Carotenuto, F. Passaro, D. Fulgione, and M. Fortelius. 2012. Ecological specialization
1102 in fossil mammals explains Cope's rule. *The American Naturalist* 179:328–37.

- Royle, J. A., and R. M. Dorazio. 2008. Hierarchical modeling and inference in ecology: the analysis
1104 of data from populations, metapopulations and communities. Elsevier, London.
- Royle, J. A., J. D. Nichols, and M. Kéry. 2005. Modelling occurrence and abundance of species
1106 when detection is imperfect. *Oikos* 110:353–359.
- Scheffer, M., and E. H. van Nes. 2006. Self-organized similarity, the evolutionary emergence of
1108 groups of similar species. *Proceedings of the National Academy of Sciences* 103:6230–6235.
- Shipley, B., D. Vile, and E. Garnier. 2006. From plant traits to plant communities: a statistical
1110 mechanistic approach to biodiversity. *Science* 314:812–814.
- Silvestro, D., A. Antonelli, N. Salamin, and T. B. Quental. 2015. The role of clade competition in
1112 the diversification of North American canids. *Proceedings of the National Academy of Sciences of*
the United States of America 112:8684–9.
- 1114 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.
- 1116 Simberloff, D., and T. Dayan. 1991. The Guild Concept and the Structure of Ecological
Communities. *Annual Review of Ecology and Systematics* 22:115–143.
- 1118 Slater, G. J. 2015. Iterative adaptive radiations of fossil canids show no evidence for
diversity-dependent trait evolution. *Proceedings of the National Academy of Sciences*
1120 112:4897–4902.
- Smith, F. A., J. Brown, J. Haskell, and S. Lyons. 2004. Similarity of mammalian body size across
1122 the taxonomic hierarchy and across space and time. *The American Naturalist* 163:672–691.
- Smith, F. A., S. K. Lyons, S. Morgan Ernest, and J. H. Brown. 2008. Macroecology: more than the
1124 division of food and space among species on continents. *Progress in Physical Geography*
32:115–138.
- 1126 Smits, P. D. 2015. Expected time-invariant effects of biological traits on mammal species duration.
Proceedings of the National Academy of Sciences 112:13015–13020.

- ¹¹²⁸ Stan Development Team. 2016. Stan Modeling Language Users Guide and Reference Manual.
- ¹¹³⁰ Strömberg, C. A. E. 2005. Decoupled taxonomic radiation and ecological expansion of open-habitat grasses in the Cenozoic of North America. *Proceedings of the National Academy of Sciences of the United States of America* 102:11980–4.
- ¹¹³² Tomiya, S. 2013. Body Size and Extinction Risk in Terrestrial Mammals Above the Species Level. *The American Naturalist* 182:196–214.
- ¹¹³⁴ Urban, M. C., M. A. Leibold, P. Amarasekare, L. De Meester, R. Gomulkiewicz, M. E. Hochberg, 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 Ecology and Evolution* 23:311–317.
- ¹¹³⁸ Valentine, J. W. 1969. Patterns of taxonomic and ecological structure of the shelf benthos during Phanerozoic time. *Paleontology* 12:684–709.
- ¹¹⁴⁰ Van Valkenburgh, B. 1990. Skeletal and dental predictors of body mass in carnivores. Pages 181–205 in J. Damuth and B. J. Macfadden, eds. *Body size in mammalian paleobiology: estimation and biological implications*. Cambridge University Press, Cambridge.
- ¹¹⁴² ———. 1999. Major patterns in the history of carnivorous mammals. *Annual Review of Earth and Planetary Sciences* 27:463–493.
- ¹¹⁴⁶ Villéger, S., P. M. Novack-Gottshall, and D. Mouillot. 2011. The multidimensionality of the niche reveals functional diversity changes in benthic marine biotas across geological time. *Ecology letters* 14:561–8.
- ¹¹⁴⁸ Wang, S. C., P. J. Everson, H. J. Zhou, D. Park, and D. J. Chudzicki. 2016. Adaptive credible intervals on stratigraphic ranges when recovery potential is unknown. *Paleobiology* 42:240–256.
- ¹¹⁵⁰ Wang, S. C., and C. R. Marshall. 2016. Estimating times of extinction in the fossil record. *Biology Letters* 12:20150989.

- 1152 Warton, D. I., B. Shipley, and T. Hastie. 2015. CATS regression - a model-based approach to
studying trait-based community assembly. *Methods in Ecology and Evolution* 6:389–398.
- 1154 Webb, C. O., D. D. Ackerly, M. a. McPeek, and M. J. Donoghue. 2002. Phylogenies and
Community Ecology. *Annual Review of Ecology and Systematics* 33:475–505.
- 1156 Weber, M. G., C. E. Wagner, R. J. Best, L. J. Harmon, and B. Matthews. 2017. Evolution in a
Community Context: On Integrating Ecological Interactions and Macroevolution. *Trends in
Ecology & Evolution* xx:1–14.
- 1158 Wilson, J. B. 1999. Guilds, functional types and ecological groups. *Oikos* 86:507–522.
- 1160 Yoder, J. B., E. Clancey, S. Des Riches, J. M. Eastman, L. Gentry, W. Godsoe, T. J. Hagey,
D. Jochimsen, B. P. Oswald, J. Robertson, B. A. J. Sarver, J. J. Schenk, S. F. Spear, and L. J.
1162 Harmon. 2010. Ecological opportunity and the origin of adaptive radiations. *Journal of
Evolutionary Biology* 23:1581–1596.
- 1164 Zachos, J. C., G. R. Dickens, and R. E. Zeebe. 2008. An early Cenozoic perspective on greenhouse
warming and carbon-cycle dynamics. *Nature* 451:279–283.
- 1166 Zachos, J. C., M. Pagani, L. Sloan, E. Thomas, and K. Billups. 2001. Trends, rhythms, and
aberrations in global climate 65 Ma to present. *Science* 292:686–693.

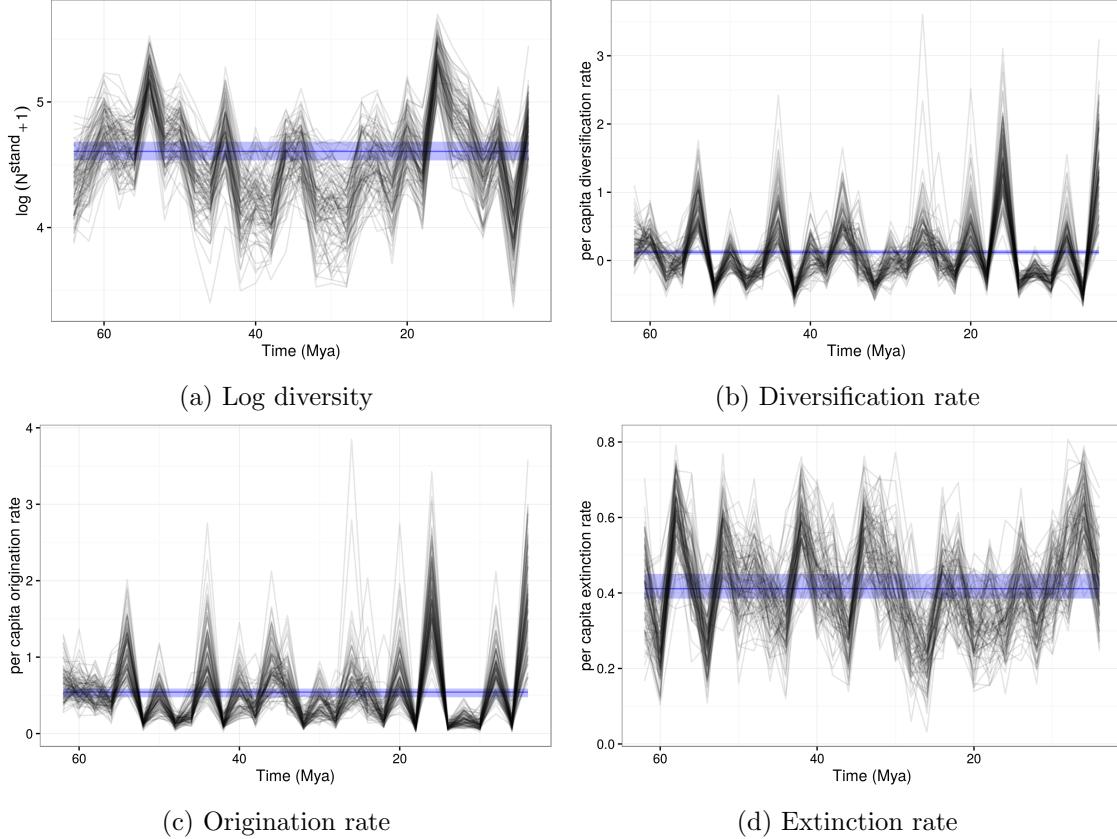


Figure 14: Posterior estimates of the time series of Cenozoic North American mammal diversity and its characteristic macroevolutionary rates; all estimates are from the birth-death model and 100 posterior draws are plotted to indicate the uncertainty in these estimates. The blue horizontal strip corresponds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

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

Time (Mya)	$P(N_t^{stand} > \bar{N}^{stand})$	$P(D_t^{rate} > \bar{D}^{rate})$
64.00	0.07	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	

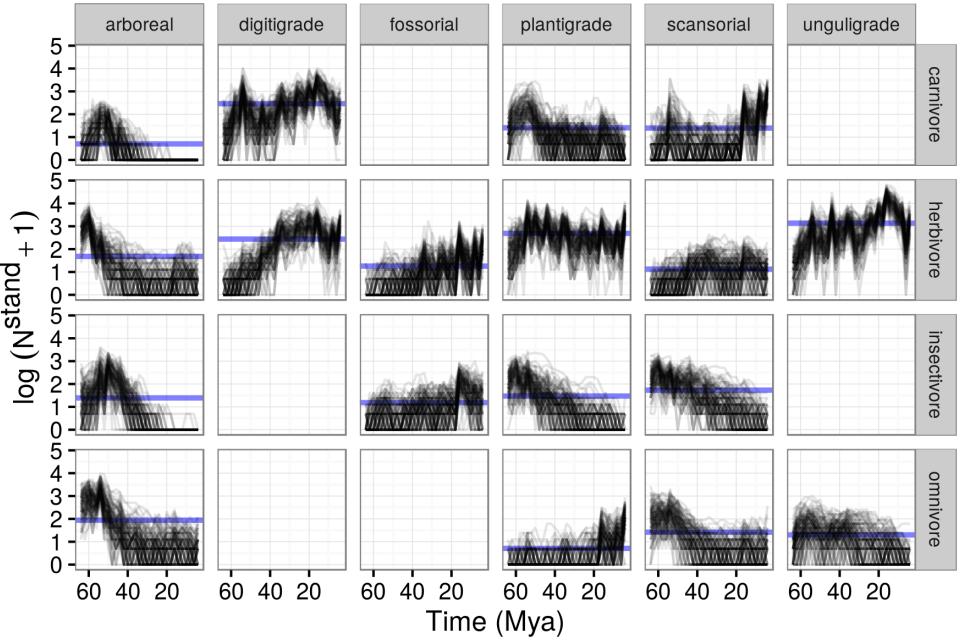


Figure 15: Posterior of standing log-diversity of North American mammals by ecotype for the Cenozoic as estimated from the birth-death model; 100 posterior draws are plotted to indicate the uncertainty in these estimates and what is technically plotted is log of diversity plus 1.

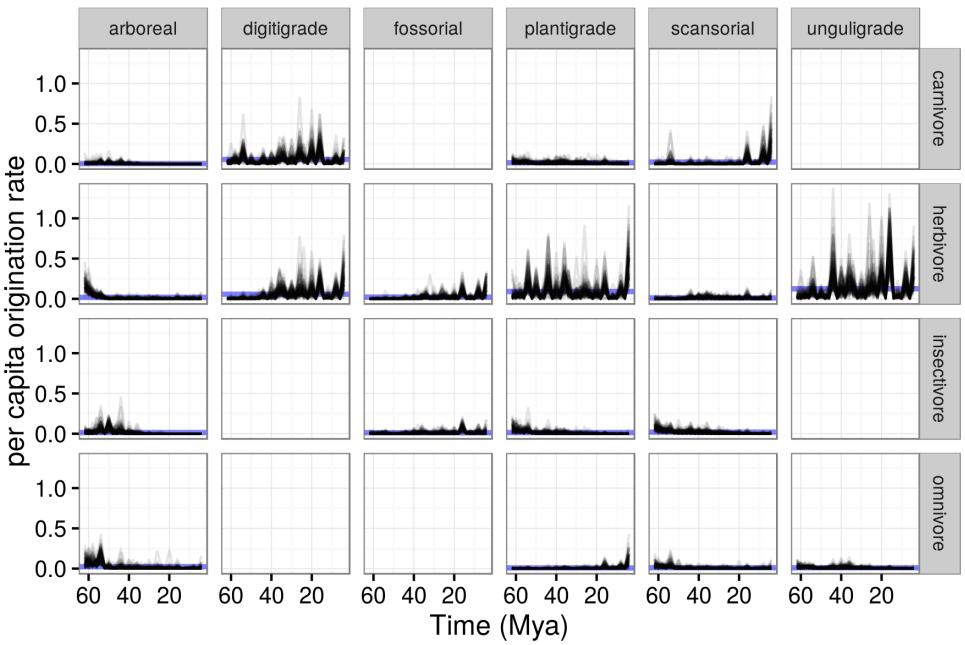


Figure 16: Posterior estimates of the per capita origination rates for each ecotype, plotted at the bin they originate from. These rates are calculated as the number of origination events for that ecotype from one time point to the next, divided by the standing diversity of all mammals at the initial time point. 100 posterior draws are plotted to indicate the uncertainty in these estimates.

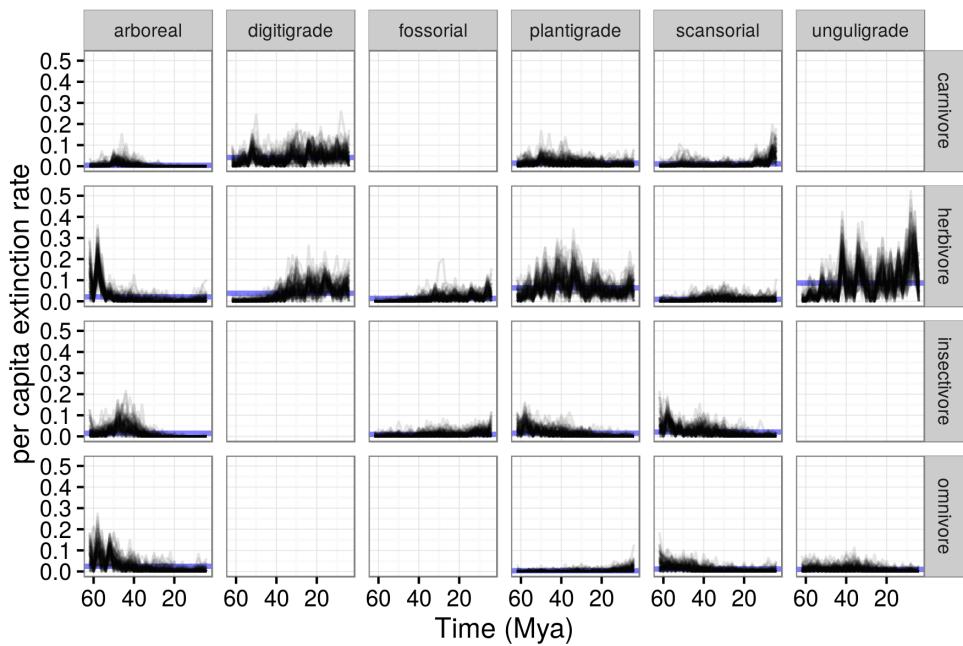


Figure 17: Posterior estimates of the per capita extinction rates for each ecotype, plotted at the bin they go extinct from. These rates are calculated as the number of extinction events for that ecotype from one time point to the next, divided by the standing diversity of all mammals at the initial time point. 100 posterior draws are plotted to indicate the uncertainty in these estimates.

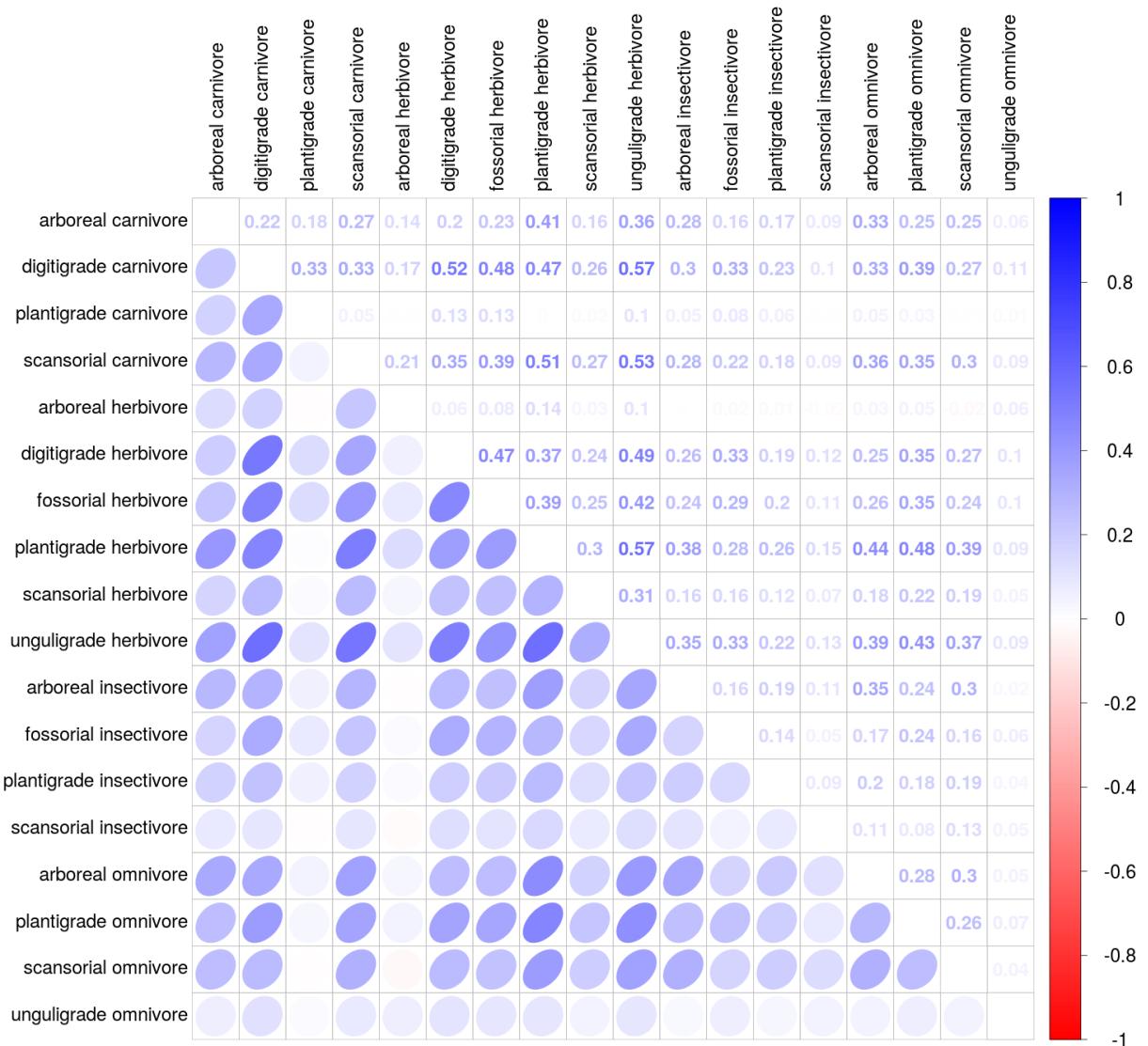


Figure 18: Posterior mean estimates of the correlations in origination probability between the mammal ecotypes. The lower triangle of the matrix is populated with ellipses corresponding to the level of correlation between the two ecotypes, while the upper triangle of the matrix corresponds to the mean estimated correlation between ecotypes. Darker values correspond to a greater magnitude of correlation with blue values corresponding to a positive correlation and red values a negative correlation.

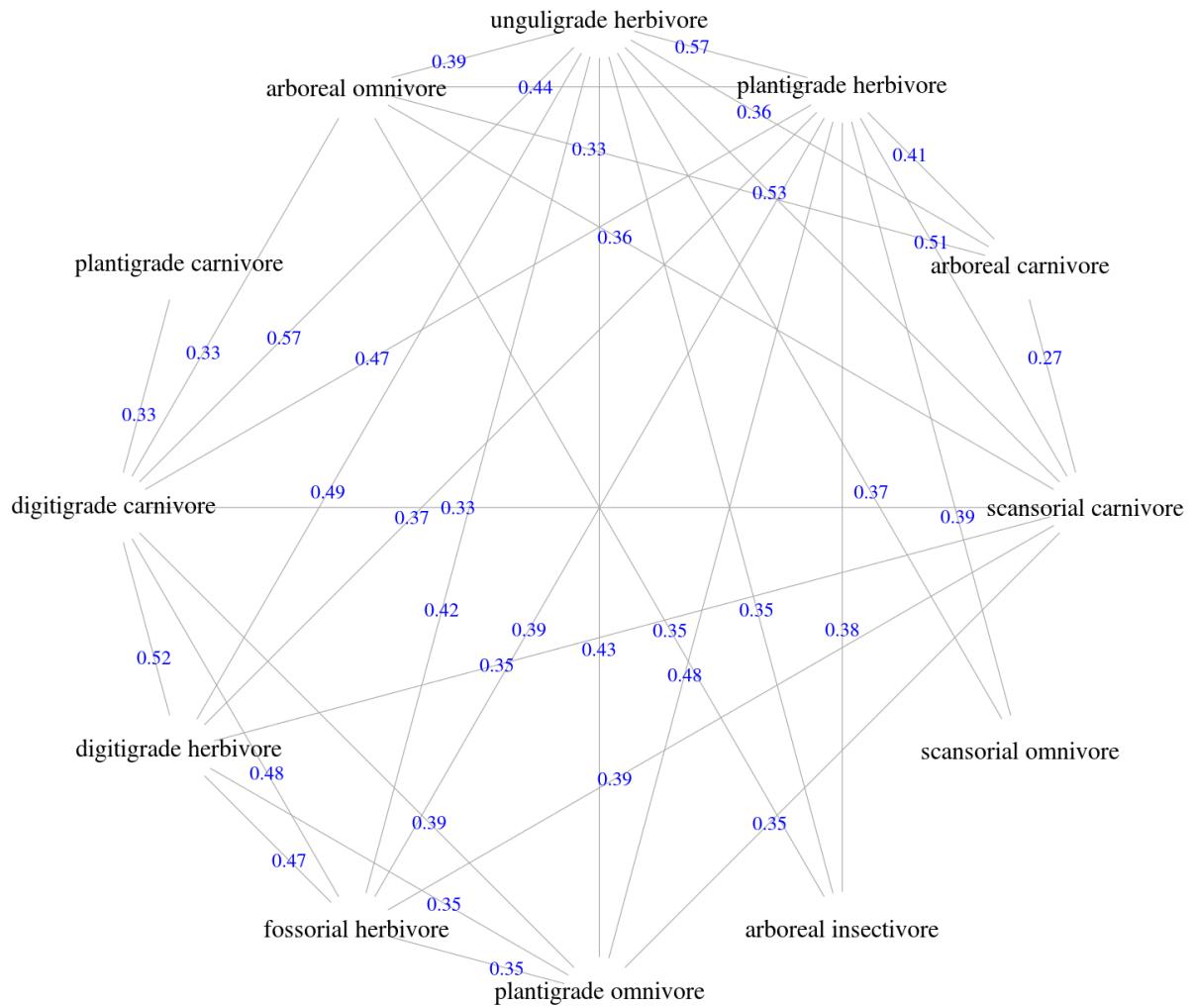


Figure 19: Ecotypes that have a strong correlation in origination probability. These ecotypes have a greater than 95% posterior probability of being positively correlated. The number plotted at the midpoint of each edge corresponds to the mean estimated correlation between those two ecotypes.

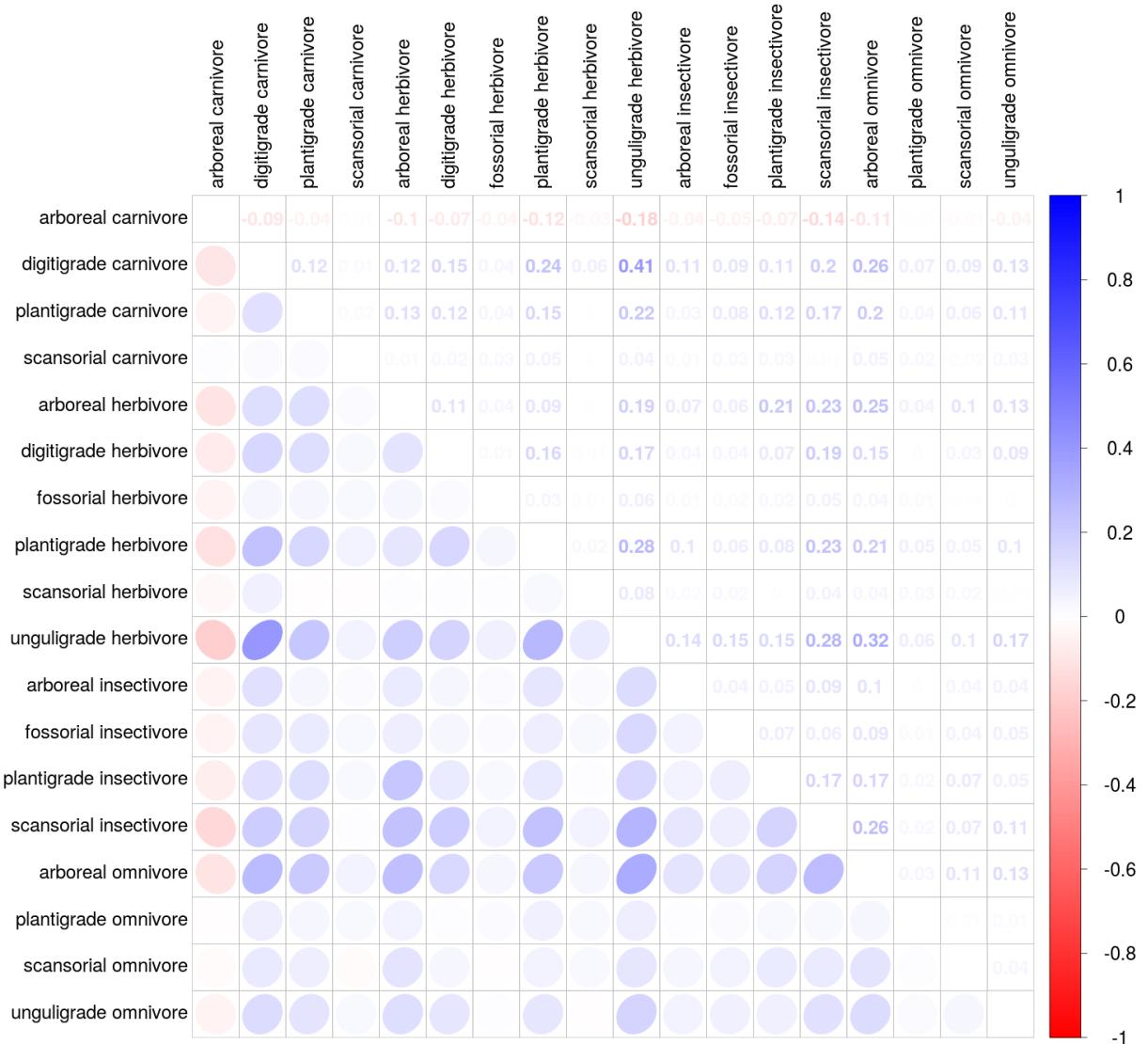


Figure 20: Posterior mean estimates of the correlations in survival probability between the mammal ecotypes. The lower triangle of the matrix is populated with ellipses corresponding to the level of correlation between the two ecotypes, while the upper triangle of the matrix corresponds to the mean estimated correlation between ecotypes. Darker values correspond to a greater magnitude of correlation with blue values corresponding to a positive correlation and red values a negative correlation.

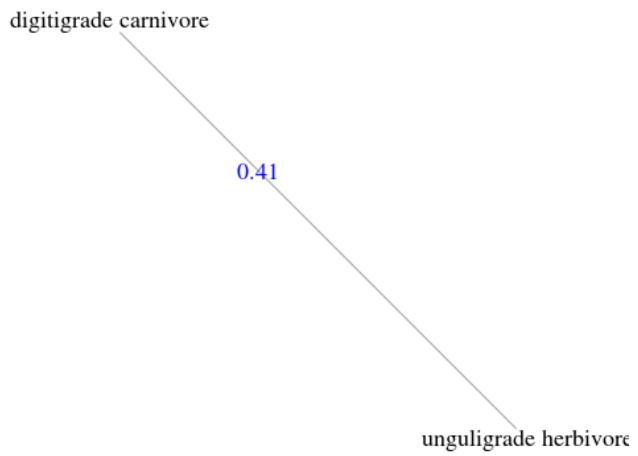


Figure 21: Ecotypes that have a strong correlation in survival probability. These ecotypes have a greater than 95% posterior probability of being positively correlated. The number plotted at the midpoint of each edge corresponds to the mean estimated correlation between those two ecotypes.