THE UNIVERSITY OF CHICAGO

REMODELING THE FOSSIL RECORD:
ANALYSIS OF EMERGENT EVOLUTIONARY AND ECOLOGICAL PATTERNS

A DISSERTATION SUBMITTED TO
THE FACULTY OF THE DIVISION OF THE BIOLOGICAL SCIENCES
AND THE PRITZKER SCHOOL OF MEDICINE
IN CANDIDACY FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

COMMITTEE ON EVOLUTIONARY BIOLOGY

BY
PETER DAVID SMITS

CHICAGO, ILLINOIS
JUNE 2017

In memory of my grandfather, George Handelman.

TABLE OF CONTENTS

LIST OF FIGURES vi

LIST OF TABLES viii

ACKNOWLEDGMENTS ix

ABSTRACT xii

1 INTRODUCTION 1
  1.1 Emergent patterns, macroevolution, and macroecology 1
  1.2 Structured data and modelling emergent patterns 6
  1.3 Study summaries 9

2 EXPECTED TIME-INVARIANT DIFFERENCES IN MAMMAL SPECIES DURATION 12
  2.1 Introduction 12
  2.2 Results 15
  2.3 Discussion 19
  2.4 Materials and Methods 21
    2.4.1 Species occurrence and covariate information 21
    2.4.2 Survival model 23
    2.4.3 Estimation 28
    2.4.4 Posterior evaluation 29
  2.5 Supplemental information for "Death and taxa" 30
    2.5.1 Supertree inference 31
    2.5.2 Modeling censored observations 32
    2.5.3 Deviance residuals 32
    2.5.4 Variance partitioning 33
    2.5.5 Widely applicable information criterion 35
    2.5.6 Results from posterior predictive checks 36
    2.5.7 Data quality concerns 37
    2.5.8 Concerns surrounding estimates of α 38

3 HOW MACROECOLOGY AFFECTS MACROEVOLUTION: THE INTERPLAY BETWEEN EXTINCTION INTENSITY AND TRAIT-DEPENDENT EXTINCTION IN BRACHIOPODS 42
  3.1 Introduction 42
    3.1.1 Factors affecting brachiopod survival 44
  3.2 Materials and Methods 47
    3.2.1 Fossil occurrence information 47
    3.2.2 Details of model 50
    3.2.3 Imputation of sampling probability 52
    3.2.4 Posterior inference and posterior predictive checks 54

3.3 Results 55
3.4 Discussion 64

4 SPECIES TRAITS AND ENVIRONMENTAL CONTEXT: THE CHANGING FUNCTIONAL COMPOSITION OF THE NORTH AMERICAN SPECIES POOL . . . 71
4.1 Introduction 71
4.1.1 Background 73
4.1.2 Foreground 77
4.2 Materials and Methods 79
4.2.1 Taxon occurrences and species-level information 79
4.2.2 Environmental and temporal covariates 87
4.2.3 Modeling species occurrence 88
4.2.4 Posterior inference and model adequacy 95
4.3 Results 99
4.3.1 Comparison of estimates from the pure-presence and birth-death models 99
4.3.2 Analysis of diversity 119
4.4 Discussion 128
4.4.1 Conclusions 136

5 CONCLUSION 138
5.1 Summary 139
5.2 Synthesis 141
5.3 Future 144
5.4 Final thoughts 145

REFERENCES 147

A TABLE OF SPECIES MASS ESTIMATES 175

LIST OF FIGURES

2.1 Posterior predictive checks . . . 15
2.2 Effect of mammal ecotypes on survival . . . 17
2.3 Partitioned variance for mammal survival . . . 18
2.4 Effect of cohort on mammal survival . . . 19
2.5 Estimate of hazard function for mammal survival . . . 30
2.6 Deviance residuals of fitted model . . . 36
2.7 Additional posterior predictive checks . . . 37
2.8 Simulation of sample size and estimates of α . . . 39

3.1 Posterior predictive check of survival . . . 55
3.2 Posterior predictive check of congruence . . . 56
3.3 Estimated relationship between environment and survival . . . 57
3.4 Cohort specific relationship between environment and survival . . . 58
3.5 Effect of environmental covariates on survival . . . 61
3.6 Correlation matrix of effects of environmental covariates . . . 62
3.7 Example imputed gap statistic distributions . . . 63

4.1 Conceptual diagram of the paleontological fourth-courner problem . . . 78
4.2 Conceptual figure of all possible occurrence histories for an observed species . . . 96
4.3 Posterior predictive check of average occurrence . . . 99
4.4 Ecotype occurrence probability estimated from the pure-presence model . . . 101
4.5 Ecotype origination probability estimated from the birth-death model . . . 102
4.6 Ecotype survival probability estimated from the birth-death model . . . 103
4.7 Estimates of the effect of mass on observation probability . . . 104
4.8 Effect of mass on probability of species occurrence as estimated from the pure-presence model . . . 105
4.9 Effect of mass on probability of species origination as estimated from the birth-death model . . . 106
4.10 Effect of mass on probability of species survival as estimated from the birth-death model . . . 107
4.11 Effects of group-level covariates on log-odds of ecotype occurrence as estimated from the pure-presence model . . . 108
4.12 Effects of group-level covariates on log-odds of ecotype origination as estimated from the birth-death model . . . 109
4.13 Effects of group-level covariates on log-odds of ecotype survival as estimated from the birth-death model . . . 110
4.14 Estimated correlations in origination probability between ecotypes . . . 111
4.15 Ecotypes with strong correlations in origination probability . . . 112
4.16 Estimated correlations in survival probability between ecotypes . . . 113
4.17 Ecotypes with strong correlations in survival probability . . . 115
4.18 Estimated mammal log-diversity and macroevolutionary rates for the Cenozoic . . . 122
4.19 Estimated mammal ecotype log-diversity for the Cenozoic . . . 123

4.20 Relative mammal ecotype log-diversity for the Cenozoic 127
4.21 Estimated per capita origination rates by mammal ecotype 129
4.22 Estimated per capita extinction rates by mammal ecotype 130

LIST OF TABLES

2.1 Posterior estimates of covariates on mammal survival 40
2.2 Cypher for ecotype assignments 40
2.3 Equations used to estimate mammal mass 41

3.1 Parameter estimates for brachiopod survival model 70

4.1 Species trait reassigments 81
4.2 Posture assignment based on taxonomy 82
4.3 Equations used to estimate mammal mass 87
4.4 Plant phase defintions 88
4.5 Transition matrices for the pure-presence and birth-death models 89
4.6 Parameters for the observation process part of the hidden Markov model. 90
4.7 Parameters for the model of occurrence in the pure-presence model 91
4.8 Parameters for the model of presence in the pure-presence model 94
4.9 Posterior probablity estimates of differences in occurrence by plant phase 114
4.10 Posterior probablity estimates of differences in origination by plant phase 114
4.11 Posterior probablity estimates of differences in survival by plant phase 115
4.12 Posterior probablity of effects of temperature on occurrence 116
4.13 Posterior probablity of effects of temperature on origination 117
4.14 Posterior probablity of effects of temperature on survival 118
4.15 Posterior probability estimates of a peak in diversity, diversification 124

A.1 Mammal body mass estimates 175

ACKNOWLEDGMENTS

First and foremost I would like to thank my advisors Kenneth D. Angielczyk and Michael Foote. Their support and advice helped guide my often erratic thinking into concrete products. Michael has been an incredible mentor and his constant pushing for me to “speak in English” has greatly improved my communication skills. Ken is an excellent collaborator and has been incredibly responsive even though we very rarely see each other in person; his help with crafting the narrative and improving the impact of my research has been an incredible boon. I look forward to continued collaboration with both of them.

My dissertation committee has been an invaluable resource for expanding my thinking and helping me translate my ideas for a broader audience than just myself. I am very lucky to have had such knowledgeable, questioning, and responsive group of readers. Thank you P. David Polly, Richard Ree, and Graham Slater for your incredible support.

I owe a big thanks to my undergraduate advisor Gregory Wilson. Greg took a risk on me and let me contribute directly to his research program, which was an incredible experience that led to my first paper. I learned a lot about fossil mammals, the Cretaceous, and field work from Greg and I owe what little “real” paleontology *bona fides* to him and for that I’m forever grateful.

My Master’s advisor Alistair Evans taught me the importance of not overwhelming your audience when trying to explain the context of scientific study, something which I still work on everyday. Al also gave me an incredible degree of freedom to explore what questions interest me, something which really shaped my approach to research; I’m nothing but independent. And while I may not have worked directly on mammal morphology for my dissertation, the dietary and functional perspective that pervades Alistair’s research has never been far from my mind. Thank you.

There have been countless colleagues, professors, and peers who have helped shaped not only my dissertation but also my thinking in general and sadly I cannot name them all. Nancy Simmons, Jim Kenagy, Jeff Bradley, and Donald Grayson all provided me with research and/or discussion opportunities which have proved invaluable for my growth as a scientist. I will never forget Nancy telling me that the most important thing I could do to increase the impact of my research was to grow my statistical knowledge; this is a lesson I've truly taken to heart. Stewart Edie has been a fantastic collaborator and I look forward to our continued work together. My officemates Dallas Krentzel and Max Winston have been an excellent sounding board for questions, frustrations, and curiosities. The graduate student community of Darwinian Sciences and the paleontology section of Geophysical Sciences have also been a source of continued strength and support; I can't count the number of times TG has saved my mood from completely nose-diving. A big thank you to you all.

My parents, Alexander Smits and Louise Handelman, have been an unending source of support through out this entire process. They never questioned my goals or career decisions, and have provided such important life advice and unconditional love that I cannot even begin to express my appreciation and gratefulness for all the support you've given me over the years. I also want to give a special thanks to my brother, James Smits. I could not have asked for a more supportive and helpful brother. Thank you; I would never have been able to do this without all of you.

As I've moved around the globe, my friends have always been an important source of distraction and happiness that I could not have dreamed surviving without. Thank you to Jessica Escott, Benjamin Frable, Colin Kyle, Connor O'Donnell, Darcy Ross, Elizabeth Sander, Julie Szymaszek, Brian Waligorski, Brian Goodrich, anyone who has every played Dungeons and Dragons with me, and countless others. You have all made my life easier and happy. Without your presence I would never have had release.

Finally, I would like to thank my wonderful partner Megan Boatright who has been an incredible anchor through out everything. Our lake-front conversations helped me coalesce around my research topics, our late-nights working together help me write both my proposal and my dissertation, and our many journeys together re-opened my eyes to the beauty and joy in the world. Our cats Annyong, Sophie, and Monty have brought us incredible joy. Providing a retirement home for Monty was one of the greatest experiences of my life and I'm so happy we shared it together. I love you with all my heart Megan.

ABSTRACT

Macroevolution and macroecology are concerned with the patterns in evolutionary and ecological data, respectively, which arise when observing multiple species over time and/or space. Species extinction and species pool functional composition are the macroevolutionary and macroecological patterns at the heart of this dissertation. The hypotheses and analyses in the three studies forming this dissertaiton were all framed in terms of how species functional traits can shape these emergent patterns.

In my first study, I analyzed the Cenozoic fossil record of North American mammals to test two long standing hypotheses: the survival of the unspecialized hypothesis, and the Law of Constant Extinction. My analysis centers around a model of species duration as a function of multiple species traits, species' phylogenetic relatedness, and species' origination cohort. My results support the conclusion that generalist species will, on average, have a greater duration than more specialized species. I also find that species extinction risk increases with species duration, a result that is counter the Law of Constant Extinction. Additionally, I find that only some of the factors associated with extinction risk for Modern mammals could be considered risk factors for mammals from the rest of the Cenozoic, indicating a difference between the modern biodiversity crisis and "normal" extinction dynamics.

My second study also deals with the survival of the unspecialized and the Law of Constant Extinction, but focuses on a different system: post-Cambrian Paleozoic brachiopods. An additional aspect of this study is an analysis of the relationship between extinction intensity and the strength of trait selection. I find support for greater survival among environmental generalists than specialists. I also find evidence that for geographic range and environmental preference, as extinction intensity increases, the selective importance of these traits increases. This result is evidence for a qualitative difference between background and mass extinction.

The final study is an analysis of the changing functional composition of the North American mammal regional species pool over the last 65 million years. The goals of this analysis are to understand when functional groups are enriched or depleted, and how changes to environmental context may shape these changes. I find that mammal diversity is more strongly shaped by changes to origination probability rather than changes to extinction probability. I also find that all arboreal ecotypes declined throughout the Paleogene and disappeared from the species pool by the Neogene. Additionally, I found that most herbivore ecotypes expand their relative contribution to functional diversity over time.

My desire with this dissertation is to present the types of analyses and results that are possible through a synthesis of macroevolution and macroecology. The first step to building any dialogue is to agree on a common language and I've emphasised an expressive statistical framework with which to phrase our questions in a common tongue. My hope is that this studies serve as an example of how to use paleontological data to unite questions about the processes underlying macroevolutionary and macroecological patterns.

CHAPTER 1

INTRODUCTION

Species traits are the bridge between evolution and ecology [196, 358]. A trait is an identifiable property of an organism, such as individual body size, while a species trait is an identifiable property of the entire species, such as the average body size or geographic range of a species [196]. A class of species traits called functional traits are those traits that describe a species means of interacting with its environment such as leaf surface area or trophic role [196]. In macroevolutionary studies, analyses are typically focused on a patterns associated with a single trait or are instead traitless analyses of diversity and the diversification process [291, 290, 230, 247, 304, 129, 130, 171, 222]. Macroecological studies are frequently concerned with describing the distribution of species and individuals over space or time, such as shifts in community composition along some gradient or axis [306, 305, 33, 32, 36, 65, 90, 151, 152]. My desire with this dissertation is to present the types of analyses and results that are possible through a synthesis of both macroevolution and macroecology; my approach is to develop inference devices (i.e. statistical models) to better understand the interactions between the effects of multiple species traits, as well as those of a species' temporal and environmental context, on diversity and differential diversification both in space and time.

1.1 Emergent patterns, macroevolution, and macroecology

An emergent pattern is one that is not observable from individual constituent parts, instead the observation of many individuals is necessary. Emergence is ubiquitous in biological systems: cells form tissues which in form organs with a complex functions, species extinction requires all individual members of that species to die for possibly unrelated reasons, and a species global geographic range is the product of many individual home ranges and their collective distribution over space. The history of a species, or set of species, over time is inherently an emergent pattern as the temporal history of a species is not knowable from an instantaneous sample. Macroevolution and macroecology are the studies of emergent patterns in evolutionary and ecological data, respectively [32, 33, 321, 320]. Traditionally, macroevolution is the study of patterns over time while macroecology is concerned with patterns over space, but I find this division overly reductive.

Both macroevolution and macroecology are disciplines concerned with emergent patterns; they both implicitly and explicitly accept a hierarchical perspective on biology as emergence is not possible without different levels of organization, however they are defined. Even if an analysis is concerned with only a single species, the species' identity is itself an emergent property or label of a collection of populations and individuals which all share a common evolutionary history. While it may be argued that the species label or identity is a non-biological construct or is simply heuristic for understanding the complexity of populations and reproductive isolation, we are still concerned with patterns associated with that construct as well as its intrinsic properties (e.g. extinction, conservation, ecosystem services) [52, 137]. My opinion is that any level which has discernible, measurable, and assignable properties is potentially worth studying, though special attention to the species level most likely affords the greatest translation between paleontological and neontological studies. In the analyses presented in this dissertation the levels of organizations thare are studied are mammal species and brachiopod genera.

Macroevolution is much more than evolution above the species level [87]; this is overly reductive and assigns too much meaning to a single level of organization rather than embracing the multitudes of possible levels of organisation. A broad definition of macroevolution, *qua* field of study, is the study of emergent evolutionary patterns; these are patterns of speciation/extinction (diversity) and trait evolution (disparity) that are observable when considering a collection of individuals. The key here is that it is the pattern which is emergent at the level of organization is a property of the distribution of the constituent members of that organization; the traits and factors which affect or cause this pattern need not be themselves emergent. In complement, macroecology is then the study of emergent ecological patterns, which means patterns in spatial distribution or community composition which are observable only when considering more than one taxon or when considering the temporal history of one or more taxa [32, 33, 306]. Because taxa inherently respond differently and individually to environmental changes, both biotic and abiotic, macroecological patterns are those due to the similarity in response across individuals [28].

Species selection is enshrined as one of the most important patterns in macroevolution [320, 321, 345, 137, 246, 295, 224]. Rabosky and McCune [246] portray species selection as the resultant phenomena due to the heritability of speciation and extinction rates. This definition is an expansion of which phenomenon fall under this category by divorcing the idea of levels of organization from species selection and removing any requirement surrounding the mechanisms behind differences in speciation and extinction rates. The result of these distinction is two fold. First, this definition avoids unnecessary semantic conflicts surrounding "effect macroevolution" versus "strict" species selection [136, 137, 344]. Second, this definition also avoids the unproductive debate surrounding the differences between species selection versus species sorting which has been the cause for a considerable amount of confusion and rhetorical fights [343, 345, 176, 224], not the least of which is the adoption of "sorting" as an important term for understanding community assembly [331, 178, 126, 51, 310, 334, 288]. One shortcoming of the definition presented by Rabosky and McCune [246] is that it provides no operational structure to understanding the causes for differences in speciation and extinction rate. Species inherit more than just speciation and extinction rates; they also inherit traits which themselves may be linked to differences in speciation or extinction rates due to their effects on species fitness. Species fitness is a concept that is vitally important to understanding and estimating the effect of species selection in biological systems [49, 220].

Cooper [49] defined fundamental fitness for "populations" as the expected time to extinction, otherwise called mean survival time or average persistence time; this is a fundamentally statistical definition: one entity is fitter than another if and only if it is statistically expected to survive longer. This definition applies equally to the fitness of alleles, genotype populations, and species; the only requirement is a definition of extinction. The key with this definition of fitness is that it is a *fundamental* measure in that most other definitions of fitness are derivable from probability of survival [49]. Again, this definition of fitness is for "populations" and not individuals. Instead, individuals are considered stochastic draws from the population distribution; this distinction serves is to separate fitness from luck. Ultimately, the logic behind this is that a unit that is fitter is more likely to occur [49].

Species fitness is an emergent property of a species because it is a property of the distribution of its constituents [136]. Given Cooper's definition of fitness [49] species fitness has two derived forms depending on the definition of extinction. If species extinction is defined strictly as when the last individual of a species dies, then species fitness is the expected duration of a species. If species extinction is instead defined as the loss of the last descendent along a lineage, then species fitness is the expected duration of a species and all of its descendent species; this is a lines-of-descent definition [49]. While the former deals only with species extinction, the latter of these aspects encompasses speciation. By increasing the number of descendent species, the lineage has an increased probability of occurring for longer and thus a greater expected time till extinction [49]; in effect, a higher speciation rate can lead to a greater time till extinction. And while the mechanisms behind why speciation rates can vary both across species and time is not well understood [243, 245, 52], this derived definition is agnostic to whatever mechanism controls the generation of descendants.

Extinction is a property of, or a phenomenon affecting, species as it requires the death of all organisms within a species which do not have to all occur for the same reason [295]. Extinction is a fundamentally emergent phenomenon that is the ultimate manifestation of selection; it is also central to macroevolutionary studies and the definition of (species) fitness used here [49]. Extinction is featured centrally in one of the few “laws” in macroevolution and paleobiology: the Law of Constant Extinction [335, 175]. This law states that a species risk of going extinct is independent of that species age [335, 175], a conclusion reached via analysis of patterns of (higher) taxon survival patterns. The Red Queen hypothesis was proposed as a process that would result in the observation underlying the Law of Constant Extinction [335], though it has obviously grown to have a life of its own [175].

Extinction also features prominently in the long standing macroevolutionary hypothesis of the survival of the unspecialized [298]. This hypothesis forms the expectation that generalist species are expected to have a greater duration than specialists. This hypothesis has received a great deal of attention over the years and appears to hold in most, but not all, situations [298, 169, 217, 218, 21, 250]. Because of the strong evidence of survival of the unspecialized as a good generalization and initial condition for all analyses, it should feature heavily in all comparative analyses of species durations, survival, and extinction. In particular, this hypothesis features prominently in the first two studies of this dissertation where generality is in the context of dietary and locomotor categories (Chapter 2) and environmental preference (Chapter 3).

The functional composition of a community or species pool is a property of that unit; observing a single species at a locality does not reveal the functional composition of the community in which it interacts. The composition of a community or species pool in terms of functional groups is a community ecology exercise. Comparing the distribution of functional groups across communities or species pools is where community ecology and macroecology intersect [196, 33, 306]. In paleobiology, a successful means of classifying marine invertebrate functional groups has been a three dimensional classification scheme called an "ecocube" which uses consistently identifiable functional traits to label both possible and observed functional groups [36, 15]. This approach also emphasizes the presence or absence of different functional groups and how functional diversity can change over time. It is this strategy that inspires the third study presented in this dissertation.

1.2 Structured data and modelling emergent patterns

An inference device is a theoretical tool for improving our knowledge by processing new information and observations [195, 149]; this device has initial conditions describing what we know (e.g. nothing), mechanisms for updating this knowledge to reflect new information, and can then produce an updated "picture" that better represents our current knowledge as well as the uncertainty surrounding this knowledge. Each inference device has a specific and narrow purpose and functionality [195]; unless the mechanisms are similar, a device for processing the rate of imperfections in the manufacturing of widgets cannot process the queuing times of callers to a help line.

We can think of the well known battery of statistical tests [311] as re-usable inference devices with very narrow utility; these are unmodifiable tools for handling very specific questions and data. All Bayesian statistical models act as inference devices because they fulfill the requirements described above: initial conditions, updating mechanism, and output as updated knowledge along with the uncertainty surrounding that knowledge [195, 149]. By developing a new model for each new question there is a precision of translation; the model actually reflects the questions at hand, something that is preferable to forcing questions and data to fit into pre-made inference devices (e.g. models, tests) that do not update knowledge in a means relevant to the actual question(s) of interest.

Structure occurs naturally in the collection of data. For example, in my analyses of mammal species and brachiopod genus durations I grouped species that originated in the same temporal unit as a “cohort” (Chapter 2, and 3), the assumption being that species originating at the same time may be more similar than species that originate at different times. By grouping these species together the differences in extinction risk associated with each cohort can be estimated. If these groupings are ignored, species are all assumed to have identical properties. Additionally, by including these cohort labels in the model of species durations, the estimates for each cohort are tempered and drawn towards the overall average cohort effect through regularization [101, 195]. By sharing information across cohorts, the predictive properties of the model to out-of-sample data are improved [101, 195]. Finally, including these groupings in the actual model better reflect the actual data at hand.

Two of the most important analytical approaches at the core of macroevolutionary study are the birth-death process for diversification in both discrete and continuous time [257, 253, 212, 214, 211, 213, 316], and the random walk heuristic for continuous trait evolution [256, 70, 29, 104, 261, 262, 263, 287, 130, 129]. All three of the studies covered by this dissertation make use of some variant to the birth-death process. The first two studies are analyses of extinction, which is a pure-death processes. The third study utilizes a discrete-time birth-death process to model species presence in a species pool.

Similar analytical foundations are harder to identify for macroecology as a whole, so instead I will focus on species distribution models (SDMs) as a powerful framework for understanding the distribution of species, due to environmental factors [64] and species traits [288]. SDMs are a class of models which attempt to operationalize the multitude of processes which result in the distribution of one or more species in space and/or time. In effect, SDMs are a means of operationalizing the concept of a species’ “realized niche” in order to understand the limits on a species distribution [64]. Typically, SDMs are used to analyze the relationship between species presence at a locality and the environmental factors which characterize that locality. From this analysis, the possible distribution of a species in space can be then estimated and compared to the observed distribution of that species [64, 11, 229]. The maximum entropy theory of community assembly, and its related model, view community assembly as an ecological sorting process where traits mediate the effects of environmental filters [288, 356]; also called "community assembly via trait selection" (CATS regression). By analyzing the composition of species at localities based on their traits, the strength and relative importance of the traits that most directly structure community composition be elucidated.

Both of these approaches to analyzing species distributions can be united in a single fourth-corner model [356, 31]. The fourth-corner problem is an old problem in community ecology originating in the multivariate analysis literature: assuming species distribution is the result of functional traits interacting with environmental factors, how do we estimate which interactions are important and their relative strengths [164, 61]? By phrasing the fourth-corner problem as a model based framework, results are much more easily interpretable and actually provide estimates of the effects of species traits and environmental factors instead of the simple significance provided from the older Monte Carlo based methodology [31, 139, 232, 231].

The third study in this dissertation makes extensive use of this framework by casting the fourth-corner problem into an additional dimension: time. By combining the fourth-corner framework outlined above [356, 31] with the birth-death process used for modeling diversification into a single unified model of species occurrence through time as a function of both species traits and changing environmental context I've developed a powerful analytical bridge between macroecology and macroevolution.

I emphasize model-based approaches to analysis as well as question or study specific models because a common language is necessary for clear, coherent, and translatable results that actually relate to the question(s) at hand. Some of the greatest limits to paleobiological, macroevolutionary, and macroecological study are a lack of strong, mechanistic predictive theories that can be expressed mathematically. Some of the greatest strides in advancing discussions of macroevolutionary and macroecological theory disputes have come from translating verbal theory into mathematical and statistical models [257, 212, 70, 130, 129, 288]. The complex realities of the biological processes which shape diversity are rarely integrated into paleobiological analyses of macroevolutionary and macroecological patterns. A move to a model-forward approach to paleobiology, heavily steeped in evolutionary and ecological theory, would be beneficial for advancement of theories in macroevolution and macroecology.

Paleobiologists historically believe that neontologists ignore their approaches and insights into macroevolution and macroecology study and theory [283, 282], but without a concerted effort to engage within the same theoretical framework and language when developing scientific questions and the related analytical tools (i.e. statistical models) this worry and resentment is all but preordained. Because the systems paleobiologists study are unknown to, have no direct impact on, or are contextualized with respect to the systems studied by neontologists, a push towards unification and synthesis most likely has to begin with the paleobiological community; luckily, it appears that the neontological community is receptive to paleobiological insight [97]. The simplest and fastest way to begin not just a dialogue but a unification is by translating verbal macroevolutionary and macroecological theories from paleobiology into statistical models that are readable by all researchers, both paleontological and neontological.

1.3 Study summaries

Each of the three studies that make up this dissertation involves developing a hierarchical model to describe structured data with the goal of making macroevolutionary and/or macroecological inference. The first two studies are decidedly macroevolutionary in bent as they are analyses of trait-based extinction patterns in mammals and brachiopods, respectively. The third study is an analysis of mammal species pool temporal dynamics and is of a strong macroecological bent, though makes use of a macroevolutionary model of diversification in order to describe species turnover.

The first study presented is an analysis of North American mammal species durations and trait-based extinction risk (Chapter 2). This analysis is principally concerned with the long standing hypothesis of the survival of the unspecialized which states that average or generalist species are expected to have a greater duration than specialists or other extreme forms [298, 169, 170]. Species duration is a proxy for species extinction risk as species with a shorter duration experience a greater extinction risk than species with a long duration. In this study, differences in species extinction risk based on multiple functional traits are estimated while also taking into account time of species origination as well as its relative phylogenetic position. Finally, the possibility of species age affecting extinction risk is also considered because while the Law of Constant Extinction is extremely hard to "test" it has never definitively been proven [335, 175].

The second study presented is also an analysis of taxon durations, but this time focuses on all post-Cambrian Paleozoic brachiopod genera (Chapter 3). The question at the center of this study is "what happens to the effects of functional traits on survival when average survival increases or decreases?" Unlike the previously described study, which focused on the average effects of functional traits on survival, this study requires estimates of how the effects of functional traits vary through time. The key parameters are those of the correlation matrix of the effects of these traits on duration and the average duration of species originating at the same time. This study also has results relevant to the survival of the unspecialized hypothesis with respect to the effect of environmental preference on survival, and the Law of Constant Extinction by allowing survival to be a function of taxon duration.

As mentioned above, the third study presented is decidedly more macroecological in focus as it is an analysis of how a regional species pool changes over time due to species turnover and a changing environmental context (Chapter 4). The fundamental question is “when are certain ecotypes enriched or depleted with respect to their diversity history?” To that end, I analyze the set of North American mammals for the Cenozoic and the changing functional composition of that species pool from nearly the beginning of the Cenozoic to almost the very recent (64-2 million years ago). In this analysis, functional composition of the species pool is described as the relative diversity of 18 different mammal ecotypes which are defined for every species as its dietary and locomotor combination. The occurrence of an ecotype, in terms of origination and survival, is modeled as a function of that species environmental context as described by the dominant plant groups in North America as well as global temperature estimates.

All three of these analyses feature a hierarchical Bayesian model developed explicitly for each study in order to clearly attempt to answer the questions at hand. Each of these studies exemplifies my earlier rhetoric of how to build and advance macroevolutionary and macroecological study and theory through the explicit phrasing of scientific questions, precision of translation from question to analysis, and the mobilization of domain specific knowledge to cast results both in terms of the system specific insights as well as the theoretical insights.

CHAPTER 2

EXPECTED TIME-INVARIANT DIFFERENCES IN MAMMAL SPECIES DURATION

Determining which biological traits influence differences in extinction risk is vital for understanding the differential diversification of life and for making predictions about species' vulnerability to anthropogenic impacts. Here I present a hierarchical Bayesian survival model of North American Cenozoic mammal species durations in relation to species-level ecological factors, time of origination, and phylogenetic relationships. I find support for the "survival of the unspecialized" as a time-invariant generalization of trait-based extinction risk. Furthermore, I find that phylogenetic and temporal effects are both substantial factors associated with differences in species durations. Finally, I find that the estimated effects of these factors are partially incongruous with how these factors are correlated with extinction risk of the extant species. This parallels previous observations that background extinction is a poor predictor of mass extinction events and suggests that attention should be focused on mass extinctions to gain insight into modern species loss.[1]

2.1 Introduction

Why extinction risk varies among species remains one of the most fundamental questions in paleobiology and conservation biology [298, 335, 255, 240, 347]. To address this issue, I test for similarities in associations between extinction risk and multiple species-level traits during times of background extinction and in the modern world; which traits have time-invariant effects on species duration; and whether extinction is age-independent. I approach

---

[1] This chapter was previously published as an article in the *Proceedings of the National Academy of Sciences* [309] and is reproduced here as allowed by publisher.

these questions together by using a model of species duration whose parameter estimates act as direct tests of these questions. Cenozoic mammals are an ideal focus for this study because their fossil record is well sampled and well resolved both temporally and spatially, and because individual species ecology and taxonomic position are generally understood [298, 240, 6, 171, 305, 329, 191].

Time-invariant factors are those that have a constant directional effect even if their magnitude varies. Because change in the magnitude of extinction risk is not necessarily the best indicator of a shift from background to mass extinction [349], it is better to look for changes in either the direction of selection, the loss of a selective pressure, or the appearance of novel selective pressures [134].

The species-level traits studied here are bioprovince occupancy, body mass, and dietary and locomotor categories. These traits are related to aspects of a species' adaptive zone such as population density, expected range size, potential prey, and dispersal ability [305, 152] and are a combination of aggregate and emergent traits [137]. It is expected that species with larger geographic ranges have lower extinction rates than species with smaller geographic ranges [134, 270]; however, how traits more directly related to species–environment interactions may affect species extinction risk is more nebulous.

Body size is a complex trait related to many life history characteristics. There are three general hypotheses of how body size may affect extinction risk: 1) positive effect where an increase in body size causes an increase in extinction risk, potentially due to associated decrease in reproductive rate or other similar life history traits [171, 172]; 2) negative effect where an increase in body size causes a decrease in extinction risk because of an expected positive relationship between body size and geographic range; and 3) no effect of body size on extinction risk [329].

The strongest expectation for the effects of dietary category on extinction risk is that omnivores will have the lowest extinction risk of all species. This expectation is based on the long standing "survival of the unspecialized" hypothesis where more generalist species (e.g. omnivores) have greater survival than specialist species (e.g. carnivores/herbivores) [298, 169]. It has also been observed that both carnivores and herbivores have greater diversification rates than omnivores, with herbivores diversifying faster than carnivores [237]. How this result translates into differences in extinction risk is currently unknown [241]. In modern taxa, higher trophic levels (e.g. carnivores versus herbivores) have been associated with greater extinction risk, most likely because of human extermination of top predators [172, 238].

Similarly, there are few expectations of how locomotor category may effect extinction risk. During the Cenozoic, there was a shift at the Paleogene/Neogene boundary from predominately closed to predominately open environments [28, 142]. Based on this observation, a prediction is that arboreal taxa will have the greatest extinction risk of all, with both scansorial and ground dwelling taxa having lower extinction risks.

I use a hierarchical Bayesian survival model of species duration as predicted by the covariates of interest along with species' temporal and phylogenetic context. Species duration, in 2 My bins, was modeled as realizations from either an exponential or Weibull distribution-based hierarchical model [101]. The exponential distribution corresponds to the Law of Constant Extinction, which states that extinction is age-independent [335]. Note that the exponential is a special case of the Weibull when its shape parameter, α, is 1. The Weibull distribution allows for extinction to be taxon-age dependent, where values of α greater than 1 corresponds to increasing risk with age and values less than 1 corresponds to decreasing risk with age. Origination cohort and phylogenetic position were modeled as independent effects. Phylogenetic effect was modeled assuming species duration may have evolved via a Brownian motion-like process [184, 128]. The results from the Weibull model are detailed here because this model has a better fit to the data the exponential (Weibull WAIC 6140.37, exponential WAIC 16697.35; Fig. 2.1, S1, S2).

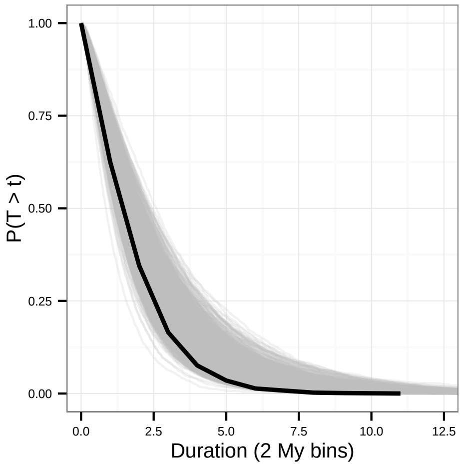


Figure 2.1: Weibull-based model estimates (grey) from 1000 posterior predictive data sets of the empirical survival function (black). The survival function is the probability that a species with duration t will not have gone extinct. Simulated data sets were generated by drawing parameter values randomly from their estimated posteriors and using the observed covariate information to estimate durations for all the observed species.

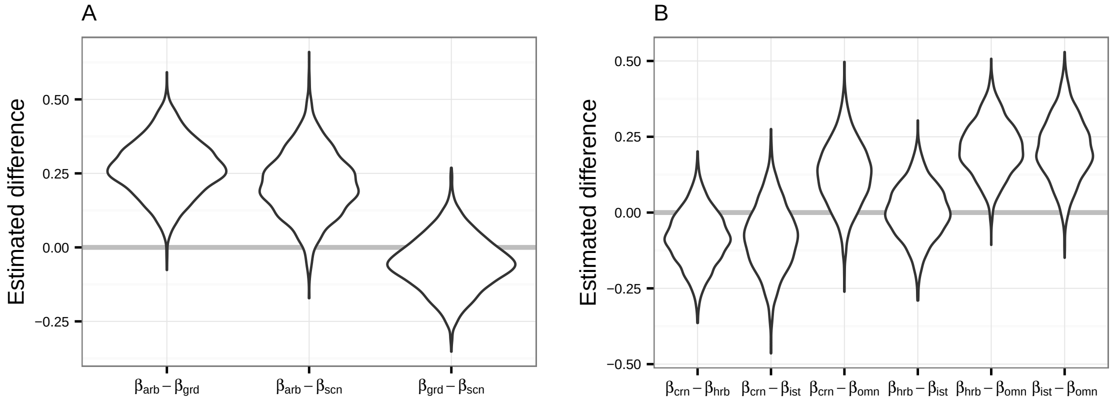
2.2 Results

A summary of the posterior distributions for the most relevant parameter estimates is presented in Table 2.1. All posterior inference is based on these estimates. For the results from the posterior predictive checks and discussion of the estimation of α, please see the accompanying Supplemental information (Section 2.5). Additionally, see the Supplemental information for discussion surrounding use of Paleobiology Database and accompanying data quality concerns.

Species with greater bioprovince occupancy are found to be associated with lower extinction risk than taxa with smaller bioprovince occupancy ($\beta_{occupancy}$mean $= -0.53$, std $= 0.08$). This is consistent with previous findings. Body size has nearly zero association with expected duration (β_{size}mean $= -0.05$, std $= 0.05$), a similar result to some previous studies [329]. However, previous studies were performed at the generic level and were unable to determine how body size may effect species-level extinction, as the effect of either extinction or speciation cannot be distinguished [171, 329].

Some clear patterns emerge from the pairwise differences in effect of each dietary category on expected duration (Fig. 2.2). Consistent with expectations from the "survival of the unspecialized" hypothesis [169, 298], omnivory appears to be associated with the lowest expected extinction risk. Carnivory is associated with a greater expected duration than either herbivory or insectivory, but a greater expected extinction risk than omnivory. Finally, herbivory and insectivory have approximately equal effects on expected duration. Given previous results, these results imply that carnivores have a greater origination rate than omnivores [237]. These results also imply that herbivores, which have the greatest extinction risk, must also have a very high origination rate in order to have the greatest diversification rate among these three categories [237].

For locomotor category, both scansoriality and ground dwelling life habitat are associated with a greater expected duration than arboreality (Fig. 2.2). Scansorial and ground dwelling life habits also have approximately equal expected effects on extinction risk. This is consistent with the expectation that arboreality will confer greater extinction risk due to the loss of associated environment with the shift from open to closed habitat at the Paleogene/Neogene boundary [28]. However, there are two possible processes which could lead to the observed pattern: arboreality confers an intrinsic difference in extinction risk or it might not be that arboreal taxa have an intrinsically higher risk but were instead "hit harder" by the environmental shift than other taxa. This analysis cannot distinguish between these two processes. Note that, while this is a study of North American Cenozoic mammals, for European Cenozoic mammals this transitionary period corresponds to the Vallesian which


(a)

(b)

Figure 2.2: Pairwise differences in effect of the locomotor (**A**) and dietary categories (**B**) on expected duration from 1000 samples from the posterior distribution. Comparisons of locomotor categories, from top to bottom (**A**), are: arboreal ($\beta_{arb} = \beta_0$) versus ground dwelling ($\beta_{grd} = \beta_0 + \beta_g$), arboreal versus scansorial ($\beta_{scn} = \beta_0 + \beta_s$), and ground dwelling versus scansorial. For dietary category, from top to bottom (**B**): carnivore ($\beta_{crn} = \beta_0$) versus herbivore ($\beta_{hrb} = \beta_0 + \beta_h$), carnivore versus insectivore ($\beta_{ist} = \beta_0 + \beta_i$), carnivore versus omnivore ($\beta_{omn} = \beta_0 + \beta_o$), herbivore versus insectivore, herbivore versus omnivore, and insectivore versus omnivore. Negative values indicate that the first category is expected to have a greater duration than the second, while positive values indicate that the first category is expected to have a shorter duration.

was a sudden shift in species demography away from arboreality [1, 209].

Of the three sources of variance present in the model, individual species variance accounts for approximately 80% of the observed, unmodeled variance (Fig. 2.3). Note that the individual variance was approximated using an simulation approach [107] because the Weibull distribution does not have a variance term. Both cohort and phylogenetic effects account for the other 20% of the observed variance. This result means that extinction risk has both temporal and phylogenetic aspects, as both contribute greater than 0% of the observed variability in the data [128].

The estimates for the individual cohort effects show a weak pattern of greater extinction risk in older Cenozoic cohorts compared to younger cohorts (Fig. 2.4). This potential slowdown

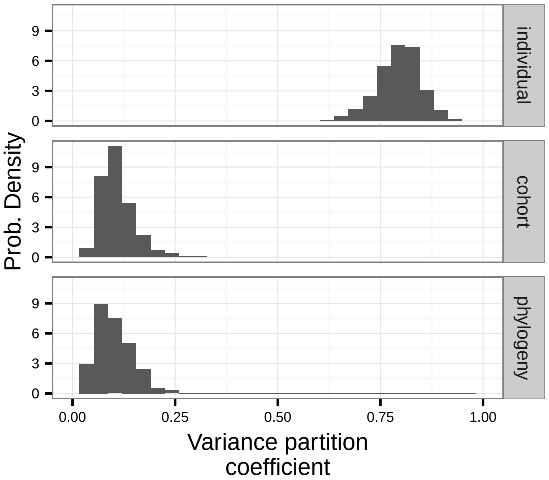


Figure 2.3: Estimates of the variance partitioning coefficients for the three different sources of variance: species, cohort, and phylogeny. Higher values correspond to greater contribution to total observed variance. Each of the estimates is a distribution of 1000 approximating simulations due to the model's non-normally distributed errors.

in extinction risk is consistent with previous analyses of marine invertebrates [258, 82] and mammals [7, 10]. There are two prevailing hypotheses as to the cause of this slowdown: 1) extinction risk is constant within, but varies between, clades so over time clades with low extinction rates increases in proportion of total diversity thus bringing down expected extinction risk; or 2) over time taxa increase in mean fitness and thus decrease in expected extinction risk [258]. The observed decrease in extinction risk with age, along with the variance partitioning results (Fig. 2.3) are consistent with both of these hypotheses with neither being more "important" than the other.

Interestingly, the shift from older cohorts with a higher extinction risk to younger cohorts with lower extinction risk is approximately at the Paleogene–Neogene boundary. Given the association with arboreality and increased extinction risk (Fig. 2.2), the decrease in expected extinction risk over time might relate to the preferential loss of arboreal taxa over the Cenozoic. However, because the model used here does not allow for time-varying effects, I cannot identify whether this boundary is associated with a shift in the direction or magnitude

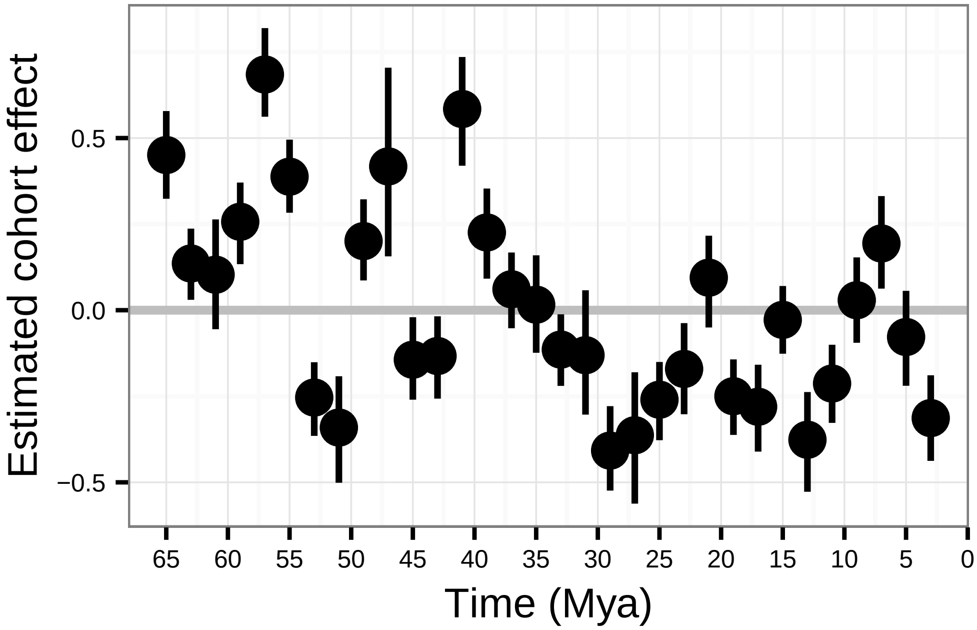


Figure 2.4: Summaries of posterior estimates of individual cohort effect depicted as medians and 80% credible intervals. High values correspond to shorter species durations while lower values correspond to greater species durations compared to the mean duration. Lines are placed at the middle of the 2 My origination cohorts.

of the expected effect of arboreality on extinction risk.

2.3 Discussion

My results indicate that Cenozoic North American mammal "generalists" are expected to have a lower extinction risk than "specialists." This implies that the diversification of specialized taxa would have required either a driven trend away from generality [199] or an increase in speciation rate relative to extinction rate [320]. This requires that specialist traits should somehow increase or be associated with increases in speciation rate, perhaps via niche partitioning or changes in habitat heterogeneity. For example, descendant species of omnivores many divide available prey items more finely or arboreal taxa may increase in both extinction and speciation rates via increases in habitat heterogeneity. Possible evidence to support this hypothesis would be to demonstrate differences in speciation rate associated with those traits analyzed here or other similar traits.

When these results are compared to factors contributing to increased extinction risk in extant mammals, there are some incongruencies. As expected, large range size is consistently associated with lower extinction risk in the modern world [172, 238, 95, 96]. While my analysis found body size to have almost no time-invariant effect on extinction risk, in extant mammals this is not necessarily the case as increased body size is associated with increased extinction risk [172, 238]. However, this pattern is partially clade dependent [95]. As stated earlier, higher trophic levels have been found to be associated with greater extinction risk in extant mammals [172, 238]. In contrast, I found that omnivores and carnivores have a lower expected extinction risk than either insectivores or herbivores (Fig. 2.2). Finally, phylogeny has been found to be a good predictor of differences in extinction risk in extant mammals as certain clades are at much higher risks than others [96]. This effect seems much greater in the Recent than for the whole Cenozoic, implying that current extinction risk is more phylogenetically concentrated than during times of background extinction levels during the Cenozoic.

Whether these incongruities are within the standard range of time-variant effects is unknown, though my comparisons do imply that current processes are different from those studied here. However, this is not a model of what makes taxa vulnerable during mass extinctions and that may account for these incongruities, assuming mass extinctions are qualitatively different than background extinction [134]. These results would also be inapplicable if the current biodiversity crisis is qualitatively different from either background or mass extinction as preserved in the fossil record.

By modeling how different ecologies and historical factors effect a species' expected extinction risk, it is possible to better understand what processes may have driven the resulting pattern of selection (i.e. diversity) while also providing a baseline for evaluating the current biodiversity crisis. This analysis finds support for the "survival of the unspecialized" hypothesis [298, 169] as a time-invariant generalization about extinction risk. I also find that there are substantial effects of both cohort and phylogeny on extinction risk, which supports the idea that the decrease in extinction risk [258] over time has both temporal and phylogenetic components. Additionally, I found evidence of increasing extinction risk with species age, the cause of which is unknown. These results show that, like prior mass extinction events in the fossil record, the current biodiversity crisis is qualitatively different from the previous period of background extinction in the fossil record [134].

2.4 Materials and Methods

2.4.1 Species occurrence and covariate information

Fossil occurrence information was downloaded from the Paleobiology Database (PBDB; http://paleodb.org/). Occurrence, taxonomic, stratigraphic, and biological information was downloaded for all North American mammals. This data set was filtered so that only occurrences identified to the species-level, excluding all "sp."-s. All aquatic and volant taxa were also excluded. Additionally, all occurrences without latitude and longitude information were excluded from the sample.

Species dietary and locomotor category assignments were done using the assignments in the PBDB, which were reassigned into coarser categories (Table 2.2). This was done to improve interpretability, increase sample size per category, and make results comparable to previous studies [152, 237].

All individual fossil occurrences were assigned to 2 My bins ranging through the entire

Cenozoic. Taxon duration was measured as the number of 2 My bins from the first occurrence to the last occurrence, inclusive. This bin size was chosen because it approximately reflects the resolution of the North American Cenozoic mammal fossil record [6, 10, 191]. Species originating in the youngest cohort, 0-2 My, were excluded from analysis because every species duration would be both left and right censored, which is illogical.

Species body size estimates in grams were sourced from a large selection of primary literature and database compilations. Databases used include the PBDB, PanTHERIA [156], and the Neogene Old World Mammal database (NOW; http://www.helsinki.fi/science/now/). Major sources of additional compiled body size estimates include [329, 30, 94, 198, 248, 307]. These were then supplemented with an additional literature search to try and fill in the remaining gaps. In many cases, species body mass was estimated using various published regression equations based on tooth or skull measurements (Table 2.3). If multiple specimens were measured, I used the mean of specimen measures as the species mean. See Appendix A for a complete list of mass estimates and sources.

Biogeographic network

Species geographic extent was measured as the mean of the relative number of bioprovinces occupied by a species for each 2 My bin the species was present. Bioprovinces were identified using a network-theoretic approach that has previously been applied to paleontological data [289, 341]. This approach relies on defining a biogeographic bipartite network of taxa and localities. In this study, taxa were defined as species and localities were grid cells from a regular lattice on a global equal-area cylinder map projection. The regular lattice was defined as a 70 x 34 global grid where each cell corresponds to approximately 250000 km^2. An advantage of this approach is that this approach reduces to occupancy when all localities are independent and to a single bioprovince when all localities are identical.

A biogeographic network was constructed for each of the 2 My bins used in this study. Emergent bioprovinces were then identified using the map equation [269, 268] as has been done before [289, 340, 341]. These bioprovinces correspond to taxa and localities that are more interconnected with each other than with other nodes.

The map projection and regular lattice were made using shape files from http://www.naturalearthdata.com/ and the `raster` package for R [122]. Bioprovince identification was done using the map equation as implemented in the `igraph` package for R [56].

Supertree

As there is no single, combined formal phylogenetic hypothesis of all Cenozoic fossils mammals from North America, it was necessary to construct a semi-formal supertree. This was done by combining taxonomic information for all the observed species and a few published phylogenies using matrix representation parsimony [25]. For further explanation of the methodology used to construct this supertree, please see the Supplementary information in Section 2.5.

2.4.2 Survival model

Presented here is the model development process used to formulate the two survival models used in this study. First, define y as a vector of length n where the ith element is the duration of species i, where $i = 1, \cdots, n$.

The simplest survival model where durations are assumed to follow an exponential distribution with a single “rate” or inverse-scale parameter λ [160]. This is written

$$
\begin{aligned}
p(y|\lambda) &= \lambda \exp(-\lambda y) \\
y &\sim \text{Exp}(\lambda). \qquad (2.1)
\end{aligned}
$$

The exponential distribution corresponds to situations where extinction risk is independent of age. To understand this, we need to define two functions: the survival function $S(t)$ and the hazard function $h(t)$. $S(t)$ is the probability that a species having existed for t 2 My bins will not have gone extinct while $h(t)$ corresponds to the instantaneous extinction rate for some taxon age t [160]. For an exponential model, $S(t)$ is

$$S(t) = \exp(-\lambda t) \qquad (2.2)$$

and $h(t)$ is defined

$$h(t) = \lambda \qquad (2.3)$$

The choice of the exponential distribution corresponds directly to the Law of Constant Extinction [335] as the right side of Eq. 2.3 does not depend on species age t.

The current sampling statement (Eq. 2.1) assumes that all species share the same rate parameter with no variation. To allow for variation in λ associated with relevant covariate information like species body size, λ is reparameterized as $\lambda_i = \exp(\sum \beta^T \mathbf{X}_i)$ with i indexing a given observation and its covariates, β is a vector of regression coefficients, and $\mathbf{X}$ is a matrix of covariates. This is a standard regression approach, where one column of $\mathbf{X}$ is all 1-s and its corresponding β coefficient is the intercept.

$\mathbf{X}$ is an $n \times K$ matrix of species-level covariates. Three of the covariates of interest are the logit of mean relative occupancy, and the logarithm of body size (g). The discrete covariate index variables of dietary and locomotor category were transformed into $n \times (k-1)$ matrices where each column is an indicator variable (0/1) for that species's category, k being the number of categories of the index variable (3 and 4, respectively). Only $k-1$ indicator variables are necessary as the intercept takes on the remaining value. For example, the difference in effect of arboreality versus scansoriality on extinction risk, given that arboreality is the reference category, is the coefficient for the scansorial indicator variable as that is the difference between between the effect of arboreality (the intercept β_0) and scansoriality (the intercept + scansorial effect β_s); Fig. 2.2). Finally, a vector of 1-s was included in the matrix $\mathbf{X}$ whose corresponding β coefficient is the intercept, making K equal eight.

β is the vector of regression coefficients. The intercept term was given a weak normal prior, $\beta_0 \sim \mathcal{N}(0, 10)$ while all of these other coefficients were slightly more informative priors, e.g. $\beta_{mass} \sim \mathcal{N}(0, 5)$. These priors were chosen because it is expected that the effect size of each variable on duration will be small, as is generally the case with binary covariates [102].

Regression coefficients are not directly comparable without first standardizing the input variables to have equal standard deviations. This is accomplished by subtracting the mean of the covariate from all values and then dividing by the standard deviation, resulting in a variable with mean of zero and a standard deviation of one. This linear transform greatly improves the interpretability of the coefficients as expected change in mean duration given a difference of one standard deviation in the covariate [275]. Additionally, this makes the intercept directly interpretable as the estimate of mean (transformed) σ (Eq. 2.7). However, because the expected standard deviation for a random binary variable is 0.5, in order to make comparisons between the binary and continuous variables, the continuous inputs were divided by twice their standard deviation [100].

Origination cohort is defined as the group of species which all originated during the same 2 My temporal bin. Because the most recent temporal bin, 0-2 My, was excluded, there are

32 total cohorts. The effect of origination cohort j was modeled with each group being a sample from a common cohort effect, η, which was considered normally distributed with mean 0, and standard deviation σ_c. The value of σ_c was then estimated from the data itself, corresponding to the amount of pooling in the individual estimates of η_j. This approach is a conceptual and statistical unification between dynamic and cohort survival analysis in paleontology [77, 252, 251, 336, 21], with σ_c acting as a measure of compromise between these two end members. The choice of the half-Cauchy prior for σ_c follows [99].

$$\begin{aligned}\eta_j &\sim \mathcal{N}(0, \sigma_c)\\ \sigma_c &\sim \mathrm{C}^{+}(0, 2.5)\end{aligned}$$

The impact of shared evolutionary history, or phylogeny, was modeled as an individual effect where each observation, i, is modeled as a multivariate normal, h, where the covariance matrix Σ is known up to a constant, σ_p^2 [184, 128]. This is written

$$\begin{aligned}h &\sim \mathrm{MVN}(0, \mathbf{\Sigma})\\ \mathbf{\Sigma} &= \sigma_p^2 \mathbf{V}_{phy}\\ \sigma_p &\sim \mathrm{C}^{+}(0, 2.5).\end{aligned}$$

$\mathbf{V}_{phy}$ is the phylogenetic covariance matrix defined as an $n \times n$ matrix where the diagonal elements are the distance from root to tip, in branch length, for each observation and the off-diagonal elements are the amount of shared history, measured in branch length, between observations i and j. σ_p was given a weakly informative half-Cauchy hyperprior. Note that because the phylogeny used here is primarily based on taxonomy, estimates of σ_p represent minimum estimates [184, 128]. Improved phylogenetic estimates of all fossil Cenozoic mammals would greatly improve this estimate.

To relax the assumption of age-independent extinction of the Law of Constant Extinction, the Weibull distribution is substituted for the exponential [160]. The Weibull distribution has a shape parameter α and scale parameter σ. Conceptually, σ is the inverse of λ. α modifies the impact of taxon age on extinction risk. When $\alpha > 1$ then $h(t)$ is a monotonically increasing function, but when $\alpha < 1$ then $h(t)$ is a monotonically decreasing function. When $\alpha = 1$ then the Weibull distribution is equivalent to the exponential.

The Weibull distribution and sampling statement were defined

$$\begin{aligned} p(y|\alpha, \sigma) &= \frac{\alpha}{\sigma} \left(\frac{y}{\sigma}\right)^{\alpha - 1} \exp\left(-\left(\frac{y}{\sigma}\right)^{\alpha}\right) \\ y &\sim \text{Weibull}(\alpha, \sigma). \end{aligned} \tag{2.4}$$

The corresponding $S(t)$ and $h(t)$ functions are defined

$$S(t) = \exp\left(-\left(\frac{t}{\sigma}\right)^{\alpha}\right) \tag{2.5}$$

$$h(t) = \frac{\alpha}{\sigma}\left(\frac{t}{\sigma}\right)^{\alpha - 1}. \tag{2.6}$$

To allow for σ to vary with a given observation's covariate information it is reparameterized in a similar fashion to λ with a few key differences. Because $\sigma = 1/\lambda$ in order to preserve the interpretation of β, while taking α into account, σ is reparameterized as

$$\sigma_i = \exp\left(\frac{-\beta}{\alpha}\right). \tag{2.7}$$

Given the above, the survival model was then fit in a Bayesian context using both exponential and Weibull distributions. The Weibull's α parameter was assumed constant across species, which is standard practice in survival analysis [160]. α was given a weakly informative half-

Cauchy (C^{+}) prior. σ was reparameterized as an exponentiated regression model (Eq. 2.7). This was further expanded (Eq. 2.8) to allow for two hierarchical factors as discussed above. This is written

$$\sigma_i = \exp\left(\frac{-(h_i + \eta_{j[i]} + \sum \beta^T \mathbf{X}_i)}{\alpha}\right) \tag{2.8}$$

where equivalent statement for the exponential distribution is defined

$$\lambda_i = \exp\left(h_i + \eta_{j[i]} + \sum \beta^T \mathbf{X}_i)\right). \tag{2.9}$$

An important part of survival analysis is the inclusion of censored observations where the failure time has not been observed [133, 160]. The most common censored observation is right censored, where the point of extinction had not yet been observed in the period of study, such as taxa that are still present in the most recent time bin (0-2 My). Left censored observations, on the other hand, correspond to observations that went extinct any time between 0 and some known point. To account for this uncertainty, the probability of a left censored observation is found by integrating over all possible durations between 0 and 1 time bin. For an explanation of how censored observations are included in the sampling statement, please see the Supplementary information in Section 2.5.

2.4.3 Estimation

Parameter posteriors were approximated using a Markov-chain Monte Carlo (MCMC) routine implemented in the Stan programming language [318]. Stan implements a version of Hamiltonian Monte Carlo called the No-U-Turn sampler [124]. Posterior approximation was done using four parallel MCMC chains run for 30000 steps, thinned to every thirtieth sample, split evenly between warm-up and sampling. Convergence was evaluated using the scale reduction factor, $\hat{R}$. Values of $\hat{R}$ close to 1, or less than or equal to 1.1, indicate approximate convergence. Convergence means that the chains are approximately stationary and the samples are well mixed [101].

2.4.4 Posterior evaluation

The most basic assessment of model fit is that simulated data generated given the model should be similar to the observed. This is the idea behind posterior predictive checks. Using the covariates from each of the observed durations, and randomly drawn parameter estimates from their marginal posteriors, a simulated data set y^{rep} was generated. This process was repeated 1000 times and the distribution of y^{rep} was compared with the observed [101]. For results from the posterior predictive tests used in this study, please see the Supplementary information in Section 2.5.

The exponential and Weibull models were compared for out-of-sample predictive accuracy using the widely-applicable information criterion (WAIC) [357]. Because the Weibull model reduces to the exponential model when $\alpha = 0$, our interest is not in choosing between these models. Instead comparison of WAIC values is useful for better understanding the effect of model complexity on out-of-sample predictive accuracy. An explanation of how WAIC is calculated is presented in the Supplementary information (Section 2.5) following the recommended "WAIC 2" formulation [101].

There are three different variance components in this model: sample component, cohort σ_c^2, and phylogenetic σ_p^2. Partitioning the variance between these sources allows the relative amount of unexplained variance of the sample to be compared. The sample component is similar to the residual variance from a normal linear regression. However, the Weibull based model used here (Eq. 2.4) does not include an estimate of the variance similar to the squared scale term of the a Normal distribution. Instead, the sample component was approximated via a simulation approach modified from [107]. For explanation of this method, please see Supplementary information (Section 2.5).

I used variance partitioning coefficients (VPC) to estimate the relative importance of the different variance components [102]. Phylogenetic heritability, h_p^2 [184, 128], is identical to the VPC of the phylogenetic effect. Phylogenetic heritability is a measure of how shared evolutionary history impacts differences in individual species trait values (e.g. duration). This is a broad sense "heritability" as it combines both genetic inheritance and other, non-genetic shared history factors. Importantly, because phylogenetic effect was estimated using a principally taxonomy based tree the estimates derived here can be considered minimum estimates of the phylogenetic effect.

2.5 Supplemental information for "Death and taxa"

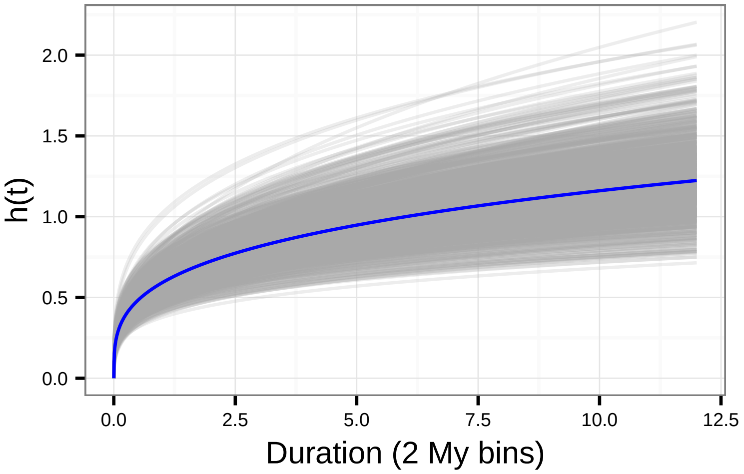


Figure 2.5: 1000 estimates of the hazard function ($h(t)$ for the observed species mean (grey), along with the median estimated hazard function (blue). $h(t)$ is an estimate of the rate at which a species of age t is expected to go extinct. Hazard functions were estimated from random draws from the estimated posterior distributions and evaluated with all covariate information set to 0, which corresponds to the expected duration of the mean species.

2.5.1 Supertree inference

As there is no single, combined formal phylogenetic hypothesis of all Cenozoic fossils mammals from North America, it was necessary to construct a semi-formal supertree. This was done by combining taxonomic information for all the observed species and a few published phylogenies.

The initial taxonomic classification of the observed species was based on the associated taxonomic information from the PBDB. This information was then updated using the Encyclopedia of Life (http://eol.org/) which collects and collates taxonomic information in a single database. This was done programatically using the `taxize` package for R [41]. Finally, this taxonomic information was further updated using a published taxonomy of fossil mammals [146, 145].

This taxonomy serves as an initial phylogenetic hypothesis which was then combined with a selection of species-level phylogenies [25, 248] in order to better constrain a minimum estimate of the actual phylogenetic relationships of the species. The supertree was inferred via matrix representation parsimony implemented in the `phytools` package for R [259]. Of the two most parsimonious trees found, I used only one for analysis.

Polytomies were resolved in order of species first appearance in order to minimize stratigraphic gaps. The resulting tree was then time scaled using the `paleotree` package via the "minimum branch length" approach with a minimum length of 0.1 My [17]. The minimum length is necessary to avoid zero-length branches which cause the phylogenetic covariance matrix not to be positive definite, which is important for computation (see below). While other time scaling approaches are possible [120, 18] this method was chosen for its simplicity and not requiring additional information about diversification rates which are the interest of this study.

2.5.2 Modeling censored observations

Censored data are modeled using the survival function of the distribution, $S(t)$, defined earlier for the Weibull distribution (Eq. 5, 6) with σ defined as above (Eq. 8, 9). $S(t)$ is the probability that an observation will survive longer than a given time t.

The likelihood of uncensored observations is evaluated as normal using equation 4 while right censored observations are evaluated at $S(t)$ and left censored observations are evaluated at $1 - S(t)$. Note, $1 - S(t)$ is equivalent to the cumulative distribution function and $S(t)$ is equivalent to the complementary cumulative distribution function [101].

The final sampling statement/likelihood for both uncensored and both right and left censored observations is then written

$$L \propto \prod_{i \in C} \text{Weibull}(y_i|\alpha, \sigma) \prod_{j \in R} S(y_j|\alpha, \sigma) \prod_{k \in L} \left(1 - S(y_k|\alpha, \sigma)\right),$$

where C is the set of uncensored observations, R is the set of right censored observations, and L is the set of left censored observations.

2.5.3 Deviance residuals

In standard linear regression, residuals are defined as $r_i = y_i - y_i^{est}$. For the model used here, this definition is inadequate. The equivalent values for survival analysis are deviance residuals. To define how deviance residuals are calculated, we first define the cumulative hazard function [160]. Given $S(t)$, we define the cumulative hazard function as

$$\Lambda(t) = -log\left(S\left(t\right)\right).$$

Next, we define martingale residuals m as

$$m_i = I_i - \Lambda(t_i).$$

I is the inclusion vector of length n, where $I_i = 1$ means the observation is completely observed and $I_i = 0$ means the observation is censored. Martingale residuals have a mean of 0, range between 1 and $-\infty$, and can be viewed as the difference between the observed number of deaths between 0 and t_i and the expected number of deaths based on the model. However, martingale residuals are asymmetrically distributed, and can not be interpreted in the same manner as standard residuals.

The solution to this is to use the deviance residuals, D. This is defined as a function of martingale residuals and takes the form

$$D_i = \text{sign}(m_i)\sqrt{-2[m_i + I_i log(I_i - m_i)]}.$$

Deviance residuals have a mean of 0 and a standard deviation of 1 by definition.

2.5.4 Variance partitioning

I calculated VPC using a resampling approach based on [107]. The procedure is as follows:

1. Simulate w (50,000) values of η; $\eta \sim \mathcal{N}(0, \sigma_c)$.
2. For a given value of $\beta^T \mathbf{X}$, calculate σ^{c*} (Eq. 7) for all w simulations, holding h constant at 0.
3. Calculate υ_c, the Weibull variance (Eq. 2.10) of each element of σ^{c*} with α drawn from the posterior estimate.

4. Simulate w values of h; $h \sim \mathcal{N}(0, \sigma_p)$.

5. For a given value of $\beta^T \mathbf{X}$, calculate σ^{p*} (Eq. 7) for all w simulations, holding η constant at 0.

6. Calculate υ_p, the Weibull variance (Eq. 2.10) of each element of σ^{p*} with α drawn from the posterior estimate.

7. $\sigma^2_{y*} = \frac{1}{2}\left(\left(\frac{1}{w}\sum_i^w \upsilon_{pi}\right) + \left(\frac{1}{w}\sum_j^w \upsilon_{cj}\right)\right)$.

8. $\sigma^2_{c*} = var(\upsilon_c)$ and $\sigma^2_{p*} = var(\upsilon_p)$.

The simulated values of h were drawn from a univariate normal distribution because each simulated value is in isolation, so there is no concern of phylogenetic autocorrelation. The chosen value for $\beta^T \mathbf{X}$ was a draw from the posterior estimate of the intercept. Because input variables were standardized prior to model fitting, the intercept corresponds to the estimated effect on survival of the sample mean.

Weibull variance is calculated as

$$var(x) = \sigma^2 \left(\Gamma\left(1 + \frac{2}{\alpha}\right) - \left(\Gamma\left(1 + \frac{1}{\alpha}\right)\right)^2\right), \tag{2.10}$$

where Γ is the gamma function.

The variance partitioning coefficients are then calculated, for example, as $VPC_{phylo} = \frac{\sigma^2_{p*}}{\sigma^2_{y*} + \sigma^2_{c*} + \sigma^2_{p*}}$ and similarly for the other components.

2.5.5 Widely applicable information criterion

WAIC can be considered fully Bayesian alternative to the Akaike information criterion, where WAIC acts as an approximation of leave-one-out cross-validation which acts as a measure of out-of-sample predictive accuracy [101]. The following explanation uses the "WAIC 2" formulation recommended by [101].

WAIC is calculated starting with the log pointwise posterior predictive density calculated as

$$\text{lppd} = \sum_{i=1}^{n} \log \left(\frac{1}{S} \sum_{s=1}^{S} p(y_i | \Theta^S) \right), \tag{2.11}$$

where n is sample size, S is the number posterior simulation draws, and Θ represents all of the estimated parameters of the model. This is similar to calculating the likelihood of each observation given the entire posterior.

A correction for the effective number of parameters is then added to lppd to adjust for overfitting. The effective number of parameters is calculated, following derivation and recommendations of [101], as

$$p_{\text{WAIC}} = \sum_{i=1}^{n} V_{s=1}^{S} (\log p(y_i | \Theta^S)). \tag{2.12}$$

where V is the sample posterior variance of the log predictive density for each data point.

Given both equations 2.11 and 2.12, WAIC is then calculated

$$\text{WAIC} = \text{lppd} - p_{\text{WAIC}}. \tag{2.13}$$

When comparing two or more models, lower WAIC values indicate better out-of-sample predictive accuracy. Importantly, WAIC is just one way of comparing models. When combined with posterior predictive checks it is possible to get a more complete understanding of model fit.

2.5.6 Results from posterior predictive checks

With all marginal posterior estimates having converged ($\hat{R} < 1.1$) it is possible to examine the quality of model fit (Table 1). If the model is an adequate descriptor of the observed data, then relatively confident inference can be made [101].

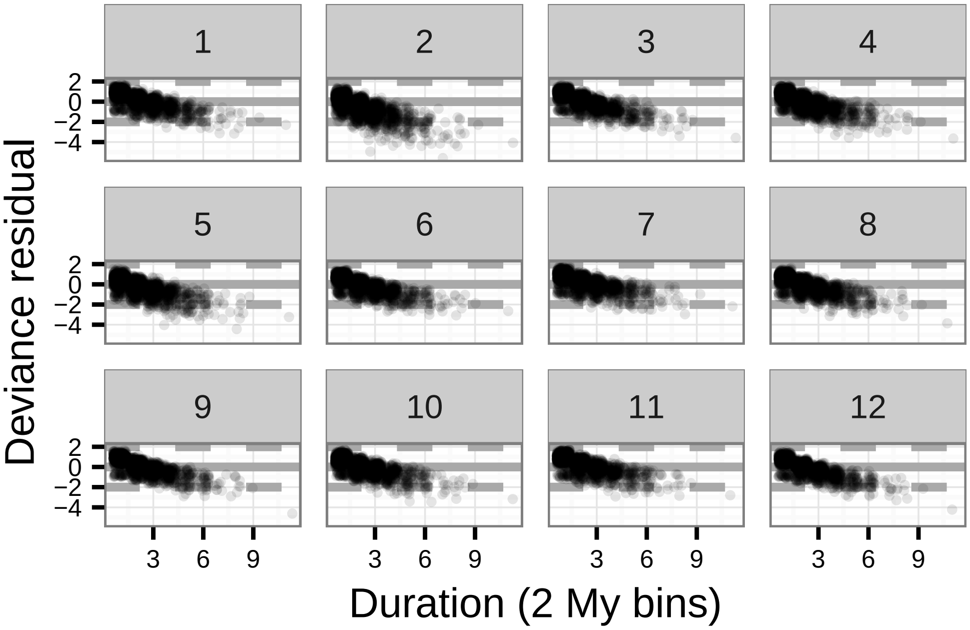


Figure 2.6: Deviance residuals from the fitted survival model compared to observed durations. Each graph depicts the residuals from single draws from the posterior distributions of all estimated parameters. Positive values indicate an underestimate of the observed duration, while negative values indicate an overestimate of the observed duration. A small amount of noise is added to each point to increase clarity. Twelve different examples are provided here to indicate consistency across multiple realizations.

Visual examination of the deviance residuals from twelve different sets of posterior predictive simulations indicates a systematic weakness estimating durations greater than 3 2-My bins (Fig. 2.6). However, the comparison of posterior predictive estimates of the 25th, 50th, and

75th quantiles to the observed indicate adequate fit. (Fig. 2.7). Importantly, this indicates that the model has approximate fit for 50+% of the data. Because, the inferred model can be inferred to be approximately adequate at capturing the observed variation.

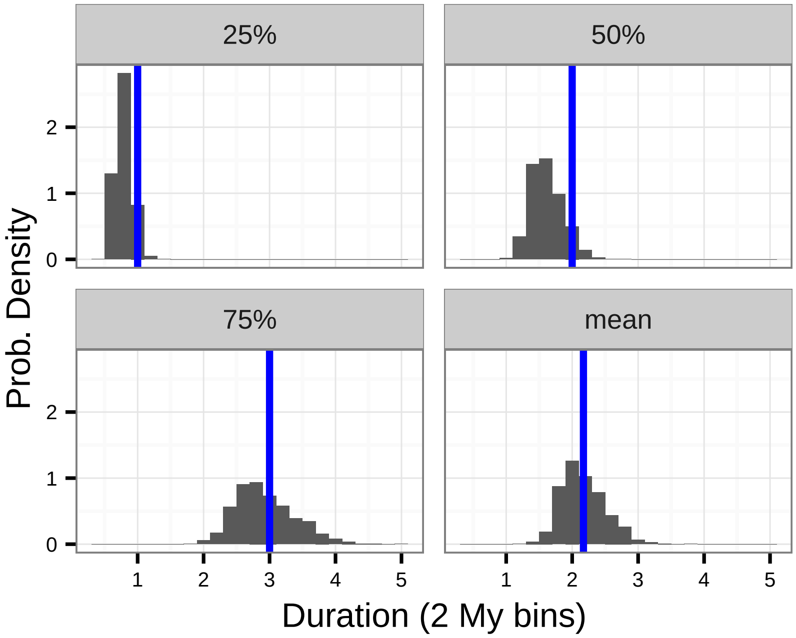


Figure 2.7: The results of additional posterior predictive checks for four summaries of the observed durations, as labeled. Blue vertical lines indicate the observed value. None of the observed values are significantly different from the posterior predictive distributions.

The Wiebull model (6140.37) also had a much lower WAIC score than the Exponential model (16697.35). This large a difference indicates that the Weibull model probably has the lower out-of-sample predictive accuracy of the two.

2.5.7 Data quality concerns

A concern with using the PBDB as a primary data source, though this concerns are general to most paleontological data, is that the results are an artifact of taxonomy or the database itself [346]. However, to obtain the results obtained in this analysis there would need to be a systematic error in assignemnts of all of diet, locomotor, and taxonomic categories for a large portion of the close to 2000 sampled species. It is important to note that species included have to have body size information, much of which is found from other sources (see Dataset S1). this means that, for many taxa, that species name has to appear in occur in more than one place. this is a strong filter for misspellings and potentially invalid taxa. Additionally, given that most mammal fossils are teeth which allows for relatively accurate dietary category assignement.

A possible concern, however, is that omnivorous taxa have feature poor morphology and thus longer durations may reflect a single anagenetic lineage as opposed to a single "species." However it is possible to consider that, from a population genetic perspective, it can be argued that a single unbranching lineage is still a single biological "unit."

2.5.8 Concerns surrounding estimates of α

The estimate of the Weibull shape parameter, α, is greater than 1 meaning that extinction risk is expected to increase with taxon age (Table 1). As the value of α is between 1 and 1.5, extinction risk for a given species only gradually increases with age (Fig. 2.5). There are three possible explanations for this result: 1) older taxa being out competed by younger taxa [347]; or 2) this is an artifact of the minimum resolution of the fossil record [284].

An additional concern is that there may be an upward bias in estimates of α at this sample size, similar to that for scale parameters [101]. The plausibility of third possibility in this example can be explored in simulation. I simulated from 10, 100, 1000, and 10000 samples from a Weibull($alpha = 1.3$, $\sigma = 1$) 100 times each. For each of these simulated datasets, I then estimated the values of α and σ in a simple maximum likelihood context in order to just get the model estimate. The modal estimates of both parameters for the simulated datasets were then compared to the known values (Fig. 2.8). The results from these simulations demonstrate that the estimates of α in the above analyses (Table 1) should not be particularity biased based on my sample size of approximately 2000 species.

The model used in this analysis, however, is unable to distinguish between the remaining two hypotheses [284, 347]. Further work on how to better constrain estimates α is necessary. A possibly is somehow incorporating these hypotheses as prior information.

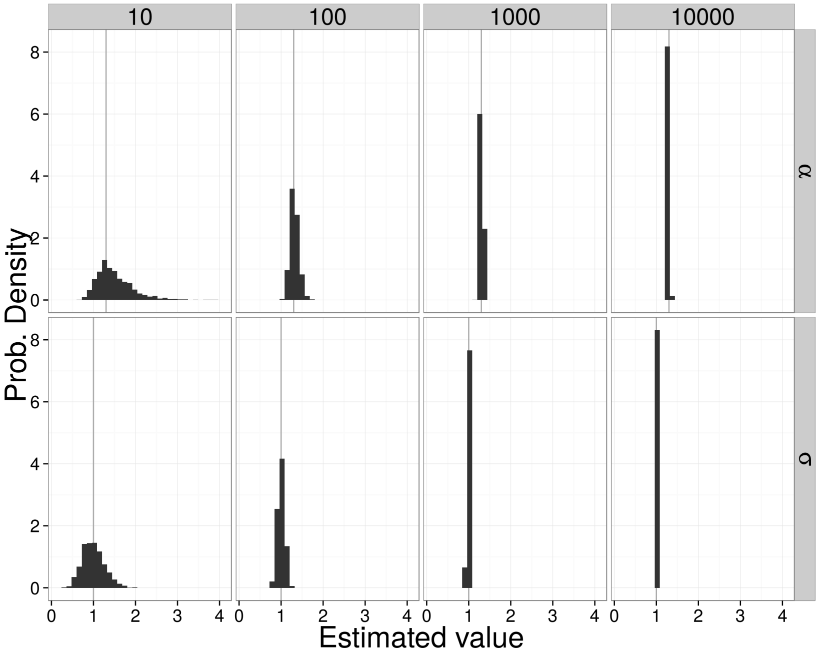


Figure 2.8: Comparison of maximum likelihood estimates of shape (α) and scale (σ) parameters from 1000 simulated data sets from 4 different sample sizes. Vertical lines are the actual parameter value used to generate the data. When sample size is approximately 100 or greater, estimates are not overly biased.

Table 2.1: Marginal posterior estimates for the parameters of interested based on 1000 posterior samples. The intercept β_0 can also be interpreted as the estimate for the mean observed species. The remaining β values can be interpreted as the effect of a trait on the expected species duration as expressed as deviation from the mean. The categorical variables are binary index variables where an observation is of that category or not. See Equation 2.6 for an explanation of the effect of α on extinction risk. $\hat{R}$ values of less than 1.1 indicate approximate chain convergence for the posterior samples.

parameter	effect	mean	sd	2.5%	25%	50%	75%	97.5%	$\hat{R}$
α	"age"	1.29	0.03	1.23	1.27	1.29	1.31	1.36	1.00
β_0	arboreal/carnivore	-0.78	0.14	-1.05	-0.87	-0.78	-0.68	-0.51	1.00
β_o	occupancy	-0.53	0.08	-0.69	-0.59	-0.53	-0.48	-0.38	1.00
β_{size}	body size	-0.05	0.05	-0.14	-0.08	-0.05	-0.01	0.05	1.00
β_g	ground dwelling	-0.28	0.10	-0.47	-0.34	-0.28	-0.21	-0.09	1.00
β_s	scansorial	-0.22	0.11	-0.43	-0.29	-0.22	-0.14	-0.00	1.00
β_h	herbivore	0.09	0.09	-0.09	0.03	0.09	0.14	0.27	1.00
β_i	insectivore	0.10	0.11	-0.11	0.03	0.10	0.17	0.31	1.00
β_o	omnivore	-0.12	0.11	-0.33	-0.19	-0.12	-0.05	0.09	1.00
σ_c	sd cohort	0.33	0.06	0.23	0.29	0.33	0.37	0.48	1.00
σ_p	sd phylogeny	0.11	0.05	0.03	0.07	0.10	0.14	0.23	1.03

Table 2.2: Species trait assignments in this study are a coarser version of the information available in the PBDB. Information was coarsened to improve per category sample size and uniformity and followed this table.

This study		PBDB categories
Diet	Carnivore	Carnivore
	Herbivore	Browser, folivore, granivore, grazer, herbivore.
	Insectivore	Insectivore.
	Omnivore	Frugivore, omnivore.
Locomotor	Arboreal	Arboreal.
	Ground dwelling	Fossorial, ground dwelling, semifossorial, saltatorial.
	Scansorial	Scansorial.

Table 2.3: 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	[165]
General	$\log(m) = 2.9677x - 5.6712$	mandible length	[91]
General	$\log(m) = 3.68x - 3.83$	skull length	[183]
Carnivores	$\log(m) = 2.97x + 1.681$	lower m1 length	[337]
Insectivores	$\log(m) = 1.628x + 1.726$	lower m1 area	[26]
Insectivores	$\log(m) = 1.714x + 0.886$	upper M1 area	[26]
Lagomorph	$\log(m) = 2.671x - 2.671$	lower toothrow area	[329]
Lagomorph	$\log(m) = 4.468x - 3.002$	lower m1 length	[329]
Marsupials	$\log(m) = 3.284x + 1.83$	upper M1 length	[108]
Marsupials	$\log(m) = 1.733x + 1.571$	upper M1 area	[108]
Rodentia	$\log(m) = 1.767x + 2.172$	lower m1 area	[165]
Ungulates	$\log(m) = 1.516x + 3.757$	lower m1 area	[201]
Ungulates	$\log(m) = 3.076x + 2.366$	lower m2 length	[201]
Ungulates	$\log(m) = 1.518x + 2.792$	lower m2 area	[201]
Ungulates	$\log(m) = 3.113x - 1.374$	lower toothrow length	[201]

CHAPTER 3

HOW MACROECOLOGY AFFECTS MACROEVOLUTION: THE INTERPLAY BETWEEN EXTINCTION INTENSITY AND TRAIT-DEPENDENT EXTINCTION IN BRACHIOPODS

As extinction intensity increases, how do the effects of traits on taxonomic survival change? Does the extinction rate associated with certain traits increase while that of others decreases? Using a hierarchical Bayesian approach, I develop a model of how the effects of biological traits on extinction risk can vary with respect to extinction intensity, origination cohort (i.e. time of origination), and in relation to each other. The emergent traits traits I analyze for each brachiopod genus are geographic range, affinity for epicontinental seas versus open ocean environments, and body size. Additionally, I estimate the effects of environmental generalization versus specialization on taxonomic survival by allowing environmental preference to have a nonlinear effect on duration. My analytical framework eschews the traditional distinction between background and mass extinction, and instead considers extinction intensity as a continuum. I find that the cohort-specific effects of geographic range and environmental preference are negatively correlated with baseline extinction intensity. Additionally, I find support for greater survival of environmental generalists versus specialists in all origination cohorts. These results support the conclusion that for Paleozoic brachiopods, as extinction intensity increases overall extinction selectivity increases.

3.1 Introduction

Extinction is one half of the diversification process [255, 321, 320], second only to speciation or origination; it can also be the ultimate manifestation of selection as a taxon with a beneficial trait should persist for longer on average than a taxon without that beneficial trait [246, 137, 255, 320].

While estimation of both trait-dependent speciation and extinction rates from phylogenies of extant taxa has grown dramatically [190, 75, 105, 106, 247, 316, 315, 317], there are two major ways to estimate trait-dependent extinction: analysis of phylogenies, and analysis of the fossil record. These two directions, phylogenetic comparative and paleobiological, are complementary and intertwined in the field of macroevolution [246, 137, 131]. In the case of extinction, analysis of the fossil record has the distinct advantage over phylogenies of only extant taxa because extinction is observable; this means that extinction rate is possible to estimate [241, 239, 174]. The approach used here is thus complementary to the analysis of trait-dependent extinction based on a phylogeny.

Jabonksi [134] observed that for bivalves at the end Cretaceous mass extinction event, the effects of some biological traits on taxonomic survival decreased. However, this pattern was not the case for the effect of geographic range on survival [134, 222]. There are multiple possible macroevolutionary mechanisms which may underlie this pattern: the effect of geographic range on survival remains constant and those of other biological traits decrease, the effect of geographic range on survival increases and those of other biological traits stay constant, or the effects of all traits decrease potentially by different degrees.

While Jablonski [134] phrased his conclusions in terms of background versus mass extinction, these states are not distinguishable in terms of extinction rate alone; my analysis treats the time period analyzed as part of the same continuum [349, 222, 296]. Additionally, in order to test the proposed macroevolutionary mechanism behind this scenario [134]; not only do the taxon trait effects need to be modeled, but the correlations between trait effects must be modeled as well.

Here I model brachiopod taxon durations because trait based differences in extinction risk should manifest as differences in taxon durations. Brachiopods are an ideal group for this study as they are are well known for having an exceptionally complete fossil record [89, 80]. I focus on the brachiopod record from the post-Cambrian Paleozoic, from the start of the Ordovician (approximately 485 My) through the end Permian (approximately 252 My) as this represents the time of greatest global brachiopod diversity [8] meaning a large sample size for this analysis.

The analysis of taxon durations, or time from origination to extinction, falls under the purview of survival analysis, a field of applied statistics commonly used in health care and engineering [160] and also has a long history in paleontology [298, 299, 335, 336, 54]. I adopt a hierarchical modeling approach [102, 101], which represents both a conceptual and statistical unification of the paleontological dynamic and cohort survival analytic approaches [335, 336, 252, 251, 77, 21, 294, 54, 68].

3.1.1 Factors affecting brachiopod survival

Conceptually, taxon survival can be considered an aspect of "taxon fitness" [49, 220]. Traits associated with taxon survival are thus examples of species (or higher-level) selection, as differences in survival are analogous to differences in fitness. The traits analyzed here are all examples of emergent and aggregate traits [137, 246]; specifically I analyze genus-level traits. Emergent traits are those which are not measurable at a lower level (e.g. species versus individual organism) such as geographic range, or even fossil sampling rate. Aggregate traits, like body size or environmental preference, are the average of a shared trait across all members of a lower level.

Geographic range is widely considered the most important biological trait for estimating differences in extinction risk at nearly all times, with large geographic range associated with low extinction risk [134, 135, 138, 222, 137, 114, 73]. This stands to reason even if extinction is completely at random; a taxon with an unrestricted range is less likely to go extinct at random than a taxon with a restricted range.

Epicontinental seas are a shallow-marine environment where the ocean has spread over the continental interior or craton with a depth typically less than 100m. In contrast, open-ocean coastline environments have much greater variance in depth, do not cover the continental craton, and can persist during periods of low sea level [203]. Because of this, a simple hypothesis that taxa which favor epicontinental seas would be at great risk during periods of low sea levels, such as during glacial periods, when epicontinental seas are drained. During the Paleozoic (approximately 541–252 My), epicontinental seas were widely spread globally but declined over the Mesozoic (approximately 252–66 My) and have nearly disappeared during the Cenozoic (approximately 66–0 My) as open-ocean coastlines became the dominant shallow-marine setting [286, 226, 203, 155]. Taxa in epicontinental environments could also have a greater extinction susceptibility than taxa in open-ocean environments due to anoxic events due to enhanced water stratification or poor water circulation [225].

Miller and Foote [203] demonstrated that during several mass extinctions taxa associated with open-ocean environments tend to have a greater extinction risk than those taxa associated with epicontinental seas. During periods of background extinction, however, they found no consistent difference between taxa favoring either environment. Miller and Foote [203] hypothesize that open-ocean taxa may have a greater extinction rate because these environments would be more strongly affected by waterborne hazards such as fallout from impacts or volcanic events which would propagate more quickly than in epicontinental environments with sluggish circulation. These two environment types represent the primary identifiable environmental dichotomy observed in ancient marine systems [203, 286]. Given these findings, I would hypothesize that as extinction risk increases, the extinction risk associated with open-ocean environments should generally increase.

Because environmental preference is defined here as the continuum between occurring exclusively in open-ocean environments versus epicontinental environments, intermediate values are considered "generalists" in the sense that they favor neither end member. A long-standing hypothesis is that generalists or unspecialized taxa will have greater survival than specialists [298, 169, 170, 217, 218, 21, 309]. Because of this, the effect of environmental preference was modeled as a quadratic function where a concave down relationship between preference and expected duration indicates that generalists are favored over specialists end-members.

Body size, measured as shell length, is also considered as a trait that may potentially influence extinction risk [223, 115]. Body size is a proxy for metabolic activity and other correlated life history traits [223]. Harnik *et al.* [116] analyzed the effect of body size selectivity in Devonian brachiopods in both a phylogenetic and non-phylogenetic context; finding that body size was not found to be associated with differences in taxonomic duration. It has also been found that, at least in the case of some bivalve subclades, body size can be as important a factor as geographic range size in determining extinction risk [115]. Given these results, I expect that if body size has any effect on brachiopod taxonomic survival it is very small.

It is well known that, given the incompleteness of the fossil record, the observed duration of a taxon is an underestimate of that taxon's true duration [313, 348, 351, 173, 9, 89]. Because of this, the concern is that a taxon's observed duration may reflect its relative chance of being sampled and not any of the effects of the covariates of interest. In this case, for sampling to be a confounding factor there must be consistent relationship between the quality of sampling of a taxon and its apparent duration (e.g. greater sampling, longer duration). If there is no relationship between sampling and duration then interpretation can be made clearly; while observed durations are obviously truncated true durations, a lack of a relationship would indicate that the amount and form of this truncation is not a major determinant of the taxon's apparent duration. By including sampling as a covariate in the model, this effect can be quantified and can be taken into account when interpreting the estimates of the effects of the other covariates.

3.2 Materials and Methods

3.2.1 Fossil occurrence information

The brachiopod dataset analyzed here was sourced from the Paleobiology Database (http://www.paleodb.org) which was then filtered based on taxonomic (Rhychonelliformea: Rhynchonellata, Chileata, Obolellida, Kutorginida, Strophomenida, Spiriferida), temporal (post-Cambrian Paleozoic), stratigraphic, and other occurrence information used in this analysis. Analyzed occurrences were restricted to those with paleolatitude and paleolongitude coordinates, assignment to either epicontinental or open-ocean environment, and belonging to a genus present in the body size dataset [223]. Epicontinental versus open-ocean assignments for each fossil occurrence are partially based on those anaylzed by Miller and Foote [203], with additional occurrences assigned similarly (Miller and Foote, personal communication). These filtering criteria are very similar to those used by Miller and Foote [88] with an additional constraint of being present in the body size data set from Payne *et al* [223]. In total, 1130 genera included in the dataset.

Fossil occurrences were analyzed at the genus level which is common for paleobiological, macroevolutionary and macroecological studies of marine invertebrates [8, 88, 114, 158, 203, 217, 218, 222, 296, 341]. While species diversity dynamics are frequently of much greater interest than those of higher taxa, though see [85, 123], the nature of the fossil record makes accurate, precise, and consistent taxonomic assignments at the species level difficult for all occurrences. As such, the choice to analyze genera as opposed to species was in order to assure a minimum level of confidence and accuracy in the data analyzed here.

Genus duration was calculated as the number of geologic stages from first appearance to last appearance, inclusive. Durations were based on geologic stages as opposed to millions of years because of the inherently discrete nature of the fossil record; dates are not assigned to individual fossils themselves but instead fossils are assigned to a geological interval which represents some temporal range. In this analysis, stages are effectively irreducible temporal intervals in which taxa may occur. Genera with a last occurrence in or after Changhsingian stage (e.g. the final stage of the study interval) were right censored at the Changhsingian; genera with a duration of only one stage were left censored [160]. How the likelihood of censored observations is calculated is detailed in section 3.2.2.

The covariates of duration included in this analysis are geographic range size (r), environmental preference (v, v^2), body size (m), and sampling (s).

Geographic range was calculated as relative occupancy corrected for incomplete sampling. First, the paleolatitude-paleolongitude coordinates for all occurrences were projected onto an equal-area cylindrical map projection. Each occurrence was then assigned to one of the cells from a 70×34 regular raster grid placed on the map. Each grid cell represents approximately 250,000 km^2. The map projection and regular lattice were made using shape files from http://www.naturalearthdata.com/ and the `raster` package for R [122]. For each stage, the total number of occupied grid cells was calculated. Then, for each temporal bin, the relative occurrence probability of the observed taxa was calculated using the JADE method developed by Chao *et al.* [42]. This method accounts for the fact that taxa with an occupancy of 0 cannot be observed which means that occupancy follows a truncated Binomial distribution. This correction is critical when comparing occupancies from different times with different geographic sampling. Finally, for each genus, the mean relative occurrence probability was calculated as the average of that genus' occurrence probabilities for all stages it was sampled to yield relative occupancy, my proxy for geographic range.

Environmental preference was defined as probability of observing the ratio of epicontinental occurrences to total occurrences ($\theta_i = e_i/E_i$) or greater given the background occurrence probability θ'_i as estimated from all other taxa occurring at the same time (e'_i/E'_i). This measure of environmental preference is expressed.

$$\begin{aligned} p\left(\theta'_i | e'_i, E'_i\right) &\propto \text{Beta}(e'_i, E'_i - e'_i)\text{Beta}(1, 1) \\ &= \text{Beta}(e'_i + 1, E'_i - e'_i + 1), \end{aligned} \tag{3.1}$$

where v is the percent of the distribution defined in equation 3.1 less than or equal to θ_i. The Beta distribution is used here because it is a continuous distribution bounded at 0 and 1, which is idea for modeling percentages.

Body size, measured as shell length, was sourced directly from Payne *et al.* [223]. These measurements were made from brachiopod taxa figured in the *Treatise on Invertebrate Paleontology* [359].

The sampling probability for individual taxa was calculated using the standard gap statistic [89, 79]. The gap statistic is calculated as the number of stages in which the taxon was sampled minus two divided by the duration of the taxon minus two. Subtracting two from both the numerator and denominator is because the first and last appearance stages are by definition sampled. Because taxa that were right censored only include a first appearance, one was subtracted from the numerator and denominator instead of two.

The minimum duration for which a gap statistic can be calculated is three stages, so I chose the impute the gap statistic for all observations with a duration less than 3. Imputation is the “filling in” of missing observations based on the observed values [273, 102]. This is fairly straightforward in a Bayesian framework because both covariates and parameters are considered random variables, meaning that the missing values of sampling can be modeled as coming from some probability distribution. The model for imputing sampling probability is described in section 3.2.3.

Prior to analysis, geographic range was logit transformed and body size was natural-log transformed; both of these transformations make these variables defined for the entire real line. Sampling probability was transformed as $(s(n-1)+0.5)/n$ where n is the sample size as recommended by Smithson and Verkuilen [308]; this serves to slightly shrink the range of the data so that there are no values of 0 or 1. All covariates except for sampling were standardized by subtracting the mean from all values and dividing by twice its standard deviation [102]. This standardization means that the associated regression coefficients are comparable as the expected change per 1-unit change in the rescaled covariates. Finally, D is defined as the total number of covariates, excluding sampling, plus one for the intercept term.

3.2.2 Details of model

Hierarchical modelling is a statistical approach which explicitly takes into account the structure of the observed data in order to model both the within and between group variance [101, 102]. The units of study (e.g. genera) each belong to a single group (e.g. origination cohort). Each group is considered a draw from a shared probability distribution (e.g. prior) of all cohorts, observed and unobserved. The group-level parameters, or the hyperparameters of this shared prior, are themselves given (hyper)prior distributions and are also estimated like the other parameters of interest (e.g. covariate effects) [101]. The subsequent estimates are partially pooled together, where parameters from groups with large samples or effects remain large while those of groups with small samples or effects are pulled towards the overall group mean. All covariate effects (regression coefficients), as well as the intercept term (baseline extinction risk), were allowed to vary by group (origination cohort). The covariance between covariate effects was also modeled.

Genus durations were assumed to follow a Weibull distribution which allows for age-dependent extinction [160]: $y \sim \text{Weibull}(\alpha, \sigma)$. The Weibull distribution has two parameters: scale σ, and shape α. When $\alpha = 1$, σ is equal to the expected duration of any taxon. α is a measure of the effect of age on extinction risk where values greater than 1 indicate that extinction risk increases with age, and values less than 1 indicate that extinction risk decreases with age. Note that the Weibull distribution is equivalent to the exponential distribution when $\alpha = 1$.

In the case of the right- and left-censored observations mentioned above, the probability of those observations has a different calculation [160]. For right-censored observations, the likelihood is calculated $p(y|\theta) = 1 - F(y) = S(y)$ where $F(y)$ is the cumulative distribution function. In contrast, the likelihood of a left-censored observation is calculated $p(y|\theta) = F(y)$.

The scale parameter σ was modeled as a regression with both varying intercept and varying slopes and the effect of sampling [161]; this is expressed

$$\sigma_i = \exp\left(\frac{-\mathbf{X}_i B_{j[i]} + \delta s_i}{\alpha}\right) \tag{3.2}$$

where i indexes across all observations where $i = 1, \ldots, n$ where n is the total number of observations, $j[i]$ is the cohort membership of the ith observation where $j = 1, \ldots, J$ where J is the total number of cohorts, X is a $N \times D$ matrix of covariates along with a column of 1's for the intercept term, B is a $J \times D$ matrix of cohort-specific regression coefficients, and δ is the regression coefficient for the effect of sampling s. δ does not vary by cohort.

Each of the rows of matrix B are modeled as realizations from a multivariate normal distribution with length D location vector μ and $J \times J$ covariance matrix Σ: $B_j \sim \text{MVN}(\mu, \Sigma)$.

The covariance matrix was then decomposed into a length J vector of scales τ and a $J \times J$ correlation matrix Ω, defined $\Sigma = \text{diag}(\tau)\Omega\text{diag}(\tau)$ where "diag" indicates a diagonal matrix.

The elements of μ were given independent normally distributed priors. The effects of geographic range size and the breadth of environmental preference were given informative priors reflecting the previous findings while the others were given weakly informative favoring no effect. The correlation matrix Ω was given an LKJ distributed prior [166] that slightly favors an identity matrix [319]. These priors are defined

$$
\begin{aligned}
\mu^{0} &\sim \mathcal{N}(0, 5) \\
\mu^{r} &\sim \mathcal{N}(-1, 1) \\
\mu^{v} &\sim \mathcal{N}(0, 1) \\
\mu^{v^{2}} &\sim \mathcal{N}(1, 1) \\
\mu^{m} &\sim \mathcal{N}(0, 1) \\
\tau &\sim \text{C}^{+}(1) \\
\Omega &\sim \text{LKJ}(2).
\end{aligned}
\tag{3.3}
$$

The log of the shape parameter α was given a weakly informative prior $\log(\alpha) \sim \mathcal{N}(0, 1)$ centered at $\alpha = 1$, which corresponds to the Law of Constant Extinction [335].

3.2.3 Imputation of sampling probability

The vector of sampling values s has two parts: the observed part s^{o}, and the unobserved part s^{u}. Of the 1130 total observations, 539 have a duration of 3 or more and have an observed gap statistic. The gap statistic for the remaining 591 observations was imputed. As stated above, the unobserved part is the imputed, or filled in, based on the observed part

s^o. Because sampling varies between 0 and 1, I chose to model it as a Beta regression with matrix W being a $N \times (D-1)$ matrix of covariates (i.e. geographic range size, environmental preference, body size) as predictors of sampling; this assumes that the sampling value for all taxa come from the same distribution. Importantly, I make no assumptions of causal structure.

Predicting sampling probability using the other covariates that are then included in the model of duration is acceptable and appropriate in the case of imputation where the sample goal is accurate prediction [273, 102]. Not including these covariates can lead to biased estimates of the imputed variable; if the covariates themselves are related, not including them will bias this correlation towards zero which then leads to improper imputation and inference [273].

The Beta regression is defined

$$s^o \sim \mathrm{Beta}(\phi = \mathrm{logit}^{-1}(X^o \gamma), \lambda), \tag{3.4}$$

where γ is a length D vector of regression coefficients, and X defined as above. The Beta distribution used in the regression is reparameterized in terms of a mean parameter

$$\phi = \frac{\alpha}{\alpha + \beta} \tag{3.5}$$

and total count parameter

$$\lambda = \alpha + \beta \tag{3.6}$$

where α and β are the characteristic parameters of the Beta distribution [101].

The next step is to then estimate $s^u | s^o, X^o, X^u, \gamma$, the posterior distribution of which are folded back into s and used as a covariate of duration (Eq. 3.2). All the elements of γ, and both δ (Eq. 3.2) and λ (Eq. 3.4) were given weakly informative priors where

$$\begin{aligned} \gamma &\sim \mathcal{N}(0, 1) \\ \delta &\sim \mathcal{N}(0, 1) \\ \lambda &\sim \text{Pareto}(0.1, 1.5). \end{aligned} \tag{3.7}$$

3.2.4 Posterior inference and posterior predictive checks

The joint posterior was approximated using a Markov-chain Monte Carlo routine that is a variant of Hamiltonian Monte Carlo called the No-U-Turn Sampler [125] as implemented in the probabilistic programming language Stan [319]. The posterior distribution was approximated from four parallel chains run for 10,000 steps each, split half warm-up and half sampling and thinned to every 10th sample for a total of 4000 posterior samples. Chain convergence was assessed via the scale reduction factor $\hat{R}$ where values close to 1 ($\hat{R} < 1.1$) indicate approximate convergence. Convergence means that the chains are approximately stationary and the samples are well mixed [101].

Model adequacy was evaluated using a couple of posterior predictive checks. Posterior predictive checks are a means for understanding model fit or adequacy where the basic idea is that replicated data sets simulated from the fitted model should be similar to the original data and systematic differences between the simulations and observations indicate weaknesses of the model fit [101]. For both approaches used here, each posterior predictive datasets were generated from a unique draw from the posterior distribution of each parameter. The two posterior predictive checks used in this analysis are a comparison of a non-parametric estimate of the survival function $S(t)$ from the empirical dataset to the non-parametric estimates of $S(t)$ from the 100 posterior predictive datasets, and comparison of the observed genus durations to the average posterior predictive estimate of $\log(\sigma)$ (Eq. 3.2). The former is to see if simulated data has a similar survival pattern to the observed, while the latter is to see if the model systematically over- or under- estimates taxon survival.

3.3 Results

Comparison of the posterior predictive estimates of $S(t)$ to the empirical estimate reveal few obvious biases except for the case of values from the far right tail of observed durations (Fig. 3.1). This result is reinforced by the additional posterior predictive comparison where most estimates are not systematically biased except for a consistent under-estimate of $\log(\sigma)$ for older taxa (Fig. 3.2). The results of both posterior predictive checks indicate that, for the majority of observations, model fit is generally not biased.

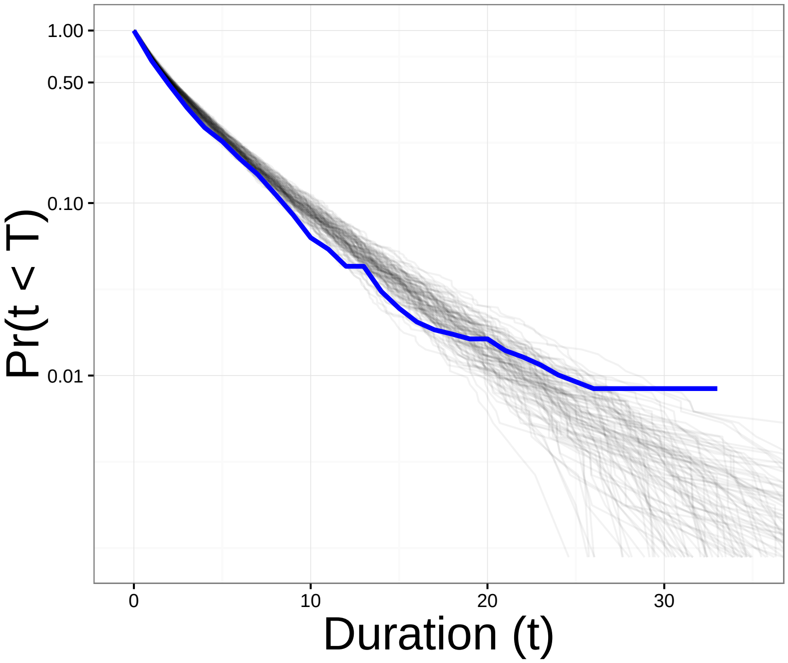


Figure 3.1: Comparison of the empirical estimate of $S(t)$ (highlighted) versus estimates from 100 posterior predictive data sets (black). $S(t)$ corresponds to the probability that the age of a genus t is less than the genus' ultimate duration T. The vertical axis is log10 transformed.

The cohort-level estimate of the effect of geographic range size indicates that as a taxon's

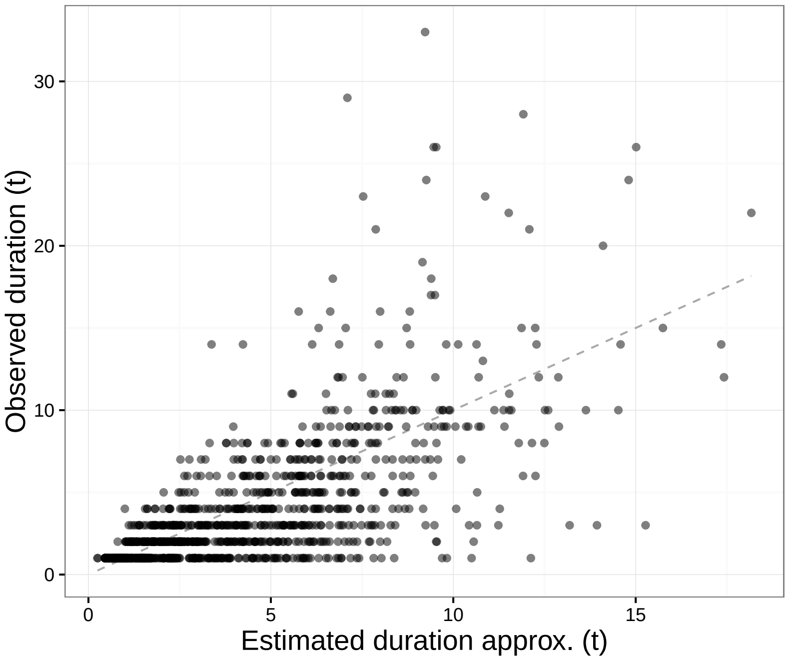


Figure 3.2: Comparison of all observed genus durations in number of geological stages to the average posterior predictive estimates of $\log(\sigma)$. The dashed, diagonal line corresponds to $x = y$.

geographic range increases, that taxon's duration is expected to increase (Table 3.1). Given the estimates of μ^r and τ^r, there is a less than 3.7% (±0.04% SD) probability that this relationships would be reversed ($\Pr\mathcal{N}(\mu^r, \tau^r) > 0)$). The between-cohort variance τ^r is the lowest of all the regression coefficients (Table 3.1).

Body size is estimated to have no effect on taxon duration, with the estimate being nearly 0 (Table 3.1). The variance between the cohort-specific estimates of the effect of body size τ^m is estimated to be greater than the variance of between-cohort estimates of the effect of geographic range size τ^r.

The group-level estimate of the effect of environmental preference is estimated from both μ^v and μ^{v^2}.

The estimate of μ^v indicates that epicontinental favoring taxa are expected to have a greater duration than open-ocean favoring taxa (Table 3.1). Additionally, given the estimate of between-cohort variance τ^v, there is approximately 18% (±7% SD) probability that, for any given cohort, taxa favoring open-ocean environments would have a greater expected duration than taxa favoring epicontinental environments ($\Pr(\mathcal{N}(\mu^v, \tau^v) > 0)$).

The estimate of μ^{v^2} indicates that the overall relationship between environmental preference and $\log(\sigma)$ is concave down (Fig. 3.3), with only a 2.7% (±3% SD) probability that any given cohort is convex up ($\Pr(\mathcal{N}(\mu^{v^2}, \tau^{v^2}) < 0)$.

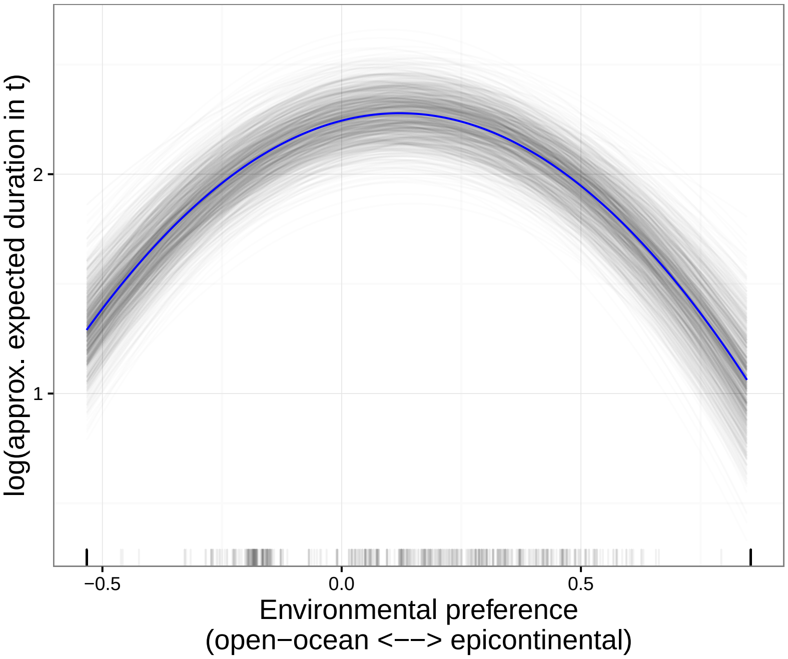


Figure 3.3: The overall expected relationship between environmental affinity v_i and a $\log(\sigma)$ when r = 0 and m = 0. The 1000 semi-transparent lines corresponds to a single draw from the posterior predictive distribution, while the highlighted line corresponds to the median of the posterior predictive distribution. The overall relationship is concave down with an optimum greater than 0, which means that taxa favoring epicontinental environments are expected to have longer durations than those favoring open-ocean environments. The tick marks along the bottom of the plot correspond to the (rescaled) observed values of environmental preference.

The cohort-specific estimates of all the regression coefficients demonstrate a lot of between cohort variance, with no obvious trends. As indicated in Table 3.1 and detectable visually (Fig. 3.4), the between-cohort estimates for β^0, β^r, and β^m all have much lower variance

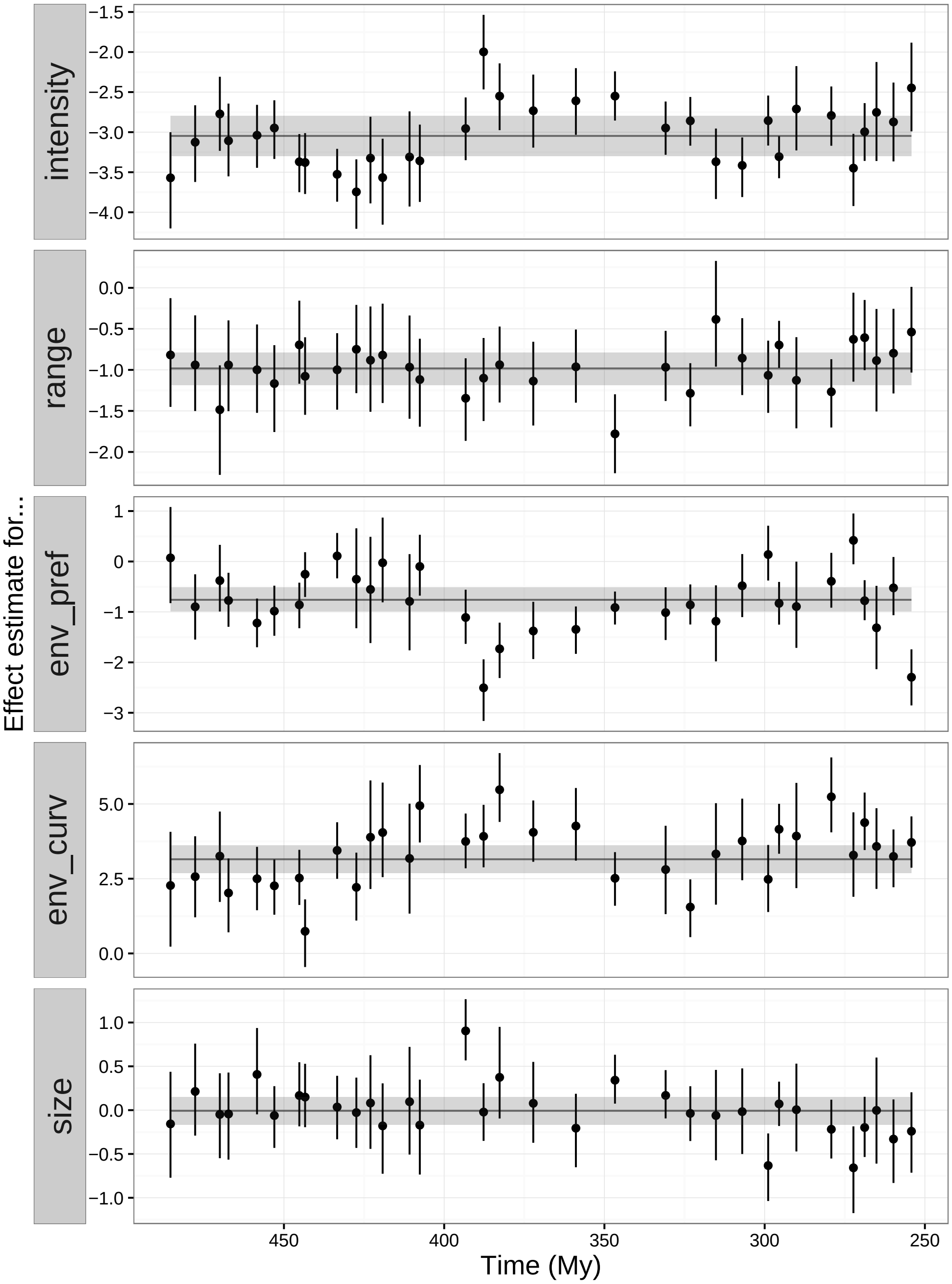


Figure 3.4: Comparison of cohort-specific estimates of β^0, the effect of geographic range on extinction risk β^r, the effect of environmental preference β^v and β^{v^2}, and body size β^m. Points correspond to the median of the cohort-specific estimate, along with 80% credible intervals. Points are plotted at the midpoint of the cohorts stage of origination in millions of years before present (My). Black, horizontal lines are the overall estimates of covariate effects along with 80% credible intervals (shaded).

than the between-cohort estimates of both β^{v} and β^{v^2}.

While most cohort-specific estimates are very similar to the overall cohort-level estimate, there are a few notable excursions away from the overall mean (Fig. 3.4). There are simultaneous excursions in both β^{0} and β^{v} for cohorts originating in the Givetian (387-382 My) and Frasnian (382-372 My) stages; both of which directly precede the late Devonian mass extinction event at the Frasnian/Famennian boundary. These cohorts are marked by both a high extinction intensity and an increase in expected duration for taxa favoring epicontinental environments over open-ocean ones.

Cohorts originating from the Silurian through the Early Devonian have a slightly lower extinction intensity than the overall mean; these cohorts are those originating in the Llandovery (443-443 My) through the Emsian (407-393 My). These time periods are when there is the lowest overall probability that epicontinental favoring taxa are expected to have greater duration than open-ocean favoring taxa. Both the Silurian and Devonian periods are notable for having been periods with a mostly "hothouse" climate, with no polar icecaps and a high sea-level [62, 153, 210].

The cohort-specific relationships between environmental preference and $\log(\sigma)$ were calculated from the estimates of β^{0}, β^{v}, and β^{v^2} (Fig. 3.5) and reflect how these three parameters act in concert and not just individually (Fig. 3.4). Beyond results already discussed above in the context of the parameters individually, it is notable that the cohort originating in the Kungurian (279-272 My) is least like the overall expected relationship and has the most sharply curved appearance due to a high estimate β^{v^2} (Fig. 3.4). This cohort has the biggest difference in extinction risk between environmental generalists and specialists. The cohorts originating during the Emsian (407-393 My) and Frasnian (382 - 372 My) are tied for second in sharpness of curvature. The least sharply curved cohorts include those originating during

Tremadocian (484-477 My), Hirnantian (445-443 My), Llandovery (443-433 My), and Ludlow (427-423 My). Except for the Tremadocian cohort, most of these cohorts originate during the Silurian through the Early Devonian range identified earlier as having lower expected extinction intensity than what is expected from the group-level estimate.

The correlations of the cohort-specific estimates of the regression coefficients are estimated as the off-diagonal elements of the correlation matrix Ω. Only two of the elements of Ω are distinguishable from 0: the correlation of β^{0} (extinction intensity) with both β^{r} and β^{v} (Fig. 3.6).

There is an approximate 90% probability that the cohort-specific estimates of baseline extinction intensity β^{0} and the effect of geographic range β^{r} are negatively correlated; this means that for cohorts experiencing a lower extinction intensity (β^{0} decreases), the magnitude of the effect of geographic range is expected to decrease as well, and *vice versa*; this is in contrast to the observation made by Jablonski [134] with regards to late Cretaceous bivalves.

Similarly, there is an approximate 97.4% probability that the cohort-specific estimates of β^{0} and β^{v} are negatively correlated; this means that as extinction intensity increases it is expected that epicontinental taxa become more favored over open-ocean environments (i.e. as β^{0} increases, β^{v} decreases).

There is only an approximate 30% probability that β^{r} and β^{v} are positively correlated. This lack of cross-correlation may be due in part to the much higher between-cohort variance of the effect of environmental preference τ^{v} than the very small between-cohort variance in the effect of geographic range τ^{r} (Table 3.1); the effect of geographic range might simply not vary enough relative to the much noisier environmental preference.

Comparison of observed values of sampling, as measured by the gap statistic, to random draws from the posterior estimates of the imputed sampling values indicate that they are

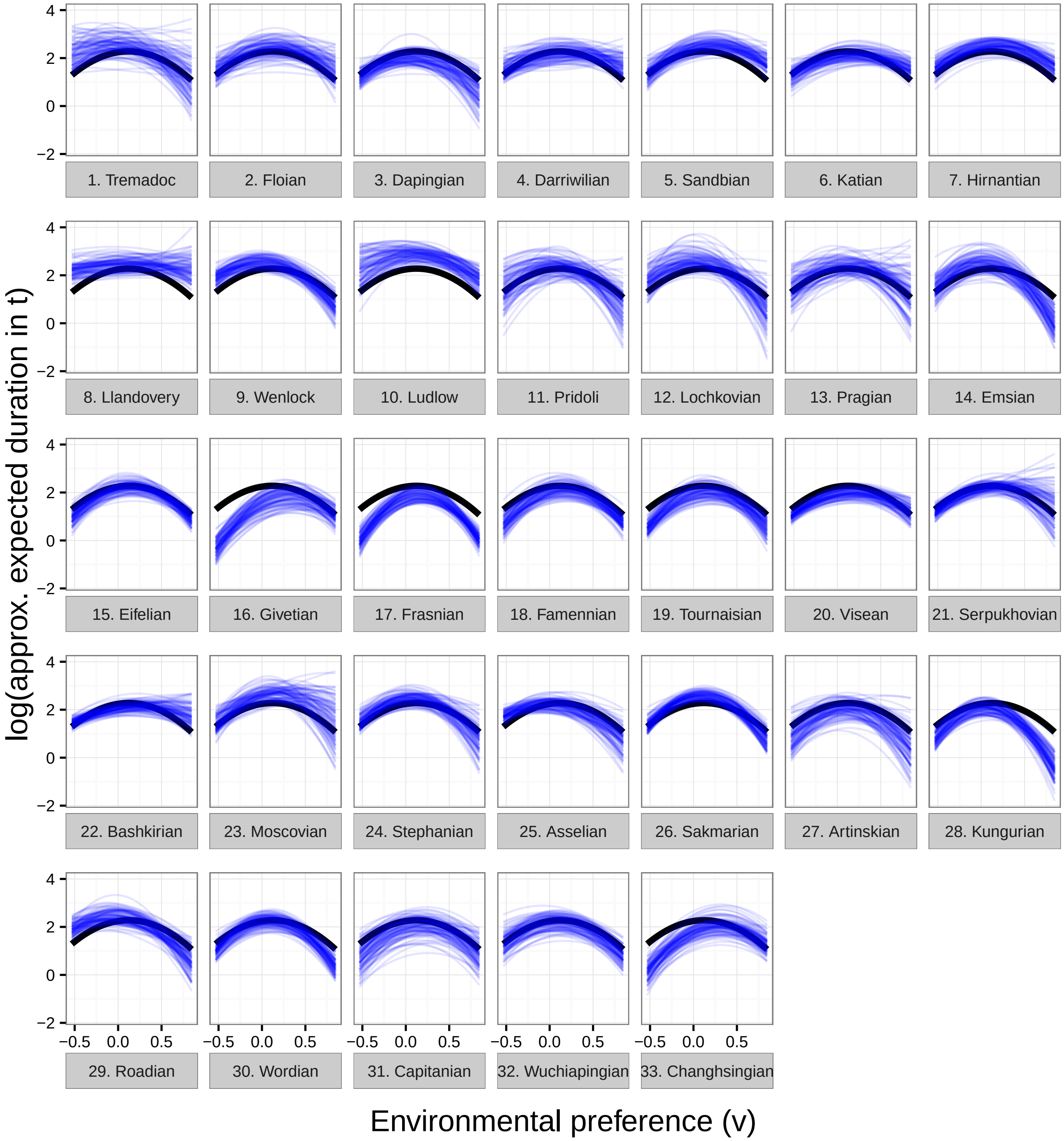


Figure 3.5: Comparison of origination cohort-specific (posterior predictive) estimates of the effect of environmental preference on $\log(\sigma)$ to the mean overall estimate of the effect of environmental preference. Cohort-specific estimates are from 100 posterior predictive simulations across the range of (transformed and rescaled) observed values of environmental preference. The oldest cohort is in the top-right and younger cohorts proceed left to right, with the youngest cohort being the right-most facet of the last row. Panel names correspond to the name of the stage in which that cohort originated.

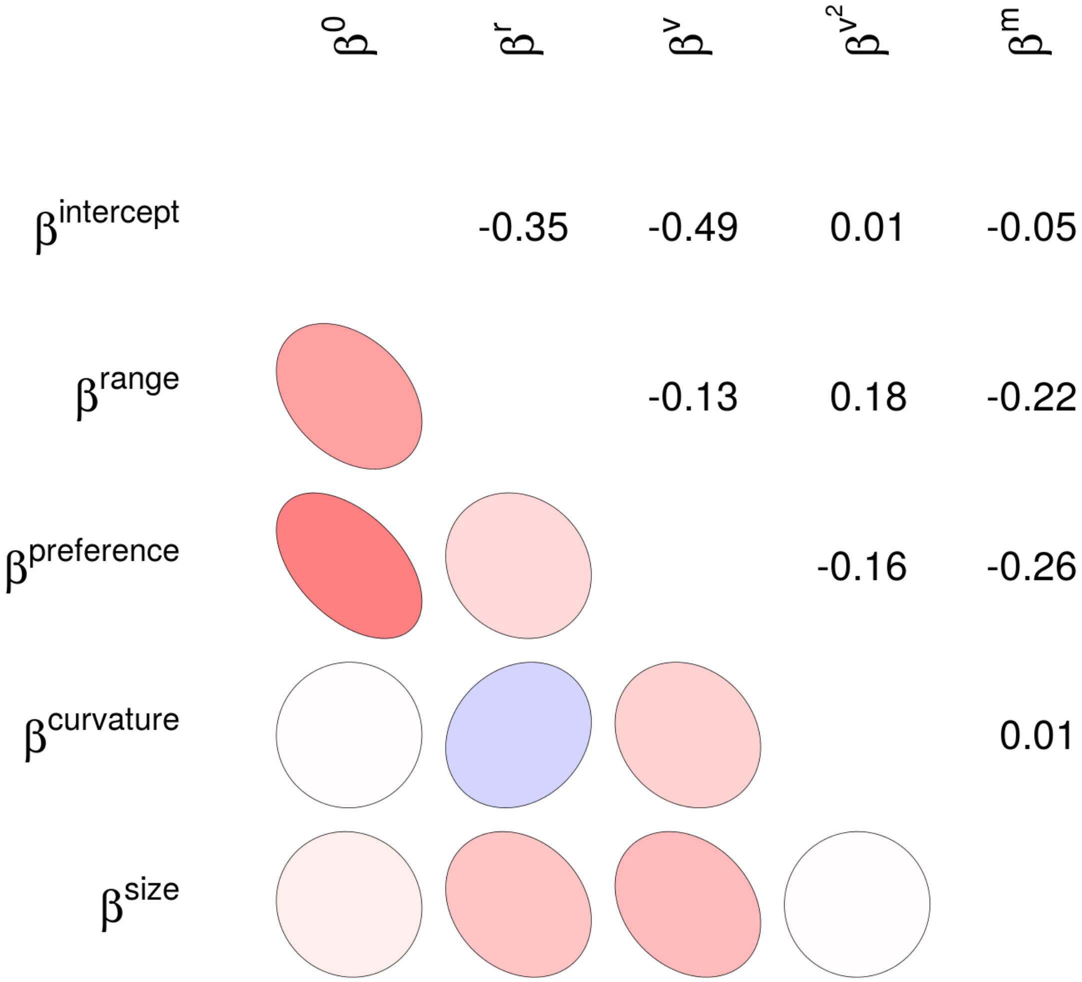


Figure 3.6: Mixed graphical and numerical representation of the correlation matrix Ω of variation in cohort-specific covariate estimates. These correlations are between the estimates of the cohort-level effects of covariates, along with intercept/baseline extinction risk. The median estimates of the correlations are presented numerically (upper-triangle) and as idealized ellipses representing that much correlation (lower-triangle). The darkness of the ellipse corresponds to the magnitude of the correlation.

very similar (Fig. 3.7. This result is very encouraging as this is the ultimate goal of multiple imputation: to fill in missing data with values similar to the observed while taking into account the randomness of that variable [273, 102]. The estimates of δ are based on the set of observed values and the entire posterior of imputed values.

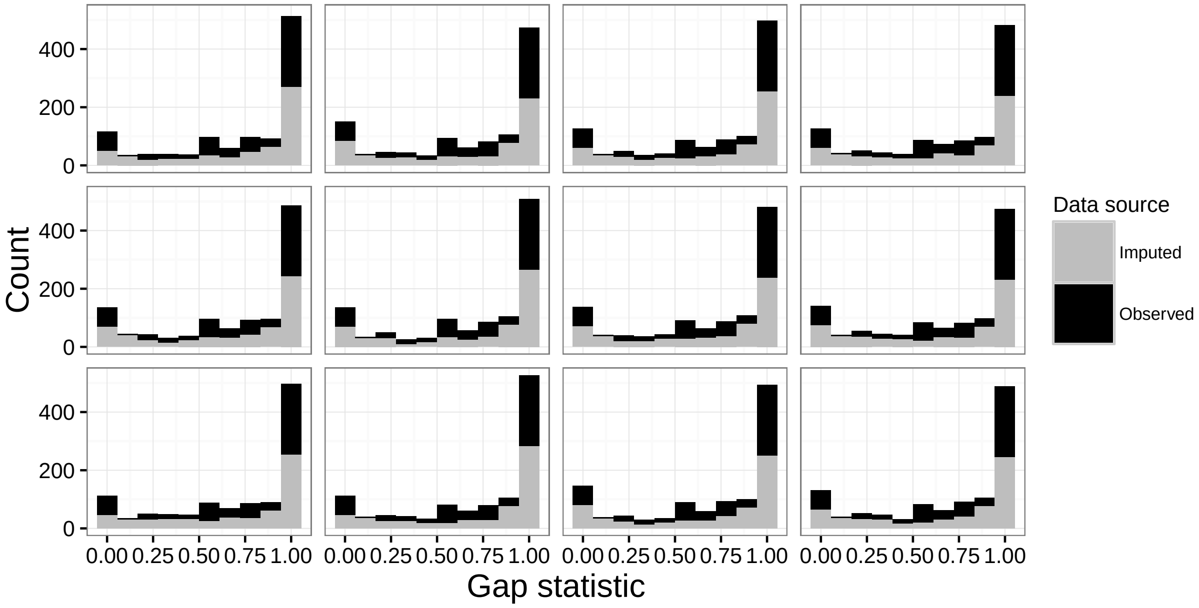


Figure 3.7: Histograms of the distribution of gap statistic values from both the observed values and the imputed values from 12 unique posterior realizations. For each panel the observed values are identical but the imputed values are from a single set of their posterior estimates.

Sampling was found to have a negative effect (positive δ) on duration: greater sampling, shorter duration (Table 3.1). While potentially counter intuitive, this result is most likely due to some long lived taxa only be sampled in the stages of the first and last appearance. Also, longer lived taxa have more opportunities to not be sampled than shorter lived taxa. These two factors will lead to this result.

While the effect of sampling appears large compared to the other regression coefficients, this is only because sampling was not standardized like the other covariates. To make the coefficients comparable, δ is multiplied by twice the posterior mean of the standard deviation of sampling probability; the transformed value of δ is then 0.642 (±0.1 SD). This effect is relatively small compared to the other covariate effects (Table 3.1). There is then a 98.6% probability that the effect of geographic range μ^r has a greater magnitude than δ. Similarly, μ^v has a 71.8% probability of having a greater magnitude of effect than δ. Finally, μ^{v^2} has a 100% probability of having a greater magnitude of effect than δ.

The Weibull shape parameter α was found to be approximately 1.36 (±0.05 SD) with a 100% probability of being greater than 1. This result is not consistent with the Law of Constant Extinction [335] and is instead consistent with accelerating extinction risk with taxon age. This may indicate that older taxa are out-competed by younger taxa, a result consistent with some empirical results [347, 240, 309] and (ironically) with a recently proposed Red Queen-like model of evolution [267]. This results, however, is not consistent with other empirical results [74, 54] and could potentially be caused by the minimum resolution of the fossil record [284]. It is thus unclear if a strong biological inference can be made from this result, which means that further work is necessary on the effect of taxon age on extinction risk.

3.4 Discussion

The generating observation behind this study was that for bivalves at the end Cretaceous mass extinction event, the only biological trait that was found to effect extinction risk was geographic range while traits that had previously been beneficial had no effect [134]. This observation raises two linked questions: how does the effect of geographic range change with changing extinction intensity, and how does the effect of other biological traits change with changing extinction intensity?

I find that as intensity increases (β^0 decreases), the magnitude of the effect of geographic range increases. I also find that as intensity increases, the effect of favoring epicontinental environments of open-ocean environments is expected to be increase; this is consistent with the results obtained by Miller and Foote [203]. There is no evidence for a correlation between the effect of geographic range and environmental preference. Additionally, the between-cohort variance in effect of geographic range is much less then the between-cohort variance of the effect of environmental preference which may underlie the lack of correlation between these two effects.

Additionally, the lower between-cohort variance in the effect of geographic range versus that higher between-cohort variance implies that for cohorts with a greater than average extinction intensity, the difference in the effect geographic range and the group-level effect of geographic range is expected to be smaller than the difference between the effect of environmental preference and the group-level effect of environmental preference.

I find consistent support for the "survival of the unspecialized," with respect to epicontinental versus open-ocean environmental preference, as a time-invariant generalization of brachiopod survival; taxa with intermediate environmental preferences are expected to have lower extinction risk than taxa specializing in either epicontinental or open-ocean environments (Fig. 3.3), though the curvature of the relationship varies from rather shallow to very peaked (Fig. 3.5). However, this relationship is not symmetric about 0, as taxa favoring epicontinental environments are expected to have a greater duration than taxa favoring open-ocean environments. This description of environment only describes one major aspect of a taxon's environmental context, with factors such as bathymetry and temperature being further descriptors of a taxon's adaptive zone [217, 114, 115, 121]; inclusion of these factors in future analyses would potentially improve our understanding of the "survival of the unspecialized" hypothesis [298].

Hopkins *et al.* [127], in their analysis of niche conservatism and substrate lithological preference in marine invertebrates, found that brachiopods were among the least "conservative" groups; taxa were found to easily change substrate preference on short time scales. While substrate preference is not the same as environmental preference (as defined here), a question does arise: are there three classes of environmental preference instead of two? These classes would be taxa with broad tolerance ("true" generalists), inflexible specialists ("true" specialists), and flexible but with a narrow tolerance. A flexible taxon is one with a narrow habitat preference at one time, but with preference that changes over time with changing environmental availability. My analysis assumes that traits are constant over the duration of the taxon meaning that this scenario is not detectable; taxa with broad tolerances and flexible taxa with narrow per-stage preference end up being treated the same way. Future work should explore how environmental preference changes over lineage duration in relation to environmental availability to estimate if the generalists–specialists continuum is actually ternary relationship.

An alternative approach for specifically modeling survival that can take into account imperfect observation than the method used here is the Cormack-Jolly-Seber (CJS) model [271, 171, 329, 173]. This model is a type of hidden Markov model with an absorbing state (i.e. extinction). In this model, survival is defined as the probability of surviving from time t to time $t + 1$. Additionally, the effect of preservation and sighting is estimated as probability of observing a taxon that is present; this can extend the duration of a taxon beyond its last occurrence. This approach is a fundamentally different from the method used in my analysis: I am estimating the biasing effect of sampling probability on taxon duration to then compare with effects of other covariates, while the CJS model estimates the pre-sampling fossil record and then estimates per-time unit survival probability.

The use of genera as the unit of the study and how to exactly interpret the effects of the biological traits is an important question. For example, if any of the traits analyzed here are associated with increases in speciation rates, this might increase the duration of genera through self-renewal [254, 255], which would be an example of the difference in biological pattern between species and genera [135, 136, 137]. This could lead to a trait appearing to decrease generic level extinction risk by that trait increasing species level origination rate instead of decreasing species level extinction risk.

The model used here could be improved through either increasing the number of analyzed traits, expanding the hierarchical structure of the model to include other major taxonomic groups of interest, and the inclusion of explicit phylogenetic relationships between the taxa in the model as an additional hierarchical effect. An example trait that may be of particular interest is the affixing strategy or method of interaction with the substrate of the taxon, which has been found to be related to brachiopod survival where, for cosmopolitan taxa, taxa that are attached to the substrate are expected to have a greater duration than those that are not [3].

It is theoretically possible to expand this model to allow for comparisons both within and between major taxonomic groups which would better constrain the brachiopod estimates while also allowing for estimation of similarities and differences in cross-taxonomic patterns. The major issue surrounding this particular expansion involves finding a similarly well sampled taxonomic group that is present during the Paleozoic. Potential groups include Crinoidea, Ostracoda, and other members of the “Paleozoic fauna” [285].

With significant updates, it would also be possible to compare the brachiopod record with with Moden groups such as bivalves or gastropods [285], though remembering that the groups may not necessarily share all cohorts with the brachiopods. This particular model expansion would act as a test of any universal cross-taxonomic patterns in the effects of emergent traits on extinction such as has been proposed for geographic range [222]. Additionally, this expanded model would also act as a test of the distinctness of the Sepkoski’s three-fauna hypothesis [285] in terms of trait-dependent extinction.

Traits like environmental preference or geographic range [135, 132] are most likely heritable. Without phylogenetic context, this analysis assumes that differences in extinction risk between taxa are independent of the shared evolutionary history of those taxa [70]. In contrast, the origination cohorts only capture shared temporal context. For example, if taxon duration is phylogenetically heritable, then closely related taxa may have more similar durations as well as more similar biological traits. Without taking into account phylogenetic similarity the effects of these biological traits would be inflated solely due to inheritance. The inclusion of phylogenetic context as an additional individual-level hierarchical effect, independent of origination cohort, would allow for determining how much of the observed variability is due to shared evolutionary history versus shared temporal context versus actual differences associated with biological traits [309].

The combination and integration of the phylogenetic comparative and paleontological approaches requires both sources of data, something which is not possible for this analysis because there is no phylogenetic hypothesis for all Paleozoic taxa, something that is frequently the case for marine invertebrates with a good fossil record. When both data sources are available has been possible, however, the analysis can more fully address the questions of interest in macroevolution [309, 303, 304, 297, 329, 302, 249, 248, 116, 97].

In summary, patterns of Paleozoic brachiopod survival were analyzed using a fully Bayesian hierarchical survival modelling approach while also eschewing the traditional separation between background and mass extinction. I find that cohort extinction intensity is negatively correlated with both the cohort-specific effects of geographic range and environmental preference. These results imply that as extinction intensity increases (β^0) it is expected that both effects will increase in magnitude. However, the change in effect of environmental preference is expected to be greater than the change in the effect of geographic range. Additionally, I find support for greater survival in environmental generalists over specialists in all origination cohorts analyzed; this is consistent with the long standing “survival of the unspecialized” hypothesis [169, 170, 298, 299, 309, 218, 217, 21]. The results of this analysis support the conclusion that for Paleozoic brachiopods, as extinction intensity increases overall extinction selectivity is expected to increase as well.

Table 3.1: Estimates of various parameters in the model used here. These include group-level estimates of the effects of biological traits on brachiopod generic survival, the standard deviation of the between-cohort effects, as well as the estimates of both the effect of sampling δ and the Weibull shape parameter α. The mean, standard deviation (SD), 10th, 50th, and 90th quantiles of the marginal posteriors are presented.

type	parameter	effect of	mean	SD	10%	50%	90%
Mean	μ^i	intercept	-3.05	0.20	-3.30	-3.05	-2.80
	μ^r	geographic range	-0.98	0.16	-1.18	-0.98	-0.79
	μ^v	environmental preference	-0.76	0.19	-0.99	-0.76	-0.52
	μ^{v^2}	environmental preference^2	3.15	0.36	2.69	3.15	3.62
	μ^m	body size	-0.01	0.13	-0.17	-0.01	0.15
Standard deviation	τ^i	intercept	0.51	0.11	0.38	0.50	0.65
	τ^r	geographic range	0.50	0.16	0.30	0.49	0.71
	τ^v	environmental preference	0.84	0.17	0.63	0.82	1.05
	τ^{v^2}	environmental preference^2	1.51	0.36	1.08	1.48	1.97
	τ^m	body size	0.47	0.13	0.32	0.46	0.64
Other	δ	sampling	0.90	0.15	0.71	0.90	1.09
	α	“time”	1.36	0.04	1.30	1.36	1.42

CHAPTER 4

SPECIES TRAITS AND ENVIRONMENTAL CONTEXT: THE CHANGING FUNCTIONAL COMPOSITION OF THE NORTH AMERICAN SPECIES POOL

4.1 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 communities are drawn from a regional species pool, or the set of all species that are present in at least one community within a region [206, 331, 117]. Species dispersal from the regional species pool to the local communities is a sorting process shaped by biotic and abiotic environmental filters which are mediated by those species' traits [288, 64, 331, 178, 51, 117]. Regional species pools are shaped by 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 the regional species pool and its constituent local communities [331, 206, 117].

Fundamentally, all species respond differently to climate and environmental change [28]. Similarities in ecological roles of species within a regional species pool can be described as a collection of guilds or functional groups [333, 14, 32, 292, 363]. Species within the same functional 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 diversity are interpretable as changes to the set of ways species within a species pool could interact with the biotic and abiotic environment.

A key question when comparing communities or regional species pools based on their functional composition is whether specific functional groups are enriched or depleted and why; what are the processes that led to a species pool having the functional composition it does [196, 358, 32, 306, 28]? Comparisons between contemporaneous regional species pools only reveal if a functional group is enriched or depleted in one species pool relative to other species pools. This type of comparison does not reveal if that functional group is enriched or depleted relative to its diversity in the regional species pool over time [28]. While a species pool may be depleted of a functional group relative to other contemporaneous species pools, that same functional group may be actually be enriched in that species pool relative to its historical diversity. Because the processes which shape regional species pool diversity (e.g. origination, extinction) operate on much longer time scales than is possible for studies of the Recent, paleontological data provides a unique opportunity to observe and estimate the changes to functional diversity and how species functional traits and environmental context can shape the enrichment or depletion of functional groups within a regional species pool [28, 306]. Being able to identify if the diversity of any functional groups are depleted relative to their long term average diversity in the species pool is particularity 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 relatively rare when compared to the functional composition of other contemporaneous species pools.

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 functional diversity changes over time. The North American mammal record is a relatively complete temporal sequence for the entire Cenozoic which is primarily, but not exclusively, based on fossil localities from the Western Interior/Great Plains region of North America [5, 10, 6]. Additionally, mammal fossils preserve a lot of important functional information, such as teeth, so that important ecological traits like the dietary/trophic category of species are easy to estimate [234, 233, 67].

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

4.1.1 Background

The diversity history of North American mammals for the Cenozoic is relatively well understood as it has been the focus of considerable study [6, 5, 143, 10, 72, 230, 93, 309, 240, 304, 290, 12, 28, 147]. Previous approaches to understanding mammal diversity, both in North America and elsewhere, fall into a number of overlapping categories: total diversity [10, 5, 72, 171], with/between guild comparisons [140, 141, 152, 147, 230, 144], within/between clade comparisons [240, 304, 290, 93, 37], and estimating the impact of environmental process on diversity [28, 147, 143, 93, 66, 12, 13, 10]. 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 [6, 5, 66, 143, 10]. Instead of being concentrated at specific time intervals, species turnover has been found to be distributed through time. It is then expected 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 [5, 10, 6]. 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 [6, 244, 113, 242]. Similarity, this question can also be asked of specific functional groups [152, 338, 290, 240].

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

The diversity history of ungulate herbivores has been characterized by more recently originating taxa having longer legs, higher crowned teeth, and a shift from graze-dominated to browse-dominated diets than their earlier originating counterparts [140, 141, 143, 144, 37, 93]. The mechanisms which drive this pattern are theorized to be some combination of tectonic activity driving environmental 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 [144, 66, 28, 13].

In contrast, the origination of modern cursorial pursuit carnivore forms was not until later in the Cenozoic; this is not to say that carnivore diversity only grew in the late Cenozoic, but that those forms were late entrants [147]. Instead, the diversity history of carnivores reflects density-dependence or some other form of self-regulation [338, 290, 304]. Specifically, it has been proposed that different canid clades have replaced each other as the dominant members of their functional group within the species pool [290, 338]. It is then expected that, for this analysis, the diversity of digitigrade and plantigrade carnivores combined (i.e. the “carnivore” guild [338]) should be relatively constant for the Cenozoic or at least have plateaued by the Neogene, though the relative diversities of digitigrade versus plantigrade forms are subject to change [235].

In a relevant study, Smits [309] 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 have a shorter duration on average than species from other locomotor categories [309]. Two possible scenarios that could yield this pattern were proposed: the extinction risk faced by arboreal species is constant and high for the entire Cenozoic or the Paleogene and Neogene represent different regimes and extinction risk increased in the Neogene, thus driving up the Cenozoic average extinction risk. These two possible explanations have clear and testable predictions with respect to the diversity history of arboreal taxa: 1) if arboreal taxa always have an elevated extinction risk when compared to other taxa, then the diversity history of arboreal taxa is expected to be constant with time, albeit possibly at low diversity; and 2) if the Paleogene and Neogene represent difference selective regimes with the former being associated with lower extinction risk than the latter, then the diversity history of arboreal taxa are expected to be present in the Paleogene but depleted or absent from the species pool during the Neogene.

The climate history of the Cenozoic can be broadly described as a gradual cooling trend, with polar ice-caps forming in the Neogene [368, 367, 53]. There are of course exceptions to this pattern such as the Eocene climatic optimum, the mid-Miocene climatic optimum, and the sudden drop in temperature at the Eocene/Oligocene boundary [368, 367]. In terms of the North American biotic environment, the Cenozoic is additionally characterized by 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

[28, 143, 141, 326]. Additionally, the landscape structure and topology of North America changed substantially over the Cenozoic with mountain uplift and other tectonic actives in Western North America [28, 66, 144, 12]. This type of geological activity affects both local climates as well as continental weather patterns while also mobilizing increased grit into the environment, something which may be responsible for increasing trend of hyposodony (high crowned teeth) among herbivores [148, 151, 58].

The effect of climate on mammal diversity and its accompanying diversification process has been the focus of considerable research with a slight consensus favoring mammal diversification being more biologically-mediated than climate-mediated [10, 72, 46]. However, differences in temporal and geographic scale seem to underlie the contrast between these two perspectives. For example when the mammal fossil record analyzed at small temporal and geographic scales a correlation between diversity and climate is observable [46]. However, when the record is analyzed at the scale of the continent and most of the Cenozoic this correlation disappears [10]. This result, however, does not go against the idea that there may be short periods of correlation between diversity and climate and 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 [72].

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 differences between species; after all, species respond differently based on their individual ecologies [28]. When analysis is based on diversity or taxonomy alone no mechanisms are possible to infer. Taxonomy, like body size, stands in for many important species traits to the point that mechanistic or process based inference is impossible. While emergent patterns might correspond to taxonomic grouping, this itself is an emergent phenomenon. Instead, by framing hypotheses in terms of species traits and their environmental context, these emergent phenomena can be observed and analyzed rather than assumed.

4.1.2 Foreground

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 interaction between traits and environment [31, 356, 232, 139]; effectively uniting climate-based species distribution modeling (SDMs) 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 [232, 139]. 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. 4.1). By analyzing assemblages over 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 and paleontologicial, 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 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.

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 [36, 15, 35, 34, 216, 342] and creodont mammals in a similar fashion [208]. Unique combinations of traits represent ecotypes, 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 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 analysis, but not as a functional trait for defining

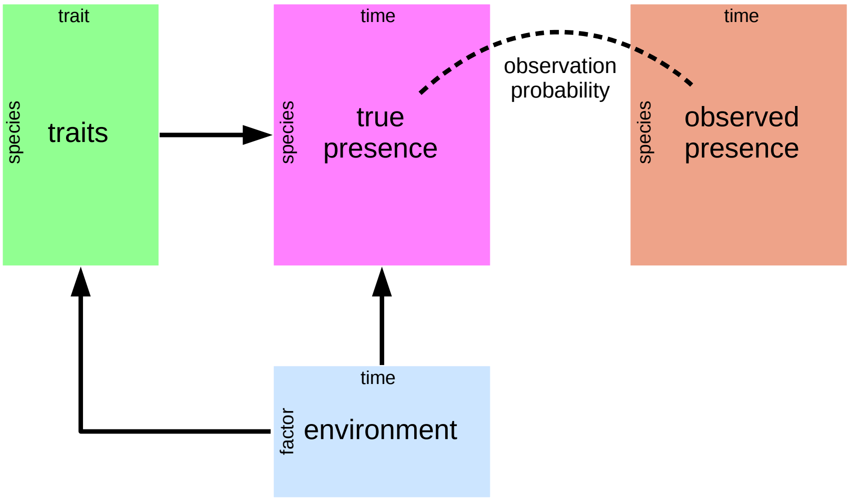


Figure 4.1: Conceptual diagram of the paleontological fourth corner problem. The observed presence matrix (orange) is the empirical presence/absence pattern for all species for all time points; this matrix is an incomplete observation of the "true" presence/absence pattern (purple). The estimated true 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.* [31] and Warton *et al.* [356].

ecotypes; instead, its inclusion is principally to control for differences in species dynamics that driven by mass and not ecotype.

The environmental factors included in this study are estimates of global temperature and the changing floral groups present in North America across the Cenozoic [53, 109]. 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. Importantly, the effects of a species ecotype on diversity are themselves modeled as functions of environmental factors (Fig. 4.1) allowing for inference as to how a species ecology can mediate selective pressures due to its environmental context.

All observations, paleontological or modern, are made with uncertainty. With presence/absence data this uncertainty comes from not knowing if an absence is a "true" absence or just a failure to observe [271, 272, 76, 81, 177, 352]. For paleontological data, the incomplete preservation and sampling of species means that the true times of origination or extinction may not be observed [76, 81, 350, 352]. The model(s) I propose below represent an attempt to translate the verbal/visual model described here (Fig. 4.1) 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 fundamental incompleteness of the fossil record.

Ultimately, the goals of this analysis are to understand when are different ecotypes enriched or 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 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 hierarchal Bayesian model combining the fourth-corner modeling approach with a model of an observation-occurrence or observation-origination-extinction process.

4.2 Materials and Methods

4.2.1 Taxon occurrences and species-level information

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 [47]. 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 [53]. 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/2slgeU`. 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/coping`.

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 Artidodactyla were reclassified as Cetartiodactyla. These re-assignments follow previous work [309] which were based on taxonomies present in the Encyclopedia of Life (`http://eol.org`) and other volumes [146, 145]. 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.

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 [309] following Table 4.1. The locomotor 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 4.2 [38]. 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 4.1: 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 4.2: Ankle posture assignment as based on taxonomy. Assignments are based on [38]. 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
	Aplodontidae	plantigrade
	Apternodontidae	scansorial
	Arctocyonidae	unguligrade
	Barbourofelidae	digitigrade
	Barylambdidae	plantigrade
	Bovidae	unguligrade
	Camelidae	unguligrade
	Canidae	digitigrade
	Cervidae	unguligrade
	Cimolodontidae	scansorial
	Coryphodontidae	plantigrade
	Cricetidae	plantigrade
	Cylindrodontidae	plantigrade
	Cyriacotheriidae	plantigrade

Continued on next page

Table 4.2 – continued from previous page

Order	Family	Stance
	Dichobunidae	unguligrade
Dinocerata		unguligrade
	Dipodidae	digitigrade
	Elephantidae	digitigrade
	Entelodontidae	unguligrade
	Eomyidae	plantigrade
	Erethizontidae	plantigrade
	Erinaceidae	plantigrade
	Esthonychidae	plantigrade
	Eutypomyidae	plantigrade
	Felidae	digitigrade
	Florentiamyidae	plantigrade
	Gelocidae	unguligrade
	Geolabididae	plantigrade
	Glyptodontidae	plantigrade
	Gomphotheriidae	unguligrade
	Hapalodectidae	plantigrade
	Heteromyidae	digitigrade
	Hyaenidae	digitigrade
	Hyaenodontidae	digitigrade
	Hypertragulidae	unguligrade
	Ischyromyidae	plantigrade
	Jimomyidae	plantigrade

Continued on next page

Table 4.2 – continued from previous page

Order	Family	Stance
Lagomorpha		digitigrade
	Leptictidae	plantigrade
	Leptochoeridae	unguligrade
	Leptomerycidae	unguligrade
	Mammutidae	unguligrade
	Megalonychidae	plantigrade
	Megatheriidae	plantigrade
	Mephitidae	plantigrade
	Merycoidodontidae	digitigrade
Mesonychia		unguligrade
	Mesonychidae	digitigrade
	Micropternodontidae	plantigrade
	Mixodectidae	plantigrade
	Moschidae	unguligrade
	Muridae	plantigrade
	Mustelidae	plantigrade
	Mylagaulidae	fossorial
	Mylodontidae	plantigrade
	Nimravidae	digitigrade
	Nothrotheriidae	plantigrade
Notoungulata		unguligrade
	Oromerycidae	unguligrade
	Oxyaenidae	digitigrade

Continued on next page

Table 4.2 – continued from previous page

Order	Family	Stance
	Palaeomerycidae	unguligrade
	Palaeoryctidae	plantigrade
	Pampatheriidae	plantigrade
	Pantolambdidae	plantigrade
	Periptychidae	digitigrade
Perissodactyla		unguligrade
	Phenacodontidae	unguligrade
Primates		plantigrade
	Procyonidae	plantigrade
	Proscalopidae	plantigrade
	Protoceratidae	unguligrade
	Reithroparamyidae	plantigrade
	Sciuravidae	plantigrade
	Sciuridae	plantigrade
	Simimyidae	plantigrade
	Soricidae	plantigrade
	Suidae	digitigrade
	Talpidae	fossorial
	Tayassuidae	unguligrade
	Tenrecidae	plantigrade
	Titanoideidae	plantigrade
	Ursidae	plantigrade
	Viverravidae	plantigrade

Continued on next page

Table 4.2 – continued from previous page

Order	Family	Stance
	Zapodidae	plantigrade

Estimates of species mass used in this study were sourced from multiple databases and papers, especially those focusing on similar macroevolutionary or macrecological questions [329, 30, 94, 198, 248, 305]; this is similar to what has been done before [309]. When species mass was not available, proxy measures were used and then used to estimate average species mass. For example, given a measurement of a mammal tooth size, it is possible and routine to estimate its mass given some regression equation (Table 4.3). The PBDB has one or more body part measures for many species. These were used as body size proxies for many species. Mass was log-transformed and then rescaled by first subtracting mean log-mass from all mass estimates, then dividing by two-times its standard deviation; this insures that the magnitude of effects for both continuous and discrete covariates are directly comparable [102, 100].

In total, 1400 mammal species occurrence histories were included in this study after applying all of the restrictions above.

All fossil occurrences from 64 to 2 million years ago (Mya) were binned into 31 two-million year (My) bins. This temporal length was chosen because it is approximately the resolution of the North American mammal fossil record [5, 10, 191, 6].

Table 4.3: 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	[165]
General	$\log(m) = 2.9677x - 5.6712$	mandible length	[91]
General	$\log(m) = 3.68x - 3.83$	skull length	[183]
Carnivores	$\log(m) = 2.97x + 1.681$	lower m1 length	[337]
Insectivores	$\log(m) = 1.628x + 1.726$	lower m1 area	[26]
Insectivores	$\log(m) = 1.714x + 0.886$	upper M1 area	[26]
Lagomorph	$\log(m) = 2.671x - 2.671$	lower toothrow area	[329]
Lagomorph	$\log(m) = 4.468x - 3.002$	lower m1 length	[329]
Marsupials	$\log(m) = 3.284x + 1.83$	upper M1 length	[108]
Marsupials	$\log(m) = 1.733x + 1.571$	upper M1 area	[108]
Rodentia	$\log(m) = 1.767x + 2.172$	lower m1 area	[165]
Ungulates	$\log(m) = 1.516x + 3.757$	lower m1 area	[201]
Ungulates	$\log(m) = 3.076x + 2.366$	lower m2 length	[201]
Ungulates	$\log(m) = 1.518x + 2.792$	lower m2 area	[201]
Ungulates	$\log(m) = 3.113x - 1.374$	lower toothrow length	[201]

4.2.2 Environmental and temporal covariates

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 occurrences of an ecotype). Additionally, these covariates are defined for temporal bins and not the species themselves; as such they predict the parts of each species occurrence history. The group-level covariates in this study are two global temperature estimates and the Cenozoic "plant phases" defined by Graham [109].

Global temperature across most of the Cenozoic was calculated from Mg/Ca isotope record from deep sea carbonates [53]. Mg/Ca based temperature estimates are preferable to the frequently used δ^{18}O temperature proxy [368, 367, 10, 72] because Mg/Ca estimates do not conflate temperature with ice sheet volume and depth/stratification changes. The former is particularly important to this analysis as the current polar ice-caps appeared and grew during the second half of the Cenozoic. These properties make Mg/Ca based temperature

Table 4.4: Definitions of the start and stop times of the three plant phases used this study as defined by Graham [109].

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

estimates preferable for macroevolutionary and macroecological studies [69]. Two aspects of the Mg/Ca-based temperature curve were included in this analysis: mean and range. Both were calculated as the mean of all respective estimates for each 2 My temporal bins. The distributions of the temperature mean and range estimates were then rescaled by subtracting their respective means from all values and then dividing by twice their respective standard deviations.

The second set of environmental factors included in this study are the Cenozoic plant phases defined by Graham [109]. Graham's plant phases are holistic descriptors of the taxonomic composition of 12 ecosystem types, which plants are present at a given time, and the relative modernity of those plant groups with younger phases representing increasingly modern taxa [109]. Graham [109] defines four intervals from the Cretaceous to the Pliocene, though only three of these intervals take place during the time frame being analyzed. Graham's plant phases was included as a series of "dummy variables" encoding the three phases included in this analysis [102]; this means that the first phase is synonymous with the intercept and subsequent phases are defined by their differences from the first phase. The temporal boundaries of these plant phases are defined in Table 4.4.

4.2.3 Modeling species occurrence

Two different models were used in this study: a pure-presence model and a birth-death model. Both models at their core are hidden Markov models where the latent process has an

		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 4.5: Transition matrices for the pure-presence (4.5a) and birth-death (4.5b) 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. 4.2, and 4.3). 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 $0_{extinct}$ [4].

absorbing state [4]. The difference between these two models lies in whether the probabilities of a species originating or surviving are considered equal or different (Table 4.5). While there are only two state "codes" in a presence-absence matrix (i.e. 0/1), there are in fact three states in a birth-death model: 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 [4]. Thus, in the transition matrices the probability of an extinct species changing states is 0 (Table 4.5). See below for parameter explanations (Tables 4.6, 4.7, and 4.8).

Observation process

The type of hidden Markov model used in this study has three characteristic probabilities: probability p of observing a species given that it is present, probability ϕ of a species surviving from one time to another, and probability π of a species first appearing [271]. In this formulation, the probability of a species becoming extinct is $1-\phi$. For the pure-presence model $\phi = \pi$, while for the birth-death model $\phi \neq \pi$.

The probability p of observing a species that is present is modeled as a logistic regression with a time-varying intercept and species mass as a covariate. The effect of species mass on p

Table 4.6: Parameters for 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

was assumed linear and constant over time. These assumptions are reflected in the structure of this part of the model being presented here:

$$\begin{aligned} y_{i,t} &\sim \text{Bernoulli}(p_{i,t} z_{i,t}) \\ p_{i,t} &= \text{logit}^{-1}(\alpha_0 + \alpha_1 m_i + r_t) \\ r_t &\sim \mathcal{N}(0, \sigma). \end{aligned} \tag{4.1}$$

The parameters associated with Equation 4.1 are described in Table 4.6.

Pure-presence process

For the pure-presence model there is only a single probability dealing with the presence of a species θ (Table 4.5a). This probability was modeled as multi-level logistic regression with both species-level and group-level covariates [102, 101]. The parameters associated with the pure-presence model are presented in Table 4.7, and the full sampling statement in Equation 4.2.

Species mass was included as a covariate with two regression coefficients allowing for a quadratic relationship with log-odds of occurrence. Because the distribution of mammal species body mass is unimodal and approximately log-normal [305], I assume that species of

Table 4.7: Parameters for the model of occurrence 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

intermediate 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 prior with most of its density below 0 (Eq. 4.2).

The values of each ecotype's intercept are themselves modeled as regressions using the group-level covariates associated with environmental context. Each of these regressions has an associated variance of possible values of each ecotype's intercept [102]. In addition, the covariances between ecotype intercepts, given this group-level regression, are modeled [102]. The prior choice for the covariance matrix separates it into a vector of scales τ and a 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 the Stan manual [319].

All parameters not modeled elsewhere were given weakly informative priors [101, 195, 319]. Weakly informative means that priors do not necessarily encode actual prior information but instead help regularize or weakly constrain posterior estimates. These priors have a concentrated probability density around and near zero; this has the effect of tempering our estimates and help prevent overfitting the model to the data [101, 195, 319]. 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.

$$
\begin{aligned}
y_{i,t} &\sim \text{Bernoulli}(p_{i,t} z_{i,t}) & \alpha_0 &\sim \mathcal{N}(0,1) \\
p_{i,t} &= \text{logit}^{-1}(\alpha_0 + \alpha_1 m_i + r_t) & \alpha_1 &\sim \mathcal{N}(1,1) \\
r_t &\sim \mathcal{N}(0,\sigma) & \sigma &\sim \mathcal{N}^{+}(1) \\
z_{i,1} &\sim \text{Bernoulli}(\rho) & b_1 &\sim \mathcal{N}(0,1) \\
z_{i,t} &\sim \text{Bernoulli}(\theta_{i,t}) & b_2 &\sim \mathcal{N}(-1,1) \\
\theta_{i,t} &= \text{logit}^{-1}(a_{t,j[i]} + b_1 m_i + b_2 m_i^2) & \gamma &\sim \mathcal{N}(0,1) \\
a &\sim \text{MVN}(u\gamma, \Sigma) & \tau &\sim \mathcal{N}^{+}(1) \\
\Sigma &= \text{diag}(\tau)\Omega\text{diag}(\tau) & \Omega &\sim \text{LKJ}(2)
\end{aligned}
\tag{4.2}
$$

Birth-death process

In the birth-death version of the model, $\phi \neq \pi$ and so each of these probabilities is modeled separately but each is handled in a similar manner to how θ is modeled in the pure-presence model (Eq. 4.2, Table 4.5b). The parameters associated with the birth-death presence model are presented in Table 4.8 and the full sampling statement, including observation (Eq. 4.1), is described in Equation 4.3:

$$
\begin{aligned}
y_{i,t} &\sim \text{Bernoulli}(p_{i,t} z_{i,t}) & \Sigma^{\phi} &= \text{diag}(\tau^{\phi})\Omega^{\phi}\text{diag}(\tau^{\phi}) \\
p_{i,t} &= \text{logit}^{-1}(\alpha_0 + \alpha_1 m_i + r_t) & \Sigma^{\pi} &= \text{diag}(\tau^{\pi})\Omega^{\pi}\text{diag}(\tau^{\pi}) \\
r_t &\sim \mathcal{N}(0, \sigma) & \rho &\sim \text{U}(0,1) \\
\alpha_0 &\sim \mathcal{N}(0, 1) & b_1^{\phi} &\sim \mathcal{N}(0,1) \\
\alpha_1 &\sim \mathcal{N}(1, 1) & b_1^{\pi} &\sim \mathcal{N}(0,1) \\
\sigma &\sim \mathcal{N}^{+}(1) & b_2^{\phi} &\sim \mathcal{N}(-1,1) \\
z_{i,1} &\sim \text{Bernoulli}(\phi_{i,1}) & b_2^{\pi} &\sim \mathcal{N}(-1,1) \\
z_{i,t} &\sim \text{Bernoulli}\left(z_{i,t-1}\pi_{i,t} + \sum_{x=1}^{t}(1 - z_{i,x})\phi_{i,t}\right) & \gamma^{\phi} &\sim \mathcal{N}(0,1) \\
& & \gamma^{\pi} &\sim \mathcal{N}(0,1) \\
\phi_{i,t} &= \text{logit}^{-1}(a_{t,j[i]}^{\phi} + b_1^{\phi} m_i + b_2^{\phi} m_i^2) & \tau^{\phi} &\sim \mathcal{N}^{+}(1) \\
\pi_{i,t} &= \text{logit}^{-1}(a_{t,j[i]}^{\pi} + b_1^{\pi} m_i + b_2^{\pi} m_i^2) & \tau^{\pi} &\sim \mathcal{N}^{+}(1) \\
a^{\phi} &\sim \text{MVN}(U\gamma^{\phi}, \Sigma^{\phi}) & \Omega^{\phi} &\sim \text{LKJ}(2) \\
a^{\pi} &\sim \text{MVN}(U\gamma^{\pi}, \Sigma^{\pi}) & \Omega^{\pi} &\sim \text{LKJ}(2).
\end{aligned}
\tag{4.3}
$$

Similar to the pure-presence model, both ϕ and π are modeled as logistic regressions with varying intercept and one covariate associated with two parameters. The possible relationships between mass and both ϕ and π are reflected in the parameterization of the model and choice of priors (Eq. 4.3).

The intercepts of ϕ and π both vary by species ecotype and those values are themselves the product of group-level regression using environmental factors as covariates (Eq. 4.3); this is identical to the pure presence model (Eq. 4.2).

Table 4.8: 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 $z_{\_,t} = 1 \vert z_{\_,t-1} = 0$; origination
π	$N \times T - 1$	probability of $z_{\_,t} = 1 \vert z_{\_,t-1} = 1$; 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}

4.2.4 Posterior inference and model adequacy

Computer programs that implement joint posterior inference for the above models (Eqs. 4.2, 4.3) were written in the probabilistic programming language Stan [319]. Both models feature a large matrix of latent discrete parameters z (Tables 4.6, 4.7, 4.8; Eqs. 4.1, 4.2, 4.3). All 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 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).

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 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 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 4.2) [319]. This process is identical, language-wise, to assuming range-through and then estimating the possibility of all possible range extension due to incomplete sampling.

The combined size of the dataset and large number of parameters in both models (Eqs. 4.2, 4.3), specifically the total number of latent parameters that are the matrix z, means that stochastic approximate posterior inference is computationally very slow even using NUTS based HMC as implemented in Stan [319]. Instead, an approximate Bayesian approach was used: variational inference. A recently developed automatic variational inference algorithm called "automatic differention variational inference" (ADVI) is implemented in Stan and was used here [163, 319]. ADVI assumes that the posterior is Gaussian but still yields a true

	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 4.2: 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.

Bayesian posterior; this assumption is similar to quadratic approximation of the likelihood function commonly used in maximum likelihood based inference [195]. The principal limitation of assuming the joint posterior is Gaussian is that the true topology of the log-posterior isn't estimated; this is a particular burden for scale parameters which are bounded to be positive (e.g. standard deviation).

Of additional concern for posterior inference is the partial identifiability of observation parameters $p_{t=1}$ and $p_{t=T}$ [271]. This issue means that the estimates of sampling probabilities at the "edges" of the time series cannot fully be estimated because there are no known "gaps" in species occurrence histories that are guaranteed to be filled. Instead, the values of the first and final columns of the "true" presence-absence matrix z for thos observations that do not already have presences in the observed presence-absence matrix y cannot be estimated [271]. 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 [101], and in fact this approach might be more analytically sound than the more ad-hoc approaches that are occasionally used to overcome this hurdle [271]. Additionally, because $p_{t=1}$ and $p_{t=T}$ are only partially identifiable, estimates of occurrence θ and origination ϕ at $t = 1$ and estimates of θ, ϕ and survival π at $t = T$ may suffer from similar edge effects. Again, the hierarchical modeling approach used here may help correct for this reality by drawing these estimates towards the overall means of those parameters.

After fitting both models (Eqs. 4.2, 4.3) using ADVI, model adequacy and quality of fit were assessed using a posterior predictive check [101]. By simulating 100 theoretical data sets from the posterior estimates of the model parameters and the observed covariate information the congruence between predictions made by the model and the observed empirical data can be assessed. These datasets are simulated by starting with the observed states of the presence-absence matrix at $t = 1$; from there, the time series roll forward as stochastic processes with covariate information given from the empirical observations. Importantly, this is fundamentally different from observing the posterior estimates of the "true" presence-absence matrix z. The posterior predictive check used in this study is to compare the observed average number of observations per species to a distribution of simulated averages; if the empirically observed value sits in the middle of the distribution then the model can be considered adequate in reproducing the observed number of occurrences per species.

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 distribution and are instead just an approximation [101]. Because of this, posterior predictive estimates are themselves only approximate checks of model adequacy. The posterior predictive check that is used in this study focuses on mean occurrence and not to any scale parameters that might be most affected by the ADVI assumptions.

Given parameter estimates, diversity and diversification rates are estimated through posterior predictive simulations. Given the observed presence-absence matrix y, estimates of the true presence-absence matrix z can be simulated and the distribution of possible occurrence histories can be analyzed. This is conceptually similar to marginalization where the probability of each possible occurrence history is estimated (Fig. 4.2), but now these occurrence histories are generated relative to their estimated probabilities.

The posterior distribution of z gives the estimate of standing diversity N_t^{stand} for all time points as

$$N_t^{stand} = \sum_{i=1}^{M} z_{i,t}. \tag{4.4}$$

Given estimates of N^{stand} for all time points, the estimated number of originations O_t is estimated as

$$O_t = \sum_{i=1}^{M} z_{i,t} = 1 | z_{i,t-1} = 0 \tag{4.5}$$

and number of extinctions E_t estimated as

$$E_t = \sum_{i=1}^{M} z_{i,t} = 0 | z_{i,t-1} = 1. \tag{4.6}$$

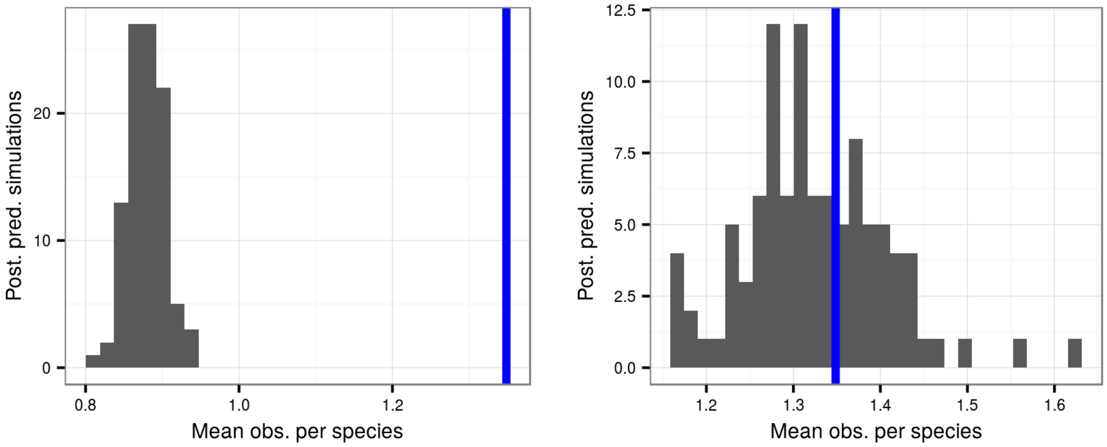
Per-capita growth D^{rate}, origination O^{rate} and extinction E^{rate} rates are then calculated as

$$\begin{aligned} O_t^{rate} &= \frac{O_t}{N_{t-1}^{stand}} \\ E_t^{rate} &= \frac{E_t}{N_{t-1}^{stand}} \\ D_t^{rate} &= O_t^{rate} - E_t^{rate}. \end{aligned} \tag{4.7}$$

4.3 Results

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

4.3.1 Comparison of estimates from the pure-presence and birth-death models



(a) Pure-presence model

(b) Birth-death model

Figure 4.3: Comparison of the average observed number of occurrences per species (blue line) to the average number of occurrences from 100 posterior predictive datasets using the posterior estimates from the pure-presence and birth-death models.

Comparison of the posterior predictive results from the pure-presence and birth-death models reveals a striking difference in performance of either model to predict the structure of the underlying data (Fig. 4.3). 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 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 data than inferences based on the pure-presence model. Further inspection of the posterior parameter estimates from both models gives further insight into the resins for this difference in posterior predictive results [101].

Increases in the occurrence probability of an ecotype is interpreted as an increase in the 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. 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 likely all but disappeared from the species pool [102]. In logistic regression, high uncertainty in the estimates of the underlying log-odds of occurrence, origination, or survival tends 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. The problem of complete separation is mitigated by the hierarchical modeling strategy used here [101, 102, 195].

Estimates of occurrence probability estimated from the pure-presence model and estimates of origination probability from the birth-death model are broadly similar (Fig. 4.4, 4.5); this is not the case for the survival probability estimates (Fig. 4.6). This result supports the idea that changes to the North American regional species pool is more likely due to changes in origination than extinction, a result to which I will return to later in the discussion of per-capita diversification, origination, and extinction rates. For most ecotypes, occurrence and origination probability estimates increase with time (Fig. 4.5). This makes sense given that, over time, all species that have at least one observed occurrence must have had that occurrence by the last time point, so our certainty in a species occurring must increase with time. Notably, ecotypes with 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. For most ecotypes, occurrence or origination probability is estimated with less uncertainty than its estimate of survival probability (Fig. 4.4, 4.5, 4.6).

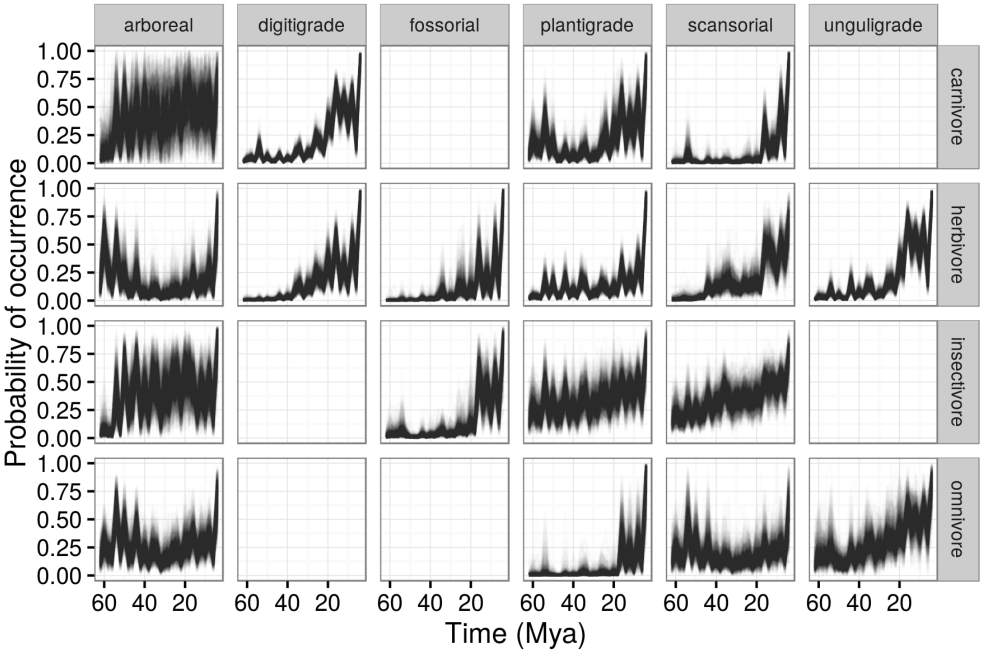


Figure 4.4: Probability of a mammal ecotype occurring over time as estimated from the pure-presence model. Each panel depicts 100 random samples from the model's posterior. The columns are by locomotor category and rows by dietary category; their intersections are the observed and analyzed ecotypes. Panels with no lines are ecotypes not observed in the dataset.

The pure-presence and birth-death models have similar estimates of the relationship between species mass and the probability of sampling a species that is present (Fig. 4.7). For both models this relationship is at least weakly positive, which means that as species body mass increases it is expected that they are more likely to be sampled if present. The estimated relationship from the pure-presence model is with greater uncertainty than that from the birth-death model (Fig. 4.7). These results are consistent with the intuition that larger fossils are easier to sample because they are more visible to the eye. In turn, this means that the observed occurrence histories of small bodied species are more likely to have gaps, where

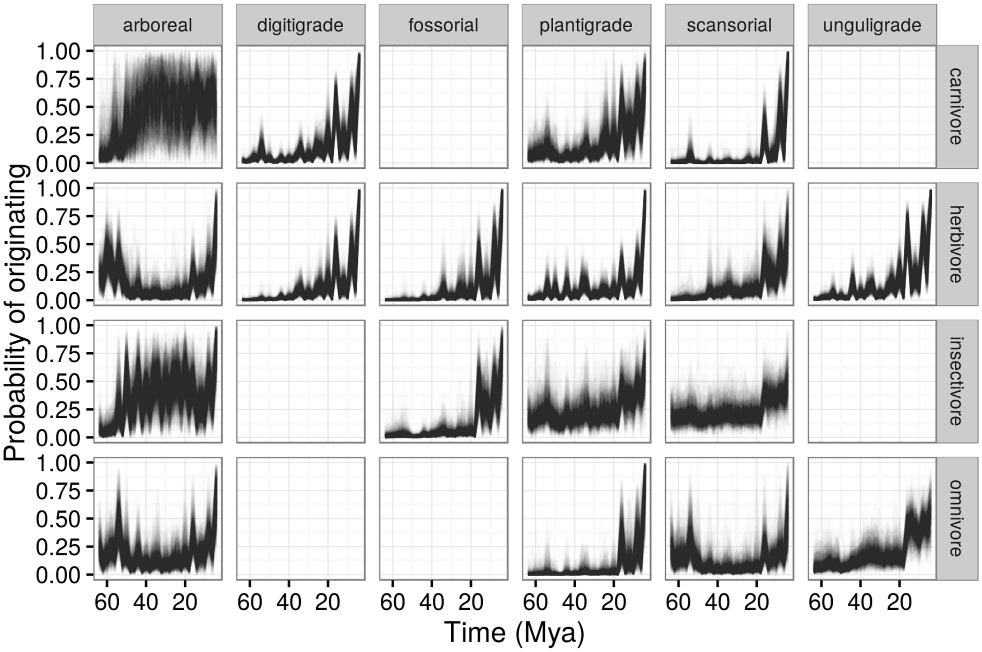


Figure 4.5: Probability of a mammal ecotype origination probabliities at each time point as estimated from the birth-death model. Each panel depicts 100 random samples from the model's posterior. The columns are by locomotor category and rows by dietary category; their intersections are the observed and analyzed ecotypes. Panels with no lines are ecotypes not observed in the dataset.

$y = 0$ for that species the true state z is 1.

There is broad congruence between the estimated effect of body mass on occurrence probability (Fig. 4.8) and the effect of species mass on body mass on origination probability (Fig. 4.9). The striking pattern is higher probability of origination for species with body sizes closer to the mean than either extremes. This result is consistent with the canonically normal distribution of mammal body sizes [305]; it is then expected that the most likely to occur species would be those from the middle of the distribution, and that species originating will on average be of average mass, especially considering species shared common ancestry [70]. All variation in estimates between ecotypes (Fig. 4.8, 4.9) is due to differences in ecotype-specific origination probabilities and the associated effects of plant phase; the effect of mass was considered constant for all ecotypes.

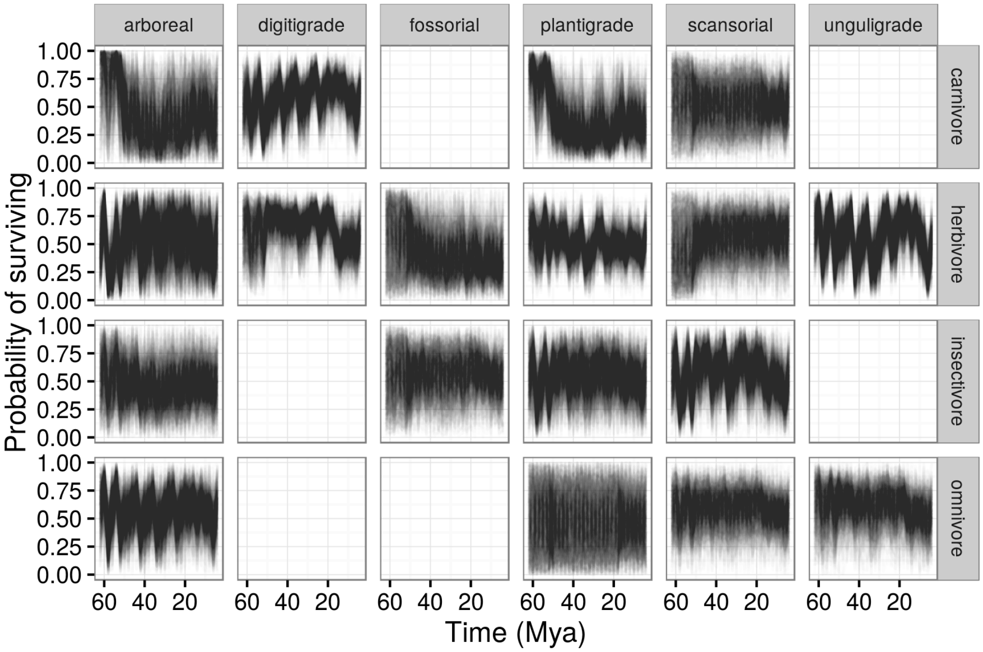
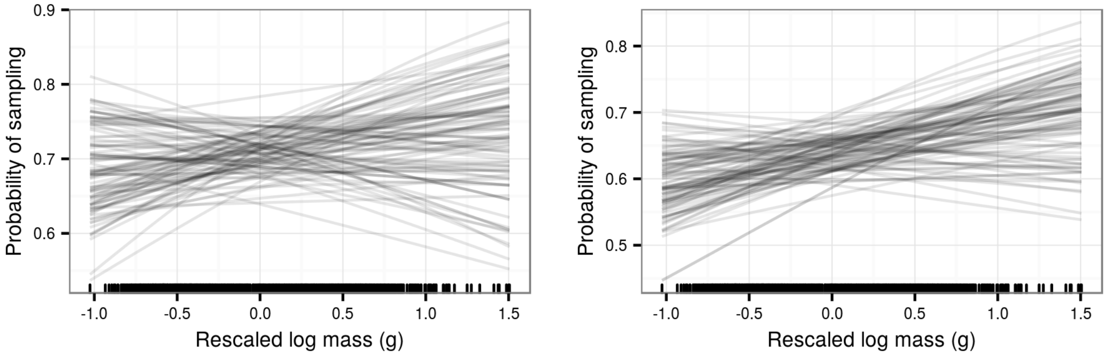


Figure 4.6: Probability of a mammal ecotype survival probabilities at each time point as estimated from the birth-death model. Each panel depicts 100 random samples from the model's posterior. The columns are by locomotor category and rows by dietary category; their intersections are the observed and analyzed ecotypes. Panels with no lines are ecotypes not observed in the dataset.

In contrast, the effect of species mass on probability of survival as estimated from the birth-death model (Fig. 4.10) is consistent with previous findings that there is little effect of mass on extinction for North American mammals for the Cenozoic [309, 329]. Note that all variation between ecotypes depicted in Figure 4.10 is due to differences in ecotype-specific survival probability and the associated effects of plant phase; the effect of mass was considered constant for all ecotypes (Eqs. 4.2, 4.3).

Similarities in parameter estimates between ecotypes may be due to a similar response to environmental factors (Fig. 4.11, 4.12, and 4.13). The estimated group-level effects on ecotype occurrence, 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. 4.11, 4.12). As demonstrated in the comparisons of the effect of body mass on occurrence



(a) Pure-presence model

(b) Birth-death model

Figure 4.7: 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.

from the pure-presence model (Fig. 4.8) with the effect of body mass on origination and survival from the birth-death model (Fig. 4.9, and 4.10), 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. 4.11), origination (Fig. 4.12), or extinction (Fig. 4.13) is interpreted to mean that the set of possible 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 mammal-plant interactions (e.g. ecological opportunity; 366, 181, 182); while adverse conditions may translate to a decreasing set of interactions or loss of appropriate environmental context. Remember that favorable versus adverse condition of a plant phase is definitionally relative to the other two plant phases.

One of the limitations to this interpretation is the almost determistic increase in probability of occurrence and origination for most ecotypes (Fig. 4.4, 4.5). This "pull of the Recent" means that interpreting the biological meaning of differences between the final plant phase

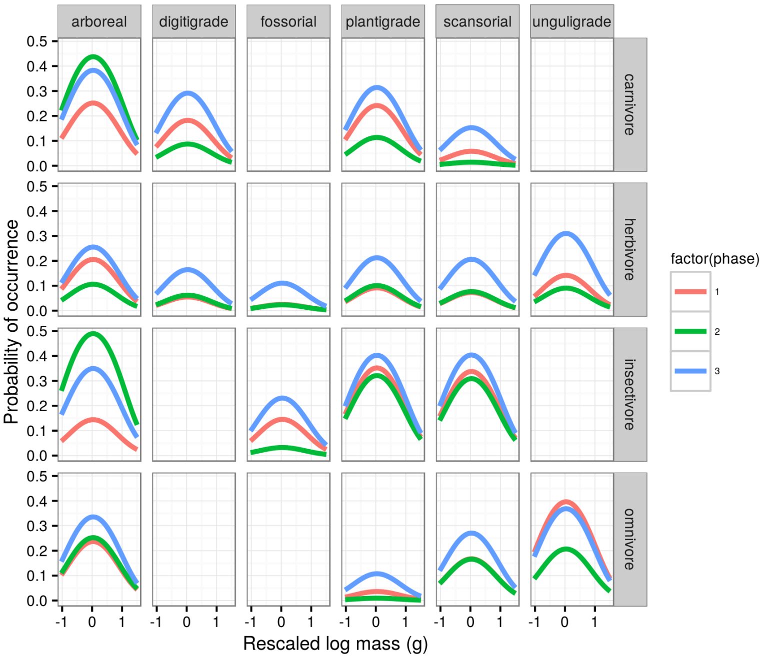


Figure 4.8: Mean estimate of the effect of species mass on the probability of a species occurrence for each of the three plant phases. The effect of mass is considered constant over time and that the only aspect of the model that changes with plant phase is the intercept of the relationship between mass 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.

and the two previous phases is difficult as the guanteed occurrence of the later taxa increases the average probability for that phase, which in turn affects the other time bins in that phase.

Plant phases are associated with large differences in log-odds for occurrence and origination probabilities (Tables 4.9, 4.10), though there is little evidence of plant phase being an important distinguishing factor in species survival, as only a few ecotypes demonstrate strong affinities with some plant phases (Table 4.11).

The effects of plant phase on occurrence and origination probabilities are broadly congruent with each other (Fig. 4.11, 4.12). The almost universal pattern of the effect of plant phase on ecotype origination is that the during first and last plant phases ecotypes have a greater

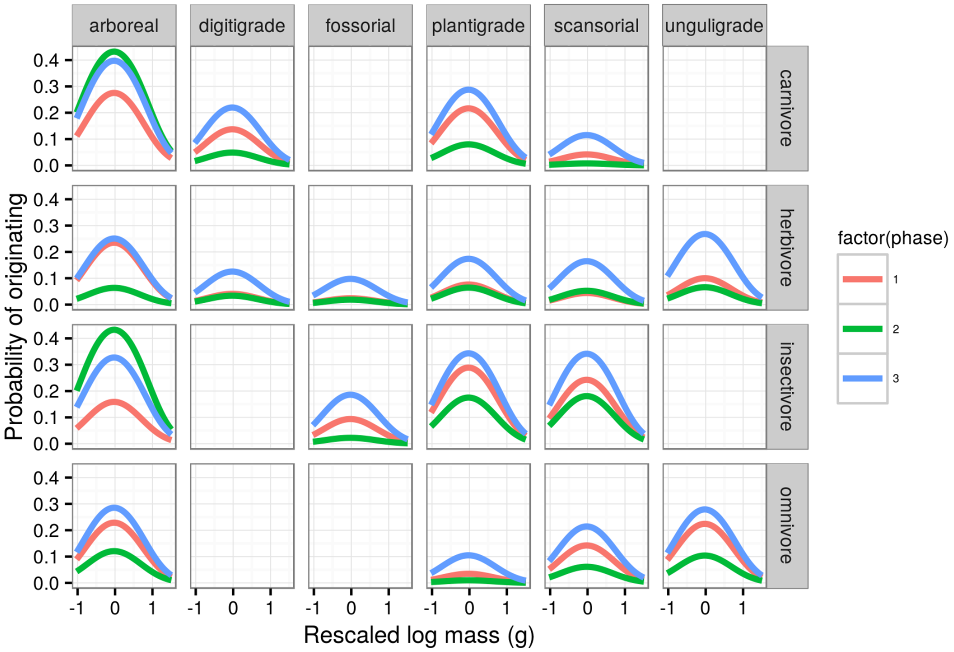


Figure 4.9: Mean estimate of the effect of species mass on the probability of a species originating for each of the three plant phases. The effect of mass is considered constant over time and that the only aspect of the model that changes with plant phase is the intercept of the relationship between 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.

log-odds of occurrence or origination than the second plant phase (Fig. 4.4, 4.5). The three ecotypes that do not follow this pattern are fossorial herbivores, scansorial herbivores, and arboreal insectivores.

Both aspects of global temperature analyzed here are estimated to have strong effects on species occurrence and origination for most mammal ecotypes (Tables 4.12, 4.13). Similarly, the probability that temperature has a large effect on species extinction is very low for all ecotypes (Table 4.14). The effects of the temperature covariates on ecotype occurrence and origination are estimated to be negative, which which means that as temperature decreases, occurrence and origination are expected to increase. In the case of survival, the only strong

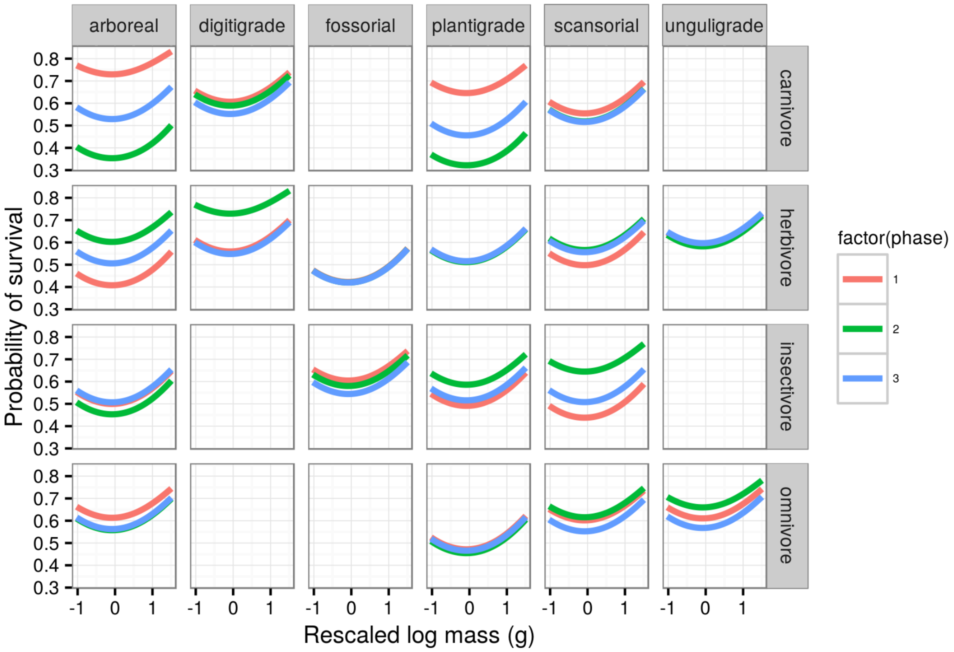


Figure 4.10: Mean estimate of the effect of species mass on the probability of a species survival for each of the three plant phases. The effect of mass is considered constant over time and that the only aspect of the model that changes with plant phase is the intercept of the relationship between mass and survival. The three plant phases are indicated by the color of the line. Mass has been log-transformed, centered, and rescaled; this means that a mass of 0 corresponds to the mean of log-mass of all observed species and that mass is in standard deviation units. For clarity, only the mean estimates of the effects of mass and plant plant are plotted.

ecotype association for either of the temperature covariates is a positive relationship between temperature range and occurrence probabilities of with plantigrade herbivores (Tab. 4.14).

The apparent similarities in origination rate of digitigrade carnivores, digitigrade herbivores, plantigrade herbivores, and unguligrade herbivores (Fig. 4.5 can be tested by inspecting the estimates of the two correlation matrices Ω^{ϕ} and Ω^{π}. The elements of these matrices are estimates 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 Gaussian, estimates of scale and correlation parameters are very approximate as these parameters tend to have decidedly non-Gaussian true posterior distributions [101].

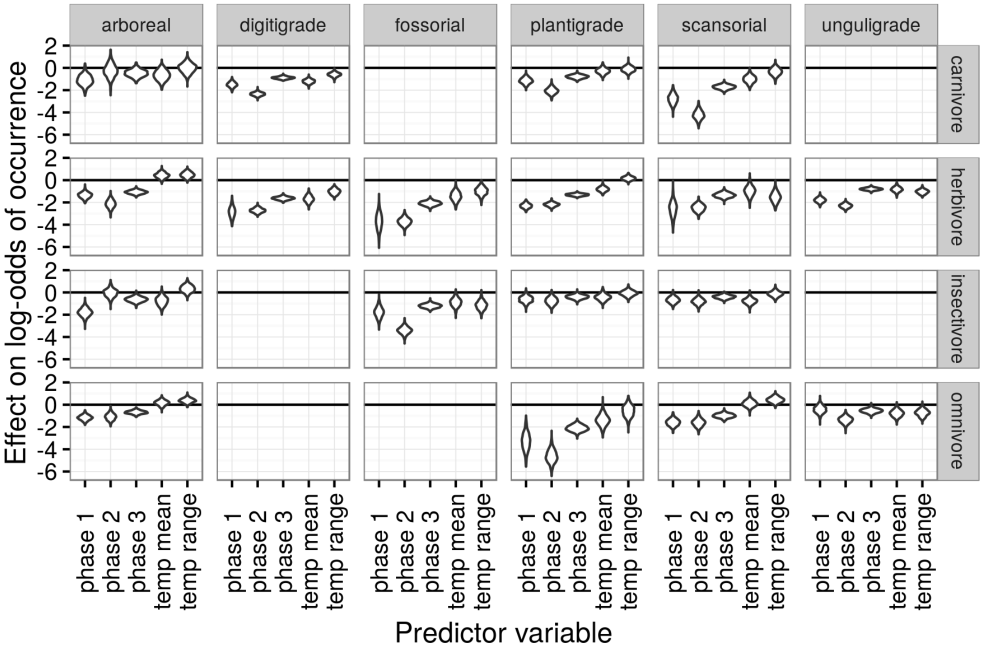


Figure 4.11: Estimated effects of the group-level covariates describing environmental context on log-odds of species occurrence. These estimates are from the pure-presence model. What is plotted is a violin of the distribution of 1000 samples from the approximate posterior. The effect of plant phase graphed here is calculated as Phase 1= $\gamma_{phase\ 1}$, Phase 2= $\gamma_{phase\ 1} + \gamma_{phase\ 2}$, and so on.

Because of this fundamental limitation, inference based on these correlation matrices are very approximate and subject to change.

Consistent with visual inspection of the ecotype origination probability time series, there are some strong positive correlations between a few ecotypes (Fig. 4.14). Given the posterior distribution of 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 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 greater than 95% posterior probability of being greater than 0 (Fig. 4.15); in total there are 35 correlations that fit this criterion. The ecotypes correlated with the most number of other ecotyes, in order from most to least, are

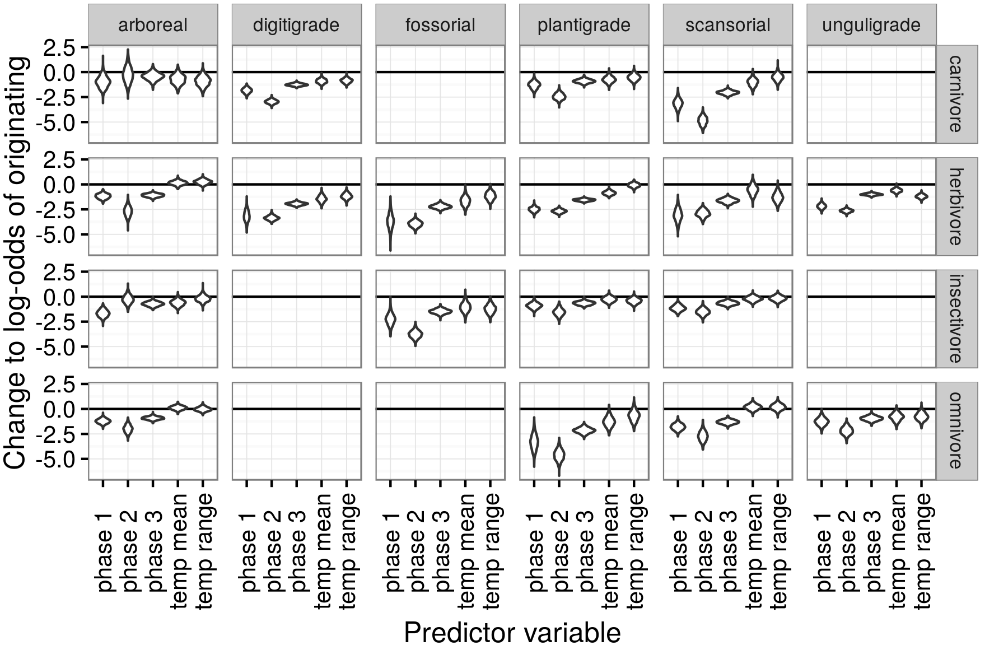


Figure 4.12: Estimated effects of the group-level covariates describing environmental context on log-odds of species origination. These estimates are from the birth-death model. What is plotted is a violin of the distribution of 1000 samples from the approximate posterior. The effect of plant phase graphed here is calculated as Phase 1= $\gamma_{phase\ 1}$, Phase 2= $\gamma_{phase\ 1} + \gamma_{phase\ 2}$, and so on.

unguligrade herbivores, plantigrade herbivores, digitigrade carnivores, scansorial carnivores, fossorial herbivores, and plantigrade omnivores. 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 lead to the simultaneous enrichment of multiple ecotypes instead of each ecotype respondingly independently. This result is most obvious for digitigrade carnivores, scansorial carnivores, digitigrade herbivores, fossorial herbivores, plantigrade herbivores, and unguligrade herbivores; these ecotypes are heavily cross-correlated with each other.

In contrast to the visually obvious correlations in ecotype origination probability, visual inspection of the ecotype-specific survival probabilities (Fig. 4.6) does not indicate that many strong correlations between ecotype survival probabilities. This conclusion is supported

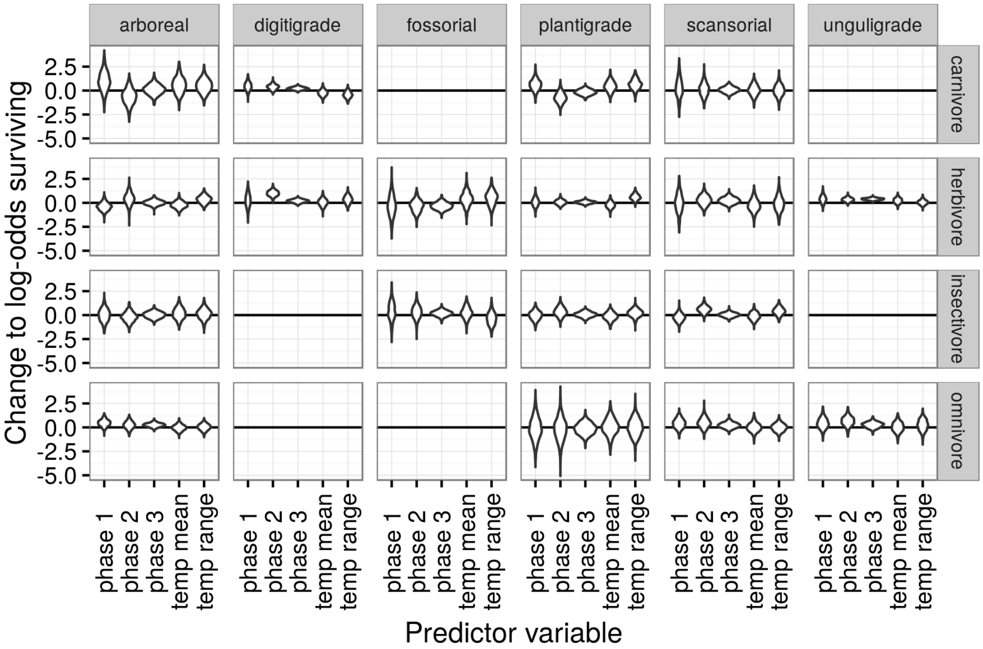


Figure 4.13: Estimated effects of the group-level covariates describing environmental context on log-odds of species survival. These estimates are from the birth-death model. What is plotted is a violin of the distribution of 1000 samples from the approximate posterior. The effect of plant phase graphed here is calculated as Phase 1= $\gamma_{phase\ 1}$, Phase 2= $\gamma_{phase\ 1} + \gamma_{phase\ 2}$, and so on.

by the estimated correlations in ecotype survival probability (Fig. 4.16). There are very 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 and unguligrade herbivores, has a greater than 95% posterior probability of being positive (Fig. 4.17). This single correlation, however, adds more nuance to the interpretations from the origination probability correlations. In addition to correlation in enrichment, these ecotypes are correlated in their depletions. This result supports the conclusion that the diversity histories of digitigrade 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.

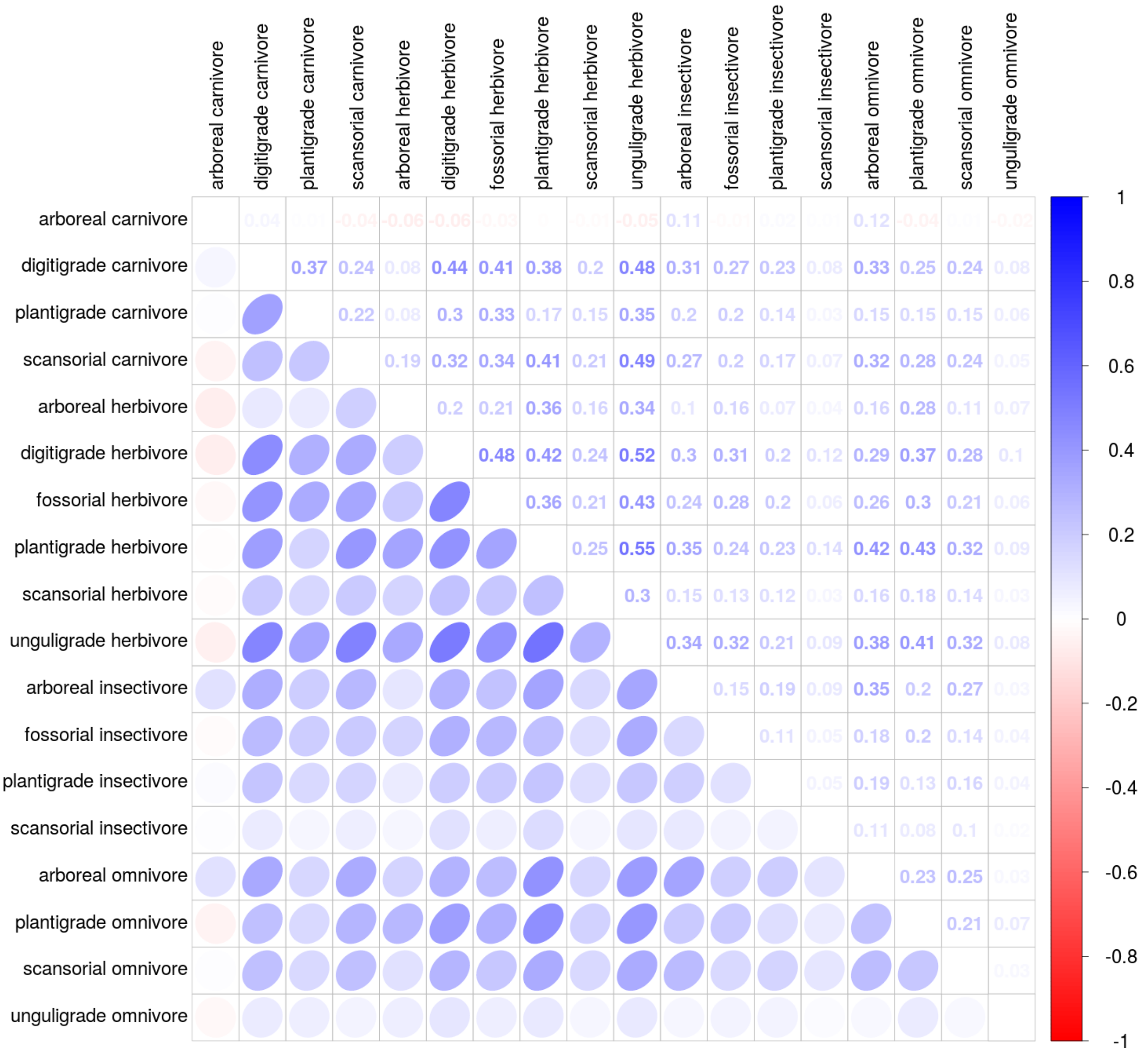


Figure 4.14: Posterior mean estimates of the correlations in origination probability between the mammal ecotypes. The lower triangle of the matrix is populated with ellipses corresponding to the level of correlation between the two ecotypes, while the upper triangle of the matrix corresponds to the mean 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.

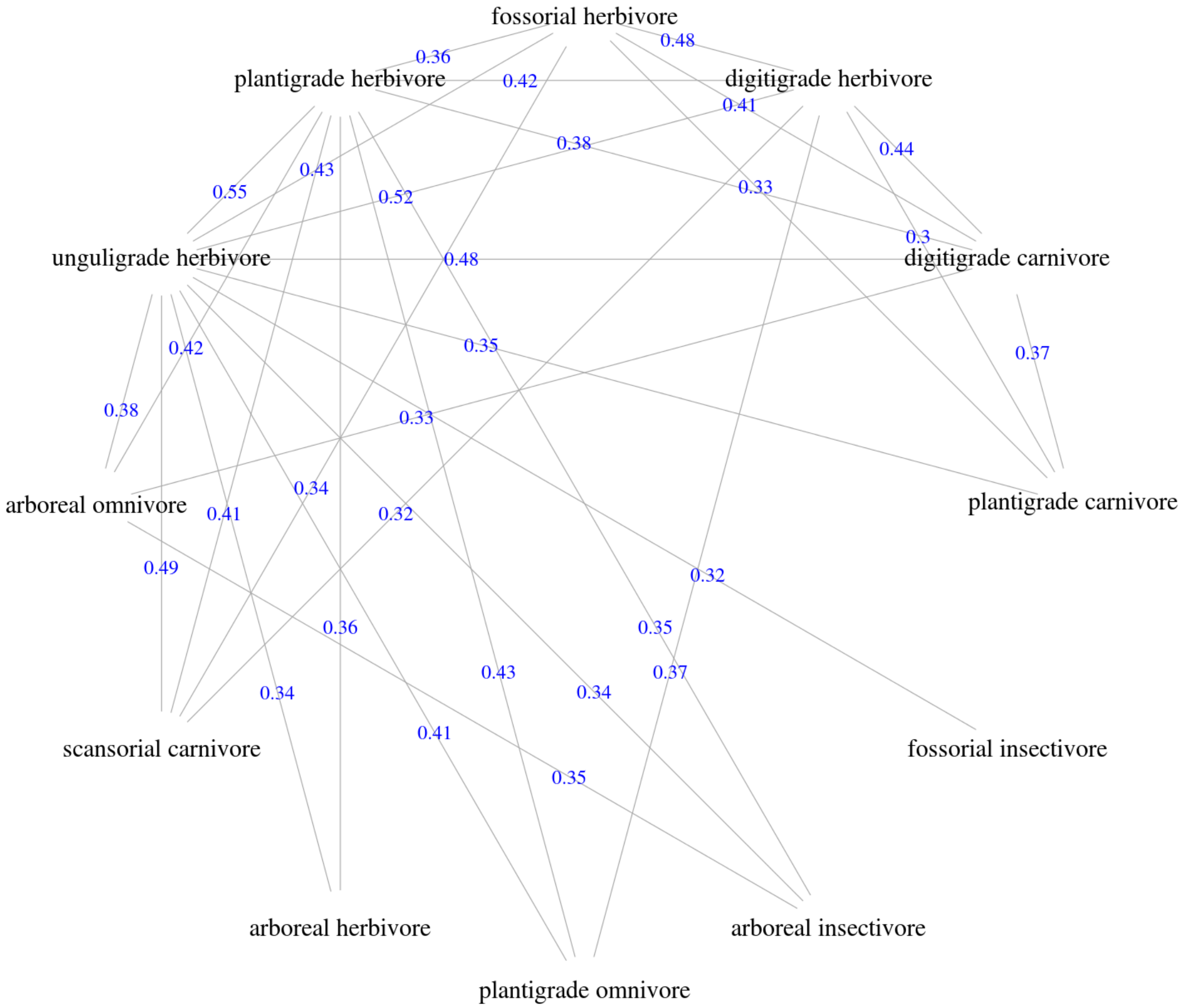


Figure 4.15: Ecotypes that have a strong correlation in origination probability. These ecotypes have a greater than 95% posterior probability of being positively correlated. The number plotted at the midpoint of each edge corresponds to the mean estimated correlation between those two ecotypes.

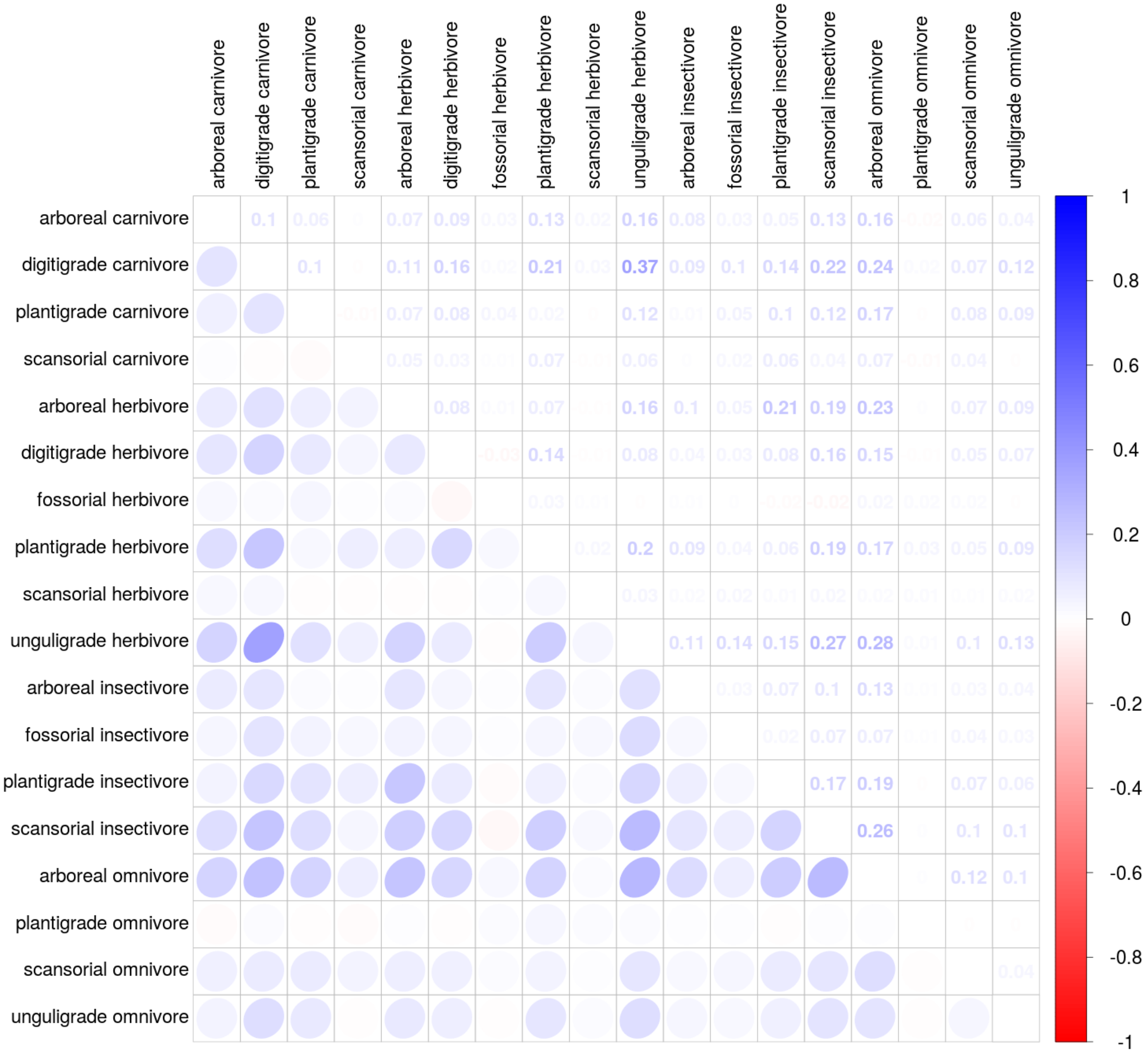


Figure 4.16: Posterior mean estimates of the correlations in survival probability between the mammal ecotypes. The lower triangle of the matrix is populated with ellipses corresponding to the level of correlation between the two ecotypes, while the upper triangle of the matrix corresponds to the mean 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.

Table 4.9: Posterior probability of the differences in the log-odds of an ecotype occurring based on plant phase. These probabilities are calculated as P(Phase 1 > 2) = $(\sum \gamma_{phase1} > \gamma_{phase1} + \gamma_{phase2})/100$ 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	0.323	0.874	0.926
digitigrade carnivore	1.000	0.000	1.000
plantigrade carnivore	1.000	0.040	1.000
scansorial carnivore	1.000	0.015	1.000
arboreal herbivore	1.000	0.515	1.000
digitigrade herbivore	1.000	0.995	1.000
fossorial herbivore	1.000	0.923	1.000
plantigrade herbivore	1.000	0.995	1.000
scansorial herbivore	1.000	0.717	1.000
unguligrade herbivore	1.000	0.000	1.000
arboreal insectivore	0.024	0.999	0.997
fossorial insectivore	1.000	0.000	1.000
plantigrade insectivore	0.923	0.558	0.985
scansorial insectivore	0.982	0.483	0.992
arboreal omnivore	0.959	0.837	1.000
plantigrade omnivore	1.000	0.247	1.000
scansorial omnivore	0.983	0.861	1.000
unguligrade omnivore	1.000	0.189	0.998

Table 4.10: Posterior probability of the differences in the log-odds of an ecotype originating based on plant phase. These probabilities are calculated as P(Phase 1 > 2) = $(\sum \gamma_{phase1} > \gamma_{phase1} + \gamma_{phase2})/100$ 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	0.405	0.777	0.905
digitigrade carnivore	1.000	0.043	1.000
plantigrade carnivore	1.000	0.053	1.000
scansorial carnivore	1.000	0.035	1.000
arboreal herbivore	1.000	0.163	1.000
digitigrade herbivore	1.000	0.998	1.000
fossorial herbivore	1.000	0.896	1.000
plantigrade herbivore	1.000	0.996	1.000
scansorial herbivore	1.000	0.884	1.000
unguligrade herbivore	1.000	0.003	1.000
arboreal insectivore	0.088	0.994	1.000
fossorial insectivore	1.000	0.020	1.000
plantigrade insectivore	0.995	0.419	1.000
scansorial insectivore	0.999	0.360	1.000
arboreal omnivore	0.999	0.317	1.000
plantigrade omnivore	1.000	0.308	1.000
scansorial omnivore	0.999	0.418	1.000
unguligrade omnivore	1.000	0.219	1.000

Table 4.11: Posterior probability of the differences in the log-odds of an ecotype surviving based on plant phase. These probabilities are calculated as P(Phase 1 > 2) = $(\sum \gamma_{phase1} > \gamma_{phase1} + \gamma_{phase2})/100$ 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	0.849	0.127	0.313
digitigrade carnivore	0.292	0.411	0.162
plantigrade carnivore	0.926	0.243	0.790
scansorial carnivore	0.409	0.544	0.453
arboreal herbivore	0.250	0.733	0.480
digitigrade herbivore	0.000	0.990	0.246
fossorial herbivore	0.547	0.680	0.812
plantigrade herbivore	0.404	0.555	0.480
scansorial herbivore	0.534	0.277	0.204
unguligrade herbivore	0.600	0.046	0.003
arboreal insectivore	0.673	0.379	0.539
fossorial insectivore	0.464	0.442	0.308
plantigrade insectivore	0.216	0.714	0.446
scansorial insectivore	0.019	0.923	0.382
arboreal omnivore	0.394	0.440	0.242
plantigrade omnivore	0.582	0.542	0.677
scansorial omnivore	0.292	0.590	0.289
unguligrade omnivore	0.212	0.555	0.183

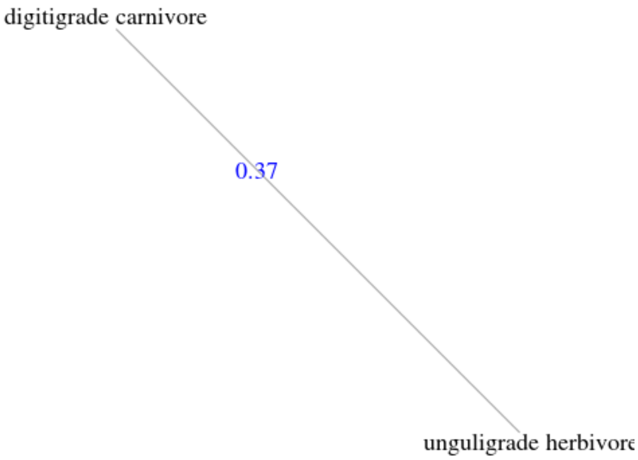


Figure 4.17: Ecotypes that have a strong correlation in survival probability. These ecotypes have a greater than 95% posterior probability of being positively correlated. The number plotted at the midpoint of each edge corresponds to the mean estimated correlation between those two ecotypes.

Table 4.12: 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	0.067	0.043
digitigrade carnivore	0.000	0.000
plantigrade carnivore	0.012	0.054
scansorial carnivore	0.007	0.086
arboreal herbivore	0.762	0.852
digitigrade herbivore	0.000	0.000
fossorial herbivore	0.000	0.002
plantigrade herbivore	0.000	0.364
scansorial herbivore	0.141	0.004
unguligrade herbivore	0.001	0.000
arboreal insectivore	0.031	0.251
fossorial insectivore	0.014	0.002
plantigrade insectivore	0.164	0.058
scansorial insectivore	0.197	0.250
arboreal omnivore	0.711	0.449
plantigrade omnivore	0.009	0.103
scansorial omnivore	0.754	0.732
unguligrade omnivore	0.015	0.022

Table 4.13: 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	0.072	0.031
digitigrade carnivore	0.000	0.001
plantigrade carnivore	0.010	0.094
scansorial carnivore	0.006	0.104
arboreal herbivore	0.785	0.881
digitigrade herbivore	0.000	0.000
fossorial herbivore	0.001	0.007
plantigrade herbivore	0.000	0.626
scansorial herbivore	0.097	0.002
unguligrade herbivore	0.010	0.000
arboreal insectivore	0.018	0.319
fossorial insectivore	0.008	0.003
plantigrade insectivore	0.223	0.084
scansorial insectivore	0.139	0.202
arboreal omnivore	0.646	0.628
plantigrade omnivore	0.013	0.120
scansorial omnivore	0.688	0.735
unguligrade omnivore	0.016	0.027

Table 4.14: 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	0.751	0.752
digitigrade carnivore	0.161	0.110
plantigrade carnivore	0.851	0.878
scansorial carnivore	0.598	0.510
arboreal herbivore	0.281	0.819
digitigrade herbivore	0.546	0.721
fossorial herbivore	0.751	0.766
plantigrade herbivore	0.316	0.960
scansorial herbivore	0.342	0.429
unguligrade herbivore	0.750	0.418
arboreal insectivore	0.572	0.579
fossorial insectivore	0.648	0.279
plantigrade insectivore	0.477	0.786
scansorial insectivore	0.303	0.845
arboreal omnivore	0.536	0.611
plantigrade omnivore	0.499	0.507
scansorial omnivore	0.544	0.488
unguligrade omnivore	0.520	0.688

4.3.2 Analysis of diversity

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. 4.3).

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 diverisity, does not fluctuate wildly and rapidly over the Cenozoic, and demonstrates no sustained directional trends (Fig. 4.18a). In broad strokes, the first 15 or so million years of the Cenozoic are characterized by first an 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). After this, there is a rapid spike in diversity followed by a slight decline in diversity up to the Recent.

The pattern exhibited by the diversity history estimated in this study (Fig. 4.18a) has some major similarities with previous mammal diversity curves [6]: both diversity history estimates begin with an increase in diversity and most of the major increases in diversity are retained including the large diversity spike during the Miocene. Unlike subsampling based approaches to estimating diversity [7], I'm able to interpolate over unsampled/poorly sampled time periods because of how the hierarchical model can share information across the different units [101]; for 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 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 [101]. Note that my time series of estimated diversity begins at a slightly different point than that used by Alroy [6] and that the time intervals used by Alroy [6] 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. 4.18a).

When viewed through the lens of diversification rate, some of the structure behind the estimated diversity history begins to take shape (Fig. 4.18b). 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. 4.18b). Other notable increases in diversification rate occur at 56, 46, 22, 18, and 6 Mya (Table 4.15), though the last of these may be due to edge effects surrounding the partial-identifiability of $p_{t=T}$. Notable decreases in diversification rate occur at 54, 50, 48, 44, 40, 34, 30, 24, 20, 16, 12, and 8 Mya (Table 4.15), meaning that 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 positive diversification rates have a greater magnitude than most periods of low or negative diversification (Fig. 4.18b). Given that diversification rate more closely resembles origination rate than extinction rate (Fig. 4.18b, 4.18c, 4.18d), these decreases in diversification rate may be indicative of "depletions" (failure to replace extinct taxa) rather than pulses of extinction.

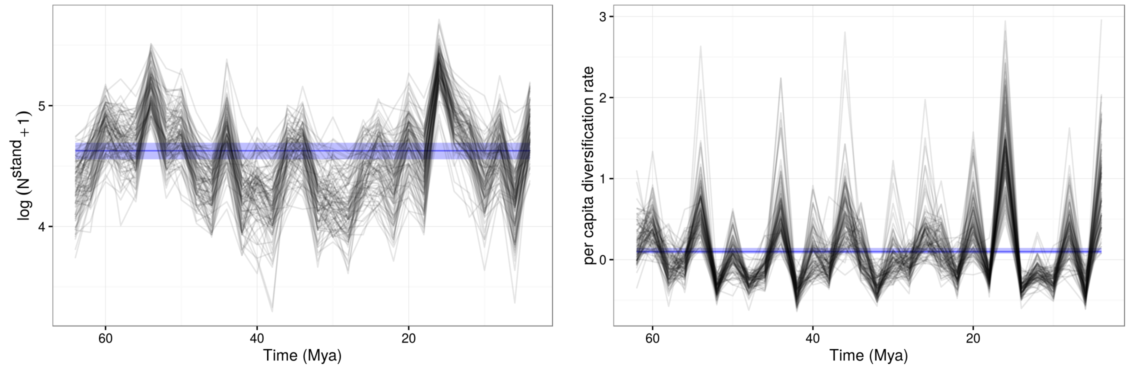
The estimates from this study of per capita origination and extinction rates for the entire species pool (Fig. 4.18c, 4.18d) are very different from the origination and extinction rates estimated by Alroy [6]. The two most striking differences are the very different estimates of extinction rate between the two studies and the very different scales of the origination rate estimates. This may be due to the fundamentally different way these rates are calculated, and how the diversification process was modeled. The per capita rates estimated in this study follow straight from the definition of a per capita rate (e.g. number of originations between time t and $t + 1$ divided by the diversity at time t) while the rates calculated by Alroy [6] are based on log ratios of standing diversity.

The comparison between per capita origination and extinction rate estimates reveals how diversification rate is formed (Fig. 4.18c, 4.18d). As expected given previous inspection of the ecotype 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. Extinction rate, on the other hand, demonstrates an almost saw-toothed pattern around a constant mean (Fig. 4.18d). These results are broadly consistent with those from previous analyses of North American mammals diversity and diversification [5, 10, 6].

Diversity partitioned by ecotype reveals a lot of the complexity behind the pattern of mammal diversity for the Cenozoic (Fig. 4.19).

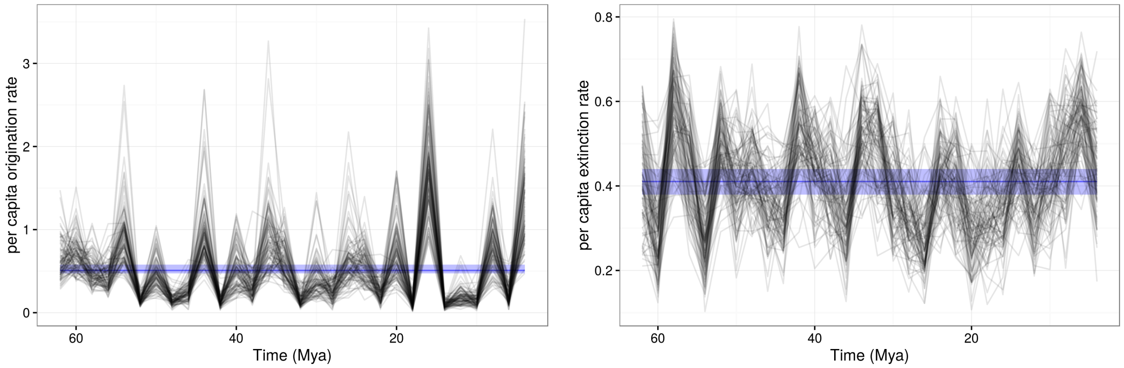
Arboreal ecotypes obtain peak diversity early in the Cenozoic and then decline for the rest of the time series, becoming increasingly rare or absent as time approaches the Recent (Fig. 4.19). Arboreal herbivores and omnivores obtain peak diversity at the beginning of the Cenozoic then go into decline while remaining a small part of the species pool, while arboreal carnivores and insectivores obtain peak diversity 52-50 Mya and then quickly decline and become extremely rare or entirely absent from the species pool. This is consistent with increasing extinction risk in the Neogene compared to the Paleogene as proposed by Smits [309].

The diversity of digitigrade and unguligrade herbivores increases over the Cenozoic (Fig. 4.19). In contrast, plantigrade herbivore diversity does not have a single, broad-strokes pattern; instead, diversity increases, decreases, and may have then increased till the Recent. In contrast, fossorial and scansorial herbivores demonstrate a much flatter history of diversity, with a slight increase in diversity that over time is more pronounced among fossorial taxa than scansorial taxa. The expansion of digitigrade and unguligrade herbivores over the Cenozoic is consistent with the gradual expansion of grasslands which these ecotypes are better adapted to than closed environments [28, 326].



(a) Log diversity

(b) Diversification rate



(c) Origination rate

(d) Extinction rate

Figure 4.18: Posterior estimates of the time series of Cenozoic North American mammal diversity and it's 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 [271]; the hierarchical modeling strategy used here helps mitigate these effects but they are still present [101, 271]. 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.

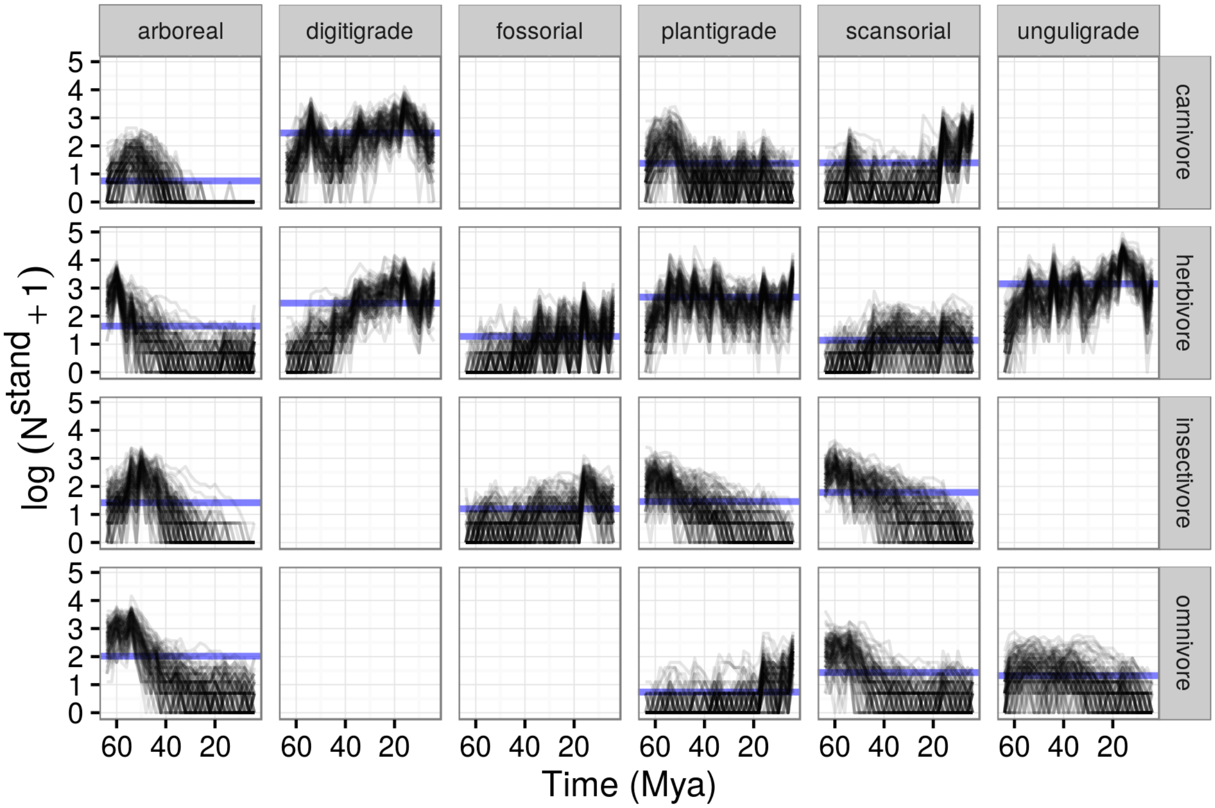


Figure 4.19: Posterior of standing log-diversity of North American mammals by ecotype for the Cenozoic as estimated from the birth-death model; 100 posterior draws are plotted to indicate the uncertainty in these estimates and what is technically plotted is log of diversity plus 1.

Table 4.15: Posterior probabilities of diversity N_t^{stand} or diversification rate D_t^{rate} being greater than average standing diversity $\overline{N^{stand}}$ or average diversification rate $\overline{D^{rate}}$ 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)	$P(N_t^{stand} > \overline{N^{stand}})$	$P(D_t^{rate} > \overline{D^{rate}})$
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	

Digitigrade carnivores have a multi-modal diversity history, with peaks at 54-52 and 12-10 Mya (Fig.4.19). Between these two peaks digitigrade carnivore diversity dips below average diversity following the first peak and then grows slowly until the second peak. Plantigrade carnivores obtain peak diversity in the early Cenozoic and then maintain a relatively stable diversity until another peak at the end of the Cenozoic. The generally flat diversity history of digitigrade carnivores lacks any sustained temporal trends and seems to reflect previous findings of limited diversity in spite of constant turnover and morphological evolution [338, 290, 304]

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 decreasing pattern with time, while fossorial insectivores have a flat diversity history with a peak approximately 10 Mya (Fig. 4.19). Arboreal and scansorial omnivores decrease in diversity from their initial peaks early in the Cenozoic, and plantigrade omnivores have a generally flat diversity history with a sudden peak in diversity late in the Cenozoic (Fig. 4.19). Unguligrade omnivores also demonstrate a possible decrease in diversity over the Cenozoic, but not as clearly as arboreal and scansorial omnivores.

The waxing and waning of the mammal ecotypes is obvious when comparing changes to estimated relative log-mean of diversity (Fig. 4.20). While ecotype diveristy does appear to change gradually, there are definite changes to the relative contributions of the ecotypes to the regional species pool. All arboreal ecotypes clearly decrease in relative diversity over the Cenozoic. In contrast the the digitigrade herbivore, fossorial herbivore, scansorial herbivore, and unguligrade herbivore ecotypes which increase in relative diversity over the Cenozoic. The digitigrade carnivore ecotype increases in relative diveristy until approximately the start of the Neogene, after which it maintains a generally constant relative diveristy; this is consistent with previous observations of constant or density-dependent diversity of the canid guild for the Neogene [338, 290, 304], a guild that overlaps with the digitigrade carnivore ecotype. Plantigrade herbivores remain a constant relative contribution to ecotypic diveristy. These results support the hypothesis of a gradual transition from the early Paleogene with a region with more avaliable habitat for aboreal taxa and less avaliable habitat for many digitigrade and unguligrade taxa, to an environment where arboreal taxa are absent from the species pool and digitigrade and unguligrade taxa are much more dominant (Fig. 4.20). It is the relative contributions of digitgrade carnivores, digitigrade herbivores, and unguligrade herbivores which really shape the regional species pool of the Neogene.

Many of the estimated ecotype-specific diversity histories share a similar increase in diversity in the late Cenozoic, 16-14 Mya (Fig. 4.19). These increases are either sustained or temporary and are seen in digitigrade carnivores, plantigrade carnivores, scansorial carnivores, unguligrade herbivores, fossorial insectivores, and plantigrade omnivores.

When ecotype diversity is decomposed into per capita origination (Fig. 4.21) and per capita extinction rates (Fig. 4.22) the way in which their diversity developed can be explained. For ecotype-specific 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.

As should be expected from the comparisons between estimates of origination and survival probability, 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. 4.18b). Importantly, the source of the massive increase in diversification rate at 16 Mya can be attributed almost solely to the origination of unguligrade herbivores (Fig. 4.21). Additionally, by decomposing origination rate by 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 increases in origination rate with unguligrade herbivores; these are all ecotypes that demonstrate an obvious increase in diversity during the Paleogene and then

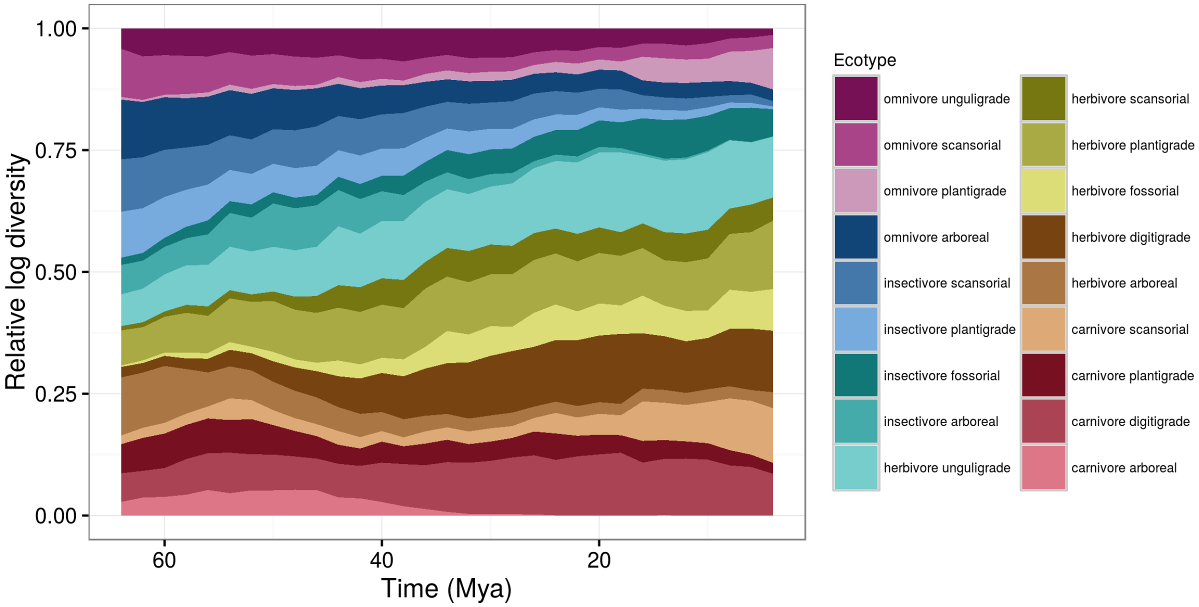


Figure 4.20: Mean posterior estimate of relative log standing diveristy of 18 North American mammal ecotypes for the Cenozoic. These estimates are calculated from 100 posterior estimates of the true occurrence matrix z as estimated from the birth-death model.

maintain relatively high diversity through out the Neogene (Fig. 4.19).

In contrast to ecotype-specific per capita origination rates which demonstrate distinct peaks, the estimates of ecotype-specific per capita rates are more of a smear (Fig. 4.22). There are few increases in extinction rate that are shared across ecotypes. The per captia extinction rates of digitigrade, 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 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 diversity pattern for these ecotypes (Fig. 4.19), 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 result highlights the distinction between extinction risk and extinction rate; an ecotype can have a high extinction risk, but if that ecotype is not present in the species pool in the first place than it has no associated extinction rate.

4.4 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 [152, 304, 338]. This implies that the size and scope of the

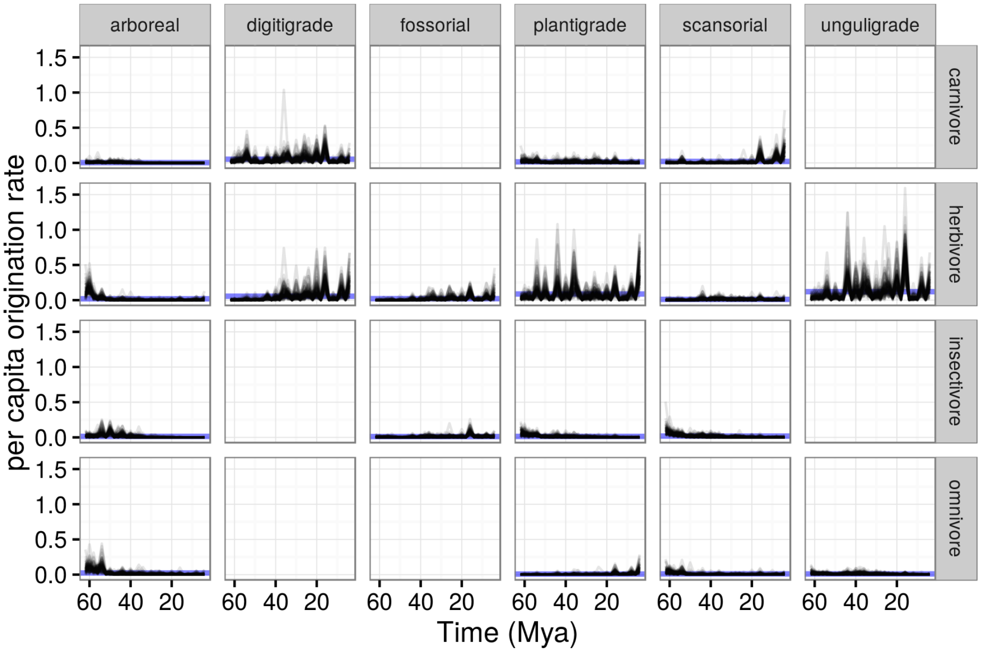


Figure 4.21: Posterior estimates of the per capita origination rates for each ecotype, plotted at the bin they originate from. These rates are calculated as the number of origination events for that ecotype from one time point to the next, divided by the standing diversity of all mammals at the initial time point. 100 posterior draws are plotted to indicate the uncertainty in these estimates.

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.” There is even evidence that functional groups are at least partially self-organizing and truly emergent [274].

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 the origination and extinction processes. Additionally, changes to the ecotypic composition and diversification rate of the North American regional species pool are driven primarily by variation in origination and not extinction (Fig. 4.18). These aspects

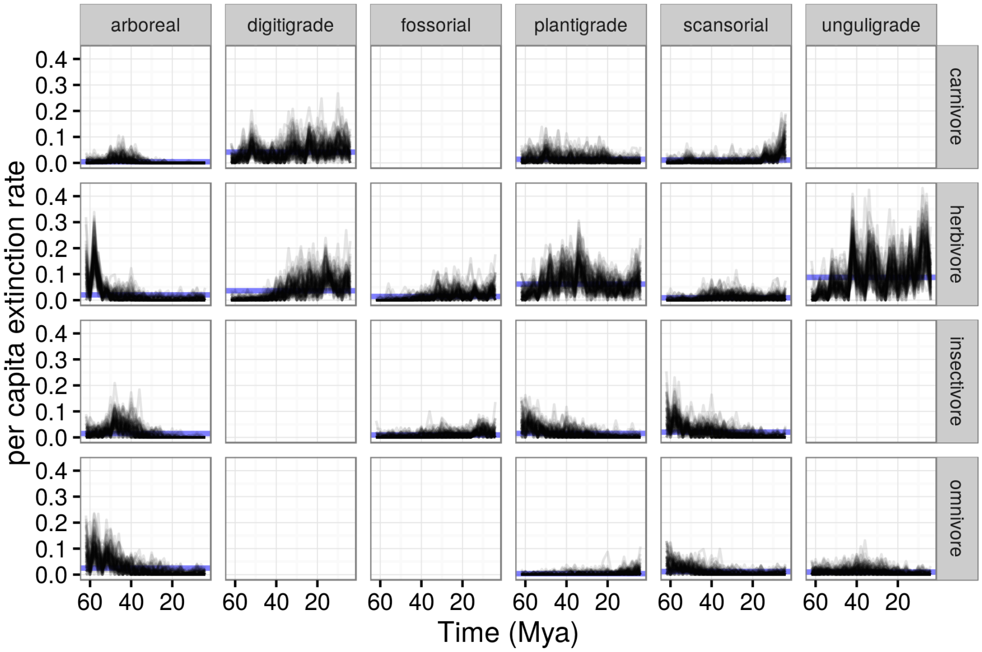


Figure 4.22: Posterior estimates of the per capita extinction rates for each ecotype, plotted at the bin they go extinct from. These rates are calculated as the number of extinction events for that ecotype from one time point to the next, divided by the standing diversity of all mammals at the initial time point. 100 posterior draws are plotted to indicate the uncertainty in these estimates.

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 [97, 236].

Extinction rate for the entire regional species pool through time is highly variable and demonstrates a saw-toothed pattern with no obvious temporal trends. While a constant mean extinction rate is consistent with previous observation [5, 10], the degree to which mammal extinction rate is actually variable may not have been equally appreciated as it has been for the marine invertebrate record [79, 80, 83, 84]. 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 [5, 10].

Comparison of the ecotype specific diversity histories adds a considerable degree of nuance to broad narrative of shifts in functional composition of the North American mammal species pool as being gradual (Fig. 4.19). Origination and extinction rates do not suddenly shift through time for most ecotypes (Fig. 4.21, 4.22). As with the diversification rate of the entire species pool, the diversification of individual ecotypes seem principally driven by origination and not extinction. Instead, while species seem to originate in waves (Fig. 4.21), they appear to leave the regional species pool in an uncoordinated and individual manner (Fig. 4.22) which could be considered consistent with the maxim that all species respond differently to environmental change [28]. Note, however, this result characterizes the entire North American mammal regional species pool and thus may not reflect the dynamics of specific local communities but instead represent the average community.

The few large-magnitude, but temporary, increases in ecotype-specific origination rate occur in digitigrade carnivores, digitigrade herbivores, plantigrade herbivores, and unguligrade hebivores. Importantly, the large peak in diversification and origination rates 16 Mya (Fig. 4.18) appears driven almost entirely by a massive increase in the origination rate of unguligrade herbivores (Fig. 4.21). Additionally, there is some evidence that the origination probabilities of these ecotypes are correlated (Fig. 4.14, 4.15). While this result does not mean that there are large and sudden cross-ecotype changes to the regional species pool, it does suggest that additions to the species pool do not occur in individual ecotypes idiosyncratically.

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 $\sim$22 Mya. This is partially consistent with one of the two possible patterns presented here and by Smits [309] that would result in arboreal taxa having a greater extinction risk than other ecotypes: the Paleogene and Neogene were different selective regimes and, while the earliest Cenozoic may have been neutral with respect to arboreal taxa, they disappeared quickly over the Cenozoic which may account for their higher extinction risk. However, these result add some nuance to this scenario as arboreal taxa were declining throughout the Paleogene instead of maintaining a flat diversity as hypothesized [309]. I interpret the decline of arboreal taxa throughout the Paleogene to mean that the shift from closed to open environments began in the Paleogene and led to increasingly hostile environments for arboreal taxa as opposed to being a sudden change in selective regime between the Paleogene and Neogene. In addition to all arboreal taxa, the diversity of plantigrade and scansorial insectivores decreases with time (Fig. 4.19).

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 overlap with the carnivore guild which has been the focus of much research [304, 338, 230, 147]. This result is consistent with some form of “control” on the diversity of this ecotype, such as diversity-dependent diversification [304, 290, 338].

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 [28, 326, 109] and may indicate a long-term shift in the interactors associated with those ecotypes leading to increased contribution to the regional species pool. This result may be comparable to the increasing percentage of hypsodont (high-crowned teeth) mammals in the Neogene of Europe being due to an enrichment of hyposodont taxa and not a depletion of non-hypsodont taxa. Smaller scale increases in fossorial herbivore species, and a lesser extent plantigrade herbivores, suggests that the increase of interactors may be associated mostly 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 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 extinctions of all herbivore ecotypes (Fig. 4.22).

An association between plant phase and differences in ecotype occurrence or origination-extinction probabilities is interpreted to mean that an ecotype enrichment or depletion is due to associations between that ecotype and whatever plants are dominant at that time. Plant phase clearly structures the occurrence and origination probability time series (Fig. 4.4, 4.5). These differences in occurrence or origination translate to the estimates of diversity and diversification rate; the largest spike in diversity, diversification rate, and origination rate all correspond to the onset of the last plant phase (Fig. 4.18). The clearest example of the diversity of an ecotype increasing at this particular transition is in scansorial carnivores (Fig. 4.19); similar shifts in other ecotypes are much more subtle, as was previously noted for fossorial insectivores.

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 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 mammal-plant interactions due to the increased modernization of taxa characteristic of the younger plant phases [109]. Alternatively, the increase in diversity associated with the third plant phase may be caused by the edge effect in origination probability that is artificially inflating the number of origination events (Fig. 4.5). However, the estimated number of origination events does not have a tremonedous spike at this transition, nor is a major increase in the number of origination events sustained (Fig. 4.21).

There are fewer, less obvious shifts in diversity surrounding the transition from the first to second plant phase, with the following ecotypes having apparent shifts in diversity at 50 My: plantigrade carnivores (down), arboreal omnivores (down), and scansorial omnivores (down). Arboreal insectivore peak diversity also occurs 50 Mya, and is then followed by a steep decline in diversity till 30 Mya when this ecotype is lost from the species pool. Because plant phase has been found to structure occurrence/origination (Fig. 4.4, 4.5), but not survival (Fig. 4.6), my interpretation of these results is that new species were not entering the system because there were fewer available interactors available for those ecotypes. Instead, these ecotypes were poorly suited for the newly available mammal-plant interactions brought upon by the changing environmental context [109].

The temperature covariates are found to have similar effects on occurrence and origination probabilities (Tables 4.12, 4.13). Temperature is found to more often affect ecotype occurrence probabilities than origination probabilities. In most cases, there is a negative association between temperature and probability of occurring or first originating; this means that if temperature decreases, we would then expect an increase in the probability of occurring or first originating. In contrast, temperature range is estimated to be a good predictor of survival in only to mammal ecotypes and only marginally for one of those (Table 4.14). Additionally, both of these cases have positive relationships, meaning that if temperature decreases it is expected that species survival will also decrease.

The result that temperature does not affect the survival probability of most ecotypes is consistent with previous analysis of mammal diversity [10]. The result that temperature affects origination probability, on the other hand, is in strong contrast to previous results [10]. An important difference between the anlayses presented here and those obtained by Alroy [10] is I am considering the effect of temperature on the probability of a species originating, assuming it hasn't originated yet while Alroy [10] analyzed the correlation between the first differences of the origination and extinction rates with an oxygen isotope curve [368]. Origination or extinction rates have very different properties than the origination probabilities estimated here brought upon by the difference both in definition and units. Origination probability is the expected probability that a species that has never been present and is not present at time t will be present at time $t + 1$; origination probability is defined for a single species. In contrast, per capita rates are defined (for origination) as the expected number of new species to have originated between time t and $t + 1$ given the total number of species present at time t; per capita rates are defined for the standing diversity. It is also important to note that even though there is an edge effect at the last time interval that causes an increase in the occurrence and origination probabilities of some ecotypes (Fig. 4.4, 4.5, the corresponding rates and population level birth/death dynamics do not share that pattern (Fig. 4.18, 4.21, 4.22). However, it is still possible that the finding that temperature has an effect on origination may simply be because as time approaches the present the number of species which have originated increases and not because of climatic forcing of origination.

All environmental factors are found to affect the occurrence and origination probabilities for most, but not all, mammal ecotypes (Fig. 4.11, 4.12). In contrast, the environmental factors probably do not affect differences in ecotype survival probability (Fig. 4.13). The 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 “environment” affects mammal diversity and diversification [10, 72]. The environment or climate are more than just global or regional temperature, it is also the set of all possible biotic and abiotic interactions that can be experienced by a member of the species pool. By including more descriptors of species’ environmental context than simple an estimate of global temperature a more complete “picture” of the diversification process is inferred.

Analysis of relationship between temperature and origination rate is probably better suited for a continuous-time birth-death or multilevel stochastic differential equation model instead of a discrete-time model because the continuous models estimate rates while discrete time models estimate probabilities [4]. For example, the `PyRate` model(s) are based on a continuous-time birth-death process [291, 290]. Unfortunately, a continuous-time model may be unsuited for most paleontological data as the fossil record is naturally discrete; fossils are assigned to temporal units, such as stages, which have age ranges. Individual fossils are not assigned individual numeric ages. This reality was in fact my one of motivations for using discrete-time birth-death model instead of one in continuous-time. There are of course exceptions to this characterization; the fossil record of graptolites from the Ordovician and Silurian [55] and the fossil record of some mammal orders from Neogene are of high enough resolution that the application of continuous-time models is appropriate and less fraught.

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 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 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 [195]. A control means that if 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. 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 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 warrant their inclusion.

4.4.1 Conclusions

These results add a considerable degree of nuance to the narrative of changes to North American diversity being gradual. My results support the conclusions that ecotypic diversity is shaped more by changes to origination than extinction and that major changes to total diversification rate can be attributed to increases in origination of only some ecotypes. There are a number of interesting estimated ecotype diversity patterns. While arboreal ecotypes are diverse in the Paleogene, by the Neogene all arboreal ecotypes had dramatically decreased in diversity and became either rare or absent in 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 on 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.

CHAPTER 5

CONCLUSION

Macroevolution and macroecology are disciplines devoted to explaining emergent patterns in evolutionary and ecological data. These disciplines are linked through the analysis of the distribution of trait values across time, space, and/or species [196, 358]. Emergent evolutionary and ecological patterns in time can require at least a million of years to observe [332]. Paleontological and phylogenetic data preserve aspects of these large scale temporal patterns. Paleontological data is unique however in being empirical observations of these dynamics while phylogenetic data only preserves the branching history leading to the diversity pattern exhibited by the tips. Additionally, phylogenies of only extant taxa are most likely unrepresentative of the evolutionary history of that clade; for example, multiple rounds of extinction and radiation can obscure the actual diversification process from possibly being inferred [205, 204, 174, 239].

In the studies presented as a part of this dissertation, I have emphasized functional traits. These are traits which directly relate to the way in which a taxon interacts with its environment [196]. When these traits are defined for a species, they are called species traits [196]; examples include species geographic range, average body size, trophic level, environmental preference, etc. Functional traits are an excellent window in macroevolution and macroecology because of their obvious selective importance: an organism which cannot interact with its environment is by definition maladapted. Additionally, the simultaneous analysis of multiple functional traints improves overall process-based inference because it more accurately accounts for the ways in which species interact.

5.1 Summary

All of the studies conducted here were analyses of fundamentally emergent patterns which are not reducible to the properties of their constituents. In each of these studies, hypotheses and analysis were framed in terms of how a species functional ecology can be associated with or shape these emergent patterns. Because of the complexity of processes which shape these emergent patterns, as well as the vagaries of the fossil record, each of these studies required the development of specific inference devices (e.g. statistical models) which attempt to estimate the actual quantities of interest to that analysis.

The emergent pattern at the heart of both the first and second studies (Chapters 2, 3) is species duration. A species endures because of the continued success of individuals of that species but the duration of that species is only knowable by integrating across all individuals. The third study (Chapter 4) deals with a fundamentally different emergent pattern: the functional composition of a regional species pool. A regional species pool is the set of species present in all communities. While the functional composition of a community depends on the set of interactors at that locality, the functional composition of a regional species pool depends on the possibility of those interactors being present in at least one constituent community.

In Chapter 2 I tested two long standing hypotheses of how species durations are structured: the survival of the unspecialized hypothesis [298], and the Law of Constant Extinction [335]. I analyzed how the distribution of mammal species durations is affected by differences in multiple species traits, species' phylogenetic relatedness, and species' origination cohort. My results supported the conclusion that generalist mammal species will, on average, have a greater duration than more specialized mammal species. I also found that phylogeny and origination cohort contribute sub-equally to variation in species duration. Additionally, I found that only some of factors associated with extinction risk for Modern mammals could be considered risk factors for mammals from the rest of the Cenozoic. For example, while there is an increased risk of extinction with increasing body size amoungst Modern fauna, this was not the normal case for Cenozoic. Finally, I found evidence of species extinction risk increasing with species duration, a result that is counter the Law of Constant Extinction.

Chapter 3 also deals with the survival of the unspecialized hypothesis as well as the Law of Constant Extinction, this time with the global record of post-Cambrian Paleozoic brachiopods. In addition to these hypotheses, I also analyzed the relationship between extinction intensity and the strength of trait selection, namely "do the selective differences between traits increase or decrease with average fitness increases or decreases?" I found a negative correlation between intensity of extinction and the effects of geographic range and environmental preference. As with Chapter 2, I also found support for greater survival among environmental generalists than specialists. These results supported the conclusion that, at least for Paleozoic brachiopods, as extinction intensity increases, the selective difference of traits increases. In this analysis this means that when average duration decreases (e.g. intensity is high) the effect of genus geographic range increases in magnitude and taxa which favor epicontinental environments are expected to have a greater duration than those which favor open-ocean environments. I also find that the change in magnitude of effect is expected to be greater for environmental preference than for geographic range as the overall effect of former have a much greater variance than that of the later.

The final study of this dissertation (Chapter 4) was an analysis of how the functional composition of the North American mammal regional species pool changed over time and in response to multiple environmental factors. The goals of this analysis were to understand when are different ecotypes enriched or depleted in the regional species pool, and to undestand how changes to environmental context may affect changes to the functional diversity of the regional species pool. By focusing on functional groups instead of taxonomic groups, the results from this study are phrased in terms of species interactions and not differences in clade diversity. My results add considerable nuance to the taxon-focused narrative of North American environmental change. There are many results and conclusions from this analysis, so I focus here on a few key results. I found that mammal diversity is more strongly shaped by changes to origination rate among a few ecotypes rather than being driven by selective extinction one or more ecotypes. I also found that all arboreal ecotypes decline throughout the Paleogene and disappear from the species pool by the Neogene. Additionally, I found that most herbivore ecotypes expand their relative contribution to functional diversity over time. Finally, I found that the environmental factors analyzed here structure differences in ecotype origination probability but not survival probability.

5.2 Synthesis

These three chapters are united by their analysis of species functional traits. Analysis of species traits is the easiest way to unite macroevolutionary and macroecological inference by using this commonality to develop and test integrated hypotheses [196, 358]. All inferences lead to more hypotheses as we realize what we do and do not know. Inferences from analysis of macroevolutionary patterns may lead to hypotheses about macroecological patterns, such as the relationship between extinction risk and diversity in arboreal mammals (Chapter 2, and 4).

The first two studies, when considered together, add a considerable degree of nuance to our understanding of multiple macroevolutionary hypotheses such as the Law of Constant Extinction and the survival of the unspecialized.

First and foremost, the results of neither study support the Law of Constant Extinction [335]. I found evidence for extinction risk increasing with taxon duration; these results are consistent with those of a “nearly-neutral” theoretical model of macroevolution [267]. While the dynamics of this model are described as “Red Queen” [267], this is not strictly true as Red Queen dynamics as described by Van Valen [335] require that extinction does not increase with taxon age. Instead, the dynamics of the nearly-neutral model are “Red Queen” in the sense that while all species increase in expected fitness, their relative fitnesses do not change. Interestingly, the decrease in absolute extinction risk towards the Modern [258, 82], which translates to an increase in expected species duration, may reflect similar dynamics to the nearly-neutral model. One of the innovations behind this result may be allowing species survival to decay non-linearly with duration by modeling durations as observations from a Weibull distribution as opposed to an exponential distribution (Chapter 2, and 3), a modeling decision that is becoming more commonplace [55, 68].

These two studies also provide broad support for the hypothesis of the survival of the unspecialized [298], at least with respect to the covariates included in either model. The survival of the unspecialized appears to be a nearly universal pattern in macroevolution [298, 169, 170, 217, 218, 21, 250]. I did find, however, that for post-Cambrian Paleozoic brachiopods when extinction intensity increases, the relationship with environmental preference and duration changes from a pattern where intermediate environmental preference are are expected to have the greatest durations to a pattern where more specialized, epicontinental-favoring taxa are expected to have the greatest durations. This result adds a degree of nuance to the survival of the unspecialized, specifically with regards to when it is expected to “hold.” My conclusion is that during periods of low intensity extinction risk or “background extinction” [134, 87], the survival of the unspecialized hypothesis will hold. However, as extinction intensity increases, this hypothesis may not accurately describe differences in extinction risk across taxa. As such, the survival of the unspecialized may serve as an excellent default for studies of trait selection and taxon extinction. This conclusion can be considered evidence for the qualitative difference between background and mass extinction [134].

In the third study I analyzed how the functional composition of a regional species pool changes over millions of years (Chapter 4). While this pattern is macroecological, many of the hypotheses and questions encompassed by this study were generated from macroevolutionary analyses. For example, the result from the first study that arboreal taxa have a greater extinction risk than other mammal locomotor groups did not have an unambiguous explanation for if and how extinction risk could have changed over the Cenozoic: always present but high risk, or higher extinction risk in the Neogene compared to the Paleogene (Chapter 2). The third study, in its analysis of the relative diversity of mammal ecotypes, was able to more fully resolve the previous results as I found that diversity of arboreal ecotypes declined through the Paleogene, becoming extremely rare or entirely absent from the species pool by the Neogene (Chapter 4). This result is much more nuanced than the either of two proposed processes (Chapter 2) and speaks to the improved insights and inferences a unified macroevolutionary and macroecological research program can provide.

The third study also, methodologically, represents the strongest unification of macroevolutionary and macroecological approaches to inference (Chapter 4). The question at the heart of this study is “when are mammal ecotypes enriched or depleted relative to their average diversity” is both macroevolutionary and macroecological, any statistical inferences needed to represent this. The model at the center of this study was a combination of a birth-death process with a fourth-corner model from ecology. The birth-death process provided a mechanism for changes to species diversity over time while the fourth-corner model recast diversification in terms of the relative contribution of functional groups to the regional species pool. The results of this study could then be phrased in how well adapted the functional groups are to their changing environmental context.

5.3 Future

There are many possible ways to expand on the analyses presented in this dissertation. There are also many unanswered questions which have been raised by each of these analyses which require future study.

A major limitation the fossil record of North American mammals is poor spatial resolution for the entire Cenozoic, something that restricts macroecological analyses (Chapter 4). While a regional species pool represents the set of all species present in a region, the individual dynamics communities give a much more complete idea of why species pool diversity changes. For example, I've been unable to estimate how changes to functional diversity of local communities scales up to the functional diversity of the regional species pool these communities are drawn from [117]. Given a fossil record with a high resolution spatial record, the types of analyses presented in Chapter 4 could be expanded to incorporate the spatial dynamics of functional diversity. The Bayesian modeling framework used throughout this dissertation makes this imminently possible given the right dataset [16]. The results of such an analysis would shed a lot of light on how the functional diversity of communities can vary spatially and how those communities respond differently to environmental change. A restricted analysis, both temporally and spatially, that focuses on a high resolution fossil record would be an appropriate place to start for this analysis; for example, the late Neogene of North America is extensively sampled as evidenced by the MioMap, Neotoma, and FaunMap projects (http://www.ucmp.berkeley.edu/neomap/use.html) and may be an ideal starting point for this type of analysis.

Preservation of the fossil record is a pervasive issue in paleontology [89, 78, 76, 81, 177, 352]. The first two studies gloss over issues of preservation by either ignoring it (Chapter 2) or including a sampling proxy as a covariate (Chapter 3); these decisions were partially due to limitations of the models underlying both those studies. In contrast, in Chapter 4 I directly modeled the preservation process and allowed preservation probability to vary with time and allowed species body size to potentially affect differences in preservation probability across taxa. However, I ignored the possible differences in preservation over time due to species functional group or changes to environmental context. I chose to limit the possible covariates which could affect of preservation because this type of analysis was beyond the scope of Chapter 4.

All of three of these studies could be improved by incorporating more information about fossil preservation such as how functional groups and environmental context can shape differences in preservation probability. For example, when studying marine invertebrates covariates such as sea-level or shell composition (e.g. aragonitic vs calcitic) are all potentially very important for better understanding the differential preservation of species [86, 227, 228, 112]. Given a more representative model of preservation, even more complete and accurate macroevolutionary and macroecological inferences can be made.

5.4 Final thoughts

The three studies presented in this dissertation are all representative of my rhetoric championing integrated macroevolutionary and macroecological study (Chapter 1). Each study emphasizes the importance of framing hypotheses in terms of ecological interactions in order to make strong inferences. By using an extremely flexible and expressive modeling strategy, I have been able to translate precise scientific questions into statistical models. The scope of inferences that can be made for each study are clear and conditioned on the explicit modeling assumptions made for each analysis. The first step to building a dialogue is agreeing on a common language and my strategy has been to emphasize a common statistical framework with which to phrase our analyses. My hope is that this approach to using paleontological data for answering questions about the processes underlying emergent patterns in evolutionary and ecological data fosters a stronger relationship between the disciplines of macroevolution and macroecology as well as paleontologists and neontologists.

REFERENCES

[1] Jordi Agustí, Lluís Cabrera, and Miguel Garcés. The Vallesian Mammal Turnover: A Late Miocene record of decoupled land-ocean evolution. *Geobios*, 46(1-2):151–157, 2013.

[2] L B Albright. *Biostratigraphy and Vertebrate Paleontology of the San Timoteo Badlands, Southern California.* University of California Press, Berkeley, 2000.

[3] Richard R Alexander. Generic longevity of articulate brachiopods in relation to the mode of stabilization on the substrate. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 21:209–226, 1977.

[4] Linda J S Allen. *An introduction to stochastic processes with applications to biology.* Chapman and Hall/CRC, Boca Raton, FL, 2 edition, 2011.

[5] John Alroy. Constant extinction, constrained diversification, and uncoordinated stasis in North American mammals. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 127:285–311, 1996.

[6] John Alroy. Speciation and extinction in the fossil record of North American mammals. In Roger K Butlin, Jon R Bridle, and Dolph Schluter, editors, *Speciation and patterns of diversity*, pages 302–323. Cambridge University Press, Cambridge, 2009.

[7] John Alroy. Fair sampling of taxonomic richness and unbiased estimation of origination and extinction rates. In John Alroy and Gene Hunt, editors, *Quantitative Methods in Paleobiology*, pages 55–80. The Paleontological Society, 2010.

[8] John Alroy. The Shifting Balance of Diversity Among Major Marine Animal Groups. *Science*, 329(5996):1191–1194, 2010.

[9] John Alroy. A simple Bayesian method of inferring extinction. *Paleobiology*, 40(4):584–607, 2014.

[10] John Alroy, Paul L Koch, and James C Zachos. Global climate change and North American mammalian evolution. *Paleobiology*, 26(1981):259–288, 2000.

[11] Mike Austin. Species distribution models and ecological theory: a critical assessment and some possible new approaches. *Ecological Modelling*, 200:1–19, 2007.

[12] Catherine Badgley and John A Finarelli. Diversity dynamics of mammals in relation to tectonic and climatic history: comparison of three Neogene records from North America. *Paleobiology*, 39(3):373–399, 2013.

[13] Catherine Badgley, Tara M Smiley, Rebecca Terry, Edward B Davis, Larisa R G Desantis, David L Fox, Samantha S B Hopkins, Tereza Jezkova, Marjorie D Matocq, Nick Matzke, Jenny L Mcguire, Andreas Mulch, Brett R Riddle, V Louise Roth,

Joshua X Samuels, Caroline A E Strömberg, and Brian J Yanites. Biodiversity and Topographic Complexity: Modern and Geohistorical Perspectives. *Trends in Ecology & Evolution*, pages 1–16, 2017.

[14] Richard K Bambach. Species richness in marine benthic habitats through the Phanerozoic. *Paleobiology*, 3(2):152–167, 1977.

[15] Richard K. Bambach, Andrew M. Bush, and Douglas H. Erwin. Autecology and the filling of ecospace: Key metazoan radiations. *Palaeontology*, 50(1):1–22, 2007.

[16] Sudipto Banerjee, Bradley P Carlin, and Alan E Gelfand. *Hierarchical modeling and analysis for spatial data*. Chapman and Hall/CRC, 2004.

[17] David W Bapst. paleotree: an R package for paleontological and phylogenetic analyses of evolution. *Methods in Ecology and Evolution*, 3:803–807, 2012.

[18] David W Bapst. A stochastic rate-calibrated method for time-scaling phylogenies of fossil taxa. *Methods in Ecology and Evolution*, 4(8):724–733, aug 2013.

[19] Jon A Baskin. Bassariscus and Probassariscus (Mammalia, Carnivora, Procyonidae) from the early Barstovian (middle Miocene). *Journal of Vertebrate Paleontology*, 24:709–720, 2004.

[20] Jon A Baskin. A new species of Cernictis (Mammalia, Carnivora, Mustelidae) from the late Miocene Bidahochi formation of Arizona, USA. *Palaeontologia Electronica*, 14(3):26A, 2011.

[21] Tomasz K Baumiller. Survivorship analysis of Paleozoic Crinoidea: effect of filter morphology on evolutionary rates. *Paleobiology*, 19(3):304–321, 1993.

[22] B L Beatty and L D Martin. The earliest North American record of the Antilocapridae (Artiodactyla, Mammalia). *PaleoBios*, 29(1):29–35, 2009.

[23] J J Becker and J A White. Late Cenozoic geomyids (Mammalia: Rodentia) from the Anza-Borrego Desert, southern California. *Journal of Vertebrate Paleontology*, 1(2):211–218, 1981.

[24] Gabe S Bever. New Record of Bassariscus Ogallalae (Carnivora: Procyonidae) From the Ogallala Group (Miocene) of Ellis County, Kansas, With Comments on Variation Within Bassariscus. *The Southwestern Naturalist*, 48(2):249–256, jun 2003.

[25] Olaf R P Bininda-Emonds, Marcel Cardillo, Kate E Jones, Ross D E Macphee, Robin M D Beck, Richard Grenyer, Samantha A Price, Rutger A Vos, John L Gittleman, and Andy Purvis. The delayed rise of present-day mammals. *Nature*, 446(7135):507–512, 2007.

[26] Jonathan I Bloch, Kenneth D Rose, and Philip D Gingerich. New species of Batodonoides (Lipotyphla, Geolabididae) from the Early Eocene of Wyoming: smallest known mammal? *Journal of Mammalogy*, 79(3):804–827, 1998.

[27] Jonathan I Bloch, Mary T Silcox, Doug M Boyer, and Eric J Sargis. New Paleocene skeletons and the relationship of plesiadapiforms to crown-clade primates. *Proceedings of the National Academy of Sciences*, 104(4):1159–64, 2007.

[28] Jessica L Blois and Elizabeth A Hadly. Mammalian Response to Cenozoic Climatic Change. *Annual Review of Earth and Planetary Sciences*, 37(1):181–208, 2009.

[29] Fred L Bookstein. Random walk and the existence of evolutionary rates. *Paleobiology*, 13(4):446–464, 1987.

[30] Barry W Brook and David M J S Bowman. The uncertain blitzkrieg of Pleistocene megafauna. *Journal of Biogeography*, 31(4):517–523, 2004.

[31] Alexandra M Brown, David I Warton, Nigel R Andrew, Matthew Binns, Gerasimos Cassis, and Heloise Gibb. The fourth-corner solution - using predictive models to understand how species traits interact with the environment. *Methods in Ecology and Evolution*, 5(4):344–352, 2014.

[32] James H Brown and Brian A Maurer. Macroecology: the division of food and space among species on continents. *Science*, 243(4895):1145–1150, 1989.

[33] James J Brown. *Macroecology.* University of Chicago Press, Chicago, 1995.

[34] A M Bush and P M Novack-Gottshall. Modelling the ecological-functional diversification of marine Metazoa on geological time scales. *Biology Letters*, 8(1):151–155, 2012.

[35] Andrew M Bush and Richard K. Bambach. Paleoecologic Megatrends in Marine Metazoa. *Annual Review of Earth and Planetary Sciences*, 39(1):241–269, 2011.

[36] Andrew M Bush, Richard K Bambach, and Gwen M Daley. Changes in theoretical ecospace utilization in marine fossil assemblages between the mid-Paleozoic and late Cenozoic. *Paleobiology*, 33(1):76–97, 2007.

[37] J L Cantalapiedra, J L Prado, and M T Alberdi. Decoupled ecomorphological evolution and diversification in Neogene-Quaternary horses. *Science*, 355:627–630, 2017.

[38] M T Carrano. What, if anything, is a cursor? Categories versus continua for determining locomotor habit in mammals and dinosaurs. *Journal of Zoology*, 247(1):29–42, 1999.

[39] Leslie N Carraway. Fossil History of Notiosorex (Soricomorpha: Soricidae) shrews with descriptions of new fossil species. *Western North American Naturalist*, 70(2):144–163, 2010.

[40] M Cassiliano. A new genus and species of Stenomylinae (Camelidae, Artiodactyla) from the Moonstone Formation (late Barstovian-early Hemphillian) of central Wyoming. *Rocky Mountain Geology*, 43(1):41–110, 2008.

[41] Scott Chamberlain and Eduard Szocs. taxize - taxonomic search and retrieval in r. *F1000Research*, 2013.

[42] Anne Chao, T C Hsieh, Robin L Chazdon, Robert K Colwell, Nicholas J Gotelli, and B D Inouye. Unveiling the species-rank abundance distribution by generalizing the Good-Turing sample coverage theory. *Ecology*, 96(5):1189–1201, 2015.

[43] Stephen G B Chester and K Christopher Beard. New Micromomyid Plesiadapiforms (Mammalia, Euarchonta) from the Late Paleocene of Big Multi Quarry, Washakie Basin, Wyoming. *Annals of Carnegie Museum*, 80(2):159–172, 2012.

[44] William A Clemens. Eoconodon (”Triisodontidae,” Mammalia) from the Early Paleocene (Puercan) of northeastern Montana, USA. *Palaeontologica Electronica*, 14(3):22A, 2011.

[45] William A Clemens and Thomas E Williamson. A new species of Eoconodon (Triisodontidae, Mammalia) from the San Juan Basin, New Mexico. *Journal of Vertebrate Paleontology*, 25(1):208–213, 2005.

[46] William C Clyde and Philip D Gingerich. Mammalian community response to the latest Paleocene thermal maximum: an isotaphonomic study in the northern Bighorn Basin, Wyoming. *Geology*, 26(11):1011–1014, 1998.

[47] K M Cohen, S C Finney, P L Gibbard, and J-X Fan. The ICS International Chronostratigraphic Chart, 2015.

[48] Margery Chalifoux Coombs. Tylocephanolyx, a new genus of North American dome-skulled Chalicoteres (Mammalia, Perissodactyla). *Bulletin of the American Museum of Natural History*, 164:1–64, 1979.

[49] William S Cooper. Expected time to extinction and the concept of fundamental fitness. *Journal of Theoretical Biology*, 107:603–629, 1984.

[50] E D Cope. On the dentition of Metalophodon. *Proceedings of the American Philosophical Society*, 12(86):542–545, 1871.

[51] Karl Cottenie. Integrating environmental and spatial processes in ecological community dynamics. *Ecology Letters*, 8(11):1175–1182, 2005.

[52] Jerry A Coyne and H Allen Orr. *Speciation.* Sinauer Associates, Sunderland, MA, 2004.

[53] Benjamin S Cramer, K G Miller, P J Barrett, and J D Wright. Late Cretaceous-Neogene trends in deep ocean temperature and continental ice volume: Reconciling records of benthic foraminiferal geochemistry (δ18O and Mg/Ca) with sea level history. *Journal of Geophysical Research: Oceans*, 116(12):1–23, 2011.

[54] James S Crampton, Roger A Cooper, Peter M Sadler, and Michael Foote. Greenhouse–icehouse transition in the Late Ordovician marks a step change in extinction regime in the marine plankton. *Proceedings of the National Academy of Sciences*, 113(6):1498–1503, 2016.

[55] James S Crampton, Roger A Cooper, Peter M Sadler, and Michael Foote. Greenhouse-icehouse transition in the Late Ordovician marks a step change in extinction regime in the marine plankton. *Proceedings of the National Academy of Sciences*, 113(6):1498–1503, 2016.

[56] Gabor Csardi and Tamas Nepusz. The igraph software package for complex network research. *InterJournal*, Complex Systems:1695, 2006.

[57] Walter W Dalquest. Early Blancan mammals of the Beck Ranch local fauna of Texas. *Journal of Mammalogy*, 59(2):269–298, 1978.

[58] John Damuth and Christine M Janis. On the relationship between hypsodonty and feeding ecology in ungulate mammals, and its utility in palaeoecology. *Biological Reviews*, 86:733–758, 2011.

[59] Mary R Dawson. Coryphodon, the northernmost Holarctic Paleogene pantodont (Mammalia), and its global wanderings. *Swiss Journal of Palaeontology*, 131(1):11–22, 2012.

[60] Mary R Dawson and K Christopher Beard. Rodents of the Family Cylindrodontidae (Mammalia) From the Earliest Eocene of the Tuscahoma Formation, Mississippi. *Annals of Carnegie Museum*, 76(3):135–144, 2007.

[61] Stéphane Dray and Pierre Legendre. Testing the species traits-environment relationships: the fourth-corner problem revisited. *Ecology*, 89(12):3400–3412, 2008.

[62] Dianne Edwards and Una Fanning. Evolution and environment in the late Silurian–early Devonian: the rise of pteridophytes. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 309:147–165, 1985.

[63] N Egi. Body mass estimates in extinct mammals from limb bone dimensions: the case of North American hyaenodontids. *Palaeontology*, 44(1990), 2001.

[64] Jane Elith and John R Leathwick. Species distribution models: ecological explanation and prediction across space and time. *Annual Review of Ecology, Evolution, and Systematics*, 40:677–697, 2009.

[65] Jussi T Eronen, Alistair R Evans, Mikael Fortelius, and Jukka Jernvall. The impact of regional climate on the evolution of mammals: a case study using fossil horses. *Evolution*, 64(2):398–408, 2009.

[66] Jussi T Eronen, Christine M Janis, C Page Chamberlain, and Andreas Mulch. 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, 2015.

[67] Jussi T Eronen, P David Polly, Marianne Fred, John Damuth, David C Frank, Volker Mosbrugger, Christoph Scheidegger, Nils Chr Stenseth, and Mikael Fortelius. Ecometrics: The traits that bind the past and present together. *Integrative Zoology*, 5(2):88–101, 2010.

[68] Thomas H G Ezard, Paul N Pearson, Tracy Aze, and Andy Purvis. The meaning of birth and death (in macroevolutionary birth-death models). *Biology Letters*, 8(1):139–42, 2012.

[69] Thomas H G Ezard, Andy Purvis, and Helene Morlon. Environmental changes define ecological limits to species richness and reveal the mode of macroevolutionary competition. *Ecology Letters*, 19(8):899–906, 2016.

[70] J Felsenstein. Phylogenies and the comparative method. *The American Naturalist*, 125(1):1–15, 1985.

[71] I Ferrusquía-Villafranca. The first Paleogene mammal record of middle America: Simojovelhyus pocitosense (Helohyidae, Artiodactyla). *Journal of Vertebrate Paleontology*, 26(4):989–1001, 2006.

[72] Borja Figueirido, Christine M Janis, Juan A Pérez-Claros, Miquel De Renzi, and Paul Palmqvist. Cenozoic climate change influences mammalian evolutionary dynamics. *Proceedings of the National Academy of Sciences*, 109(3):722–727, jan 2012.

[73] S Finnegan, N A Heim, S E Peters, and W W Fischer. Climate change and the selective signature of the Late Ordovician mass extinction. *Proceedings of the National Academy of Sciences*, 109:6829–6834, 2012.

[74] Seth Finnegan, Jonathan L Payne, and Steve C Wang. The Red Queen revisited: reevaluating the age selectivity of Phanerozoic marine genus extinctions. *Paleobiology*, 34(3):318–341, 2008.

[75] R G Fitzjohn. Quantitative Traits and Diversification. *Systematic Biology*, 59(6):619–633, 2010.

[76] M Foote and J J Sepkoski. Absolute measures of the completeness of the fossil record. *Nature*, 398(6726):415–7, 1999.

[77] Michael Foote. Survivorship analysis of Cambrian and Ordovician Trilobites. *Paleobiology*, 14(3):258–271, 1988.

[78] Michael Foote. Estimating taxonomic durations and preservation probability. *Paleobiology*, 23(3):278–300, 1997.

[79] Michael Foote. Origination and extinction components of taxonomic diversity: general problems. *Paleobiology*, 26(sp4):74–102, 2000.

[80] Michael Foote. Origination and extinction components of taxonomic diversity: Paleozoic and post-Paleozoic dynamics. *Paleobiology*, 26(4):578–605, 2000.

[81] Michael Foote. Inferring temporal patterns of preservation, origination, and extinction from taxonomic survivorship analysis. *Paleobiology*, 27(4):602–630, 2001.

[82] Michael Foote. Origination and extinction through the Phanerozoic: a new approach. *Journal of Geology*, 111:125–148, 2003.

[83] Michael Foote. Substrate affinity and diversity dynamics of Paleozoic marine animals. *Paleobiology*, 32(3):345–366, 2006.

[84] Michael Foote. The geologic history of biodiversity. In M A Bell, Douglas J Futuyma, W F Eanes, and Jeffery S Levinton, editors, *Evolution since Darwin: the first 150 years*, pages 479–510. Sinauer Associates, Sunderland, MA, 2010.

[85] Michael Foote. Environmental controls on geographic range size in marine animal genera. *Paleobiology*, 40(3):440–458, 2014.

[86] Michael Foote, James S Crampton, Alan G Beu, and Campbell S Nelson. Aragonite bias, and lack of bias, in the fossil record: lithological, environmental, and ecological controls. *Paleobiology*, 41(2):245–265, 2015.

[87] Michael Foote and Arnold I Miller. *Principles of Paleontology.* Freeman, New York, third edition, 2007.

[88] Michael Foote and Arnold I. Miller. Determinants of early survival in marine animal genera. *Paleobiology*, 39(2):171–192, 2013.

[89] Michael Foote and DM Raup. Fossil preservation and the stratigraphic ranges of taxa. *Paleobiology*, 22(2):121–140, 1996.

[90] Mikael Fortelius, Jussi Eronen, Jukka Jernvall, Liping Liu, Diana Pushkina, Juhani Rinne, Alexey Tesakov, Inesa Vislobokova, Zhaoqun Zhang, and Liping Zhou. Fossil mammals resolve regional patterns of Eurasian climate change over 20 million years. *Evolutionary Ecology Research*, 4:1005–1016, 2002.

[91] John R Foster. Preliminary body mass estimates for mammalian genera of the Morrison Formation (Upper Jurassic, North America). *PaleoBios*, 28:114–122, 2009.

[92] Richard C Fox and Craig S Scott. A New, Early Puercan (Earliest Paleocene) Species of Purgatorius (Plesiadapiformes, Primates) from Saskatchewan, Canada. *Journal of Paleontology*, 85(3):537–548, 2011.

[93] Danielle Fraser, Root Gorelick, and Natalia Rybczynski. Macroevolution and climate change influence phylogenetic community assembly of North American hoofed mammals. *Biological Journal of the Linnean Society*, 114(3):485–494, 2015.

[94] Matthijs Freudenthal and Elvira Martín-Suárez. Estimating body mass of fossil rodents. *Scripta Geologica*, 145:1–130, 2013.

[95] Susanne A Fritz, Olaf R P Bininda-Emonds, and Andy Purvis. Geographical variation in predictors of mammalian extinction risk: big is bad, but only in the tropics. *Ecology Letters*, 12(6):538–49, 2009.

[96] Susanne A Fritz and Andy Purvis. Selectivity in mammalian extinction risk and threat types: a new measure of phylogenetic signal strength in binary traits. *Conservation Biology*, 24(4):1042–51, aug 2010.

[97] Susanne A Fritz, Jan Schnitzler, Jussi T Eronen, Christian Hof, Katrin Böhning-Gaese, and Catherine H Graham. Diversity in time and space: wanted dead and alive. *Trends in Ecology & Evolution*, 28(9):509–16, 2013.

[98] C L Gazin. A Tertiary vertebrate fauna from the upper Cuyama drainage basin, California. *Carnegie Institution of Washington*, 404:55–76, 1930.

[99] Andrew Gelman. Prior distributions for variance parameters in hierarchical models. *Bayesian Analysis*, 1(3):515–533, 2006.

[100] Andrew Gelman. Scaling regression inputs by dividing by two standard deviations. *Statistics in Medicine*, pages 2865–2873, 2008.

[101] Andrew Gelman, John B Carlin, Hal S Stern, David B Dunson, Aki Vehtari, and Donald B Rubin. *Bayesian data analysis*. Chapman and Hall, Boca Raton, FL, 3 edition, 2013.

[102] Andrew Gelman and Jennifer Hill. *Data Analysis using Regression and Multilevel/Hierarchical Models*. Cambridge University Press, New York, NY, 2007.

[103] James Williams Gidley. Pleistocene peccaries from the Cumberland Cave deposit. *Proceedings of the United States National Museum*, 57(2324):651–678, 1920.

[104] PD Gingerich. Quantification and comparison of evolutionary rates. *American Journal of Science*, 293:453–478, 1993.

[105] Emma E Goldberg, Lesley T Lancaster, and Richard H Ree. Phylogenetic inference of reciprocal effects between geographic range evolution and diversification. *Systematic Biology*, 60(4):451–65, 2011.

[106] Emma E Goldberg, Kaustuv Roy, Russell Lande, and David Jablonski. Diversity, endemism, and age distributions in macroevolutionary sources and sinks. *The American Naturalist*, 165(6):623–33, 2005.

[107] Harvey Goldstein, William Browne, and Jon Rasbash. Partitioning variation in multilevel models. *Understanding Statistics*, 1(4):1–12, 2002.

[108] Cynthia L Gordon. A First Look at Estimating Body Size in Dentally Conservative Marsupials. *Journal of Mammalian Evolution*, 10(1/2):1–21, 2003.

[109] Alan Graham. *A natural history of the New World: the ecology and evolution of plants in the Americas*. University of Chicago Press, Chicago, 2011.

[110] Camille Grohé, Yaowalak Chaimanee, Louis de Bonis, Chotima Yamee, Cécile Blondel, and Jean-Jacques Jaeger. New data on Mustelidae (Carnivora) from Southeast Asia: Siamogale thailandica, a peculiar otter-like mustelid from the late middle Miocene Mae Moh Basin, northern Thailand. *Die Naturwissenschaften*, 97(11):1003–15, 2010.

[111] E R Hall. Three new genera of Mustelidae from the later Tertiary of North America. *Journal of Mammalogy*, 11(2):146–155, 1930.

[112] Bjarte Hannisdal and Shanan E Peters. Phanerozoic Earth system evolution and marine biodiversity. *Science*, 334:1121–1124, 2011.

[113] Luke J Harmon and Susan Harrison. Species Diversity Is Dynamic and Unbounded at Local and Continental Scales. *The American Naturalist*, 185(5):584–593, mar 2015.

[114] P G Harnik, C Simpson, and J L Payne. Long-term differences in extinction risk among the seven forms of rarity. *Proceedings of the Royal Society B: Biological Sciences*, 279(1749):4969–4976, 2012.

[115] Paul G Harnik. Direct and indirect effects of biological factors on extinction risk in fossil bivalves. *Proceedings of the National Academy of Sciences*, 108(33):13594–13599, aug 2011.

[116] Paul G Harnik, Paul C Fitzgerald, Jonathan L Payne, and Sandra J Carlson. Phylogenetic signal in extinction selectivity in Devonian terebratulide brachiopods. *Paleobiology*, 40(4):675–692, 2014.

[117] Susan Harrison and Howard Cornell. Toward a better understanding of the regional causes of local community richness. *Ecology Letters*, 11:969–979, 2008.

[118] O P Hay. Descriptions of some fossil vertebrates found in Texas. *Bulletin of the University of Texas*, 71:2–24, 1916.

[119] OP Hay. Notes on Some Fossil Horses, with Descriptions of Four New Species. *Proceedings of The United States National Museum*, 44:569–594, 1969.

[120] Matthew M Hedman. Constraints on clade ages from fossil outgroups. *Paleobiology*, 36(1):16–31, 2010.

[121] Noel A Heim and Shanan E Peters. Regional environmental breadth predicts geographic range and longevity in fossil marine genera. *PloS one*, 6(5):e18946, jan 2011.

[122] Robert J Hijmans. *raster: Geographic data analysis and modeling*, 2015. R package version 2.3-24.

[123] Kenneth B Hoehn, Paul G Harnik, and V Louise Roth. A framework for detecting natural selection on traits above the species level. *Methods in Ecology and Evolution*, pages doi: 10.111/2041–210X.12461, 2015.

[124] Matthew D Hoffman and Andrew Gelman. The no-U-turn sampler: Adaptively setting path lengths in Hamiltonian Monte Carlo. *arXiv*, 1111(4246), 2011.

[125] Matthew D Hoffman and Andrew Gelman. The No-U-Turn Sampler: Adaptively Setting Path Lengths in Hamiltonian Monte Carlo. *Journal of Machine Learning Research*, 15:1351–1381, 2014.

[126] Robert D Holt. Emergent neutrality. *Trends in Ecology and Evolution*, 21(10):531–533, 2006.

[127] Melanie J Hopkins, Carl Simpson, and Wolfgang Kiessling. Differential niche dynamics among major marine invertebrate clades. *Ecology Letters*, 17(3):314–323, 2014.

[128] Elizabeth A Housworth, P Martins, and Michael Lynch. The Phylogenetic Mixed Model. *The American Naturalist*, 163(1):84–96, 2004.

[129] G Hunt. The relative importance of directional change, random walks, and stasis in the evolution of fossil lineages. *Proceedings of the National Academy of Sciences*, 104:18404–18408, 2007.

[130] Gene Hunt. Fitting and comparing models of phyletic evolution: random walks and beyond. *Paleobiology*, 32(4):578–601, 2006.

[131] Gene Hunt and Daniel L Rabosky. Phenotypic Evolution in Fossil Species: Pattern and Process. *Annual Review of Earth and Planetary Sciences*, 42(1):421–441, 2014.

[132] Gene Hunt, Kaustuv Roy, and David Jablonski. Species-level heritability reaffirmed: a comment on "on the heritability of geographic range sizes". *The American naturalist*, 166(1):129–135, jul 2005.

[133] Joseph G Ibrahim, Ming-Hui Chen, and Debajyoti Sinha. *Bayesian Survival Analysis.* Springer, New York, 2001.

[134] David Jablonski. Background and mass extincitons: the alternation of macroevolutionary regimes. *Science*, 231(4734):129–133, 1986.

[135] David Jablonski. Heritability at the species level: analysis of geographic ranges of cretaceous mollusks. *Science*, 238(4825):360–363, 1987.

[136] David Jablonski. Scale and hierarchy in macroevolution. *Palaeontology*, 50:87–109, 2007.

[137] David Jablonski. Species Selection: Theory and Data. *Annual Review of Ecology, Evolution, and Systematics*, 39(1):501–524, 2008.

[138] David Jablonski and Kaustuv Roy. Geographical range and speciation in fossil and living molluscs. *Proceedings of the Royal Society B: Biological Sciences*, 270(1513):401–406, 2003.

[139] Tahira Jamil, Wim A Ozinga, Michael Kleyer, and Cajo J F Ter Braak. Selecting traits that explain species-environment relationships: A generalized linear mixed model approach. *Journal of Vegetation Science*, 24(6):988–1000, 2013.

[140] C Janis, J Damuth, and J M Theodor. The species richness of Miocene browsers, and implications for habitat type and primary productivity in the North American grassland biome. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 207(3-4):371–398, 2004.

[141] C M Janis, J Damuth, and J M Theodor. Miocene ungulates and terrestrial primary productivity: where have all the browsers gone? *Proceedings of the National Academy of Sciences*, 97(14):7899–904, 2000.

[142] Christine M Janis. Tertiary mammal evolution in the context of changing climates, vegetation, and tectonic events. *Annual Review of Ecology and Systematics*, 24:467–500, 1993.

[143] Christine M Janis. Tertiary mammal evolution in the context of changing climates, vegetation, and tectonic events. *Annual Review of Ecology and Systematics*, 24:467–500, 1993.

[144] Christine M Janis. An evolutionary history of browsing and grazing ungulates. In Iain J Gordon and Herbert H T Prins, editors, *The Ecology of Browsing and Grazing*, pages 21–45. Springer-Verlag, 2008.

[145] Christine M Janis, Gregg F Gunnell, and Mark D Uhen. *Evolution of Tertiary mammals of North America. Vol. 2. Small mammals, xenarthrans, and marine mammals.* Cambridge University Press, Cambridge, 2008.

[146] Christine M Janis, K M Scott, and L L Jacobs. *Evolution of Tertiary mammals of North America. Vol. 1. Terrestrial carnivores, ungulates, and ungulatelike mammals.* Cambridge University Press, Cambridge, 1998.

[147] Christine M Janis and Patricia Brady Wilhelm. Were there mammalian pursuit predators in the tertiary? Dances with wolf avatars. *Journal of Mammalian Evolution*, 1(2):103–125, 1993.

[148] Phillip E Jardine, Christine M Janis, Sarda Sahney, and Michael J Benton. Grit not grass: concordant patterns of early origin of hyposdonty in Great Plains ungulates and Glires. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 365-366:1–10, 2012.

[149] Edwin T Jaynes. *Probability theory: the logic of science.* Cambridge Univ Press, Cambridge, 2003.

[150] G L Jepsen. Rubulodon taylori, a Wind River Eocene Tubulidentate from Wyoming. *Proceedings of the American Ph*, 71(5):255–274, 1932.

[151] Jukka Jernvall and Mikael Fortelius. Common mammals drive the evolutionary increase of hypsodonty in the Neogene. *Nature*, 417(6888):538–40, 2002.

[152] Jukka Jernvall and Mikael Fortelius. Maintenance of trophic structure in fossil mammal communities: site occupancy and taxon resilience. *The American Naturalist*, 164(5):614–624, 2004.

[153] M M Joachimski, S Breisig, W Buggisch, J A Talent, R Mawson, M Gereke, J R Morrow, J Day, and K Weddige. Devonian climate and reef evolution: Insights from oxygen isotopes in apatite. *Earth and Planetary Science Letters*, 284(3-4):599–609, 2009.

[154] Zerina Johanson. New marsupial from the Fort Union Formation, Swain Quarry, Wyoming. *Journal of Paleontology*, 70(6):1023–1031, 1996.

[155] J G Johnson. Extinction of Perched Faunas. *Geology*, 2:479–482, 1974.

[156] Kate E Jones, Jon Bielby, Marcel Cardillo, Susanne A Fritz, Justin O'Dell, C David L Orme, Kamran Safi, Wes Sechrest, E Boakes, C Carbone, C Connolly, M J Cutts, J K Foster, R Grenyer, M Habib, C A Plaster, S A Price, E A Rigby, J Rist, Amber Teacher, Olaf R P Bininda-Emonds, John L Gittleman, Georgina M Mace, and Andy Purvis. PanTHERIA : a species-level database of life history , ecology , and geography of extant and recently extinct mammals. *Ecology*, 90(9):2648, 2009.

[157] D R Kelley and A E Wood. The Eocene mammals from the Lysite member, Wind River Formation of Wyoming. *Journal of Paleontology*, 28(3):337–366, 1954.

[158] Wolfgang Kiessling and Martin Aberhan. Environmental determinants of marine benthic biodiversity dynamics through Triassic Jurassic time. *Paleobiology*, 33(3):414–434, 2007.

[159] E Christopher Kirk and Blythe A Williams. New adapiform primate of Old World affinities from the Devil's Graveyard Formation of Texas. *Journal of human Evolution*, 61(2):156–68, 2011.

[160] John P Klein and Melvin L Moeschberger. *Survival Analysis: Techniques for Censored and Truncated Data.* Springer, New York, 2nd edition, 2003.

[161] D G Kleinbaum and M Klein. *Survival analysis: a self-learning text*. Springer, New York, NY, 2 edition, 2005.

[162] W W Korth. Miosicista angulus, a new sicistine rodent (Zapodidae, Rodentia) from the Barstovian (Miocene) of Nebraska. *Transactions of the Nebraska Academy of Sciences*, 20:97–101, 1993.

[163] Alp Kucukelbir, Rajesh Ranganath, Andrew Gelman, and David M Blei. Automatic Variational Inference in Stan. In *NIPS*, volume 28, pages 568–576, 2015.

[164] Pierre Legendre, René Galzin, and Mireille L Harmelin-Vivien. Relating behavior to habitat: solutions to the fourth-corner problem. *Ecology*, 78(2):547–562, 1997.

[165] Serge Legendre. Analysis of mammalian communities from the Late Eocene and Oligocene of Southern France. *Paleovertebrata*, 16(4):191–212, 1986.

[166] Daniel Lewandowski, Dorota Kurowicka, and Harry Joe. Generating random correlation matrices based on vines and extended onion method. *Journal of Multivariate Analysis*, 100(9):1989–2001, 2009.

[167] Jason A Lillegraven. Small rodents (Mammalia) from Eocene deposits of San Diego County, California. *Bulletin of the American Museum of Natural History*, 158:223–260, 1977.

[168] J D Lim, L D Martin, and R W Wilson. A new species of Leptarctus (Carnivora, Mustelidae) from the late Miocene of Texas. *Journal of Paleontology*, 75(5):1043–1046, 2001.

[169] Lee Hsiang Liow. A test of Simpson's "rule of the survival of the relatively unspecialized" using fossil crinoids. *The American Naturalist*, 164(4):431–43, oct 2004.

[170] Lee Hsiang Liow. Does versatility as measured by geographic range, bathymetric range and morphological variability contribute to taxon longevity? *Global Ecology and Biogeography*, 16(1):117–128, 2007.

[171] Lee Hsiang Liow, Mikael Fortelius, E Bingham, K Lintulaakso, H Mannila, L Flynn, and N C Stenseth. Higher origination and extinction rates in larger mammals. *Proceedings of the National Academy of Sciences*, 105(16):6097–6102, 2008.

[172] Lee Hsiang Liow, Mikael Fortelius, Kari Lintulaakso, Heikki Mannila, and Nils Chr Stenseth. Lower Extinction Risk in SleeporHide Mammals. *The American Naturalist*, 173(2):264–272, 2009.

[173] Lee Hsiang Liow and James D Nichols. Estimating rates and probabilities of origination and extinction using taxonomic occurrence data: Capture-mark-recapture (CMR) approaches. In John Alroy and Gene Hunt, editors, *Quantitative Methods in Paleobiology*, pages 81–94. The Paleontological Society, 2010.

[174] Lee Hsiang Liow, Tiago B Quental, and Charles R Marshall. When can decreasing diversification rates be detected with molecular phylogenies and the fossil record? *Systematic biology*, 59(6):646–59, 2010.

[175] Lee Hsiang Liow, Leigh Van Valen, and Nils Chr Stenseth. Red Queen: from populations to taxa and communities. *Trends in Ecology & Evolution*, 26(7):349–358, 2011.

[176] Elisabeth A Lloyd and Stephen J Gould. Species selection on variability. *Proceedings of the National Academy of Sciences*, 90:595–599, 1993.

[177] G T Lloyd, J R Young, and A B Smith. Taxonomic Structure of the Fossil Record is Shaped by Sampling Bias. *Systematic Biology*, 61(1):80–89, 2011.

[178] Nicolas Loeuille and Mathew a Leibold. Evolution in metacommunities: on the relative importance of species sorting and monopolization in structuring communities. *The American naturalist*, 171(6):788–99, 2008.

[179] F B Loomis. The camels of the Harrison beds, with three new species. *American Journal of Science*, 31:65–70, 1911.

[180] F B Loomis. Two new Miocene entelodonts. *Journal of Mammalogy*, 13(4):358–362, 1932.

[181] Jonathan B Losos. Adaptive radiation, ecological opportunity, and evolutionary determinism. *The American Naturalist*, 175(6):623–39, 2010.

[182] Jonathan B Losos and D Luke Mahler. Adaptive radiation: the interaction of ecological opportunity, adaptation, and speciation. In M. A. Bell, D. J. Futuyma, W. F. Eanes, and J. S. Levinton, editors, *Evolution since Darwin: the first 150 years*, chapter 15, pages 381–420. Sinauer Associates, Sunderland, MA, 2010.

[183] Zhe-Xi Luo, Alfred W Crompton, and Ai-Lin Sun. A New Mammaliaform from the Early Jurassic and Evolution of Mammalian Characteristics. *Science*, 292:1535–1540, 2001.

[184] Michael Lynch. Methods for the analysis of comparative data in evolutionary biology. *Evolution*, 45(5):1065–1080, 1991.

[185] Giles Ternan Mac Intyre. The Miacidae (Mammalia, Carnivora) Part 1. The systematics of Ictidopappus and Protictis. *Bulletin of the American Museum of Natural History*, 131:115–210, 1966.

[186] J R Macdonald. Additions to the Whitneyan fauna of South Dakota. *Journal of Paleontology*, 25(3):257–265, 1951.

[187] J R Macdonald. The North American Antracotheres. *Journal of Paleontology*, 30(3):615–645, 1956.

[188] B J MacFadden. Fossil horses from "Eohippus" (Hyracotherium) to Equus: scaling, Cope's Law, and the evolution of body size. *Paleobiology*, 12(4):355–369, 1986.

[189] Cary T Madden and John E Storer. The Proboscidea from the Middle Miocene Wood Mountain Formation, Saskatchewan. *Canadian Journal of Earth Sciences*, 22(9):1345–1350, 1985.

[190] W P Maddison, P E Midford, and S P Otto. Estimating a binary character's effect on speciation and extinction. *Systematic Biology*, 56(5):701, 2007.

[191] Jonathan D Marcot. The fossil record and macroevolutionary history of North American ungulate mammals: standardizing variation in intensity and geography of sampling. *Paleobiology*, 40(2):237–254, 2014.

[192] R A Martin, H T Goodwin, and J O Farlow. Late Neogene (Late Hemphillian) rodents from the Pipe Creek Sinkhole, Grant County, Indiana. *Journal of Vertebrate Paleontology*, 22:137–151, 2002.

[193] W D Matthew. Additional observations on the Creodonta. *Bulletin of the American Museum of Natural History*, 14:1–38, 1901.

[194] H G McDonald. Paleoecology of extinct xenarthrans and the Great American Biotic Interchange. *Bulletin of the Florida Museum of Natural History*, 45:313–333, 2005.

[195] Richard McElreath. *Statistical rethinking: a Bayesian course with examples in R and Stan*. CRC Press, Boca Raton, FL, 2016.

[196] Brian J McGill, Brian J Enquist, Evan Weiher, and Mark Westoby. Rebuilding community ecology from functional traits. *Trends in Ecology & Evolution*, 21(4):178–185, 2006.

[197] Paul O McGrew. A new Amphicyon from the Deep River Miocene. *Geological Series of Field Museum of Natural History*, 6(23):341–350, 1939.

[198] Ryan Thomas McKenna. *Potential for Speciation in Mammals Following Vast , Late Miocene Volcanic Interruptions in the Pacific Northwest*. Masters, Portland State University, 2011.

[199] Daniel W McShea. Mechanisms of Large-Scale Evolutionary Trends. *Evolution*, 48(6):1747–1763, 1994.

[200] James S Mellett. A skull of Hemipsalodon (Mammalia, Deltatheridia) from the Clarno Formation of Oregon. *American Museum Novitates*, 2387:1–19, 1969.

[201] M Mendoza, C M Janis, and P Palmqvist. Estimating the body mass of extinct ungulates: a study on the use of multiple regression. *Journal of Zoology*, 270:90–101, 2006.

[202] Matthew C Mihlbachler and Nikos Solounias. Coevolution of Tooth Crown Height and Diet in Oreodonts (Merycoidodontidae, Artiodactyla) Examined with Phylogenetically Independent Contrasts. *Journal of Mammalian Evolution*, 13(1):11–36, 2006.

[203] Arnold I Miller and Michael Foote. Epicontinental seas versus open-ocean settings: the kinetics of mass extinction and origination. *Science*, 326(5956):1106–9, 2009.

[204] J S Mitchell and P J Makovicky. Low ecological disparity in Early Cretaceous birds. *Proceedings of the Royal Society B: Biological Sciences*, 281(1787):20140608–20140608, 2014.

[205] Jonathan S Mitchell. Extant-only comparative methods fail to recover the disparity preserved in the bird fossil record. *Evolution*, 69(9):2414–2424, 2015.

[206] Gary G. Mittelbach and Douglas W. Schemske. Ecological and evolutionary perspectives on community assembly. *Trends in Ecology and Evolution*, 30(5):241–247, 2015.

[207] Cesar Alberto Laurito Mora and Ana Lucia Valerio Zamora. First record of Rhynochotherium blicki (Frick, 1933) for the late Cenozoic of Costa Rica. *Revista Geologica de America Central*, 33:75–82, 2005.

[208] Michael Morlo. Niche structure and evolution in Crodont (Mammalia) faunas of the European and North American Eocene. *Geobios*, 32(2):297–305, 1999.

[209] Isaac Casanovas-Vilar Salvador Moyà-Solà, Jordi Agustí, and Meike Kohler. 9 The geography of a faunal turnover : tracking the Vallesian Crisis. In Ashraf M T Elewa, editor, *Migration of Organisms: Climate, geography, ecology*, pages 247–300. Springer, Berlin, 2005.

[210] Axel Munnecke, Mikael Calner, David A T Harper, and Thomas Servais. Ordovician and Silurian sea-water chemistry, sea level, and climate: A synopsis. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 296(3-4):389–413, 2010.

[211] S Nee. Inferring speciation rates from phylogenies. *Evolution*, 55(4):661–8, 2001.

[212] S Nee, A O Mooers, and P H Harvey. Tempo and mode of evolution revealed from molecular phylogenies. *Proceedings of the National Academy of Sciences*, 89(17):8322–6, 1992.

[213] Sean Nee. Birth-Death Models in Macroevolution. *Annual Review of Ecology, Evolution, and Systematics*, 37(1):1–17, dec 2006.

[214] Sean Nee, R M May, and P H Harvey. The reconstructed evolutionary process. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 344:305–311, 1994.

[215] Michael J Novacek. A review of Paleocene and Eocene Leptictidae (Eutheria: Mammalia) from North America. *PaleoBios*, 24:1–42, 1977.

[216] Philip M Novack-Gottshall. Using a theoretical ecospace to quantify the ecological diversity of Paleozoic and modern marine biotas Using a theoretical ecospace to quantify the ecological diversity of Paleozoic and modern marine biotas. *Paleobiology*, 33(2):273–294, 2007.

[217] Sabine Nürnberg and Martin Aberhan. Habitat breadth and geographic range predict diversity dynamics in marine Mesozoic bivalves. *Paleobiology*, 39(3):360–372, 2013.

[218] Sabine Nürnberg and Martin Aberhan. Interdependence of specialization and biodiversity in Phanerozoic marine invertebrates. *Nature Communications*, 6:6602, 2015.

[219] Henry Fairfield Osborn. Sebelodon burnhami, a new shovel-tusker from California. *American Museum Novitates*, 639:1–5, 1933.

[220] Michael E Palmer and Marcus W Feldman. Survivability is more fundamental than evolvability. *PloS one*, 7(6):e38025, 2012.

[221] Thomas Hudson Patton and Beryl E Taylor. The Protoceratinae (Mammalia, Tylopoda, Protoceratidae) and the systematics of the Protoceratidae. *Bulletin of the American Museum of Natural History*, 150:347–414, 1973.

[222] Jonathan L Payne and Seth Finnegan. The effect of geographic range on extinction risk during background and mass extinction. *Proceedings of the National Academy of Sciences*, 104:10506–11, jun 2007.

[223] Jonathan L Payne, Noel A Heim, Matthew L Knope, and Craig R Mcclain. Metabolic dominance of bivalves predates brachiopod diversity decline by more than 150 million years. *Proceedings of the Royal Society B: Biological Sciences*, 281:20133122, 2014.

[224] Matthew W Pennell, Luke J Harmon, and Josef C Uyeda. Is there room for punctuated equilibrium in macroevolution? *Trends in Ecology & Evolution*, 29(1):23–32, 2014.

[225] Shanan E Peters. The problem with the Paleozoic. *Paleobiology*, 33(2):165–181, 2007.

[226] Shanan E Peters. Environmental determinants of extinction selectivity in the fossil record. *Nature*, 454(7204):626–629, jul 2008.

[227] Shanan E Peters and Michael Foote. Determinants of extinction in the fossil record. *Nature*, 416(6879):420–4, mar 2002.

[228] Shanan E Peters and Noel A Heim. The geological completeness of paleontological sampling in North America. *Paleobiology*, 36(1):61–79, 2010.

[229] Steven J Phillips, Robert P Anderson, and Robert E Schapire. Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, 190:231–259, 2006.

[230] Mathias M Pires, Daniele Silvestro, and Tiago B Quental. Continental faunal exchange and the asymmetrical radiation of carnivores. *Proceedings of the Royal Society B: Biological Sciences*, 282:20151952, 2015.

[231] Laura J Pollock, Michael J Bayly, and Peter A Vesk. The Roles of Ecological and Evolutionary Processes in Plant Community Assembly: The Environment, Hybridization, and Introgression Influence Co-occurrence of Eucalyptus. *The American Naturalist*, 185(6):784–796, mar 2015.

[232] Laura J. Pollock, William K. Morris, and Peter A. Vesk. The role of functional traits in species distributions revealed through a hierarchical model. *Ecography*, 35(8):716–725, 2012.

[233] P D Polly, J T Eronen, Marianne Fred, Gregory P Dietl, Volker Mosbrugger, Christoph Scheidegger, David C Frank, John Damuth, Nils C Stenseth, and Mikael Fortelius. History matters: ecometrics and integrative climate change biology. *Proceedings of the Royal Society B: Biological Sciences*, 278(1709):1131–1140, 2011.

[234] P David Polly, A Michelle Lawing, Jussi T Eronen, and Jan Schnitzler. Processes of ecometric patterning: modelling functional traits, environments, and clade dynamics in deep time. *Biological Journal of the Linnean Society*, pages n/a–n/a, 2015.

[235] Paul David Polly, P David Polly, Jesualdo Fuentes-Gonzalez, A Michelle Lawing, Allison K Bormet, and Robert G Dundas. Clade sorting has a greater effect than local adaptation on ecometric patterns in Carnivora. *Evolutionary Ecology Research*, 18:61–95, 2017.

[236] S A Price and L Schmitz. A promising future for integrative biodiversity research: an increased role of scale-dependency and functional biology. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 371:20150228, 2016.

[237] Samantha A Price, Samantha S B Hopkins, Kathleen K Smith, and V Louise Roth. Tempo of trophic evolution and its impact on mammalian diversification. *Proceedings of the National Academy of Sciences*, 109(18):7008–12, 2012.

[238] A Purvis, J L Gittleman, G Cowlishaw, and G M Mace. Predicting extinction risk in declining species. *Proceedings of the Royal Society B: Biological Sciences*, 267(1456):1947–52, 2000.

[239] Tiago B Quental and Charles R Marshall. Extinction during evolutionary radiations: reconciling the fossil record with molecular phylogenies. *Evolution*, 63(12):3158–3167, 2009.

[240] Tiago B Quental and Charles R Marshall. How the Red Queen Drives Terrestrial Mammals to Extinction. *Science*, 341(6143):290–292, 2013.

[241] Daniel L Rabosky. Extinction rates should not be estimated from molecular phylogenies. *Evolution*, 64(6):1816–1824, jun 2010.

[242] Daniel L Rabosky. Diversity-Dependence, Ecological Speciation, and the Role of Competition in Macroevolution. *Annual Review of Ecology, Evolution, and Systematics*, 44:1–22, 2013.

[243] Daniel L Rabosky. Reproductive isolation and the causes of speciation rate variation in nature. *Biological Journal of the Linnean Society*, 2015.

[244] Daniel L Rabosky and Allen H Hurlbert. Species Richness at Continental Scales Is Dominated by Ecological Limits. *The American Naturalist*, 185(5):000–000, 2015.

[245] Daniel L Rabosky and Daniel R Matute. Macroevolutionary speciation rates are decoupled from the evolution of intrinsic reproductive isolation in Drosophila and birds. *Proceedings of the National Academy of Sciences*, 110(38):15354–15359, 2013.

[246] Daniel L Rabosky and Amy R McCune. Reinventing species selection with molecular phylogenies. *Trends in Ecology & Evolution*, 25(2):68–74, 2010.

[247] Daniel L Rabosky, Francesco Santini, Jonathan Eastman, Stephen A Smith, Brian Sidlauskas, Jonathan Chang, and Michael E Alfaro. Rates of speciation and morphological evolution are correlated across the largest vertebrate radiation. *Nature Communications*, 4:1–8, 2013.

[248] P Raia, F Carotenuto, F Passaro, D Fulgione, and M Fortelius. Ecological specialization in fossil mammals explains Cope's rule. *The American Naturalist*, 179(3):328–37, 2012.

[249] P Raia, F Carotenuto, F Passaro, P Piras, D Fulgione, L Werdelin, J Saarinen, and Mikael Fortelius. Rapid action in the Palaeogene, the relationship between phenotypic and taxonomic diversification in Coenozoic mammals. *Proceedings of the Royal Society B: Biological Sciences*, 280:20122244, 2013.

[250] Pasquale Raia, Francesco Carotenuto, A Mondanaro, S Castiglione, F Passaro, F Saggese, M Melchionna, C Seiro, L Alessio, D Silvestro, and M Fortelius. Progress to extinction: increased specialisation causes the demise of animal clades. *Scientific Reports*, 6:30965, 2016.

[251] David M Raup. Taxonomic survivorship curves and Van Valen's Law. *Paleobiology*, 1(1):82–96, 1975.

[252] David M Raup. Cohort Analysis of generic survivorship. *Paleobiology*, 4(1):1–15, 1978.

[253] David M Raup. Mathematical models of cladogenesis. *Paleobiology*, 11(1):42–52, 1985.

[254] David M Raup. *Extinction: Bad Genes or Bad Luck?* Norton, New York, 1991.

[255] David M Raup. The role of extinction in evolution. *Proceedings of the National Academy of Sciences*, 91(July):6758–6763, 1994.

[256] David M Raup and Stephen Jay Gould. Stochastic simulation and evolution of morphology – towards a nomothetic paleontology. *Systematic Zoology*, 23(3):305–322, 1974.

[257] David M Raup, Stephen Jay Gould, Thomas J M Schopf, and Daniel S Simberloff. Stochastic models of phylogeny and the evolution of diversity. *The Journal of Geology*, 81(5):525–542, 1973.

[258] David M Raup and J John Sepkoski. Mass Extinctions in the Marine Fossil Record. *Science*, 215(4539):1501–1503, 1982.

[259] Liam J Revell. phytools: An r package for phylogenetic comparative biology (and other things). *Methods in Ecology and Evolution*, 3:217–223, 2012.

[260] P Robinson. Fossil Mammalia of the Huerfano Formation, Eocene, of Colorado. *Peabody Museum of Natural History Bulletin*, 21:1–95, 1966.

[261] Peter D Roopnarine. The description and classification of evolutionary mode: a computational approach. *Paleobiology*, 27(3):446–465, 2001.

[262] Peter D Roopnarine. Analysis of rates of morphologic evolution. *Annual Review of Ecology, Evolution, and Systematics*, 34:605–632, 2003.

[263] Peter D Roopnarine, Gabe Byars, and Paul Fitzgerald. Anagenetic evolution, stratophenetic patterns, and random walk models. *Paleobiology*, 25(1):41–57, 1999.

[264] K D Rose and D W Krause. Cyriacotheriidae, a new family of early Tertiary pantodonts from western North America. *Proceedings of the American Philosophical Society*, 126(1):26–50, 1982.

[265] Kenneth D Rose, Stephen G B Chester, Rachel H Dunn, Doug M Boyer, and Jonathan I Bloch. New fossils of the oldest North American euprimate Teilhardina brandti (Omomyidae) from the paleocene-eocene thermal maximum. *American Journal of Physical Anthropology*, 146(2):281–305, 2011.

[266] Kenneth D. Rose, Kishor Kumar, Rajendra S. Rana, Ashok Sahni, and Thierry Smith. New Hypsodont Tillodont (Mammalia, Tillodontia) from the Early Eocene of India. *Journal of Paleontology*, 87(5):842–853, 2013.

[267] James Rosindell, Luke J Harmon, and Rampal S Etienne. Unifying ecology and macroevolution with individual-based theory. *Ecology Letters*, 18(5):472–482, 2015.

[268] Martin Rosvall, D Axelsson, and CT Bergstrom. The map equation. *The European Physical Journal Special Topics*, 178(14):13–24, 2009.

[269] Martin Rosvall and Carl T Bergstrom. Maps of random walks on complex networks reveal community structure. *Proceedings of the National Academy of Sciences*, 105(4):1118–23, 2008.

[270] Kaustuv Roy, Gene Hunt, David Jablonski, Andrew Z Krug, and James W Valentine. A macroevolutionary perspective on species range limits. *Proceedings of the Royal Society B: Biological Sciences*, 276:1485–1493, 2009.

[271] J Andrew Royle and Robert M Dorazio. *Hierarchical modeling and inference in ecology: the analysis of data from populations, metapopulations and communities.* Elsevier, London, 2008.

[272] J Andrew Royle, James D Nichols, and Marc Kéry. Modelling occurrence and abundance of species when detection is imperfect. *Oikos*, 110(January):353–359, 2005.

[273] Donald B Rubin. Multiple imputation after 18+ years. *Journal of the American Statistical Assocaition*, 91(434):473–489, 1996.

[274] Marten Scheffer and Egbert H van Nes. Self-organized similarity, the evolutionary emergence of groups of similar species. *Proceedings of the National Academy of Sciences*, 103(16):6230–6235, 2006.

[275] Holger Schielzeth. Simple means to improve the interpretability of regression coefficients. *Methods in Ecology and Evolution*, 1(2):103–113, 2010.

[276] C S Scott. A new species of the ptilodontid multituberculate Prochetodon (Mammalia, Allotheria) from the Paleocene Paskapoo Formation of Alberta, Canada. *Canadian Journal of Earth Sciences*, 41:237–246, 2004.

[277] Craig S Scott. Late Torrejonian (Middle Paleocene) mammals from South Central Alberta, Canada. *Journal of Paleontology*, 77(4):745–768, 2003.

[278] Craig S Scott, Daniel N Spivak, Arthur R Sweet, and Hans Sues. First mammals from the Paleocene Porcupine Hills Formation of southwestern Alberta, Canada 1. *Canadian Journal of Earth Sciences*, 50(3):355–378, 2013.

[279] William Berryman Scott and Glenn Lowell Jepsen. The mammalian fauna of the White River Oligocene: Part IV. Artiodactyla. *Transactions of the American Philosophical Society*, 28(4):363–746, 1940.

[280] William Berryman Scott, Glenn Lowell Jepsen, and Albert Elmer Wood. The mammalian fauna of the White River Oligocene: Part II. Rodentia. *Transactions of the American Philosophical Society*, 28(2):155–269, 1937.

[281] Ross Secord. *The Tiffanian Land-Mammal Age (Middle and Late Paleocene) in the Northern Bighorn Basin , Wyoming.* University of Michicagn, 2008.

[282] David Sepkoski. *Rereading the fossil record: the growth of paleobiology as an evolutionary discipline.* University of Chicago Press, Chicago, 2015.

[283] David Sepkoski and Michael Ruse. *The paleobiological revolution: essays on the growth of modern paleontology.* University of Chicago Press, Chicago, 2009.

[284] J John Sepkoski. Stratigraphic biases in the analysis of taxonomic survivorship. *Paleobiology*, 1(4):343–355, 1975.

[285] J John Sepkoski. A factor analytic description of the Phanerozoic marine fossil record. *Paleobiology*, 7(1):36–53, 1981.

[286] P M Sheehan. The late Ordovician mass extinction. *Annual Review of Earth and Planetary Sciences*, 29:331–364, 2001.

[287] H David Sheets and Charles E Mitchell. Uncorrelated change produces the apparent dependence of evolutionary rate on interval. *Paleobiology*, 27(3):429–445, 2001.

[288] Bill Shipley, Denis Vile, and Eric Garnier. From plant traits to plant communities: a statistical mechanistic approach to biodiversity. *Science*, 314:812–814, 2006.

[289] Christian A Sidor, Daril A Vilhena, Kenneth D Angielczyk, Adam K Huttenlocker, Sterling J Nesbitt, Brandon R Peecook, J Sébastien Steyer, Roger M H Smith, and Linda A Tsuji. Provincialization of terrestrial faunas following the end-Permian mass extinction. *Proceedings of the National Academy of Sciences*, 110(20):8129–33, 2013.

[290] Daniele Silvestro, Alexandre Antonelli, Nicolas Salamin, and Tiago B Quental. The role of clade competition in the diversification of North American canids. *Proceedings of the National Academy of Sciences*, 112(28):8684–9, 2015.

[291] Daniele Silvestro, Jan Schnitzler, Lee Hsiang Liow, Alexandre Antonelli, and Nicolas Salamin. Bayesian estimation of speciation and extinction from incomplete fossil occurrence data. *Systematic Biology*, 63(3):349–67, 2014.

[292] Daniel Simberloff and Tamar Dayan. The Guild Concept and the Structure of Ecological Communities. *Annual Review of Ecology and Systematics*, 22:115–143, 1991.

[293] Elwyn L Simons. The Paleocene Pantodonta. *Transactions of the American Philosophical Society*, 50(6):3–99, 1960.

[294] Carl Simpson. *Levels of selection and large-scale morphological trends*. PhD thesis, University of Chicago, 2006.

[295] Carl Simpson. The case for species selection. *bioRxiv*, 2016.

[296] Carl Simpson and Paul G. Harnik. Assessing the role of abundance in marine bivalve extinction over the post-Paleozoic. *Paleobiology*, 35(4):631–647, dec 2009.

[297] Carl Simpson, Wolfgang Kiessling, Heike Mewis, Rosemarie C Baron-Szabo, and Johannes Müller. Evolutionary diversification of reef corals: a comparison of the molecular and fossil records. *Evolution*, 65(11):3274–3284, nov 2011.

[298] George Gaylord Simpson. *Tempo and Mode in Evolution*. Columbia University Press, New York, 1944.

[299] George Gaylord Simpson. *The Major Features of Evolution*. Columbia University Press, New York, 1953.

[300] W J Sinclair. Additions to the fauna of the lower Pliocene Snake Creek beds (results of the Princeton University 1914 expedition to Nebraska). *Proceedings of the American Philosophical Society*, 54(217):73–95, 1915.

[301] Morris F Skinner and Claude W Hibbard. Early Pleistocene pre-glacial and glacial rocks and faunas of North-Central Nebraska. *Bulletin of the American Museum of Natural History*, 148:1–148, 1972.

[302] G J Slater, L J Harmon, and M E Alfaro. Integrating fossils with molecular phylogenies improves inference of trait evolution. *Evolution*, 66(12):3931–3944, 2012.

[303] Graham J Slater. Phylogenetic evidence for a shift in the mode of mammalian body size evolution at the Cretaceous-Palaeogene boundary. *Methods in Ecology and Evolution*, 4(8):734–744, 2013.

[304] Graham J Slater. Iterative adaptive radiations of fossil canids show no evidence for diversity-dependent trait evolution. *Proceedings of the National Academy of Sciences*, 112(16):4897–4902, 2015.

[305] F A Smith, J Brown, J Haskell, and S Lyons. Similarity of mammalian body size across the taxonomic hierarchy and across space and time. *The American Naturalist*, 163:672–691, 2004.

[306] F A Smith, S K Lyons, S K Morgan Ernest, and J H Brown. Macroecology: more than the division of food and space among species on continents. *Progress in Physical Geography*, 32(2):115–138, apr 2008.

[307] Felisa A Smith, James H Brown, John P Haskell, S Kathleen Lyons, John Alroy, Eric L Charnov, Tamar Dayan, Brian J Enquist, S K Morgan Ernest, Elizabeth A Hadly, Kate E Jones, Dawn M Kaufman, Pablo A Marquet, Brian A Maurer, Karl J Niklas, Warren P Porter, Bruce Tiffney, and Michael R Willig. Similarity of Mammalian Body Size across the Taxonomic Hierarchy and across Space and Time. *The American Naturalist*, 163(5):672–691, 2004.

[308] Michael Smithson and Jay Verkuilen. A better lemon squeezer? Maximum-likelihood regression with beta-distributed dependent variables. *Psychological Methods*, 11(1):54–71, 2006.

[309] Peter D Smits. Expected time-invariant effects of biological traits on mammal species duration. *Proceedings of the National Academy of Sciences*, 112(42):13015–13020, 2015.

[310] Janne Soininen. A quantitative analysis of species sorting across organisms and ecosystems. *Ecology*, 95(12):3284–3292, 2014.

[311] R R Sokal and F J Rohlf. *Biometry*. W. H. Freeman, New York, 4 edition, 2011.

[312] Christophe Soligo and Robert D Martin. Adaptive origins of primates revisited. *Journal of Human Evolution*, 50(4):414–30, 2006.

[313] A R Solow and Woollcott Smith. On fossil preservation and the stratigraphic ranges of taxa. *Paleobiology*, 23(3):271–277, 1997.

[314] Boris Sorkin. A biomechanical constraint on body mass in terrestrial mammalian predators. *Lethaia*, 41(4):333–347, dec 2008.

[315] T Stadler. Mammalian phylogeny reveals recent diversification rate shifts. *Proceedings of the National Academy of Sciences*, 108(15):6187–6192, 2011.

[316] T Stadler. Recovering speciation and extinction dynamics based on phylogenies. *Journal of Evolutionary Biology*, 26:1203–1219, 2013.

[317] Tanja Stadler and Folmer Bokma. Estimating speciation and extinction rates for phylogenies of higher taxa. *Systematic Biology*, 62(2):220–30, 2013.

[318] Stan Development Team. Stan: A c++ library for probability and sampling, version 2.5.0, 2014.

[319] Stan Development Team. Stan Modeling Language Users Guide and Reference Manual, 2016.

[320] S M Stanley. A theory of evolution above the species level. *Proceedings of the National Academy of Sciences*, 72(2):646–650, 1975.

[321] S M Stanley. *Macroevolution: pattern and process*. W. H. Freeman, San Francisco, 1979.

[322] R A Stirton. An association of horn cores and upper molars of the antelope Sphenophalos nevadanus from the lower Pliocene of Nevada. *American Journal of Science*, 24:46–51, 1932.

[323] C Stock. An Eocene titanothere from San Diego County, California, with remarks on the age of the Poway Conglomerate. *Proceedings of the National Academy of Sciences*, 23:48–53, 1937.

[324] Chester Stock. Restos de Tejon (Taxidea) Pliocenico del Occidente de Chihuahua. *Boletin de la Sociedad Geologica Mexicana*, 13:69–76, 1948.

[325] S G Strait. Dietary reconstruction of small-bodied omomyoid primates. *Journal of Vertebrate Paleontology*, 21:322–334, 2001.

[326] Caroline A E Strömberg. Decoupled taxonomic radiation and ecological expansion of open-habitat grasses in the Cenozoic of North America. *Proceedings of the National Academy of Sciences*, 102(34):11980–4, 2005.

[327] Beryl E Taylor and S David Webb. Miocene Leptomerycidae (Artiodactyla, Ruminantia) and their relationships. *American Museum Novitates*, 2596:1–22, 1976.

[328] Richard H Tedford, L G Barnes, and C E Ray. The early Miocene littoral ursoid carnivoran Kolponomos: systematics and mode of life. *Proceedings of the San Diego Society of Natural History*, 29:11–32, 1994.

[329] Susumu Tomiya. Body Size and Extinction Risk in Terrestrial Mammals Above the Species Level. *The American Naturalist*, 182:196–214, 2013.

[330] Zhijie J Tseng, Jingmai K O'Connor, Xiaoming Wang, and Donald R Prothero. The first Old World occurrence of the North American mustelid Sthenictis (Mammalia, Carnivora). *Geodiversitas*, 31(4):743–751, 2009.

[331] Mark C Urban, Mathew A Leibold, Priyanga Amarasekare, Luc De Meester, Richard Gomulkiewicz, Michael E. Hochberg, Christopher A Klausmeier, Nicolas Loeuille, Claire de Mazancourt, Jon Norberg, Jelena H Pantel, Sharon Y Strauss, Mark Vellend, and Michael J Wade. The evolutionary ecology of metacommunities. *Trends in Ecology and Evolution*, 23(6):311–317, 2008.

[332] J C Uyeda, T F Hansen, S J Arnold, and J Pienaar. The million-year wait for macroevolutionary bursts. *Proceedings of the National Academy of Sciences*, 108(38):15908–15913, 2011.

[333] James W Valentine. Patterns of taxonomic and ecological structure of the shelf benthos during Phanerozoic time. *Paleontology*, 12(4):684–709, 1969.

[334] Katleen Van der Gucht, Karl Cottenie, Koenraad Muylaert, Nele Vloemans, Sylvie Cousin, Steven Declerck, Erik Jeppesen, Jose-Maria Conde-Porcuna, Klaus Schwenk, Gabriel Zwart, Hanne Degans, Wim Vyverman, and Luc De Meester. The power of species sorting: local factors drive bacterial community composition over a wide range of spatial scales. *Proceedings of the National Academy of Sciences of the United States of America*, 104(51):20404–20409, 2007.

[335] Leigh Van Valen. A new evolutionary law. *Evolutionary Theory*, 1:1–30, 1973.

[336] Leigh Van Valen. Taxonomic survivorship curves. *Evolutionary Theory*, 4:129–142, 1979.

[337] Blair Van Valkenburgh. Skeletal and dental predictors of body mass in carnivores. In John Damuth and Bruce J Macfadden, editors, *Body size in mammalian palebiology: estimation and biological implications*, pages 181–205. Cambridge University Press, Cambridge, 1990.

[338] Blaire Van Valkenburgh. Major patterns in the history of carnivorous mammals. *Annual Review of Earth and Planetary Sciences*, 27:463–493, 1999.

[339] Blaire Van Valkenburgh. Deja vu: the evolution of feeding morphologies in the Carnivora. *Integrative and comparative biology*, 47(1):147–63, jul 2007.

[340] Daril A Vilhena. *Boundaries and dynamics of biomes.* PhD thesis, University of Washington, 2013.

[341] Daril A Vilhena, Elisha B Harris, Carl T Bergstrom, Max E Maliska, Peter D Ward, Christian A Sidor, Caroline A E Strömberg, and Gregory P Wilson. Bivalve network reveals latitudinal selectivity gradient at the end-Cretaceous mass extinction. *Scientific Reports*, 3:1790, 2013.

[342] Sébastien Villéger, Philip M Novack-Gottshall, and David Mouillot. The multidimensionality of the niche reveals functional diversity changes in benthic marine biotas across geological time. *Ecology Letters*, 14(6):561–8, 2011.

[343] Elisabeth S Vrba. What is species selection? *Systematic Zoology*, 33(3):318–328, 1984.

[344] Elisabeth S Vrba and Niles Eldredge. Individuals, hierarchies and processes: towards a more complete evolutionary theory. *Paleobiology*, 10(2):146–171, 1984.

[345] Elisabeth S Vrba and Stephen Jay Gould. The hierarchical expansion of sorting and selection: sorting and selection cannot be equated. *Paleobiology*, 12(2):217–228, 1986.

[346] P J Wagner, M Aberhan, A Hendy, and W Kiessling. The effects of taxonomic standardization on sampling-standardized estimates of historical diversity. *Proceedings of the Royal Society B: Biological Sciences*, 274(1608):439, 2007.

[347] Peter J Wagner and George F Estabrook. Trait-based diversification shifts reflect differential extinction among fossil taxa. *Proceedings of the National Academy of Sciences*, 111:16419–16424, 2014.

[348] Peter J Wagner and Jonathan D Marcot. Modelling distributions of fossil sampling rates over time, space and taxa: assessment and implications for macroevolutionary studies. *Methods in Ecology and Evolution*, 4(8):703–713, 2013.

[349] Steve C Wang. On the continuity of background and mass extinction. *Paleobiology*, 29(4):455–467, 2003.

[350] Steve C Wang, Philip J Everson, Heather Jianan Zhou, Dasol Park, and David J Chudzicki. Adaptive credible intervals on stratigraphic ranges when recovery potential is unknown. *Paleobiology*, 42(2):240–256, 2016.

[351] Steve C Wang and C R Marshall. Improved confidence intervals for estimating the position of a mass extinction boundary. *Paleobiology*, 30(1):5–18, 2004.

[352] Steve C Wang and Charles R Marshall. Estimating times of extinction in the fossil record. *Biology Letters*, 12(4):20150989, 2016.

[353] Xiaoming Wang. Phylogenetic systematics of the Hesperocyoninae (Carnivora: Canidae). *Bulletin of the American Museum of Natural History*, 221:1–207, 1994.

[354] Xiaoming Wang, Óscar Carranza-Castañeda, and José Jorge Aranda-Gómez. A transitional skunk, Buisnictis metabatos sp. nov. (Mephitidae, Carnivora), from Baja California Sur and the role of southern refugia in skunk evolution. *Journal of Systematic Palaeontology*, 12(3):291–302, 2014.

[355] Xiaoming Wang, Richard H Tedford, and Beryl E Taylor. Phylogenetic systematics of the Borophaginae (Carnivora, Canidae). *Bulletin of the American Museum of Natural History*, 243:2–391, 1999.

[356] David I Warton, Bill Shipley, and Trevor Hastie. CATS regression - a model-based approach to studying trait-based community assembly. *Methods in Ecology and Evolution*, 6(4):389–398, 2015.

[357] Sumio Watanabe. Asymptotic Equivalence of Bayes Cross Validation and Widely Applicable Information Criterion in Singular Learning Theory. *Journal of Machine Learning Research*, 11:3571–3594, 2010.

[358] Marjorie G Weber, Catherine E Wagner, Rebecca J Best, Luke J Harmon, and Blake Matthews. Evolution in a Community Context: On Integrating Ecological Interactions and Macroevolution. *Trends in Ecology & Evolution*, xx:1–14.

[359] C Williams, C H C Brunton, and S J Carlson. *Treatise on invertebrate paleontology. Part H, Brachiopoda.* Geological Society of America, Boulder, Colorado, 2007.

[360] Thomas E Williamson and Stephen L Brusatte. New specimens of the rare taeniodont Wortmania (Mammalia: Eutheria) from the San Juan Basin of New Mexico and comments on the phylogeny and functional morphology of "archaic" mammals. *PloS one*, 8(9):e75886, 2013.

[361] Thomas E Williamson, Stephen L Brusatte, Thomas D Carr, Anne Weil, and Barbara R Standhardt. The phylogeny and evolution of CretaceousPalaeogene metatherians: cladistic analysis and description of new early Palaeocene specimens from the Nacimiento Formation, New Mexico. *Journal of Systematic Palaeontology*, 10(4):625–651, 2012.

[362] Gregory P Wilson, Alistair R Evans, Ian J Corfe, Peter D Smits, Mikael Fortelius, and Jukka Jernvall. Adaptive radiation of multituberculate mammals before the extinction of dinosaurs. *Nature*, 483:457–460, 2012.

[363] J Bastow Wilson. Guilds, functional types and ecological groups. *Oikos*, 86(3):507–522, 1999.

[364] A E Wood. The early Tertiary rodents of the family Paramyidae. *Transactions of the American Philosophical Society*, 52(1):3–261, 1962.

[365] J L Wortman and C Earle. Ancestors of the tapir from the lower Miocene of Dakota. *Bulletin of the American Museum of Natural History*, 5:159–180, 1893.

[366] J B Yoder, 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 Harmon. Ecological opportunity and the origin of adaptive radiations. *Journal of Evolutionary Biology*, 23:1581–1596, 2010.

[367] James C Zachos, Gerald R Dickens, and Richard E Zeebe. An early Cenozoic perspective on greenhouse warming and carbon-cycle dynamics. *Nature*, 451(7176):279–283, 2008.

[368] James C Zachos, M Pagani, L Sloan, E Thomas, and K Billups. Trends, rhythms, and aberrations in global climate 65 Ma to present. *Science*, 292:686–693, 2001.

[369] S P Zack, T A Penkrot, D W Krause, and M C Maas. A new apheliscine "condylarth" mammal from the late Paleocene of Montana and Alberta and the phylogeny of "hyopsodontids". *Acta Palaeontologica Polonica*, 50:809–830, 2005.

[370] R J Zakrzewski. New species of Blancan woodrat (Cricetidae) from north-central Kansas. *Journal of Mammalogy*, 72(1):104–109, 1991.

[371] John-Paul Zonneveld and Gregg F Gunnell. A new species of cf. Dilophodon (Mammalia; Perissodactyla) from the early Bridgerian of southwestern Wyoming. *Journal of Vertebrate Paleontology*, 23(3):652–658, 2003.

APPENDIX A

TABLE OF SPECIES MASS ESTIMATES

Table A.1: Body mass estimates in grams for all mammal species included in this study. Also included is the source of this measurement. PBDB = Paleobiology Database, NOW = Neogene Old World Database, EOL = Encyclopedia of Life, ADW = Animal Diversity Web. PBDB + regression indicates that body mass was estimated from a measurement on the PBDB using one of the many regression equations listed in Table 2.3.

Species	Mass (g)	Source
Aaptoryctes ivyi	215.70	PBDB + regression
Absarokius abbotti	176.39	PBDB + regression
Absarokius metoecus	166.29	PBDB + regression
Acarictis ryani	291.90	PBDB + regression
Achaenodon robustus	492900.83	PBDB + regression
Achlyoscapter longirostris	11.94	[329]
Acmeodon secans	325.36	PBDB + regression
Acritohippus isonesus	135944.23	[329]
Acritohippus quinni	178082.11	[329]
Acritoparamys atwateri	190.07	PBDB + regression
Acritoparamys francesi	125.17	PBDB + regression
Acritoparamys pattersoni	432.28	PBDB
Acritoparamys wyomingensis	274.75	PBDB
Adeloblarina berklandi	12.68	[329]
Adelphailurus kansensis	33189.87	[329]
Adilophontes brachykolos	119372.01	[329]
Adjidaumo burkei	4.91	PBDB
Adjidaumo craigi	7.29	PBDB + regression
Adjidaumo intermedius	10.39	PBDB + regression

Continued on next page

Table A.1 – continued from previous page

Species	Mass (g)	Source
Adjidaumo maximus	20.14	PBDB + regression
Adjidaumo minimus	7.29	PBDB + regression
Adjidaumo minutus	15.90	PBDB + regression
Adunator ladae	17.36	PBDB + regression
Aelurodon asthenostylus	22026.47	[329]
Aelurodon ferox	26370.47	[329]
Aelurodon mcgrewi	22247.84	[329]
Aelurodon montanensis	27722.51	[329]
Aelurodon stirtoni	20537.34	[329]
Aelurodon taxoides	29436.77	[329]
Aepinacodon americanus	423622.49	[20]
Aepycamelus bradyi	516464.33	[59]
Aepycamelus giraffinus	499050.27	PBDB
Aepycamelus robustus	420836.64	[329]
Aepycamelus stocki	348014.70	[329]
Aethomylos simplicidens	30.03	PBDB + regression
Ageitodendron matthewi	331.60	PBDB + regression
Agriochoerus antiquus	56387.34	[329]
Agriochoerus guyotianus	59000.00	[198]
Agriotherium schneideri	355045.06	[329]
Alagomys russelli	4.67	PBDB + regression
Aletodon conardae	576.81	PBDB + regression
Aletodon gunnelli	615.51	PBDB + regression
Aletodon quadravus	365.67	PBDB + regression
Aletomeryx gracilis	27173.57	[329]

Continued on next page

Table A.1 – continued from previous page

Species	Mass (g)	Source
Alforjas taylori	408399.03	[329]
Alilepus vagus	225.12	PBDB + regression
Alilepus wilsoni	111.45	PBDB + regression
Allomys cristabrevis	239.72	PBDB + regression
Allomys simplicidens	122.00	[198]
Allomys storeri	92.85	PBDB + regression
Alluvisorex arcadentes	5.70	[329]
Alphagaulus pristinus	521.00	[198]
Alphagaulus vetus	523.22	[329]
Alticonus gazini	1140.52	PBDB + regression
Alveojunctus minutus	25.80	PBDB + regression
Alveugena carbonensis	815.39	PBDB
Alwoodia harkseni	294.29	PBDB + regression
Alwoodia magna	226.00	[198]
Amebelodon floridanus	37020787.90	PBDB + regression
Amelotabes simpsoni	57.16	PBDB + regression
Ammospermophilus hanfordi	58.00	[198]
Ammospermophilus junturensis	53.52	[329]
Amphechinus horncloudi	175.91	[329]
Amphicaenopus platycephalus	2397650.84	[329]
Amphicyon frendens	245241.81	[329]
Amphicyon galushai	138690.48	[329]
Amphicyon ingens	600000.00	[314]
Amphicyon longiramus	113550.16	[329]
Amphicyon riggsi	418398.16	PBDB

Continued on next page

Table A.1 – continued from previous page

Species	Mass (g)	Source
Amphimachairodus coloradensis	157147.46	PBDB + regression
Ampliconus antoni	2906.21	PBDB + regression
Amynodon advenus	1987010.17	PBDB + regression
Amynodon reedi	579004.85	PBDB
Amynodontopsis bodei	608129.44	PBDB + regression
Anaptomorphus aemulus	95.39	PBDB + regression
Anaptomorphus westi	186.75	PBDB + regression
Anasazia williamsoni	186.23	[361]
Anchitherium clarencei	230960.04	[329]
Anconodon cochranensis	57.00	[362]
Anemorhysis natronensis	32.59	PBDB + regression
Anemorhysis pattersoni	57.16	PBDB + regression
Anemorhysis pearcei	32.00	[2]
Anemorhysis wortmani	42.74	PBDB + regression
Angustidens vireti	18.17	[329]
Anisonchus oligistus	388.07	PBDB + regression
Anisonchus onostus	616.78	PBDB + regression
Anisonchus sectorius	1619.16	PBDB + regression
Ankalagon saurognathus	28250.27	PBDB + regression
Ankylodon annectens	30.27	[329]
Anomoemys lewisi	207.64	[293]
Ansomys hepburnensis	44.70	[329]
Ansomys nevadensis	51.94	[329]
Ansomys nexodens	92.85	PBDB + regression
Antecalomys phthanus	22.87	[329]

Continued on next page

Table A.1 – continued from previous page

Species	Mass (g)	Source
Antecalomys valensis	13.60	[329]
Antecalomys vasquezi	10.70	PBDB
Antiacodon pygmaeus	464.26	PBDB + regression
Antiacodon venustus	2156.43	PBDB + regression
Apataelurus kayi	10833.16	PBDB + regression
Apatemys bellulus	17.36	PBDB + regression
Apatemys bellus	21.80	PBDB + regression
Apatemys chardini	14.63	PBDB + regression
Apatemys downsi	65.02	PBDB + regression
Apatemys hendryi	12.08	PBDB + regression
Apatemys uintensis	28.30	PBDB + regression
Apatosciuravus bifax	35.35	PBDB + regression
Apatosciuravus jacobsi	29.87	PBDB + regression
Apheliscus chydaeus	47.27	PBDB + regression
Apheliscus insidiosus	60.38	PBDB + regression
Apheliscus nitidus	77.18	PBDB + regression
Apheliscus wapitiensis	26.62	PBDB + regression
Aphelops malacorhinus	3541284.24	[329]
Aphelops megalodus	1689595.99	[329]
Aphelops mutilus	4325334.34	[329]
Aphronorus fraudator	112.44	PBDB + regression
Aphronorus orieli	427.05	PBDB + regression
Aphronorus ratatoski	176.14	PBDB + regression
Apletotomeus crassus	8.78	PBDB + regression
Apternodus gregoryi	90.14	PBDB + regression

Continued on next page

Table A.1 – continued from previous page

Species	Mass (g)	Source
Apternodus iliffensis	87.48	PBDB + regression
Arapahovius advena	42.74	PBDB + regression
Archaeocyon falkenbachi	2116.88	[322]
Archaeocyon leptodus	3533.34	[329]
Archaeocyon pavidus	2275.60	[329]
Archaeohippus blackbergi	33189.87	[329]
Archaeohippus mannulus	48917.80	PBDB + regression
Archaeohippus mourningi	54176.36	[329]
Archaeohippus penultimus	71682.36	[329]
Archaeohippus stenolophus	63678.98	PBDB
Archaeolagus acaricolus	578.25	[329]
Archaeolagus emeraldensis	2344.90	[329]
Archaeolagus ennisianus	1064.22	[329]
Archaeolagus macrocephalus	1826.21	[329]
Archaeolagus primigenius	1540.71	[329]
Archaeotherium lemleyi	1361176.27	PBDB + regression
Archaeotherium mortoni	240808.54	PBDB + regression
Archaeotherium trippensis	3353590.86	PBDB + regression
Arctocyon montanensis	14771.98	PBDB + regression
Arctocyon mumak	57080.05	PBDB
Arctodontomys nuptus	318.11	PBDB + regression
Arctodontomys simplicidens	219.44	PBDB + regression
Arctodontomys wilsoni	236.68	PBDB + regression
Arctodus pristinus	299916.25	[305]
Arctonasua eurybates	15994.50	[329]

Continued on next page

Table A.1 – continued from previous page

Species	Mass (g)	Source
Arctonasua gracilis	8866.19	[329]
Arctonasua minima	7044.48	[329]
Arctostylops steini	512.92	PBDB + regression
Ardynomys occidentalis	202.01	PBDB + regression
Arfia junnei	1188.60	PBDB + regression
Arfia opisthotoma	11400.50	PBDB + regression
Arfia shoshoniensis	9550.20	PBDB + regression
Arfia zele	2236.53	PBDB + regression
Arretotherium acridens	179871.86	[329]
Arretotherium fricki	138690.48	[329]
Arretotherium leptodus	252848.32	PBDB
Artimonius australis	133.09	PBDB + regression
Artimonius nocerai	66.71	PBDB + regression
Artimonius witteri	91.55	PBDB + regression
Astrohippus stockii	134877.06	PBDB
Aulolithomys bounites	68.53	PBDB + regression
Aulolithomys vexilliames	17.97	PBDB + regression
Australocamelus orarius	100709.96	[329]
Auxontodon pattersoni	6261.76	PBDB + regression
Avunculus didelphodonti	60.38	PBDB + regression
Aycrossia lovei	124.23	PBDB + regression
Aztlanolagus agilis	1999.86	[305]
Azygonyx ancylion	15969.85	PBDB + regression
Azygonyx grangeri	24602.41	PBDB + regression
Azygonyx xenicus	7316.20	PBDB + regression

Continued on next page

Table A.1 – continued from previous page

Species	Mass (g)	Source
Baioconodon denverensis	4195.64	PBDB + regression
Baioconodon nordicus	1675.41	PBDB + regression
Baiomys rexroadi	7.29	PBDB + regression
Baiotomeus douglassi	192.03	PBDB + regression
Baiotomeus rhothonion	12.82	PBDB + regression
Barbourofelis fricki	255250.32	[329]
Barbourofelis loveorum	160091.76	PBDB + regression
Barbourofelis morrisi	90219.42	[329]
Barbourofelis osborni	162309.29	[192]
Barbourofelis whitfordi	77652.58	[329]
Barbouromeryx trigonocorneus	33860.35	[329]
Barylambda faberi	255787.00	PBDB + regression
Barylambda jackwilsoni	43914.63	PBDB + regression
Basirepomys pliocenicus	72.36	PBDB + regression
Basirepomys robertsi	54.13	PBDB + regression
Bassariscus antiquus	1881.83	[329]
Bassariscus casei	1652.43	[329]
Bassariscus minimus	1183.06	[260]
Bassariscus ogallalae	1724.26	[103]
Bassariscus parvus	1685.81	[329]
Bathygenys alpha	6155.81	PBDB + regression
Bathygenys reevesi	5297.18	PBDB + regression
Batodonoides powayensis	3.14	PBDB + regression
Batodonoides vanhouteni	1.26	PBDB + regression
Beckiasorex hibbardi	6.56	PBDB + regression

Continued on next page

Table A.1 – continued from previous page

Species	Mass (g)	Source
Bensonomys arizonae	23.19	PBDB
Bensonomys baskini	26.95	PBDB
Bensonomys elachys	8.78	[159]
Bensonomys gidleyi	13.90	PBDB
Bensonomys lindsayi	8.78	PBDB + regression
Bensonomys meadensis	19.00	PBDB
Bensonomys winklerorum	9.09	PBDB
Bensonomys yazhi	9.98	[266]
Betonnia tsosia	50.22	[45]
Bisonalveus browni	90.14	PBDB + regression
Bisonalveus holtzmani	146.12	PBDB + regression
Blacktops latidens	344.57	PBDB
Blacktops longinares	255.48	PBDB
Blarina brevicauda	16.40	PBDB
Blarina carolinensis	13.49	[305]
Blastomeryx gemmifer	10938.02	[329]
Blastomeryx pristinus	38955.58	PBDB
Blickomylus galushai	16983.54	[329]
Boreameryx braskerudi	9643.88	PBDB + regression
Borophagus diversidens	34891.55	[329]
Borophagus dudleyi	94433.22	[57]
Borophagus hilli	29143.87	[329]
Borophagus littoralis	23388.51	[329]
Borophagus orc	16814.55	[329]
Borophagus parvus	19341.34	[329]

Continued on next page

Table A.1 – continued from previous page

Species	Mass (g)	Source
Borophagus pugnator	24100.79	[329]
Borophagus secundus	24100.79	[329]
Bothriodon rostratus	865704.50	[40]
Bouromeryx americanus	68186.37	[329]
Bouromeryx submilleri	50011.09	[329]
Brachycrus buwaldi	250196.03	[329]
Brachycrus laticeps	359952.66	PBDB
Brachycrus rusticus	117343.54	PBDB
Brachycrus siouense	145801.30	[329]
Brachyerix hibbardi	78.19	[44]
Brachyerix incertis	79.84	[329]
Brachyerix macrotis	131.63	[329]
Brachyerix richi	340.36	[329]
Brachyhyops viensis	202519.16	PBDB + regression
Brachyhyops wyomingensis	97111.99	PBDB + regression
Brachypsalis modicus	512.86	[329]
Brachypsalis obliquidens	63169.09	PBDB
Brachypsalis pachycephalus	487.85	[329]
Brachyrhynchocyon dodgei	13906.88	PBDB + regression
Brachyrhynchocyon montanus	7397.01	PBDB + regression
Brontops tyleri	571500.00	PBDB
Buisnictis breviramus	849.04	PBDB + regression
Buisnictis burrowsi	817.07	PBDB
Buisnictis schoffi	22.42	[329]
Bunomeryx montanus	3196.73	PBDB + regression

Continued on next page

Table A.1 – continued from previous page

Species	Mass (g)	Source
Bunophorus etsagicus	13323.34	PBDB + regression
Bunophorus grangeri	13969.78	PBDB + regression
Bunophorus macropternus	10378.10	PBDB + regression
Bunophorus pattersoni	6332.74	[187]
Bunophorus robustus	5163.91	PBDB + regression
Bunophorus sinclairi	16312.82	PBDB + regression
Caenolambda jepseni	40906.63	PBDB + regression
Calippus cerasinus	81633.91	[329]
Calippus elachistus	43044.94	[329]
Calippus hondurensis	71682.36	[329]
Calippus martini	119372.01	[329]
Calippus placidus	79221.26	[329]
Calippus proplacidus	64860.88	[329]
Calippus regulus	45251.90	[329]
Camelops hesternus	1099005.84	[305]
Camelops traviswhitei	1027000.94	PBDB + regression
Campestrallomys annectens	80.29	PBDB + regression
Campestrallomys dawsonae	298.87	[329]
Campestrallomys siouxensis	159.17	[329]
Canis armbrusteri	30333.26	[329]
Canis edwardii	79873.95	PBDB + regression
Canis latrans	11765.00	PBDB
Canis lepophagus	14617.87	[329]
Canis rufus	15566.00	PBDB
Cantius abditus	2798.04	PBDB + regression

Continued on next page

Table A.1 – continued from previous page

Species	Mass (g)	Source
Cantius angulatus	417.80	PBDB + regression
Cantius frugivorus	913.71	PBDB + regression
Cantius mckennai	902.97	PBDB + regression
Cantius nunienus	1212.97	PBDB + regression
Cantius ralstoni	738.98	PBDB + regression
Cantius simonsi	3858.47	PBDB + regression
Cantius torresi	410.27	PBDB + regression
Cantius trigonodus	2000.00	[312]
Capricamelus gettyi	361557.35	PBDB + regression
Capromeryx tauntonensis	15835.35	[329]
Cardiolophus radinskyi	32305.76	PBDB + regression
Cardiolophus semihians	33364.14	PBDB + regression
Carpocristes cygneus	25.80	PBDB + regression
Carpocristes hobackensis	33.00	[312]
Carpocyon compressus	15214.44	[329]
Carpocyon robustus	19341.34	[329]
Carpocyon webbi	20537.34	[329]
Carpodaptes hazelae	51.17	PBDB + regression
Carpodaptes stonleyi	32.59	PBDB + regression
Carpolestes nigridens	87.00	[277]
Carpolestes simpsoni	27.99	PBDB + regression
Carpomegodon jepseni	166.29	PBDB + regression
Catopsalis alexanderi	3415.85	PBDB + regression
Catopsalis calgariensis	22815.78	PBDB + regression
Catopsalis foliatus	7371.53	PBDB + regression

Continued on next page

Table A.1 – continued from previous page

Species	Mass (g)	Source
Catopsalis joyneri	2435.00	[362]
Cedromus wardi	246.28	PBDB + regression
Centetodon aztecus	17.36	PBDB + regression
Centetodon bembicophagus	10.87	PBDB + regression
Centetodon chadronensis	23.37	PBDB + regression
Centetodon divaricatus	30.57	[329]
Centetodon hendryi	17.36	PBDB + regression
Centetodon kuenzii	26.62	PBDB + regression
Centetodon magnus	33.45	[329]
Centetodon neashami	53.67	PBDB + regression
Centetodon pulcher	30.03	PBDB + regression
Centimanomys major	239.72	PBDB + regression
Ceratogaulus hatcheri	1490.63	[40]
Cernictis hesperus	177.68	[329]
Cernictis repenningi	8800.69	[111]
Chacomylus sladei	151.65	PBDB + regression
Chadrolagus emryi	63.45	PBDB + regression
Chalicomomys willwoodensis	4.06	PBDB + regression
Chasmaporthetes ossifragus	107821.85	PBDB + regression
Chipetaia lamporea	417.80	PBDB + regression
Chiromyoides caesor	219.44	PBDB + regression
Chiromyoides minor	70.04	PBDB + regression
Chiromyoides potior	202.80	PBDB + regression
Chriacus badgleyi	1200.76	PBDB + regression
Chriacus baldwini	1550.07	PBDB + regression

Continued on next page

Table A.1 – continued from previous page

Species	Mass (g)	Source
Chriacus gallinae	1468.97	PBDB + regression
Chriacus pelvidens	5047.26	PBDB + regression
Chumashius balchi	111.45	PBDB + regression
Churcheria baroni	50.76	PBDB + regression
Cimexomys minor	27.99	PBDB + regression
Cimolestes incisus	204.60	PBDB + regression
Colodon cingulatus	222816.55	PBDB
Colodon kayi	49060.48	PBDB + regression
Colodon occidentalis	110573.45	PBDB + regression
Colodon stovalli	57498.66	PBDB + regression
Colodon woodi	85748.60	PBDB + regression
Conacodon cophater	236.68	PBDB + regression
Conacodon delphae	3318.67	PBDB + regression
Conacodon entoconus	2268.93	PBDB + regression
Conacodon kohlbergeri	590.00	PBDB + regression
Conoryctes comma	16724.87	PBDB + regression
Copecion brachypternus	4501.11	PBDB + regression
Copecion davisi	2798.04	PBDB + regression
Copedelphys innominata	23.70	PBDB + regression
Copedelphys titanelix	6.11	PBDB + regression
Copelemur australotutus	1633.14	PBDB + regression
Copelemur tutus	1864.20	PBDB + regression
Copemys barstowensis	32.14	[329]
Copemys esmeraldensis	27.94	[329]
Copemys lindsayi	14.88	[329]

Continued on next page

Table A.1 – continued from previous page

Species	Mass (g)	Source
Copemys longidens	26.84	[329]
Copemys loxodon	28.79	[329]
Copemys mariae	31.50	[329]
Copemys pagei	15.18	[329]
Copemys russelli	24.29	[329]
Copemys shotwelli	10.39	PBDB + regression
Copemys tenuis	23.34	[329]
Coriphagus encinensis	254.29	PBDB + regression
Coriphagus montanus	130.39	PBDB + regression
Cormocyon copei	4817.45	[329]
Cormocyon haydeni	4359.01	[329]
Coryphodon armatus	351055.71	PBDB + regression
Coryphodon eocaenus	218518.57	PBDB + regression
Coryphodon lobatus	570461.25	PBDB + regression
Coryphodon proterus	479331.19	PBDB + regression
Coryphodon radians	339220.51	PBDB + regression
Cosoryx cerroensis	16814.55	[329]
Cosoryx furcatus	13678.75	PBDB
Cranioceras clarendonensis	89936.76	PBDB + regression
Cranioceras teres	96761.07	[329]
Cranioceras unicornis	134591.56	[329]
Cratogeomys sansimonensis	140.29	PBDB + regression
Crucimys milleri	36.60	[329]
Crypholestes vaughni	18.68	PBDB
Cryptotis adamsi	13.33	[329]

Continued on next page

Table A.1 – continued from previous page

Species	Mass (g)	Source
Cryptotis kansasensis	33.60	PBDB + regression
Cryptotis parva	4.10	PBDB
Cupidinimus avawatzensis	20.09	[329]
Cupidinimus bidahochiensis	27.39	[329]
Cupidinimus boronensis	14.59	[329]
Cupidinimus eurekensis	2.98	PBDB
Cupidinimus halli	10.39	PBDB + regression
Cupidinimus lindsayi	7.61	[329]
Cupidinimus madisonensis	10.56	PBDB
Cupidinimus magnus	22.41	PBDB + regression
Cupidinimus nebraskensis	9.30	[329]
Cupidinimus prattensis	21.33	[329]
Cupidinimus tertius	16.44	[329]
Cupidinimus whitlocki	16.78	[329]
Cuvieronius tropicus	56818220.68	PBDB + regression
Cuyamalagus dawsoni	639.06	[329]
Cylindrodon fontis	88.57	PBDB + regression
Cylindrodon nebraskensis	167.14	PBDB + regression
Cynarctoides acridens	2921.93	[329]
Cynarctoides gawnae	2643.87	[329]
Cynarctoides harlowi	1863.11	[329]
Cynarctoides lemur	2321.57	[329]
Cynarctoides luskensis	2392.27	[329]
Cynarctoides roii	1826.21	[329]
Cynarctus crucidens	4964.16	[329]

Continued on next page

Table A.1 – continued from previous page

Species	Mass (g)	Source
Cynarctus galushai	9228.02	[329]
Cynarctus saxatilis	10097.06	[329]
Cynelos caroniavorus	16647.24	[329]
Cynelos idoneus	105873.47	[329]
Cynelos sinapius	213202.99	[329]
Cynodesmus martini	14185.85	[329]
Cynodesmus thooides	9228.02	[329]
Cynorca occidentale	20537.34	[329]
Cynorca sociale	34624.09	PBDB + regression
Cyriacotherium psamminum	3551.95	PBDB
Dakotallomys lillegraveni	196.53	PBDB
Dakotallomys pelycomyoides	244.01	[202]
Daphoenodon falkenbachi	137310.49	[329]
Daphoenodon notionastes	43477.55	[329]
Daphoenodon superbus	77652.58	[329]
Daphoenus hartshornianus	13329.17	PBDB + regression
Daphoenus lambei	7397.01	PBDB + regression
Daphoenus ruber	12767.60	PBDB + regression
Daphoenus socialis	13000.00	[198]
Daphoenus vetus	19535.72	[329]
Dartonius jepseni	10.52	[119]
Dasypus bellus	44977.99	[305]
Delahomeryx browni	38948.67	[329]
Desmatippus avus	179871.86	[329]
Desmatippus texanus	102452.28	PBDB

Continued on next page

Table A.1 – continued from previous page

Species	Mass (g)	Source
Desmatochoerus hesperus	239142.20	[197]
Desmatochoerus megalodon	335100.00	[198]
Desmatoclaenus hermaeus	11400.50	PBDB + regression
Desmatolagus schizopetrus	198.46	PBDB
Desmocyon matthewi	8103.08	[329]
Desmocyon thomsoni	6974.39	[329]
Diacocherus meizon	23.70	PBDB + regression
Diacocherus minutus	45.48	PBDB + regression
Diacodexis gracilis	788.63	PBDB + regression
Diacodexis kelleyi	1550.07	PBDB + regression
Diacodexis metsiacus	1176.50	PBDB + regression
Diacodexis minutus	555.14	PBDB + regression
Diacodexis primus	1389.84	PBDB + regression
Diacodexis secans	2124.77	PBDB + regression
Diacodon alticuspis	324.82	PBDB + regression
Diceratherium annectens	864580.76	[329]
Diceratherium armatum	3541284.24	[329]
Diceratherium gregorii	946518.90	PBDB + regression
Diceratherium niobrarense	2105366.25	[329]
Diceratherium tridactylum	965112.54	[329]
Didelphodus absarokae	361.10	PBDB + regression
Didelphodus altidens	246.37	PBDB + regression
Didelphodus rheos	139.74	PBDB + regression
Didelphodus serus	133.48	PBDB + regression
Didymictis altidens	30566.06	PBDB

Continued on next page

Table A.1 – continued from previous page

Species	Mass (g)	Source
Didymictis leptomylus	3213.72	PBDB + regression
Didymictis protenus	7205.96	PBDB + regression
Didymictis proteus	3605.48	PBDB + regression
Didymictis vancleveae	65421.88	[276]
Dikkomys matthewi	47.47	[329]
Dillerlemur pagei	133.09	PBDB + regression
Dilophodon destitutus	369866.51	[300]
Dilophodon minusculus	27074.75	PBDB + regression
Dinaelurus crassus	37320.00	[198]
Dinofelis palaeoonca	70032.50	PBDB + regression
Dinohippus interpolatus	257815.63	[329]
Dinohippus leardi	392385.48	[329]
Dinohippus leidyanus	229900.00	[188]
Dinohippus mexicanus	609259.77	[329]
Dinohippus spectans	536500.00	[198]
Dinohyus hollandi	2032979.00	PBDB + regression
Diplobunops matthewi	100751.34	PBDB + regression
Dipodomys compactus	49.20	[305]
Dipodomys gidleyi	17.97	PBDB + regression
Dipodomys hibbardi	17.81	[329]
Diprionomys agrarius	29.87	PBDB + regression
Diprionomys parvus	11.00	[198]
Dipsalidictis aequidens	42263.45	PBDB + regression
Dipsalidictis platypus	8936.27	PBDB + regression
Dipsalidictis transiens	22310.00	PBDB + regression

Continued on next page

Table A.1 – continued from previous page

Species	Mass (g)	Source
Dipsalodon churchillorum	26647.76	PBDB + regression
Dipsalodon matthewi	66505.41	PBDB + regression
Dissacus navajovius	6299.67	PBDB + regression
Dissacus praenuntius	14804.22	PBDB + regression
Domnina dakotensis	28.50	[329]
Domnina gradata	36.60	[329]
Domnina greeni	33.78	[329]
Domnina thompsoni	20.28	PBDB + regression
Domninoides hessei	149.90	[329]
Domninoides knoxjonesi	31.40	PBDB
Domninoides mimicus	135.64	[329]
Domninoides riparensis	56.83	[329]
Dorraletes diminutivus	37.32	PBDB + regression
Douglassciurus jeffersoni	360.63	[202]
Downsimus chadwicki	56.26	[329]
Drepanomeryx falciformis	90219.42	[329]
Dromomeryx borealis	144350.55	[329]
Dryomomys dulcifer	4.58	[215]
Dryomomys szalayi	6.07	[215]
Duchesneodus uintensis	2406372.12	PBDB + regression
Dyseohyus fricki	21807.30	[329]
Dyseolemur pacificus	70.04	PBDB + regression
Ectocion cedrus	4839.14	PBDB + regression
Ectocion collinus	5869.11	PBDB + regression
Ectocion major	12433.55	PBDB + regression

Continued on next page

Table A.1 – continued from previous page

Species	Mass (g)	Source
Ectocion mediotuber	5451.18	PBDB + regression
Ectocion osbornianus	6800.01	PBDB + regression
Ectocion parvus	3109.34	PBDB + regression
Ectocion superstes	13800.02	PBDB + regression
Ectoconus ditrigonus	22355.77	PBDB + regression
Ectoganus gliriformes	51205.04	[215]
Ectoganus gliriformis	51205.04	PBDB + regression
Ectopocynus antiquus	8266.78	[329]
Ectopocynus intermedius	12456.53	[329]
Ectypodus aphronorus	18.00	[362]
Ectypodus childei	14.42	PBDB + regression
Ectypodus lovei	14.42	PBDB + regression
Ectypodus musculus	30.25	PBDB + regression
Ectypodus powelli	19.74	PBDB + regression
Ectypodus szalayi	17.88	PBDB + regression
Ectypodus tardus	16.11	PBDB + regression
Edworthia lerbekmoi	23.70	PBDB + regression
Ekgmowechashala philotau	2079.74	[329]
Elomeryx armatus	157944.66	[329]
Elphidotarsius florencae	17.88	PBDB + regression
Elphidotarsius russelli	23.70	PBDB + regression
Elphidotarsius wightoni	27.99	PBDB + regression
Elpidophorus elegans	829.40	PBDB + regression
Elpidophorus minor	151.65	PBDB + regression
Elwynella oreas	99.30	PBDB + regression

Continued on next page

Table A.1 – continued from previous page

Species	Mass (g)	Source
Elymys complexus	5.92	PBDB + regression
Enhydritherium terraenovae	22180.76	PBDB + regression
Enhydrocyon basilatus	20332.99	[329]
Enhydrocyon crassidens	18582.95	[329]
Enhydrocyon pahinsintewakpa	14764.78	[329]
Enhydrocyon stenocephalus	14044.69	[329]
Entoptychus grandiplanus	59.74	[329]
Entoptychus individens	84.00	[198]
Entoptychus planifrons	134.29	[329]
Entoptychus sheppardi	94.63	[329]
Entoptychus wheelerensis	84.00	[198]
Eoconodon gaudrianus	3093.31	[371]
Eoconodon hutchisoni	9450.73	PBDB
Eoconodon nidhoggi	1665.06	PBDB + regression
Eohaplomys matutinus	1536.44	PBDB + regression
Eohaplomys serus	953.63	PBDB + regression
Eohaplomys tradux	772.04	PBDB + regression
Eomoropus amarorum	67651.40	PBDB + regression
Eoryctes melanus	53.67	PBDB + regression
Eotitanops borealis	240072.90	PBDB + regression
Eotitanops minimus	43277.45	PBDB + regression
Eotitanotherium osborni	1484895.52	PBDB
Eotylopus reedi	38563.77	PBDB + regression
Epeiromys spanius	101.65	PBDB + regression
Epicyon haydeni	41772.77	[329]

Continued on next page

Table A.1 – continued from previous page

Species	Mass (g)	Source
Epicyon saevus	27722.51	[329]
Epihippus gracilis	18366.29	PBDB + regression
Epihippus intermedius	21835.02	PBDB
Epitriplopus uintensis	32600.00	[188]
Eporeodon occidentalis	118300.00	[198]
Equus complicatus	399944.75	[305]
Equus conversidens	306196.34	[305]
Equus cumminsii	314354.39	PBDB + regression
Equus francisci	217178.22	PBDB + regression
Equus fromanius	172311.00	[198]
Equus giganteus	399944.75	[305]
Equus idahoensis	629100.30	PBDB + regression
Equus leidyi	291150.23	PBDB + regression
Equus occidentalis	574116.46	[305]
Equus scotti	554625.71	[305]
Equus simplicidens	296558.57	[329]
Eremotherium eomigrans	2584400.00	[194]
Eremotherium laurillardi	799834.26	[305]
Erethizon bathygnathum	9163.63	PBDB + regression
Erethizon kleini	3854.02	PBDB + regression
Escavadodon zygus	137.63	PBDB + regression
Esthonyx acutidens	7159.09	PBDB + regression
Esthonyx bisulcatus	3066.57	PBDB + regression
Esthonyx spatularius	4505.96	PBDB + regression
Eucyon davisi	10509.13	[329]

Continued on next page

Table A.1 – continued from previous page

Species	Mass (g)	Source
Eudaemonema cuspidata	529.62	PBDB + regression
Eumys brachyodus	126.47	[329]
Eumys elegans	106.19	PBDB + regression
Euoplocyon brachygnathus	11271.13	[329]
Euoplocyon spissidens	9798.65	[329]
Eusmilus cerebralis	804.32	[329]
Eusmilus sicarius	34789.94	PBDB + regression
Eutypomys acares	84.39	PBDB + regression
Eutypomys hibernodus	729.27	PBDB + regression
Eutypomys inexpectatus	708.27	PBDB + regression
Eutypomys montanensis	943.88	[329]
Eutypomys obliquidens	202.01	PBDB + regression
Eutypomys parvus	273.26	PBDB + regression
Eutypomys thomsoni	578.37	PBDB + regression
Fanimus clasoni	87.14	PBDB
Fanimus ultimus	164.02	[329]
Felis rexroadensis	30333.26	[329]
Ferinestrix vorax	29333.85	PBDB
Florentiamys agnewi	84.77	[329]
Florentiamys kinseyi	157.59	[329]
Florentiamys loomisi	151.41	[329]
Florentiamys tiptoni	117.92	[329]
Floridachoerus olseni	35596.41	[329]
Floridameryx floridanus	8184.52	[329]
Floridatragulus dolichanthereus	43477.55	[329]

Continued on next page

Table A.1 – continued from previous page

Species	Mass (g)	Source
Floridatragulus texanus	61083.68	[329]
Fouchia elyensis	10126.74	PBDB + regression
Galbreathia bettae	165.67	[329]
Galbreathia novellus	271.54	PBDB
Galecyon mordax	1984.90	PBDB + regression
Gazinius amplus	875.00	[312]
Gazinocyon vulpeculus	5070.63	PBDB + regression
Gelastops joni	183.10	PBDB + regression
Gelastops parcus	149.35	PBDB + regression
Geomys carranzai	29.87	PBDB + regression
Geringia gloveri	26.84	[329]
Geringia mcgregori	41.26	[329]
Gigantocamelus spatulus	2615036.35	PBDB + regression
Glossotherium chapadmalense	310540.00	[194]
Glyptotherium arizonae	789680.00	[194]
Gomphotherium obscurum	50047676.55	PBDB
Gomphotherium osborni	24506452.87	[354]
Goniodontomys disjunctus	44.26	[329]
Gracilocyon winkleri	690.80	PBDB + regression
Grangeria anarsius	65434.40	PBDB + regression
Gregorymys curtus	106.70	[329]
Gregorymys formosus	77.48	[329]
Gregorymys riograndensis	43.82	[329]
Gripholagomys lavocati	507.76	[329]
Griphomys alecer	15.90	PBDB + regression

Continued on next page

Table A.1 – continued from previous page

Species	Mass (g)	Source
Griphomys toltecus	25.61	[202]
Guanajuatomys hibbardi	167.14	PBDB + regression
Guildayomys hibbardi	82.27	[329]
Hapalodectes anthracinus	352.26	PBDB + regression
Hapalodectes leptognathus	617.19	PBDB + regression
Haplaletes disceptatrix	84.86	PBDB + regression
Haplaletes pelicatus	130.39	PBDB + regression
Haploconus angustus	1200.76	PBDB + regression
Haplohippus texanus	23265.89	PBDB + regression
Haplolambda quinni	53092.09	PBDB
Haplolambda simpsoni	138271.36	[328]
Haplomylus bozemanensis	87.48	PBDB + regression
Haplomylus palustris	72.22	PBDB + regression
Haplomylus scottianus	115.36	PBDB + regression
Haplomylus simpsoni	162.53	PBDB + regression
Haplomylus speirianus	98.28	PBDB + regression
Haplomys galbreathi	38.23	PBDB + regression
Haplomys liolophus	39.00	[198]
Harpagolestes leotensis	138880.62	PBDB + regression
Harpagolestes uintensis	106686.23	PBDB + regression
Harrymys irvini	83.93	[329]
Harrymys magnus	50.40	[329]
Harrymys woodi	43.52	[159]
Heliscomys hatcheri	7.19	[362]
Heliscomys ostranderi	3.56	PBDB + regression

Continued on next page

Table A.1 – continued from previous page

Species	Mass (g)	Source
Heliscomys senex	4.67	PBDB + regression
Heliscomys vetus	4.67	PBDB + regression
Heliscomys woodi	10.39	PBDB + regression
Hemiacodon engardae	538.07	PBDB + regression
Hemiacodon gracilis	563.76	PBDB + regression
Hemiauchenia gracilis	157598.86	PBDB + regression
Hemiauchenia macrocephala	109900.58	[305]
Hemiauchenia minima	85819.37	[329]
Hemiauchenia vera	243143.10	PBDB + regression
Hemipsalodon grandis	436751.47	[280]
Hemithlaeus harbourae	1283.98	[180]
Hendryomeryx defordi	5742.28	PBDB
Hendryomeryx esulcatus	6619.36	PBDB + regression
Hendryomeryx wilsoni	3001.58	PBDB + regression
Heptacodon pellionis	88438.62	PBDB + regression
Heptodon calciculus	26551.76	PBDB + regression
Herpetotherium fugax	19.74	PBDB + regression
Herpetotherium knighti	51.17	PBDB + regression
Herpetotherium merriami	25.80	PBDB + regression
Herpetotherium valens	57.16	PBDB + regression
Herpetotherium youngi	40.09	PBDB + regression
Hesperhys pinensis	68186.37	[329]
Hesperhys vagrans	73865.41	[329]
Hesperocyon gregarius	3533.34	[329]
Hesperolagomys fluviatilis	169.02	[329]

Continued on next page

Table A.1 – continued from previous page

Species	Mass (g)	Source
Hesperolagomys galbreathi	149.90	[329]
Hesperoscalops mcgrewi	186.62	PBDB + regression
Heteraletes leotanus	14016.35	PBDB + regression
Heteromeryx dispar	51072.86	PBDB + regression
Heteropliohippus hulberti	302549.45	[329]
Hexacodus pelodes	625.82	PBDB + regression
Hexameryx simpsoni	30638.11	[329]
Hexobelomeryx fricki	59621.29	PBDB + regression
Hibbarderix obfuscatus	33.60	PBDB + regression
Hibbardomys fayae	61.14	PBDB + regression
Hibbardomys marthae	84.39	PBDB + regression
Hibbardomys skinneri	61.14	PBDB + regression
Hibbardomys voorhiesi	61.14	PBDB + regression
Hipparion forcei	194852.86	[329]
Hipparion tehonense	100709.96	[329]
Hippotherium emsliei	201622.66	PBDB + regression
Hippotherium ingenuum	109258.44	PBDB + regression
Hippotherium plicatile	165091.29	PBDB + regression
Hippotherium quinni	422193.86	PBDB + regression
Hitonkala macdonaldtau	24.53	[329]
Holmesina floridanus	68696.59	PBDB + regression
Homacodon vagans	6691.69	PBDB + regression
Homogalax protapirinus	19815.58	PBDB + regression
Homotherium crusafonti	233968.19	PBDB
Homotherium idahoensis	257900.00	[198]

Continued on next page

Table A.1 – continued from previous page

Species	Mass (g)	Source
Homotherium johnstoni	124571.99	PBDB + regression
Hoplophoneus mentalis	48420.73	PBDB + regression
Hoplophoneus primaevus	20243.49	PBDB + regression
Huerfanodon polecatensis	29749.42	PBDB + regression
Huerfanodon torrejonius	20823.53	PBDB + regression
Hyaenodon brevirostrus	37331.34	PBDB + regression
Hyaenodon crucians	14771.98	PBDB + regression
Hyaenodon horridus	91750.41	PBDB + regression
Hyaenodon montanus	28357.49	PBDB + regression
Hyaenodon mustelinus	8079.01	PBDB + regression
Hyaenodon raineyi	1093.34	PBDB + regression
Hyaenodon venturae	2552.88	PBDB + regression
Hyaenodon vetus	33333.28	PBDB + regression
Hylomeryx quadricuspis	4966.19	PBDB + regression
Hyopsodus fastigatus	1894.05	PBDB + regression
Hyopsodus lepidus	625.82	PBDB + regression
Hyopsodus loomisi	488.31	PBDB + regression
Hyopsodus lysitensis	924.51	PBDB + regression
Hyopsodus mentalis	1482.35	PBDB + regression
Hyopsodus minor	440.75	PBDB + regression
Hyopsodus minusculus	440.75	PBDB + regression
Hyopsodus paulus	768.59	PBDB + regression
Hyopsodus pauxillus	181.54	PBDB + regression
Hyopsodus powellianus	1416.00	PBDB + regression
Hyopsodus simplex	729.22	PBDB + regression

Continued on next page

Table A.1 – continued from previous page

Species	Mass (g)	Source
Hyopsodus tonksi	464.26	PBDB + regression
Hyopsodus uintensis	1495.78	PBDB + regression
Hyopsodus walcottianus	6172.52	PBDB + regression
Hyopsodus wortmani	798.73	PBDB + regression
Hypertragulus calcaratus	8880.00	[198]
Hypertragulus hesperius	4572.00	[198]
Hypohippus affinis	442413.39	[329]
Hypohippus equinus	271034.12	[329]
Hypohippus osborni	299539.03	[329]
Hypolagus edensis	665.14	[329]
Hypolagus fontinalis	1211.97	[329]
Hypolagus furlongi	678.58	[329]
Hypolagus gidleyi	1998.20	[329]
Hypolagus oregonensis	2344.90	[329]
Hypolagus parviplicatus	1702.75	[329]
Hypolagus ringoldensis	319.00	[198]
Hypolagus vetus	2892.86	[329]
Hypolagus voorhiesi	319.00	[198]
Hypsiops bannackensis	146895.45	[355]
Hypsiops breviceps	156373.08	[329]
Hyrachyus affinis	111138.67	PBDB + regression
Hyrachyus eximius	198513.08	PBDB + regression
Hyrachyus modestus	217770.66	PBDB + regression
Hyracodon leidyanus	459564.25	PBDB + regression
Hyracodon medius	102488.96	PBDB + regression

Continued on next page

Table A.1 – continued from previous page

Species	Mass (g)	Source
Hyracodon nebraskensis	211081.59	[329]
Hyracodon petersoni	313951.52	PBDB + regression
Hyracodon priscidens	351463.36	PBDB + regression
Hyracotherium cristatum	23898.63	PBDB
Hyracotherium vasacciense	24900.00	[188]
Ictidopappus mustelinus	354.83	PBDB + regression
Ignacius fremontensis	19.74	PBDB + regression
Ignacius frugivorus	73.44	PBDB + regression
Ignacius graybullianus	128.62	PBDB + regression
Indarctos nevadensis	376215.79	PBDB + regression
Indarctos oregonensis	302549.45	[329]
Ischyrocyon gidleyi	282095.23	[329]
Ischyromys typus	761.25	PBDB + regression
Ischyromys veterior	477.49	PBDB + regression
Jacobsomys verdensis	50.91	[329]
Janimus dawsonae	17.97	PBDB + regression
Jaywilsonomys ojinagaensis	895.10	PBDB + regression
Jemezius szalayi	95.39	PBDB + regression
Jepsenella praepropera	18.80	PBDB + regression
Jimomys labaughi	69.41	[329]
Jimomys lulli	28.00	[198]
Kansasimys dubius	135.64	[329]
Kansasimys wilsoni	228.43	[193]
Kimbetohia mziae	72.00	[362]
Knightomys cremneus	27.28	PBDB + regression

Continued on next page

Table A.1 – continued from previous page

Species	Mass (g)	Source
Knightomys cuspidatus	48.29	PBDB
Knightomys depressus	64.79	PBDB + regression
Knightomys huerfanensis	135.17	PBDB + regression
Knightomys minor	27.28	PBDB + regression
Knightomys reginensis	10.39	PBDB + regression
Knightomys senior	44.30	PBDB + regression
Kolponomos clallamensis	103303.51	[330]
Kolponomos newportensis	876418.72	[278]
Kyptoceras amatorum	369534.73	[329]
Labidolemur kayi	33.60	PBDB + regression
Labidolemur serus	67.38	PBDB + regression
Labidolemur soricoides	14.63	PBDB + regression
Lambdotherium popoagicum	29620.91	PBDB + regression
Lambertocyon eximius	6045.26	PBDB + regression
Lambertocyon ischyrus	2015.61	PBDB + regression
Lanthanotherium sawini	90.14	PBDB + regression
Laredochoerus edwardsi	5523.91	PBDB + regression
Laredomys riograndensis	5.92	PBDB + regression
Leidymys cerasus	41.68	[329]
Leidymys korthi	35.35	PBDB + regression
Leipsanolestes siegfriedti	26.62	PBDB + regression
Lemoynea biradicularis	45.60	[329]
Lepoides lepoides	3640.95	[329]
Leptacodon munusculum	6.56	PBDB + regression
Leptacodon packi	14.63	PBDB + regression

Continued on next page

Table A.1 – continued from previous page

Species	Mass (g)	Source
Leptacodon tener	8.61	PBDB + regression
Leptarctus mummorum	8673.78	PBDB + regression
Leptarctus oregonensis	2580.00	[198]
Leptarctus primus	149.90	[329]
Leptarctus supremus	7725.15	PBDB
Leptauchenia decora	20130.67	[329]
Leptauchenia major	30946.03	[329]
Leptictis dakotensis	644.52	PBDB + regression
Leptochoerus elegans	8623.90	PBDB + regression
Leptochoerus spectabilis	9021.94	PBDB + regression
Leptochoerus supremus	19397.68	PBDB + regression
Leptocyon mollis	3300.00	[198]
Leptocyon vafer	5377.61	[329]
Leptodontomys douglassi	4.67	PBDB + regression
Leptodontomys stirtoni	7.32	[323]
Leptolambda schmidti	90362.43	[369]
Leptomeryx blacki	4578.72	PBDB + regression
Leptomeryx elissae	6511.38	PBDB + regression
Leptomeryx evansi	9426.05	PBDB + regression
Leptomeryx mammifer	18828.04	PBDB + regression
Leptomeryx speciosus	9223.24	PBDB + regression
Leptomeryx yoderi	10716.49	PBDB + regression
Leptoreodon edwardsi	13277.57	PBDB + regression
Leptoreodon leptolophus	9223.24	PBDB + regression
Leptoreodon major	32615.84	PBDB + regression

Continued on next page

Table A.1 – continued from previous page

Species	Mass (g)	Source
Leptoreodon marshi	20660.42	PBDB + regression
Leptoreodon pusillus	3957.21	PBDB + regression
Leptoreodon stocki	20980.32	PBDB + regression
Leptoromys wilsoni	110.81	PBDB + regression
Leptotomus caryophilus	872.14	PBDB + regression
Leptotomus leptodus	2852.74	PBDB + regression
Leptotomus parvus	941.80	PBDB + regression
Leptotragulus clarki	5364.24	PBDB + regression
Leptotragulus medius	13323.34	PBDB + regression
Leptotragulus proavus	11797.49	PBDB + regression
Lepus californicus	2288.00	PBDB
Lignimus austridakotensis	14.88	[329]
Lignimus montis	38.23	PBDB + regression
Limaconyssus habrus	18.80	PBDB + regression
Limnoecus niobrarensis	7.17	[329]
Limnoecus tricuspis	5.10	[329]
Liodontia alexandrae	198.34	[329]
Litaletes disjunctus	681.34	PBDB + regression
Litocherus lacunatus	84.86	PBDB + regression
Litocherus notissimus	37.32	PBDB + regression
Litocherus zygeus	55.87	PBDB + regression
Litolagus molidens	103.28	PBDB + regression
Litolestes ignotus	18.80	PBDB + regression
Litomylus dissentaneus	124.29	PBDB + regression
Litomylus orthronepius	115.36	PBDB + regression

Continued on next page

Table A.1 – continued from previous page

Species	Mass (g)	Source
Litoyoderimys auogoleus	47.48	PBDB + regression
Longirostromeryx clarendonensis	13226.80	[329]
Longirostromeryx wellsi	17500.77	[329]
Lophiparamys debequensis	29.87	PBDB + regression
Lophiparamys murinus	88.57	PBDB + regression
Loveina minuta	45.48	PBDB + regression
Loveina zephyri	66.71	PBDB + regression
Loxolophus criswelli	2140.57	PBDB + regression
Loxolophus hyattianus	1984.90	PBDB + regression
Loxolophus pentacus	14847.99	PBDB + regression
Loxolophus priscus	3899.92	PBDB + regression
Loxolophus schizophrenus	881.66	PBDB + regression
Loxolophus spiekeri	18323.92	PBDB + regression
Lutravus halli	454.86	[329]
Lycophocyon hutchisoni	4865.03	PBDB + regression
Lynx proterolyncis	15677.78	[329]
Lynx rufus	10482.00	PBDB
Machaeromeryx gilchristensis	4536.90	[329]
Machaeromeryx tragulus	4536.90	[329]
Macrocranion junnei	15.97	PBDB + regression
Macrocranion nitens	33.60	PBDB + regression
Macrogenis crassigenis	106737.50	PBDB
Macrognathomys gemmacolis	5.92	PBDB + regression
Macrognathomys nanus	5.16	[329]
Macrotarsius montanus	1152.46	PBDB + regression

Continued on next page

Table A.1 – continued from previous page

Species	Mass (g)	Source
Macrotarsius roederi	1640.00	[312]
Macrotarsius siegerti	1188.60	PBDB + regression
Mahgarita stevensi	358.28	[365]
Malaquiferus tourteloti	14816.10	PBDB + regression
Mammacyon obtusidens	70262.96	[329]
Mammut americanum	4528975.80	[305]
Mammut cosoensis	29076021.42	[281]
Mammut furlongi	29535033.05	PBDB + regression
Mammuthus columbi	7998342.55	[305]
Mammut raki	41795334.69	PBDB + regression
Marmota korthi	2038.56	[329]
Marmota vetus	658.52	[329]
Martes gazini	1767.00	[198]
Martes parviloba	5831.67	[186]
Martinogale alveodens	17.99	[329]
Mattimys kalicola	32.56	PBDB + regression
Megacamelus merriami	1905014.16	[329]
Megadelphus lundeliusi	3614.01	PBDB + regression
Megahippus matthewi	882046.45	[329]
Megalagus abaconis	1118.79	[329]
Megalagus brachyodon	373.57	PBDB + regression
Megalagus primitivus	3165.29	[329]
Megalagus turgidus	3071.74	[329]
Megalesthonyx hopsoni	15975.00	[118]
Megalictis ferox	1587.63	[329]

Continued on next page

Table A.1 – continued from previous page

Species	Mass (g)	Source
Megalictis frazieri	33591.60	PBDB
Megalonyx curvidens	185050.00	[194]
Megalonyx leptostomus	279404.72	PBDB + regression
Megalonyx wheatleyi	252445.15	PBDB + regression
Megantereon hesperus	67507.91	[329]
Megapeomys bobwilsoni	411.58	[329]
Megasminthus tiheni	55.70	[329]
Megatylopus cochrani	884486.95	PBDB + regression
Megatylopus matthewi	1383324.20	[329]
Megatylopus primaevus	884295.95	PBDB + regression
Meniscomys hippodus	69.41	[329]
Meniscomys uhtoffi	70.00	[198]
Meniscotherium chamense	17413.77	PBDB + regression
Meniscotherium tapiacitum	1864.20	PBDB + regression
Menoceras arikarense	597195.61	[329]
Menoceras barbouri	1251683.50	[329]
Menops bakeri	3993312.11	PBDB
Mentoclaenodon acrogenius	37446.92	PBDB + regression
Mephitis mephitis	2055.00	PBDB
Merriamoceros coronatus	15214.44	[329]
Merychippus brevidontus	124243.67	[329]
Merychippus calamarius	196811.17	[329]
Merychippus californicus	86681.87	[329]
Merychippus goorisi	222888.24	PBDB + regression
Merychippus gunteri	92041.97	[329]

Continued on next page

Table A.1 – continued from previous page

Species	Mass (g)	Source
Merychippus insignis	125492.34	[329]
Merychippus primus	95798.28	[329]
Merychippus relictus	60471.00	[198]
Merychippus sejunctus	75357.60	[329]
Merychyus arenarum	52892.75	PBDB + regression
Merychyus crabilli	6038.43	[154]
Merychyus elegans	45706.69	[329]
Merychyus medius	108012.26	[329]
Merychyus minimus	29620.91	PBDB + regression
Merychyus relictus	20923.07	PBDB
Merychyus smithi	135340.58	PBDB + regression
Merycochoerus carrikeri	313334.13	[265]
Merycochoerus chelydra	289450.77	PBDB
Merycochoerus magnus	331041.82	[329]
Merycochoerus matthewi	59736.09	PBDB + regression
Merycochoerus proprius	265667.29	[329]
Merycochoerus superbus	438400.00	[198]
Merycodus prodromus	8738.23	[157]
Merycodus sabulonis	10509.13	[329]
Merycodus warreni	12597.43	[19]
Merycoides harrisonensis	90219.42	[329]
Merycoides longiceps	56672.65	[324]
Merycoides pariogonus	77000.00	[198]
Merycoidodon bullatus	109500.00	[198]
Merycoidodon culbertsoni	116748.58	PBDB + regression

Continued on next page

Table A.1 – continued from previous page

Species	Mass (g)	Source
Merycoidodon major	62200.00	[198]
Mescalerolemur horneri	193.21	[185]
Mesocyon brachyops	7942.63	[329]
Mesocyon coryphaeus	10198.54	[329]
Mesocyon temnodon	7186.79	[329]
Mesodma ambigua	115.64	PBDB + regression
Mesodma formosa	32.59	PBDB + regression
Mesodma garfieldensis	37.51	PBDB + regression
Mesodma pygmaea	8.53	PBDB + regression
Mesodma thompsoni	54.13	PBDB + regression
Mesogaulus paniensis	926.80	PBDB
Mesohippus bairdi	39223.29	PBDB + regression
Mesohippus exoletus	47569.37	PBDB + regression
Mesohippus texanus	35643.43	PBDB + regression
Mesohippus westoni	39355.66	PBDB + regression
Mesonyx obtusidens	147850.73	PBDB
Mesoreodon chelonyx	133078.52	[24]
Mesoreodon floridensis	104820.01	[329]
Mesoreodon minor	159700.00	[198]
Mesoscalops montanensis	85.63	[329]
Mesoscalops scopelotemos	95.58	[329]
Metadjidaumo hendryi	10.38	[329]
Metaliomys sevierensis	15.90	PBDB + regression
Metanoiamys agorus	12.11	PBDB + regression
Metanoiamys fantasma	20.14	PBDB + regression

Continued on next page

Table A.1 – continued from previous page

Species	Mass (g)	Source
Metanoiamys korthi	13.95	PBDB + regression
Metanoiamys lacus	10.39	PBDB + regression
Metanoiamys marinus	8.78	PBDB + regression
Metarhinus pater	1098772.19	PBDB
Metatomarctus canavus	10938.02	[329]
Metechinus amplior	567.90	[328]
Miacis deutschi	968.28	PBDB + regression
Miacis exiguus	1775.93	PBDB + regression
Miacis hookwayi	2350.87	[221]
Miacis latidens	2236.53	PBDB + regression
Miacis parvivorus	2384.02	PBDB + regression
Miacis petilus	1591.36	PBDB + regression
Michenia agatensis	99707.88	[329]
Michenia exilis	65512.75	[329]
Microcosmodon conus	16.11	PBDB + regression
Microcosmodon rosei	27.99	PBDB + regression
Microeutypomys karenae	12.11	PBDB + regression
Microeutypomys tilliei	22.41	PBDB + regression
Micromomys antelucanus	5.04	PBDB
Microparamys cheradius	57.59	PBDB + regression
Microparamys dubius	20.14	PBDB + regression
Microparamys hunterae	14.33	PBDB
Microparamys minutus	24.79	PBDB + regression
Microparamys nimius	29.87	PBDB + regression
Microparamys sambucus	41.22	PBDB + regression

Continued on next page

Table A.1 – continued from previous page

Species	Mass (g)	Source
Microparamys tricus	57.59	PBDB + regression
Microparamys woodi	32.56	PBDB + regression
Micropternodus borealis	41.19	PBDB + regression
Micropternodus montrosensis	103.85	PBDB + regression
Micropternodus morgani	72.24	[329]
Microsus cuspidatus	2597.76	PBDB + regression
Microsyops angustidens	352.30	PBDB + regression
Microsyops annectens	1522.82	PBDB + regression
Microsyops cardiorestes	186.75	PBDB + regression
Microsyops elegans	738.98	PBDB + regression
Microsyops knightensis	521.24	PBDB + regression
Microsyops kratos	1455.64	PBDB + regression
Microsyops latidens	472.21	PBDB + regression
Microsyops scottianus	634.92	PBDB + regression
Microtomarctus conferta	8866.19	[329]
Mictomys vetus	76.28	PBDB + regression
Mimatuta morgoth	512.92	PBDB + regression
Mimetodon silberlingi	21.68	PBDB + regression
Mimomys mcknighti	110.70	PBDB
Mimomys panacaensis	84.39	PBDB + regression
Mimomys parvus	99.10	PBDB
Mimomys primus	163.60	PBDB
Mimomys taylori	80.90	PBDB
Mimoperadectes labrus	285.50	PBDB + regression
Mimotricentes fremontensis	2674.19	PBDB + regression

Continued on next page

Table A.1 – continued from previous page

Species	Mass (g)	Source
Mimotricentes subtrigonus	1984.90	PBDB + regression
Miniochoerus affinis	35835.68	PBDB
Miniochoerus gracilis	15437.78	PBDB
Minippus index	4819.64	PBDB
Mioclaenus turgidus	4412.84	PBDB + regression
Miocyon scotti	41776.02	PBDB + regression
Miocyon vallisrubrae	14244.90	PBDB + regression
Mioheteromys amplissimus	41.26	[329]
Miohippus anceps	41650.00	[198]
Miohippus assiniboiensis	78519.40	PBDB + regression
Miohippus equiceps	41650.00	[198]
Miohippus equinanus	35954.16	[329]
Miohippus gidleyi	129635.34	PBDB + regression
Miohippus grandis	76252.18	PBDB
Miohippus intermedius	82454.34	[329]
Miohippus obliquidens	52575.21	[329]
Miolabis fricki	240385.70	[329]
Miomustela madisonae	13.46	[329]
Mionictis elegans	6652.51	PBDB
Mionictis incertus	9269.73	PBDB
Mionictis letifer	134.29	[329]
Mionictis pristinus	32167.60	PBDB + regression
Miosciurus ballovianus	36.60	[329]
Miosicista angulus	8.15	PBDB
Miospermophilus lavertyi	106.70	[329]

Continued on next page

Table A.1 – continued from previous page

Species	Mass (g)	Source
Miospermophilus wyomingensis	83.10	[329]
Miotapirus harrisonensis	275463.83	PBDB + regression
Miotylopus gibbi	48050.12	[329]
Miotylopus leonardi	43044.94	[329]
Miotylopus taylori	73865.41	[329]
Miracinonyx inexpectatus	41653.42	PBDB + regression
Miracinonyx studeri	25529.71	PBDB + regression
Mithrandir gillianus	644.09	PBDB + regression
Mixodectes malaris	538.07	PBDB + regression
Mixodectes pungens	1563.78	PBDB + regression
Mojavemys galushai	28.22	[329]
Mojavemys lophatus	15.19	PBDB
Mojavemys magnumarcus	56.26	[329]
Montanatylopus matthewi	176770.65	PBDB + regression
Mookomys altifluminis	13.07	[329]
Mookomys formicarum	11.86	PBDB
Mookomys thrinax	11.94	[329]
Moropus elatus	707858.86	[329]
Moropus merriami	173767.00	[198]
Moropus oregonensis	189094.09	[329]
Mustela frenata	146.89	[305]
Mustela meltoni	895.71	PBDB + regression
Mustela rexroadensis	9.58	[329]
Mylanodon rosei	27.99	PBDB + regression
Mylohyus elmorei	168990.43	PBDB + regression

Continued on next page

Table A.1 – continued from previous page

Species	Mass (g)	Source
Myrmecoboides montanensis	45.21	PBDB + regression
Myrmecophaga tridactyla	32544.00	PBDB
Mysops parvus	54.13	PBDB + regression
Mystipterus martini	29.37	[329]
Mystipterus pacificus	14.88	[329]
Mytonolagus petersoni	213.83	PBDB + regression
Mytonolagus wyomingensis	84.09	PBDB + regression
Mytonomeryx scotti	5465.38	PBDB + regression
Mytonomys burkei	2757.96	PBDB + regression
Mytonomys mytonensis	1177.52	PBDB + regression
Mytonomys robustus	2795.71	PBDB + regression
Namatomys lloydi	24.79	PBDB + regression
Nannippus aztecus	106272.86	PBDB + regression
Nannippus beckensis	137564.87	PBDB + regression
Nannippus peninsulatus	68871.66	[329]
Nannippus westoni	62317.65	[329]
Nannodectes gidleyi	521.24	PBDB + regression
Nannodectes intermedius	133.09	PBDB + regression
Nannodectes simpsoni	402.81	PBDB + regression
Nanodelphys hunti	7.27	PBDB + regression
Nanotragulus loomisi	2724.39	[329]
Nanotragulus ordinatus	5541.39	[329]
Nanotragulus planiceps	3678.00	[198]
Navajovius kohlhaasae	12.82	PBDB + regression
Nebraskomys mcgrewi	46.06	[329]
Continued on next page		

Table A.1 – continued from previous page

Species	Mass (g)	Source
Nebraskomys rexroadensis	47.48	PBDB + regression
Nekrolagus progressus	1510.20	[329]
Neohipparion affine	167711.41	[329]
Neohipparion eurystyle	133252.35	[329]
Neohipparion leptode	233281.23	[329]
Neohipparion trampasense	120571.71	[329]
Neoliotomus conventus	1894.05	PBDB + regression
Neoliotomus ultimus	2252.70	PBDB + regression
Neoplagiaulax donaldorum	52.00	[362]
Neoplagiaulax grangeri	93.00	[362]
Neoplagiaulax hazeni	95.39	PBDB + regression
Neoplagiaulax hunteri	42.74	PBDB + regression
Neoplagiaulax macrotomeus	14.42	PBDB + regression
Neoplagiaulax mckennai	60.27	PBDB + regression
Neotoma cinerea	299.23	[305]
Neotoma fossilis	161.61	PBDB + regression
Neotoma leucopetrica	389.66	[110]
Neotoma quadriplicata	477.49	PBDB + regression
Neotoma sawrockensis	99.48	[329]
Neotoma taylori	190.07	PBDB + regression
Neotoma vaughani	190.07	PBDB + regression
Nerterogeomys anzensis	80.55	PBDB
Nerterogeomys garbanii	157.02	PBDB
Nerterogeomys minor	38.23	PBDB + regression
Nerterogeomys persimilis	49.62	[110]

Continued on next page

Table A.1 – continued from previous page

Species	Mass (g)	Source
Nexuotapirus marslandensis	144350.55	[329]
Nexuotapirus robustus	302549.45	[329]
Niglarodon koerneri	68.03	[329]
Nimravides galiani	159532.03	[329]
Nimravides thinobates	9582364.97	PBDB
Nimravus brachyops	3789.54	[329]
Nimravus sectator	112444.51	PBDB + regression
Niptomomys doreenae	17.88	PBDB + regression
Niptomomys thelmae	19.74	PBDB + regression
Nonomys gutzleri	20.14	PBDB + regression
Nonomys simplicidens	15.90	PBDB + regression
Notharctus pugnax	3034.76	PBDB + regression
Notharctus robinsoni	3318.67	PBDB + regression
Notharctus robustior	6900.00	[312]
Notharctus tenebrosus	4325.36	PBDB + regression
Notharctus venticolus	2691.73	PBDB + regression
Nothokemas floridanus	100709.96	[329]
Nothokemas waldropi	22925.38	[329]
Nothotylopus camptognathus	334368.85	[329]
Nothrotheriops shastensis	613762.01	[30]
Nothrotheriops texanus	124768.59	PBDB + regression
Notiosorex jacksoni	11.14	[189]
Notiosorex repenningi	17.08	[167]
Notiotitanops mississippiensis	4455686.16	PBDB + regression
Notolagus lepusculus	372.41	[329]

Continued on next page

Table A.1 – continued from previous page

Species	Mass (g)	Source
Nototamias hulberti	17.12	[329]
Nototamias quadratus	35.87	[329]
Nyctitherium serotinum	12.08	PBDB + regression
Nyctitherium velox	24.97	PBDB + regression
Ochotona spanglei	188.67	[329]
Odocoileus virginianus	52607.00	PBDB
Ogmodontomys poaphagus	131.63	[329]
Ogmodontomys sawrockensis	140.40	PBDB
Oklahomalagus whisenhunti	387.61	[329]
Oligobunis floridanus	24237.19	PBDB + regression
Oligoryctes cameronensis	4.73	PBDB + regression
Oligoscalops galbreathi	15.04	PBDB + regression
Oligospermophilus douglassi	115.51	PBDB + regression
Omomys carteri	171.31	PBDB + regression
Omomys lloydi	66.71	PBDB + regression
Onychomys hollisteri	20.14	PBDB + regression
Onychomys martini	18.36	[329]
Onychomys pedroensis	36.23	[329]
Oregonomys magnus	20.14	PBDB + regression
Oregonomys pebblespringsensis	26.58	[329]
Oregonomys sargenti	26.05	[329]
Oreodontoides oregonensis	25336.47	[329]
Oreolagus colteri	48.28	PBDB + regression
Oreolagus nebrascensis	336.97	[329]
Oreolagus nevadensis	78.00	[198]

Continued on next page

Table A.1 – continued from previous page

Species	Mass (g)	Source
Oreolagus wallacei	354.25	[329]
Oreotalpa florissantensis	10.87	PBDB + regression
Orohippus pumilus	11144.64	PBDB + regression
Orohippus sylvaticus	10210.29	PBDB + regression
Oromeryx plicatus	8545.03	PBDB + regression
Oropyctis pediasius	161.61	PBDB + regression
Osbornoceros osborni	17326.63	[329]
Osbornodon brachypus	12282.74	[43]
Osbornodon fricki	25848.30	[329]
Osbornodon iamonensis	13904.95	[329]
Osbornodon renjiei	9337.04	PBDB + regression
Osbornodon scitulus	11849.01	[329]
Osbornodon sesnoni	8349.86	[329]
Otarocyon cooki	1826.21	[329]
Otarocyon macdonaldi	1270.84	PBDB + regression
Ottoceros peacevalleyensis	13493.99	[329]
Ourayia hopsoni	590.00	PBDB + regression
Ourayia uintensis	758.66	PBDB + regression
Oxetocyon cuspidatus	2440.60	[329]
Oxyacodon agapetillus	352.30	PBDB + regression
Oxyacodon apiculatus	1116.82	PBDB + regression
Oxyacodon ferronensis	616.78	PBDB + regression
Oxyacodon priscilla	808.90	PBDB + regression
Oxyaena forcipata	49289.23	PBDB + regression
Oxyaena gulo	42202.38	PBDB + regression

Continued on next page

Table A.1 – continued from previous page

Species	Mass (g)	Source
Oxyaena intermedia	44676.22	PBDB + regression
Oxyclaenus cuspidatus	760.59	PBDB + regression
Oxyclaenus pugnax	1812.53	PBDB + regression
Oxyclaenus simplex	566.68	PBDB + regression
Oxydactylus longipes	112420.32	[329]
Oxydactylus lulli	140198.81	PBDB
Oxyprimus erikseni	133.09	PBDB + regression
Pachyaena gigantea	71758.33	PBDB + regression
Pachyaena gracilis	38119.28	PBDB + regression
Pachyaena ossifraga	460896.74	PBDB
Pachyarmatherium leiseyi	15420.00	[194]
Paciculus montanus	68.53	PBDB + regression
Paciculus nebraskensis	80.64	[329]
Paciculus woodi	32.63	[162]
Paenemarmota barbouri	10301.04	[329]
Paenemarmota mexicana	10977.77	PBDB + regression
Paenemarmota nevadensis	7644.00	[198]
Paenemarmota sawrockensis	5943.18	[329]
Palaechthon alticuspis	51.17	PBDB + regression
Palaechthon woodi	27.99	PBDB + regression
Palaeictops bicuspis	143.19	[293]
Palaeictops bridgeri	139.74	PBDB + regression
Palaeictops multicuspis	149.35	PBDB + regression
Palaeogale dorothiae	607.89	[329]
Palaeogale minuta	400.00	NOW

Continued on next page

Table A.1 – continued from previous page

Species	Mass (g)	Source
Palaeogale sectoria	716.63	PBDB
Palaeolagus burkei	307.97	[329]
Palaeolagus hemirhizis	128.62	PBDB + regression
Palaeolagus hypsodus	678.58	[329]
Palaeolagus philoi	1211.97	[329]
Palaeolagus primus	119.90	PBDB + regression
Palaeolagus temnodon	171.31	PBDB + regression
Palaeonictis occidentalis	36870.62	PBDB + regression
Palaeonictis peloria	60916.03	PBDB + regression
Palaeoryctes cruoris	62.68	PBDB + regression
Palaeoryctes puercensis	24.97	PBDB + regression
Palaeosyops laevidens	1052429.44	PBDB + regression
Palaeosyops paludosus	1388030.44	PBDB + regression
Palaeosyops robustus	1619966.85	PBDB + regression
Palenochtha minor	9.87	PBDB + regression
Palenochtha weissae	9.87	PBDB + regression
Paleotomus junior	171.31	PBDB + regression
Paleotomus radagasti	1351.50	PBDB
Panthera onca	100000.00	PBDB
Pantolambda bathmodon	63252.15	PBDB
Pantolambda cavirictus	35087.78	PBDB + regression
Pantolambda intermedius	19134.73	PBDB + regression
Parablastomeryx galushi	19345.66	PBDB + regression
Paracosoryx furlongi	13493.99	[329]
Paracosoryx wilsoni	15473.78	PBDB

Continued on next page

Table A.1 – continued from previous page

Species	Mass (g)	Source
Paracryptotis gidleyi	28.22	[329]
Paracryptotis rex	42.10	[329]
Paracynarctus kelloggi	8349.86	[329]
Paracynarctus sinclairi	8022.46	[329]
Paradaphoenus cuspigerus	4023.87	[329]
Paradaphoenus minimus	3314.52	PBDB + regression
Paradaphoenus tooheyi	3498.19	[329]
Paradjidaumo alberti	15.90	PBDB + regression
Paradjidaumo hypsodus	29.87	PBDB + regression
Paradjidaumo reynoldsi	20.14	PBDB + regression
Paradjidaumo spokanensis	54.13	PBDB + regression
Paradjidaumo trilophus	32.56	PBDB + regression
Paradjidaumo validus	61.14	PBDB + regression
Paradomnina relictus	23.81	[329]
Paraenhydrocyon josephi	7942.63	[329]
Paraenhydrocyon wallovianus	14185.85	[329]
Parahippus leonensis	94845.07	[329]
Parahippus pawniensis	99707.88	[329]
Parahippus tyleri	105994.41	PBDB + regression
Parahippus wyomingensis	98400.00	[188]
Parahyus vagus	185903.42	PBDB + regression
Paralabis cedrensis	75771.66	PBDB + regression
Parallomys americanus	162.39	[329]
Paramiolabis taylori	131926.47	[329]
Paramylodon harlani	1185306.75	PBDB + regression

Continued on next page

Table A.1 – continued from previous page

Species	Mass (g)	Source
Paramys adamus	50.76	PBDB + regression
Paramys atavus	47.48	PBDB + regression
Paramys compressidens	1746.66	PBDB + regression
Paramys copei	540.72	PBDB + regression
Paramys delicatior	708.27	PBDB + regression
Paramys delicatus	1793.22	PBDB + regression
Paramys excavatus	259.62	PBDB + regression
Paramys taurus	316.00	PBDB + regression
Paranamatomys storeri	7.59	PBDB
Parapliohippus carrizoensis	80821.64	[329]
Parapliosaccomys oregonensis	24.29	[329]
Parapliosaccomys transversus	45.58	PBDB
Parapotos tedfordi	3533.34	[329]
Pararyctes pattersoni	17.36	PBDB + regression
Paratomarctus euthos	14472.42	[329]
Paratomarctus temerarius	11498.82	[329]
Paratylopus labiatus	47357.63	PBDB + regression
Paratylopus primaevus	29814.41	PBDB + regression
Parectypodus clemensi	23.70	PBDB + regression
Parectypodus corystes	40.09	PBDB + regression
Parectypodus laytoni	11.30	PBDB + regression
Parectypodus lunatus	21.68	PBDB + regression
Parectypodus simpsoni	37.51	PBDB + regression
Parectypodus sinclairi	19.00	[362]
Parectypodus sylviae	12.82	PBDB + regression

Continued on next page

Table A.1 – continued from previous page

Species	Mass (g)	Source
Parectypodus trovessartianus	87.78	PBDB + regression
Pareumys boskeyi	88.57	PBDB + regression
Pareumys grangeri	57.59	PBDB + regression
Pareumys guensburgi	140.29	PBDB + regression
Pareumys milleri	76.28	PBDB + regression
Parictis parvus	1594.41	PBDB + regression
Parictis personi	2218.53	PBDB + regression
Paromomys depressidens	57.16	PBDB + regression
Paromomys maturus	197.38	PBDB + regression
Paronychomys alticuspis	17.12	[329]
Paronychomys lemredfieldi	20.09	[329]
Paronychomys tuttlei	38.86	[329]
Paroreodon parvus	51150.00	[198]
Parvericius montanus	41.26	[329]
Parvericius voorhiesi	28.79	[329]
Patriofelis ferox	22418.00	[198]
Patriofelis ulta	48311.29	PBDB + regression
Patriolestes novaceki	246.37	PBDB + regression
Pauromys exallos	17.97	PBDB + regression
Pauromys lillegraveni	12.11	PBDB + regression
Pauromys simplex	13.95	PBDB + regression
Pauromys texensis	17.97	PBDB + regression
Pediomeryx hemphillensis	167711.41	[329]
Pelycodus jarrovii	3222.75	PBDB + regression
Pelycomys brulanus	214.27	PBDB + regression

Continued on next page

Table A.1 – continued from previous page

Species	Mass (g)	Source
Pelycomys rugosus	330.84	PBDB + regression
Penetrigonias hudsoni	601751.47	PBDB
Pentacemylus leotensis	5601.22	PBDB + regression
Pentacemylus progressus	7131.24	PBDB + regression
Pentacodon inversus	604.36	PBDB + regression
Pentacodon occultus	1560.69	PBDB + regression
Peraceras hessei	936589.16	[329]
Peraceras profectum	2326789.55	[329]
Peraceras superciliosum	1639660.88	[329]
Peradectes californicus	6.11	PBDB + regression
Peradectes chesteri	5.04	PBDB + regression
Peradectes elegans	11.30	PBDB + regression
Peradectes minor	11.23	[327]
Peradectes protinnominatus	12.82	PBDB + regression
Peratherium comstocki	87.78	PBDB + regression
Peratherium marsupium	91.55	PBDB + regression
Perchoerus probus	24602.41	PBDB + regression
Peridiomys halis	63.43	[329]
Peridiomys oregonensis	35.35	PBDB + regression
Peridiomys rusticus	69.41	[329]
Periptychus carinidens	30641.27	PBDB + regression
Periptychus coarctatus	22723.44	PBDB + regression
Perognathus ancenensis	11.47	[329]
Perognathus coquorum	35.35	PBDB + regression
Perognathus dunklei	7.92	[329]

Continued on next page

Table A.1 – continued from previous page

Species	Mass (g)	Source
Perognathus furlongi	9.68	[329]
Perognathus gidleyi	11.82	[329]
Perognathus maldei	10.39	PBDB + regression
Perognathus mclaughlini	8.33	[329]
Perognathus minutus	5.93	[329]
Perognathus pearlettensis	6.89	[329]
Perognathus rexroadensis	17.97	PBDB + regression
Perognathus trojectioansrum	3.63	[329]
Peromyscus antiquus	26.84	[329]
Peromyscus brachygnathus	10.39	PBDB + regression
Peromyscus complexus	25.46	PBDB
Peromyscus cragini	15.49	[329]
Peromyscus dentalis	12.11	PBDB + regression
Peromyscus hagermanensis	19.89	[329]
Peromyscus minimus	4.67	PBDB + regression
Peromyscus nosher	15.90	PBDB + regression
Peromyscus polionotus	14.30	PBDB
Peromyscus sarmocophinus	17.97	PBDB + regression
Petauristodon jamesi	307.97	[329]
Petauristodon mathewsi	214.86	[329]
Petauristodon pattersoni	336.97	[329]
Pewelagus dawsonae	447.84	[150]
Pewelagus mexicanus	186.75	PBDB + regression
Phelosaccomys annae	26.31	[329]
Phelosaccomys hibbardi	46.53	[329]

Continued on next page

Table A.1 – continued from previous page

Species	Mass (g)	Source
Phelosaccomys neomexicanus	19.89	[329]
Phelosaccomys shotwelli	20.49	[329]
Phenacocoelus typus	52052.08	[329]
Phenacodaptes sabulosus	109.55	PBDB + regression
Phenacodus bisonensis	17047.49	PBDB + regression
Phenacodus grangeri	28510.66	PBDB + regression
Phenacodus intermedius	45430.80	PBDB + regression
Phenacodus magnus	72471.82	PBDB + regression
Phenacodus matthewi	5446.34	[50]
Phenacodus trilobatus	64494.29	PBDB + regression
Phenacodus vortmani	10636.06	PBDB + regression
Phenacolemur fortior	186.75	[364]
Phenacolemur mcgrewi	63.45	PBDB + regression
Phenacolemur praecox	151.65	PBDB + regression
Phenacolemur simonsi	40.09	PBDB + regression
Phenacomys gryci	228.00	[198]
Philotrox condoni	11968.10	[329]
Phlaocyon achoros	2951.30	[329]
Phlaocyon annectens	3498.19	[329]
Phlaocyon latidens	2779.43	[329]
Phlaocyon leucosteus	3827.63	[329]
Phlaocyon minor	3498.19	[329]
Phlaocyon taylori	1939.14	[329]
Phlaocyon yatkolai	9604.62	[329]
Phoberocyon johnhenryi	179871.86	[329]

Continued on next page

Table A.1 – continued from previous page

Species	Mass (g)	Source
Picrodus calgariensis	12.82	PBDB + regression
Picrodus canpacius	48.28	PBDB + regression
Picrodus silberlingi	48.28	PBDB + regression
Pipestoneomys bisulcatus	47.48	PBDB + regression
Plagioctenodon krausae	4.73	PBDB + regression
Plagioctenodon rosei	10.87	PBDB + regression
Plagiomene accola	359.33	PBDB + regression
Plagiomene multicuspis	850.13	PBDB + regression
Planisorex dixonensis	12.08	PBDB + regression
Platygonus bicalcaratus	114741.07	PBDB + regression
Platygonus oregonensis	40134.84	[329]
Platygonus pearcei	55826.28	[329]
Platygonus vetus	65463.62	[30]
Plesiadapis anceps	291.90	PBDB + regression
Plesiadapis churchilli	871.09	PBDB + regression
Plesiadapis cookei	4434.83	PBDB + regression
Plesiadapis dubius	417.80	PBDB + regression
Plesiadapis fodinatus	709.89	PBDB + regression
Plesiadapis gingerichi	2726.96	PBDB + regression
Plesiadapis praecursor	192.03	PBDB + regression
Plesiadapis rex	690.80	PBDB + regression
Plesiocolopirus hancocki	42441.52	PBDB + regression
Plesiogulo lindsayi	4628.55	[329]
Plesiogulo marshalli	3133.79	[329]
Plesiolestes nacimienti	233.00	[312]

Continued on next page

Table A.1 – continued from previous page

Species	Mass (g)	Source
Plesiolestes problematicus	115.64	PBDB + regression
Plesiolestes wilsoni	366.42	PBDB + regression
Plesiosminthus clivosus	9.30	[329]
Plesiosorex coloradensis	192.48	[329]
Plesiosorex donroosai	685.40	[329]
Pleurolicus dakotensis	60.95	[329]
Pleurolicus exiguus	24.79	PBDB + regression
Pleurolicus sellardsi	53.28	[370]
Pleurolicus sulcifrons	83.93	[329]
Pliocyon medius	172818.99	[329]
Pliocyon robustus	176310.16	[329]
Pliogale furlongi	456.36	[364]
Pliogale manka	2986.47	PBDB + regression
Pliogeomys parvus	10.28	[329]
Pliogeomys russelli	15.18	[329]
Pliohippus fossulatus	257815.63	[329]
Pliohippus pernix	198789.15	[329]
Pliohippus tehonensis	164875.58	PBDB + regression
Pliometanastes galushai	8636.25	PBDB
Pliometanastes protistus	88139.12	PBDB + regression
Plionarctos edensis	56954.05	[329]
Plionarctos harroldorum	47351.79	PBDB + regression
Plionictis ogygia	36.97	[329]
Pliophenacomys dixonensis	50.76	PBDB + regression
Pliophenacomys finneyi	64.79	PBDB + regression

Continued on next page

Table A.1 – continued from previous page

Species	Mass (g)	Source
Pliophenacomys meadensis	54.13	PBDB + regression
Pliophenacomys osborni	86.49	[329]
Pliophenacomys primaevus	61.56	[329]
Pliosaccomys dubius	27.66	[329]
Pliosaccomys higginsensis	18.17	[329]
Pliotaxidea garberi	4440.12	PBDB + regression
Pliotaxidea nevadensis	130.32	[329]
Pliotomodon primitivus	107.77	[329]
Pliozapus solus	18.73	[329]
Plithocyon ursinus	189094.09	[329]
Poabromylus golzi	11885.48	PBDB + regression
Poabromylus kayi	23487.74	PBDB + regression
Poebrotherium eximium	50423.02	PBDB + regression
Poebrotherium wilsoni	60615.67	PBDB + regression
Pogonodon eileenae	138453.42	[92]
Pratifelis martini	215345.72	[329]
Pratilepus kansasensis	972.63	[329]
Premnoides douglassi	40.09	PBDB + regression
Presbymys lophatus	72.36	PBDB + regression
Presbytherium rhodorugatus	2901.91	PBDB + regression
Princetonia yalensis	1703.86	PBDB + regression
Probassariscus matthewi	2396.46	PBDB + regression
Probathyopsis harrisorum	96406.85	PBDB + regression
Probathyopsis praecursor	119830.36	PBDB + regression
Problastomeryx primus	14913.17	[329]

Continued on next page

Table A.1 – continued from previous page

Species	Mass (g)	Source
Procamelus grandis	400312.19	[329]
Procamelus occidentalis	189094.09	[329]
Procerberus formicarum	67.38	PBDB + regression
Prochetodon cavus	213.83	PBDB + regression
Prochetodon foxi	202.80	PBDB + regression
Prochetodon speirsae	417.88	PBDB
Prochetodon taxus	480.23	PBDB + regression
Procranioceras skinneri	169396.94	[329]
Procynodictis progressus	3665.64	PBDB + regression
Procyon lotor	5814.00	PBDB
Procyon rexroadensis	9632.47	PBDB + regression
Prodiacodon concordiarcensis	43.19	PBDB + regression
Prodiacodon crustulum	125.81	[168]
Prodiacodon furor	66.91	[280]
Prodiacodon puercensis	225.49	[200]
Prodiacodon tauricinerei	82.27	PBDB + regression
Prodipodomys centralis	21.26	[23]
Prodipodomys idahoensis	22.42	[329]
Prodipodomys kansensis	12.81	[329]
Prodipodomys timoteoensis	20.14	PBDB + regression
Prohesperocyon wilsoni	3787.93	PBDB + regression
Proheteromys fedti	17.97	PBDB + regression
Proheteromys floridanus	5.37	[329]
Proheteromys gremmelsi	27.97	[23]
Proheteromys ironcloudi	10.59	[329]

Continued on next page

Table A.1 – continued from previous page

Species	Mass (g)	Source
Proheteromys maximus	87.36	[329]
Proheteromys nebraskensis	21.54	PBDB
Proheteromys sulculus	15.90	PBDB + regression
Proheteromys toledoensis	59.15	[329]
Prolapsus junctionis	72.36	PBDB + regression
Prolapsus sibilatoris	178.45	PBDB + regression
Prolimnocyon antiquus	1442.37	PBDB + regression
Prolimnocyon atavus	2467.80	PBDB + regression
Prolimnocyon haematus	1058.55	PBDB + regression
Promartes darbyi	3029.14	[60]
Promartes gemmarosae	1945.16	[60]
Promartes lepidus	6776.78	[207]
Promioclaenus acolytus	410.27	PBDB + regression
Promioclaenus pipiringosi	748.79	PBDB + regression
Promioclaenus thnetus	260.57	[98]
Promylagaulus riggsi	85.63	[329]
Pronodens silberlingi	22247.84	[329]
Pronothodectes gaoi	254.50	PBDB + regression
Pronothodectes jepi	242.55	PBDB + regression
Pronothodectes matthewi	142.24	PBDB + regression
Pronotolagus apachensis	445.86	[329]
Pronotolagus nevadensis	60.34	[329]
Pronotolagus whitei	1436.55	[329]
Proscalops miocaenus	28.30	PBDB + regression
Proscalops secundus	72.97	[329]

Continued on next page

Table A.1 – continued from previous page

Species	Mass (g)	Source
Proscalops tertius	96.54	[329]
Prosciurus magnus	106.19	PBDB + regression
Prosciurus parvus	68.53	PBDB + regression
Prosciurus relictus	64.79	PBDB + regression
Prosigmodon chihuahuensis	72.36	PBDB + regression
Prosigmodon ferrusquiai	48.33	PBDB
Prosigmodon holocuspis	113.30	[329]
Prosigmodon oroscoi	27.28	PBDB + regression
Prosomys mimus	32.56	PBDB + regression
Prosthennops niobrarensis	43044.94	[329]
Prosthennops serus	53637.30	[329]
Prosthennops xiphodonticus	23860.99	[329]
Prosynthetoceras francisi	134591.56	[329]
Prosynthetoceras orthrionanus	40134.84	[329]
Protadjidaumo pauli	17.97	PBDB + regression
Protadjidaumo typus	12.11	PBDB + regression
Protapirus obliquidens	168000.80	[264]
Protapirus simplex	189283.62	PBDB + regression
Protepicyon raki	23623.56	[329]
Proterix bicuspis	194.51	[48]
Proterix loomisi	523.22	PBDB + regression
Proterixoides davisi	242.45	PBDB + regression
Prothryptacodon albertensis	464.26	PBDB + regression
Prothryptacodon furens	1351.03	PBDB + regression
Prothryptacodon hilli	1954.40	PBDB + regression

Continued on next page

Table A.1 – continued from previous page

Species	Mass (g)	Source
Protictis haydenianus	2074.37	PBDB + regression
Protictis microlestes	722.34	PBDB
Protictis minor	1270.84	PBDB + regression
Protictis paralus	354.83	PBDB + regression
Protictis paulus	298.35	PBDB
Protictis simpsoni	3202.65	PBDB + regression
Protitanops curryi	142477.00	[198]
Protitanotherium superbum	2858807.95	PBDB + regression
Protoceras celer	105068.04	PBDB + regression
Protoceras skinneri	81023.96	PBDB
Protohippus gidleyi	164390.50	[329]
Protohippus perditus	135944.23	[329]
Protohippus supremus	167711.41	[329]
Protohippus vetus	123329.89	PBDB
Protolabis coartatus	110194.25	[329]
Protolabis heterodontus	219695.99	[329]
Protomarctus optatus	11271.13	[329]
Protoreodon pacificus	34181.30	PBDB + regression
Protoreodon parvus	33051.65	PBDB + regression
Protoreodon pearcei	38169.88	PBDB + regression
Protoreodon petersoni	18828.04	PBDB + regression
Protoreodon pumilus	51870.94	PBDB + regression
Protoreodon walshi	59926.66	PBDB + regression
Protorohippus venticolus	14768.64	PBDB + regression
Protosciurus mengi	401.25	PBDB + regression

Continued on next page

Table A.1 – continued from previous page

Species	Mass (g)	Source
Protosciurus tecuyensis	478.19	[329]
Protoselene griphus	1619.16	PBDB + regression
Protoselene opisthacus	1703.86	PBDB + regression
Protospermophilus kelloggi	202.35	[329]
Protospermophilus malheurensis	103.54	[329]
Protospermophilus oregonensis	450.34	[329]
Protospermophilus quatalensis	273.14	[329]
Protospermophilus vortmani	230.44	[329]
Prototomus deimos	504.65	PBDB + regression
Prototomus martis	3573.98	PBDB + regression
Prototomus phobos	2691.73	PBDB + regression
Prototomus robustus	2434.13	PBDB + regression
Prototomus secundarius	1522.82	PBDB + regression
Protungulatum donnae	546.57	PBDB + regression
Protylopus annectens	14816.10	PBDB + regression
Protylopus pearsonensis	38366.65	PBDB + regression
Protylopus petersoni	18468.57	PBDB + regression
Protylopus robustus	17655.81	PBDB + regression
Protylopus stocki	15437.78	PBDB + regression
Proviverroides piercei	7343.84	PBDB + regression
Psalidocyon marianae	8777.97	[329]
Pseudaelurus aeluroides	31571.18	[329]
Pseudaelurus intrepidus	40945.61	[329]
Pseudaelurus marshi	33189.87	[329]
Pseudaelurus stouti	5767.53	[329]

Continued on next page

Table A.1 – continued from previous page

Species	Mass (g)	Source
Pseudhipparion curtivallum	58104.59	[329]
Pseudhipparion gratum	108012.26	[329]
Pseudhipparion hessei	75357.60	[329]
Pseudhipparion retrusum	86681.87	[329]
Pseudhipparion simpsoni	48050.12	[329]
Pseudhipparion skinneri	54176.36	[329]
Pseudoblastomeryx advena	10097.06	[329]
Pseudoceras skinneri	10938.02	[329]
Pseudocylindrodon lateriviae	101.65	PBDB + regression
Pseudocylindrodon medius	44.30	PBDB + regression
Pseudocylindrodon neglectus	88.57	PBDB + regression
Pseudocylindrodon pintoensis	194.89	MIOMAP
Pseudodiplacodon progressum	1263102.01	PBDB
Pseudolabis dakotensis	59874.14	[329]
Pseudoparablastomeryx francescita	4900.87	PBDB
Pseudoparablastomeryx scotti	5884.05	[329]
Pseudoprotoceras longinaris	35005.16	PBDB + regression
Pseudoprotoceras minor	19973.07	PBDB + regression
Pseudotheridomys cuyamensis	12.43	[329]
Pseudotheridomys hesperus	14.15	[329]
Pseudotheridomys pagei	8.17	[329]
Pseudotomus californicus	3660.37	PBDB + regression
Pseudotomus eugenei	14363.43	PBDB + regression
Pseudotomus hians	2299.39	PBDB
Pseudotomus horribilis	2971.99	[39]

Continued on next page

Table A.1 – continued from previous page

Species	Mass (g)	Source
Pseudotomus johanniculi	12241.50	PBDB + regression
Pseudotomus littoralis	2500.02	PBDB + regression
Pseudotomus petersoni	4209.29	[39]
Pseudotomus robustus	4052.00	PBDB + regression
Pseudotrimylus mawbyi	142.92	PBDB + regression
Ptilodus fractus	105.00	[362]
Ptilodus gnomus	87.78	PBDB + regression
Ptilodus kummae	208.28	PBDB + regression
Ptilodus mediaevus	311.46	PBDB + regression
Ptilodus montanus	373.57	PBDB + regression
Ptilodus wyomingensis	147.00	[362]
Puercolestes simpsoni	138.73	PBDB + regression
Puma concolor	48009.00	PBDB
Puma lacustris	17724.85	PBDB + regression
Pyrocyon dioctetus	1834.56	PBDB + regression
Quadratomus grandis	1595.31	PBDB + regression
Quadratomus grossus	4624.34	PBDB + regression
Quadrodens wilsoni	46.06	[329]
Rakomeryx sinclairi	111301.72	[329]
Rapamys fricki	1050.61	PBDB + regression
Raphictis gausion	261.58	PBDB + regression
Reithrodontomys galushai	10.39	PBDB + regression
Reithrodontomys rexroadensis	7.29	PBDB + regression
Reithrodontomys wetmorei	9.03	[329]
Reithroparamys debequensis	190.07	PBDB + regression

Continued on next page

Table A.1 – continued from previous page

Species	Mass (g)	Source
Reithroparamys delicatissimus	460.05	PBDB + regression
Reithroparamys huerfanensis	338.37	PBDB + regression
Reithroparamys sciuroides	482.77	[353]
Repomys arizonensis	54.13	PBDB + regression
Repomys gustelyi	81.45	[329]
Repomys maxumi	122.73	[329]
Repomys panacaensis	38.47	[329]
Rhizocyon oregonensis	3361.02	[329]
Rhynchotherium falconeri	54956347.86	PBDB + regression
Russellagus vonhofi	214.86	[329]
Sanctimus falkenbachi	151.41	[329]
Sanctimus stouti	120.30	[329]
Sanctimus stuartae	100.48	[329]
Satherium piscinarium	934.49	[329]
Saxonella naylori	45.48	PBDB + regression
Scalopoides isodens	32.14	[329]
Scalopoides ripafodiator	27.39	[329]
Scalopus aquaticus	39.60	PBDB
Scapanoscapter simplicidens	58.56	[329]
Scapanus hagermanensis	7.00	[198]
Scapanus latimanus	55.00	PBDB
Scapanus proceridens	65.37	[329]
Scapanus shultzi	97.00	[198]
Scapanus townsendii	141.58	[305]
Scaphohippus sumani	103960.05	PBDB

Continued on next page

Table A.1 – continued from previous page

Species	Mass (g)	Source
Scenopagus curtidens	26.62	PBDB + regression
Scenopagus edenensis	72.22	PBDB + regression
Scenopagus priscus	17.36	PBDB + regression
Schaubeumys galbreathi	19.54	[179]
Schaubeumys grangeri	20.14	PBDB + regression
Schaubeumys sabrae	17.54	[193]
Schizodontomys amnicolus	111.05	[329]
Schizodontomys greeni	94.00	[198]
Schizodontomys harkseni	105.64	[329]
Sciuravus bridgeri	41.22	PBDB + regression
Sciuravus nitidus	167.14	PBDB + regression
Sciuravus popi	338.37	PBDB + regression
Sciuravus powayensis	88.57	PBDB + regression
Sciuravus wilsoni	84.39	PBDB + regression
Sciurion campestre	27.28	PBDB + regression
Sciurus carolinensis	518.00	PBDB
Sciurus olsoni	22.73	PBDB
Scottimus exiguus	68.99	[215]
Scottimus longiquus	111.05	[329]
Scottimus lophatus	132.82	PBDB
Scottimus viduus	44.30	PBDB + regression
Selenaletes scopaeus	8232.01	PBDB + regression
Serbelodon barbourensis	11462314.59	[281]
Sespedectes singularis	28.30	PBDB + regression
Sespedectes stocki	23.37	PBDB + regression

Continued on next page

Table A.1 – continued from previous page

Species	Mass (g)	Source
Sespemys thurstoni	290.03	[329]
Sespia californica	3604.72	[329]
Sespia nitida	2394.82	[339]
Shoshonius bowni	142.24	PBDB + regression
Shoshonius cooperi	57.16	PBDB + regression
Sifrhippus aemulor	6498.52	[219]
Sifrhippus grangeri	11578.47	PBDB + regression
Sifrhippus sandrae	10631.54	PBDB + regression
Sigmodon curtisi	92.85	PBDB + regression
Sigmodon hudspethensis	68.53	PBDB + regression
Sigmodon minor	52.98	[329]
Simidectes magnus	3787.93	PBDB + regression
Simidectes medius	2882.11	PBDB + regression
Simidectes merriami	6833.74	PBDB + regression
Similisciurus maxwelli	259.82	[329]
Simimeryx hudsoni	4171.22	PBDB + regression
Simimeryx minutus	1736.45	PBDB + regression
Simimys landeri	38.23	PBDB + regression
Simimys simplex	13.95	PBDB + regression
Simocyon primigenius	70000.00	NOW
Simojovelhyus pocitosense	19763.18	[355]
Simpsonictis pegus	245.14	PBDB + regression
Simpsonictis tenuis	126.87	PBDB + regression
Simpsonlemur citatus	171.00	[312]
Simpsonlemur jepseni	121.00	[312]

Continued on next page

Table A.1 – continued from previous page

Species	Mass (g)	Source
Simpsonodus chacensis	1262.40	PBDB + regression
Sinclairella dakotensis	139.74	PBDB + regression
Sinopa major	7566.66	PBDB + regression
Sinopa rapax	3299.38	PBDB + regression
Smilodectes gracilis	1035.64	PBDB + regression
Smilodectes mcgrewi	1212.97	PBDB + regression
Smilodectes sororis	935.37	PBDB + regression
Smilodon gracilis	57206.56	PBDB + regression
Sminthosinis bowleri	100.48	[329]
Sorex cinereus	3.80	PBDB
Sorex edwardsi	6.84	PBDB
Sorex hagermanensis	7.56	PBDB + regression
Sorex meltoni	3.46	[329]
Sorex palustris	13.46	[305]
Sorex powersi	7.77	[329]
Sorex rexroadensis	8.00	[198]
Sorex yatkolai	4.91	PBDB
Spermophilus argonautus	131.63	[329]
Spermophilus bensoni	184.93	[329]
Spermophilus boothi	498.11	[281]
Spermophilus cragini	578.25	[329]
Spermophilus dotti	270.43	[329]
Spermophilus fricki	231.96	[293]
Spermophilus gidleyi	700.00	[198]
Spermophilus howelli	139.77	[329]

Continued on next page

Table A.1 – continued from previous page

Species	Mass (g)	Source
Spermophilus jerae	87.36	[329]
Spermophilus matachicensis	220.52	PBDB + regression
Spermophilus matthewi	399.62	[22]
Spermophilus meadensis	100.48	[329]
Spermophilus rexroadensis	345.98	PBDB + regression
Spermophilus russelli	166.00	[198]
Spermophilus shotwelli	204.00	[198]
Spermophilus tephrus	89.12	[329]
Spermophilus wellingtonensis	295.89	[329]
Spermophilus wilsoni	247.15	[329]
Sphacorhysis burntforkensis	40.09	PBDB + regression
Sphenophalos nevadanus	8789.20	PBDB
Spilogale microdens	1270.84	[43]
Spilogale putorius	341.19	[305]
Spilogale rexroadi	895.71	PBDB + regression
Stegomastodon mirificus	35283150.07	PBDB + regression
Steinius annectens	397.75	[325]
Steinius vespertinus	119.90	PBDB + regression
Stelocyon arctylos	2132.95	PBDB + regression
Stenoechinus tantalus	48.42	[329]
Stenomylus gracilis	44801.64	[329]
Stenomylus hitchcocki	38948.67	[329]
Stenomylus taylori	47980.97	PBDB
Sthenictis dolichops	665.14	[329]
Sthenictis junturensis	330.30	[329]

Continued on next page

Table A.1 – continued from previous page

Species	Mass (g)	Source
Stibarus montanus	1152.46	PBDB + regression
Stibarus obtusilobus	1416.00	PBDB + regression
Stibarus quadricuspis	2093.31	PBDB + regression
Stockia powayensis	161.34	PBDB + regression
Stratimus strobeli	20.09	[329]
Strigorhysis bridgerensis	91.55	PBDB + regression
Strigorhysis huerfanensis	176.39	PBDB + regression
Stygimys gratus	128.62	PBDB + regression
Stygimys jepseni	84.00	[362]
Stygimys kuszmauli	291.90	PBDB + regression
Stylinodon mirus	47987.28	PBDB + regression
Subdromomeryx antilopinus	59874.14	[329]
Subhyracodon mitis	458341.99	[279]
Subhyracodon occidentalis	662391.44	PBDB + regression
Sunkahetanka geringensis	11158.98	[329]
Swaindelphys cifellii	24.57	PBDB
Symmetrodontomys simplicidens	26.84	[329]
Syndyoceras cooki	73865.41	[329]
Tachylagus gawneae	146.91	PBDB + regression
Taeniolabis taoensis	75621.27	PBDB + regression
Talpavoides dartoni	8.53	PBDB + regression
Talpavus conjunctus	45.48	PBDB + regression
Talpavus duplus	17.88	PBDB + regression
Talpavus nitidus	11.30	PBDB + regression
Tamias ateles	32.56	PBDB + regression

Continued on next page

Table A.1 – continued from previous page

Species	Mass (g)	Source
Tanymykter brachyodontus	102744.44	[329]
Tapiravus validus	64860.88	[329]
Tapirus simpsoni	369534.73	[329]
Tapochoerus egressus	9671.40	PBDB + regression
Tapochoerus mcmillini	4357.69	PBDB + regression
Tapocyon dawsonae	17454.68	PBDB + regression
Tapocyon robustus	21673.71	PBDB + regression
Tardontia nevadans	157.59	[329]
Tardontia occidentale	184.22	PBDB + regression
Tarka stylifera	3858.47	PBDB + regression
Tatmanius szalayi	46.59	[71]
Taxidea mexicana	10434.28	[27]
Taxidea taxus	7112.14	[305]
Tayassu protervus	45119.24	PBDB
Teilhardina americana	66.71	PBDB + regression
Teilhardina crassidens	57.16	PBDB + regression
Teleoceras meridianum	2022813.66	[329]
Teletaceras mortivallis	89575.08	PBDB + regression
Telmatherium altidens	2087486.17	PBDB + regression
Telmatherium cultridens	627910.03	PBDB + regression
Telmatherium manteoceras	478340.42	PBDB + regression
Temnocyon altigenis	32532.67	[329]
Temnocyon percussor	68871.66	[329]
Tenudomys bodei	43.38	[329]
Tenudomys macdonaldi	79.84	[329]

Continued on next page

Table A.1 – continued from previous page

Species	Mass (g)	Source
Tephrocyon rurestris	13095.19	[329]
Tetonius ambiguus	92.52	PBDB
Tetonius matthewi	107.32	PBDB + regression
Tetonius mckennai	57.16	PBDB + regression
Tetraclaenodon puercensis	9425.94	PBDB + regression
Tetrapassalus mckennai	5.04	PBDB
Texomys ritchiei	97.51	[329]
Thinobadistes segnis	645890.00	[194]
Thinocyon velox	1924.12	PBDB + regression
Thinohyus lentus	101400.00	[198]
Thisbemys corrugatus	965.53	PBDB + regression
Thisbemys elachistos	140.29	PBDB + regression
Thisbemys perditus	587.96	PBDB + regression
Thisbemys uintensis	1777.64	PBDB + regression
Thomomys bottae	114.82	[305]
Thomomys carsonensis	29.87	PBDB + regression
Thomomys gidleyi	37.34	[329]
Thryptacodon antiquus	4110.15	PBDB + regression
Thryptacodon australis	2744.65	PBDB + regression
Thryptacodon orthogonius	1536.41	PBDB + regression
Thryptacodon pseudarctos	6070.61	PBDB + regression
Thylacaelurus campester	35.01	PBDB + regression
Thylacaelurus montanus	63.45	PBDB + regression
Thylacodon pusillus	48.28	PBDB + regression
Ticholeptus zygomaticus	106937.52	[329]

Continued on next page

Table A.1 – continued from previous page

Species	Mass (g)	Source
Tillodon fodiens	49838.54	PBDB + regression
Tillomys senex	84.39	PBDB + regression
Tinimomys graybulliensis	11.30	PBDB + regression
Tinimomys tribos	5.32	[301]
Titanoides gidleyi	48022.06	PBDB + regression
Titanoides nanus	37406.72	PBDB + regression
Titanoides primaevus	72645.54	PBDB + regression
Tomarctus brevirostris	17500.77	[329]
Tomarctus hippophaga	13766.59	[329]
Torrejonia sirokyi	440.75	PBDB + regression
Toxotherium hunteri	91951.16	PBDB + regression
Tregosorex holmani	26.84	[329]
Tremarctos floridanus	149968.48	[305]
Trigenicus profectus	28723.88	PBDB + regression
Trigonias osborni	599072.86	PBDB + regression
Trigonias yoderensis	264875.92	PBDB + regression
Trigonictis cookii	7397.01	PBDB + regression
Trigonictis macrodon	419.89	[329]
Triisodon quivirensis	22180.76	PBDB + regression
Trilaccogaulus ovatus	103.54	[329]
Triplopus cubitalis	34118.20	PBDB + regression
Triplopus implicatus	64570.33	PBDB + regression
Triplopus obliquidens	95152.03	PBDB + regression
Triplopus rhinocerinus	88700.44	PBDB
Triplopus woodi	69515.35	PBDB

Continued on next page

Table A.1 – continued from previous page

Species	Mass (g)	Source
Tritemnodon agilis	10780.00	[63]
Tritemnodon strenuus	4702.60	PBDB + regression
Trogolemur amplior	60.27	PBDB + regression
Trogolemur myodes	32.59	PBDB + regression
Trogomys rupinimenthae	12.55	[329]
Trogosus castoridens	25945.98	PBDB + regression
Trogosus grangeri	42730.29	PBDB + regression
Trogosus latidens	86402.00	PBDB + regression
Tubulodon atopum	186.75	PBDB + regression
Tubulodon taylori	113.11	PBDB
Tuscahomys medius	84.22	PBDB
Tuscahomys minor	44.37	PBDB
Tylocephalonyx skinneri	1696671.69	PBDB
Uintaceras radinskyi	516493.43	PBDB + regression
Uintacyon asodes	6096.02	PBDB + regression
Uintacyon massetericus	3203.72	PBDB + regression
Uintacyon rudis	1864.20	PBDB + regression
Uintanius ameghini	48.28	PBDB + regression
Uintanius rutherfurdi	51.17	PBDB + regression
Uintasorex montezumicus	2.40	PBDB + regression
Uintasorex parvulus	6.11	PBDB + regression
Uintatherium anceps	523092.59	PBDB + regression
Untermannerix copiosus	121.51	[329]
Unuchinia dysmathes	114.44	PBDB
Uriscus californicus	130.13	PBDB + regression

Continued on next page

Table A.1 – continued from previous page

Species	Mass (g)	Source
Urocyon cinereoargenteus	3829.00	PBDB
Ursavus brevirhinus	80000.00	NOW
Ursavus pawniensis	61697.58	[329]
Ursavus primaevus	90000.00	NOW
Ursus abstrusus	43499.17	PBDB + regression
Ursus americanus	93431.00	PBDB
Utahia carina	17.88	PBDB + regression
Utahia kayi	37.51	PBDB + regression
Valenia wilsoni	279.17	PBDB + regression
Vassacyon promicrodon	5427.03	PBDB + regression
Viverravus acutus	483.44	PBDB + regression
Viverravus gracilis	849.04	PBDB + regression
Viverravus laytoni	241.14	PBDB + regression
Viverravus lutosus	602.46	PBDB + regression
Viverravus minutus	602.46	PBDB + regression
Viverravus politus	1009.45	PBDB + regression
Viverravus rosei	148.81	PBDB + regression
Viverravus sicarius	1594.41	PBDB + regression
Vulpavus australis	1746.94	PBDB + regression
Vulpavus palustris	2833.89	PBDB + regression
Vulpavus profectus	3388.44	[360]
Vulpes stenognathus	7331.97	[329]
Vulpes velox	2197.86	[305]
Washakius insignis	111.45	PBDB + regression
Washakius izetti	73.44	PBDB + regression

Continued on next page

Table A.1 – continued from previous page

Species	Mass (g)	Source
Washakius woodringi	48.28	PBDB + regression
Wilsoneumys planidens	51.70	PBDB
Worlandia inusitata	73.44	PBDB + regression
Wyolestes apheles	1594.41	PBDB + regression
Wyolestes iglesius	1288.90	PBDB + regression
Wyonycteris chalix	8.61	PBDB + regression
Xenicohippus craspedotum	10895.00	[198]
Xenicohippus grangeri	12284.22	PBDB + regression
Ysengrinia americana	110194.25	[329]
Yumaceras figginsi	293607.76	[329]
Yumaceras hamiltoni	247706.54	[329]
Yumaceras ruminalis	314896.72	[329]
Zapus burti	21.33	[329]
Zapus rinkeri	27.28	PBDB + regression
Zapus sandersi	18.92	[329]
Zemiodontomys burkei	120.30	PBDB + regression
Zetamys nebraskensis	60.34	[329]