Palaeomacroecology: Large Scale Patterns in Species Diversity Through the Fossil Record

Matthew J. Vavrek

Doctor of Philosophy

Department of Biology

McGill University

Montreal,Quebec

2010-06-31

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Doctor of Philosophy

DEDICATION

I would like to dedicate this thesis to my family: Meaghan, Callum and the little one to come.

ACKNOWLEDGEMENTS

I would like to thank all the people in my lab for their help with this project in so many ways, with a special thanks to Luke Harrison. My supervisors, Brian McGill and Hans Larsson, both deserve my thanks as well for guiding the thesis through the waters of academia. I would especially like to thank my wife Meaghan, who helped so much to make this thesis better, through not just editing but also through all her emotional support.

ABSTRACT

Palaeomacroecology is the study of large scale patterns of species diversity in the fossil record, encompassing a variety of subtopics. This thesis also addresses a variety of these subtopics, making it difficult to define under one heading.

The first portion of the thesis deals with a new package of software tools for the analysis of large scale datasets, with a specific focus towards palaeoecology and palaeogeography. These software tools have been combined into a package called `fossil` that has been released on the Comprehensive R Archive Network (CRAN), and is already being used by other palaeoecologists. While the majority of these tools had a basis in previous statistical methods, I have also independently developed a clustering algorithm for use with biogeographic datasets. This clustering algorithm is relational, non-Euclidean and non-hierarchical and as such is called Non-Euclidean Relational Clustering (NERC). NERC eliminates several of the assumptions common to most other clustering methods that are often violated by biogeographic data.

The next portion of my thesis describes a new Triassic aged flora from Axel Heiberg Island in Nunavut. Macroecological studies typically use large databases compiled from individual samples; therefore, these individual samples represent the foundation on which macroecological analyses rest, and collection and description of new fossil bearing sites is vital to the advancement of palaeomacroecology. Chapter 5 is an analysis of the provinciality and beta diversity of dinosaurs in the Late Cretaceous of North America. This analysis found that contrary to previous

studies, dinosaur genera were widespread across the continent and not restricted to small geographic ranges. Chapter 6 is the final culmination of my thesis, and where I see palaeomacroecology headed in the future. It is an analysis of how latitudinal diversity gradients in plants have changed through time. The analysis assesses the impact of changing climate in creating and sustaining the latitudinal diversity gradient, and lends support to the idea that temperatures are important drivers of the gradient.

The final chapter is a summary of where palaeomacroecology has been, and where its future work might be best focused. While the field of palaeontology is vital to our understanding of large scale, especially temporally, patterns of species diversity, the field of palaeontology has an opportunity to advance our understanding at an even more rapid pace provided we ask the appropriate questions of our data.

ABRÉGÉ

La palaeomacroecology est l'étude des modèles à grande échelle de la diversité des espèces dans les archives fossiles, et inclue une variété de sous-thèmes. Cette thèse adresse aussi une variété de ces sous-thèmes, ce qui en fait diffucult de définir sous une seule rubrique.

La première partie de la thèse discute d'un nouvel ensemble d'outils logiciels pour l'analyse des ensembles de données à grande échelle, avec une attention particulière à la paléoécologie et la paléogéographie. Ces outils logiciels ont été combinés dans un paquet appelé `fossil` qui a été publié sur le réseau Comprehensive R Archive Network (CRAN), et est déjà utilisé par d'autres palaeoecologists. Bien que la majorité de ces outils avait une base en preious méthodes statistiques, j'ai aussi développé indépendamment un algorithm de regroupement pour une utilisation avec des bases de données biogéographiques. Cet algorithme de regroupement est relationnelle, non-euclidienne et non-hiérarchique et en tant que telle est appelé Non-Euclidean Relational Clustering (NERC). NERC élimine plusieurs des hypothèses communes à la plupart des autres méthodes de classification, et qui sont souvent violées par des données biogéographiques.

La partie suivante de ma thèse décrit une nouvelle flore du Trias à l'île Axel Heiberg, au Nunavut. Les études macroécologiques utilisent généralement de grandes bases de données compilées à partir des échantillons individuels et, par conséquent, ces échantillons individuels représentent le fondement del'analyse

macroécologique, et la collecte et la description des nouveaux sites fossiliféres est indispensable à l'avancement de la palaeomacroecologie.

Le Chapitre 5 est une analyse du provincialisme et de la diversité bêta des dinosaures aux Crétacé supérieur en Amérique du Nord. Contrairement aux études précédentes, cette analyse a révélé que les genres de dinosaures ont été beaucoup plus répondus á travers le continent et ne se limitement pas à de petites aires geographiques. Le Chapitre 6 est l'aboutissement final de ma thèse, où je vois dans quelle direction se dirigé à la palaeomacroecologie. Il s'agit d'une analyse de la façon dont les gradients de diversité des plantes ont changé au fil du temps. L'analyse évalue le rôle des changements climatiques dans la création et le maintien du gradient latitudinal de diversité, et soutient l'idée que les températures sont d'importants moteurs de ce gradient.

Le dernier chapitre résume l'évolution palaeomacroecologie dans quelle direction les travaux futurs devraient être orientés. Bien que le domaine de la paléontologie sait vital pour notre compréhension des modeles de la diversité des especés à grande échelle, en particulier celle temporelle, le domaine de la paléontologie a une occasion de faire progresser notre compréhension à un rythme encore plus rapide, à condition de poser les bonnes questions.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Contribution of Authors

This thesis is a compilation of related manuscripts, four of which are being prepared to be or have been submitted, and two previously published papers. For all of these papers I am listed as the primary author, as I was the one in charge of data collection, experimental design and preparation of the manuscript. For several of the manuscripts, Hans Larsson is listed as a second author, as he was actively engaged in developing several of the ideas for the papers, as well as providing a useful help in writing the papers. Also, for Chapter 4 (Triassic Flora) Natalia Rybczynski is listed as third author for her role in the fossil collection and excavation.

CHAPTER 1
Introduction to and Overview of the Field of Palaeomacroecology

1.1 Introduction

Macroecology can be broadly defined as the study of ecology over large temporal, spatial and taxonomic scales (Brown 1995; Blackburn and Gaston 2002). It is an amalgamation of several aspects in more traditional areas of study such as community ecology, biogeography and palaeontology (Blackburn and Gaston 1998). Though the study of macroecology was only clearly articulated in 1989 by Brown and Maurer (1989), the beginnings of research in the field can be traced to many of the great naturalists of the 19$^{\text{th}}$ Century (Darwin 1859; Watson 1859; Wallace 1878; Smith et al. 2008). The study of macroecology has rapidly expanded over the last two decades as evidenced by a steadily increasing number of papers being published on the subject every year (Smith et al. 2008) and new journals founded dedicated to the science (Brown 1995; Blackburn and Gaston 2002). In large part the expansion has occurred so quickly due to the coincident appearance of large multivariate taxonomic datasets and large amounts of computing power, allowing for analyses today that were simply not possible even several years ago.

1.2 The Importance of Palaeomacroecology

Palaeomacroecology, or the study of macroecology in the fossil record, plays an important role in determining how patterns of species diversity have changed throughout the evolution of life (MacFadden 2005). At present, many areas of macroecology focus on patterns rather than process due in large part to the non-experimental nature of macroecology (Platt 1964; Willig et al. 2003). Manipulative experiments at the scales needed for macroecology are both unfeasible and unethical, and make testing the mechanics of macroecological drivers difficult

(Blackburn and Gaston 1998; Willig et al. 2003). As well, the most long term ecological experiments are typically no more than a few decades old, making it difficult to observe long term ecological patterns (Silvertown et al. 2006). However, the fossil record offers the ability to study the important dimension of time. Observations of the changing ecologic patterns throughout geologic history can be compared to large shifts in geography and climate. In this way, it is possible to study the mechanism and relative importance of these shifting drivers in creating the large scale patterns seen today.

1.3 Scale in Palaeontology and Palaeoecology

Macroecology uses both extant and fossil data to answer questions of ecology over large scales; however, there are several fundamental differences between the two types of data. The largest difference between extant and fossil data is possibly that of temporal scale. Extant datasets have the advantage of looking at fine temporal scales; for example, days, months, and years. Fossil datasets, on the other hand, can often only be refined to hundreds of thousands of years due to the difficulty of precise dating and correlating fossil specimens. Fossils can often only be identified to within a million years of their true age, which means that fossil sites that are considered contemporaneous may have not actually existed at the same time.

Another difference between extant and fossil data is the availability of additional samples. In an extant study new samples are always theoretically, if not practically, accessible if more data is needed. However, fossil localities are often isolated in their exposure and so obtaining a well distributed set of data points can be quite difficult. Collecting additional specimens from even a known locality may be impossible if the locality has already been exhausted by previous sampling, as is often the case where fossils are restricted to isolated, depositional lenses.

Despite these setbacks, it would be foolish to dismiss palaeontology as having little benefit to our understanding of ecological matters. As the macroecological literature has shown, studies over large scales may show different trends and processes that would not have been known if only small spatial scales were studied. For example, while the latitudinal gradient in species diversity (i.e., the presence of more species at the equators than the poles) is well known, the pattern only manifests itself over large (i.e. $>$ 300 km) spatial scales (Willig et al. 2003; Mittelbach et al. 2007). Likewise, long time scales are needed for speciation and extinction events to manifest. These events are the fundamental cause of patterns of species richness yet they typically only occur over millions of years (Stanley 1985; Magallón and Sanderson 2001).

Palaeomacroecology is a logical extension of much of the work going on in macroecology today. As defined, macroecology encompasses studies over large scales and can be thought of in three orthogonal dimensions: space, taxonomy, and time. While extant ecological data can span large spatial and taxonomic scales, there is simply no possible way for us to conduct observational studies over truly long temporal scales. However, the fossil record does preserve these long temporal scales, and by using fossils we can expand our knowledge of macroecology into this third dimension.

1.4 The Adequacy of the Fossil Record*

*after Donovan and Paul (1998)

The fossil record is often faulted for gaps and lack of information. However, based simply on Occam's razor, the fossil record should preserve faithfully, if incompletely, the history of life (Donovan and Paul 1998; Benton et al. 2000). A similar situation of completeness and adequacy can be drawn in neontology [the study of recent life; Paul (1998)]. Despite the limited ability to know about all living organisms and their interactions, attempts are made that make broad conclusions based on this incomplete, although accurate, knowledge. It is important to note that differences exist between data collection in palaeontology and neontology. However, so long as these differences are recognized and accounted for, there is no reason that the information obtained from the fossil record is not more than adequate to answer important ecological questions (Paul 1998).

One factor that must be accounted for in palaeoecological data is taphonomy (Markwick and Lupia 2002). Taphonomy is the study of what happens to an organism after death until the time it is discovered as a fossil (Martin 1999). This includes both biostratinomy, or what happens to the organism before it is buried, and diagenesis, or how an organism becomes fossilized and its subsequent travel through the rock record. It is important to keep both processes in mind when analyzing the diversity of the fossil record, as both affect the type, quantity and quality of the fossils found. In general, biostratinomy will affect the types and numbers of organisms within each locality and diagenesis can determine which localities are preserved.

In regards to biostratinomy, when an organism dies or sheds parts (i.e. teeth, leaves, branches) these are subject to various modes of transport, sorting, burial, and degradation. However, biostratinomy can affect different remains in different ways. For example, resilient material, such as logs, can be transported great distances and are highly resistant to abrasion and degradation. Such resilient material can avoid degradation on the order of hundreds of years, and are therefore quite commonly preserved. On the contrary, flowers and soft animal tissue are typically very fragile and degrade quite quickly, often within weeks, and are much rarer in the fossil record (Burnham 1993).

With more resilient remains, biostratinomic processes can alter the composition of fossil assemblages through an effect called time averaging (Markwick and Lupia 2002). Time averaging occurs when organic remains or fossils from two separate times become mixed together (Behrensmeyer and Hook 1992; Behrensmeyer and Chapman 1993). For example, resilient remains such as pollen and mollusc shells can accumulate over long periods of time, meaning that in some cases organisms may be found together even though they may not have lived at the same time or place. One way to avoid the bias of time-averaging is by using fossils that are not as resilient and prone to the effect, such as flattened (compression-impression) leaf assemblages (Burnham 1993; Wing and Dimichele 1995; Martin 1999). Compression-impression assemblages are composed of flattened organs, primarily leaves, that have abscissed or otherwise fallen from the parent plant and are then buried in generally horizontal planes (Martin 1999). They typically occur in slower, fluvial environments (Martin 1999). As leaves usually decompose in less than 6 months if not buried, these assemblages usually represent a single growing season (Burnham 1993). Although there will be time averaging within a growing season, this is random between sites and has relatively minor effects (Wing and Dimichele 1995). Another way of dealing with time averaging is to assume its effects are non-significant and affect each locality in the same way. Even extant ecological studies cannot make all observations simultaneously, and so studies involving large numbers of observations will use data that has been averaged over some amount of time.

Another common criticism of fossil assemblages is the uncertain area that they represent. Organisms can be transported by wind, water, or biological vectors to their final resting place, which at times may be hundreds of kilometres from their original home (Martin 1999). Assemblages in which the remains have been transported from the site of death and out of the original habitat are called allochthonous (Behrensmeyer and Hook 1992). Not only do allochthonous collections occur away from the habitat in which they occurred, but due to the species-area effect (discussed further in subsection 1.5.1) the diversity at these sites is often much higher than local, autochthonous (preserved at the exact site they were discarded) assemblages.

Although allochthonous assemblages can be a major concern with resistant remains, there is two different ways the issue can be avoided. First, if non-resistant remains are used, there is less chance that the material can be transported before breaking down. Leaf assemblages are often not strictly autochthonous but are typically parautochthonous; this means that they may have moved from the site of death but are still within the original habitat (Spicer 1981; Burnham et al. 1992; Burnham 1993; Wing and Dimichele 1995). Though there is no way to directly measure the size of the area that the fossils originally came from, actualistic studies on extant ecosystems have found that leaf litter is typically derived from the surrounding 1000 to 3000 square metres (Chaney 1959; Scheihing and Pfefferkorn 1984; Spicer and Wolfe 1987; Burnham et al. 1992; Burnham 1993). Chaney (1959) collected plant remains from 19 pools in a river in California and determined the plants represented came from within 15 m (50 feet) of the site of deposition. Although some argue that massive storm events and high flow river systems may transport debris further, a study of higher energy storm deposits by Spicer and Wolfe (1987) found that the plant remains deposited represented well the locally growing species, with only one example of 1 x 10^6 specimens from distant upland areas. Such studies suggest that more regular events are just as important for plant deposition and that plant remains are highly degraded and abraded when transported over long distances (Scheihing and Pfefferkorn 1984). The impact of this phenomenon means that only locally deposited remains will likely be identifiable, thus removing the possible bias of allocthonous assemblages. This bias can be further recognized from examination of the local depositional setting, which is typically included in the description of the flora; sites that are indicative of massive flooding events may be excluded when comparing sites.

Although we could limit fossil studies to non-resistant remains, doing so would eliminate a large portion of the fossil record. A second way of dealing with the issue of long distance transport is by expanding the grain [sensu Palmer and

White (1994)] to encompass a larger area. Individual rock formations are typically confined to single depositional basins, and so even if an organism is transported some distance it would still remain within this basin. Many recent studies of the latitudinal diversity gradient have a grain on the order of thousands of kilometres squared, similar in extent to many rock formations.

Though not strictly taphonomy, a final consideration when using the fossil record is the abundances of species found within deposits. Larger plants typically produce more organic debris; for example, large deciduous trees can have tens of thousands of leaves, while small ferns may have fewer than 20. Although different plants produce differing amounts of debris, studies have shown that the rank-abundance of common vegetational elements is typically preserved in even small leaf deposits (Burnham et al. 1992). Small samples of 350 specimens in extant leaf litter were found to consistently contain all the common species in an area and suggest that relative abundances can be captured in small, restricted samples (Burnham et al. 1992).

The fossil record does not completely record the history of life, but the signals it does preserve are an accurate picture of biodiversity through time (Benton et al. 2000). Fossil assemblages, and especially leaf compression-impression collections, can be thought of as 'snapshots' of a moment in time (Wing and Dimichele 1995). These assemblages are relatively comparable to one another and extant equivalents when their biases are properly understood, and provide "the best hope for understanding global diversity patterns" (Johnson 2003).

1.5 Biases in Sampling and Analyzing Ecological Data

In any ecological analysis there are problems to be aware of, and many of which are common to both neoecology and palaeoecology. Any large-scale study, including those in palaeoecology, need to use clearly defined terminology. To illustrate, species richness is a general measure of biodiversity, but diversity is closely related to area, time, and taxonomic scales. Any comparison of biodiversity values across samples that differ in any of these three scales are not valid. Each type of scale will be discussed below.

1.5.1 Area

The species-area relationship is one of the oldest recognized relationships in species diversity (Arrhenius 1921; Gleason 1922; Preston 1960; Rosenzweig 1995). Simply put, it states that the larger an area surveyed, the more species will be found. This relationship is often expressed as $S = cA^z$, where S is the number of species, A is the area, and c and z are constants. Many studies, both theoretical and empirical, have shown the z-value of this equation to be approximately 0.25 (Rosenzweig 1995; Crawley and Harral 2001), although different values are also common. For example, data accumulated within a biogeographic province generally has a slope less than 0.25, while studies across continental biogeographic provinces have shown values closer to 0.9 (Barnosky et al. 2005).

There is strong evidence in the fossil record of a species-area effect (Smith 2001; Barnosky et al. 2005), although it is not always directly due to area *per se.* The species-area effect in fossil assemblages can occur because of differences in the amount of exposed rock; time periods with more rock exposure have higher species diversity (Raup 1976). The difference in the amount of rock exposure can come from several sources, with the "pull of the recent" the best known and understood (Raup and Crick 1979; Rosenzweig 1995; Martin 1999). This "pull" refers to the fact that older rocks are more likely to be eroded or subject to some type of alteration, destroying the fossils preserved, when compared to younger rocks.

The simplest way of countering the species-area effect for my analyses is to avoid the problem as much as possible. For example, in a study on latitudinal diversity gradients (Chapter 6), individual localities were used, rather than combining localities when determining species richness (Ziegler et al. 1993). By using individual localities, the area encompassed is comparable between all individual data points.

1.5.2 Time

Both a boon and bane to palaeoecological analyses are their use of vast temporal scales. Because the fossil record is so long, one can examine how speciation and extinction create patterns of species diversity at a global scale (Rosenzweig 1995). On the other hand, studies of extant diversity tend to look at time scales in years or decades, often too short to observe speciation/extinction events. However, because palaeoecology often investigates patterns on time scales of millions of years, extinction and speciation can have unintended effects on estimates of biodiversity. For example, if we compare two temporal units, one containing a single species and the other containing two species, we could conclude that the unit with two species had a higher diversity. However, if the unit with one species spanned one million years while the unit with two species spanned 20 million years, the former unit would likely be considered more diverse. Wherever possible, palaeontological data should be compared with equivalent time units to eliminate such problems, although the use of such time-averaged faunas is a necessary constraint on Mesozoic global biogeographic syntheses (Raup and Jablonski 1993; Jablonski and Raup 1995; Crame 2002; Markwick and Lupia 2002).

For the analytical studies in this thesis, I have attempted to make the time slices as narrow as possible while still retaining enough localities per time slice. In an analysis of dinosaurian beta diversity (Chapter 5), I used the same duration of time as previous studies in order to keep my results directly comparable. For an analysis of fossil latitudinal diversity gradients (Chapter 6), geologic ages and epochs were used, as too many fossil localities would have been eliminated otherwise. However, the time divisions used are still as good or better than many other large scale studies of species diversity in the fossil record.

1.5.3 Sampling Methods

Closely related to the criticism of the adequacy of the fossil record is the sampling methods that accompany it. Fossil localities are not always easy to come by, and so palaeontologists must make the most of those that exist. This results in an uneven sampling of species diversity across time and space. Some studies have used methods which grouped, or "binned," localities based on latitudinal or continental delineations for various methodological reasons (Ziegler et al. 1993; Rees et al. 2004; Leighton 2005). This method of binning combined with the uneven sampling can lead to false signals because bins of equal size may have been sampled very differently. For example, even though Australia has relatively few dinosaurs known, this is likely due more to the lack of appropriate rock than to an actual lack of dinosaurs. Although we could then compare the diversity of Australian dinosaurs to other continents, it would give misleading results.

This suggests that the best way to analyze large palaeontological datasets is through analyzing large numbers of individual localities (Alroy et al. 2001). This requires the availability of large databases of species occurrences. Where it is available, abundance data is also beneficial, as local abundances are a reasonable proxy for sampling intensity and can be used for estimating species richness (Vermeij and Herbert 2004; Jackson et al. 1999; Jackson and Johnson 2001; Kidwell 2001). In the chapter on dinosaur beta diversity (Chapter 5), I use abundance data for this exact purpose. However, in some cases species occurrence data is all that is available and previous studies have shown useful results from occurrence-only data sets (Cecca et al. 2005). In the chapter on latitudinal gradients (Chapter 6), I use occurrence data only, because there is not enough abundance data available for fossil plant localities at this time without using coarser time divisions.

Another common problem in many data sets is the quality of the taxonomic data sampled. Because large databases typically contain primary data collected by multiple researchers, there is often a concern for consistency within the data. This concern is equally prevalent in neoecology as it is in palaeoecology. However, some experimental data has shown that patterns present within a data set can be very resistant to error. Sheppard (1998) showed that basic patterns of clustering remained the same for coral reef samples even when 25% of the data was randomly altered. As well, a study of Cambrian trilobites found that results gathered from published records with up to 70% inaccuracy shared the same major trends as independent, collected field data (Westrop and Adrain 2001). Although finer scale patterns may be obscured by taxonomic errors, large scale patterns stand up well to any errors (Adrain and Westrop 2000; Westrop and Adrain 2001).

In addition to errors of taxonomic misassignment, the way in which Recent and fossil species are determined is very different. There are many species definitions with the biological species concept the most widespread, however the fossil record is forced to use a morphological species concept. Morphological differences are generally apparent in extant biological species, but may be cryptic in extinct species. For example, for all the plumage and call differences in North American finches, there are few skeletal differences that would be preserved in the fossil record. This dependence on a morphological species concept surely underestimates the biological species diversities of extinct life. Fossil vertebrate genera also tend to be either monospecific or otherwise lower in species number than comparable extant genera. To avoid some of these biases, biodiversity measures are derived from genera in this thesis, rather than species. Although pure alpha diversities are still expected to be lower for fossils than extant systems, the beta diversities are expected to be more comparable because this metric is derived from a ratio, rather than absolute diversities.

Area, time, and sampling methods are all important considerations in any ecological analysis. In order to reduce the amount of bias in the data used, it is important to explicitly state the scales and methods used. This also makes the replication of analyses with new data sets simpler, creating reproducible and comparable research.

1.6 Palaeomacroecology: A Thesis

I have titled this thesis under a more general heading – Palaeomacroecology – because it addresses several related but significantly different aspects of the field. The goal of my thesis was to create a framework where large scale patterrns of species diversity in the fossil record could be first quantified and compared to shifting drivers of diversity over geologic time scales. Chapters 2 and 3 describe the software and statistics that were developed in order to carry out the analyses for the remainder of the chapters. I have tried to make the methods easy to find and use, so that others need not duplicate the work already done. The software is already available online, and is being actively used within the palaeontological community. Chapter 4 describes a new flora from the High Arctic. Since the best way to study species diversity in the fossil record is with individual localities, this requires palaeontologists to find and describe new localities. This flora was collected from Triassic-aged rocks on Axel Heiberg Island in Nunavut, and is one of the most speciose sites known from this time and region. Chapter 5 presents an analysis of dinosaurian biogeography using both estimates of beta diversity and cluster analysis. Previous studies suggested distinct regional dinosaur faunas in North America, however my work found no statistical support for distinct faunas. This new work suggests that dinosaurs were relatively mobile animals that were able to colonize vast tracts of land and were not restricted in their ranges as once thought. The main focus of my thesis however is articulated in Chapter 6, an analysis of how the latitudinal diversity gradient has changed through time in plant assemblages of North America. By examining the gradient over a time series, we can begin to correlate changes in the intensity of the gradient and how this relates to shifts in driving forces. The latitudinal diversity gradient appears to change in response to changing climate, although there are likely other driving factors that need to be investigated. Finally, Chapter 7 presents a brief overview of the broader field of palaeomacroecology, and gives a number of suggestions as to how palaeontologists can make the most impact by integrating their work more closely with modern ecological work.

Bridging Text

Chapter 2 has been submitted and accepted for publication to Palaeontologia Electronica. The paper describes the "fossil" package, a suite of palaeoecological and palaeogeographical software tools that I developed in order to carry out most of the statistical analyses for my thesis. The paper provides a general introduction to statistics using the R Statistical Language (R Development Core Team 2010) and discusses the implementation of the various functions. The functions include a large number of species diversity measures, such as species similarity measures (Sorenson, Jaccard) and species richness estimators (Chao 1 and 2, Jackknife, ICE), and geographic analysis tools, such such as spherical surface area calculators. The different functions are illustrated with an example data set and sample graphs.

Creating this package allowed me to become much more efficient while doing the various data analyses in my thesis. Instead of having multiple copies of the same lines of code, putting together a package that contained most of the functions I used meant that I did not have to waste time retyping, and more importantly, debugging the functions themselves. I feel that being able to demonstrate a proficiency in basic computer programming is becoming an essential tool for virtually any scientist, but it is especially true when the work is primarily *in silico* and on large, complex, multivariate data sets.

Full citation:

Vavrek, M.J. fossil: Palaeoecological and Palaeogeographical Analysis Tools. Palaeontologia Electronica, in press.

CHAPTER 2
fossil: Palaeoecological and Palaeogeographical Analysis Tools

2.1 Abstract

The `fossil` software package is a collection of analytical tools to synthetically analyse ecological and geographical data sets. The software is designed to be used with the R Statistical Language, and is under an Open Source license, making it free to download, use or modify. The package includes functions for estimating species richness, shared species/beta diversity, species area curves and geographic distances and areas. The package also contains extensive documentation and examples of how to use all of the functions.

2.2 Introduction

Multivariate analyses in palaeontology have become an increasing focus of many palaeontological research programs, especially with the development over the past decade of large datasets (i.e., Paleobiology Database; Carrasco et al. 2005) and readily available computing power. A variety of statistical programs and software has been used and developed by and for palaeontologists, ecologists, and evolutionary biologists as these massive data sets have become more commonplace (e.g., Hammer et al. 2001; Colwell 2009; Harrison and Larsson 2008; Maddison and Maddison 2009).

Large databases necessarily involve large numbers of collaborators, which may lead to an issue of heterogeneity and incompatibility of computing platforms and file formats. Despite the large number of freely available programs, there are few truly cross platform solutions available. One statistical environment gaining recognition over the last decade with its ability to perform intensive statistical analyses has been the R Statistical Language (R Development Core Team 2010;

Ezard and Purvis 2009). This software is cross platform, freely available (Open Source) and has an extensive installed user and contributor base. While the base software when installed can perform many common statistical procedures, the software is easily extensible through packages, such as phylogenetic analysis (Paradis et al. 2004), time series analysis (Hunt 2008) and palaeobiological phylogenies (Ezard and Purvis 2009). These packages are available through a central repository called the Comprehensive R Archive Network, or CRAN. Additionally, data from virtually any source can be used, from plain text and Microsoft Excel tables to images and GIS shapefiles, and graphs and figures can be output in virtually any format. This flexibility and availability is what has made it a growing success in the field of statistics and database analysis.

Here I present a new package that has been developed to enable a selection of ecological and geographic analysis tools to be added to the base R environment. The package was originally developed with palaeontologists in mind, and is appropriately entitled `fossil`. As of this writing, it is in version 0.3.2, and although there are planned additions to the code, the functions already present allow for a large number of analyses to be performed.

Reasons for developing `fossil` are many fold. The underlying impetus was to create a single package to examine large datasets with up to date methods of biodiversity estimators and ecological pattern recognition that can be used in conjunction with geographic data over long time scales. Macroecological analyses in palaeontology are a growing field, and have the real opportunity to answer modern questions of biodiversity distributions, thanks in large part to the deep time of the fossil record. By providing powerful tools that integrate well, we can spend more time on the questions rather than the methods.

A number of the functions that have been implemented in `fossil` can also be found in the excellent package `vegan` (Oksanen et al. 2010). Many of the species diversity and species estimator functions are implemented in both packages. However, the fossil package was implemented to cover a number of use cases that `vegan` did not cover. Initially, the primary function that was needed was a way to estimate species diversity using a number of functions all at once. As well, the function to create distance matrices with user defined measures was at the time more difficult to use, and so I have tried to implement a more easily extensible method. The fossil package also implements a number of spatial analysis and export tools that are not found within vegan, such as methods to calculate geographic distances and areas from a set of points.

For example, the fossil record, while accurate, is by no means complete (Benton et al. 2000) yet can still provide important information on biogeographic patterns. Using fossil, we can compare sparse ecological data with a number of ecological similarity indices (i.e. Chao-Jaccard, Chao-Sorenson, Simpson) and then observe the patterns of connectivity using various types of neighbour joining techniques. These patterns can then be visualised in ecological space, using ordinations to group similar sites, and in geographic space, placing localities on a map and observing how this ecological connectivity relates to geography. Combining spatial, ecological and temporal data can provide a more complete picture of the evolution of the biosphere than any one factor alone.

2.3 What is R and why sould we use it?

`fossil` is constructed for use with the R Statistical Language. R owes it's origins to the S Language, a program initiated at Bell Labs in the 1970's as a way to implement a computational statistical language (Becker et al. 1988). The S Language has been the basis for another well known statistical program, S-PLUS. In 1991 Ross Ihaka and Robert Gentleman at the University of Auckland began developing a statistical language for their teaching laboratory as no adequate commercial solution existed at the time. Their work mimicked many of the styles and methods of S, and eventually this evolved into the R Language for Statistical Computing (Ihaka and Gentleman 1996). Since it's origins, R has been open-sourced under the GNU Public License, meaning that anyone who chooses to use, redistribute or improve the software is free to do so provided they allow others the same rights (Stallman 1999). The program was originally written for a Macintosh system, but it has since been ported to virtually every computing architecture, both legacy and modern. This makes it an ideal candidate for a statistical system in many modern laboratories, where every researcher possesses their own (if not multiple) computers, often with different operating systems.

Many other statistical programs encourage their users to manually select their data and choose the analyses to be run with a mouse cursor. At first glance this is a much simpler way of interacting with the data, but it suffers from a major drawback; analyses of this type are not truly reproducible (Leisch and Rossini 2003; Green 2003). Although descriptions of statistical procedures used in refereed papers is a must, trying to record exact mouse clicks and button selections is virtually impossible. R on the other hand encourages users to record each and every step of the process used. Most users of R will write their methods of analysis out in a text editor of some kind and then proceed to run this code in the R environment, with every step, from analysis through to figure creation, fully documented.

The deeper benefits of this method may not at first be obvious either. I have personally experienced situations where mistakes were made early on in the process of data analysis and not found until much later. While in a graphical, mouse driven environment trying to repeat all the steps necessary is often time consuming, well written R code can be easily modified and re-run with minimal fuss. Further, as the program is consistent across platforms, collaborators can run the code on their platform of choice, without having to worry if their version of a program has the same available functions. This benefit also extends to other scientists, who by taking other researchers' code can re-run published findings exactly, without having to purchase software of any kind.

What follows is not an in depth introduction to R; there have already been many books written on the subject. For a good start, the original text by Becker et al. (1988) and a more recent text by Braun and Murdoch (2008) are highly recommended. Rather, the focus of this paper is the use of the functions found within the `fossil` package.

2.4 Setting up the Environment

R is available for virtually any platform and can be installed from the R Project website, `http://www.r-project.org/`. Please note that throughout this paper all R commands are distinguished from the text using a `monospace` font. All commands are preceded by a chevron (`>`) that does not need to be entered, but simply represents the beginning of a new command.

Throughout much of this paper, I use a theoretical data set called `fdata`, consisting of three parts. `fdata.list` is a table with each row representing an individual species occurrence, and columns for locality name, species name, species abundance, latitude and longitude. `fdata.mat` is a matrix (12 by 12) with each unique species as a row and each locality as a column. The last part is `fdata.lats`, a SpatialPoints object containing the longitude and latitude for each locality. All of this data is found as part of the `fossil` package. As well, the entirety of the code used to analyse the data and create figures for this paper is available as an appendix, along with full instructions on how to use it.

To begin using the `fossil` package in an interactive session, you must first ensure the package has been installed on your computer. It is available online from CRAN, and can be downloaded from within an R session by typing `install.packages('fossil')` at the command prompt. You will be prompted to choose a download location; simply try to choose one closest to your location. Once the fossil package is available on your computer, you can load it in to R using the command `library(fossil)`. Every time you start a new session, you will have to load the package again using the `library()` command as extra libraries are not loaded by default to keep the memory use as low as possible.

```
> library(fossil)
```

2.5 Loading your data in R

Large databases used in palaeoecology studies are often simply tables, whether in plain text files or Excel tables, where every row consists of a unique observation, usually of a species at some location in space and time. However, the species, locations and times in these lists are rarely unique, and often consolidation of the data into usable matrices of species versus location is needed. There are two functions that aid in the conversion of lists of points into two types of matrices that will be referred to throughout the remainder of the paper. The first function is the `create.matrix()` function, which is able to take a list of species and their occurrences and convert it to a matrix of species (rows) by localities (columns). With the commands

```
> data(fdata.list)
> create.matrix(fdata.list, tax.name = "species", locality = "locality")
```

we can create an occurrence matrix from the `fdata.list` example data set; alternatively, if we wish to create an abundance matrix, we use virtually the same command, but include the option `abund = TRUE` and give the name of the abundance column (in this case, `'abundance'`) for the `abund.col` option. This method will give us an abundance matrix identical to `fdata.mat`.

```
> data(fdata.list)
> create.matrix(fdata.list, tax.name = "species", locality = "locality",
+     abund = TRUE, abund.col = "abundance")
```

For the fossil package, data follows the convention of species as rows and localities as columns. Data that is in matrix format already but with species as columns and localities ass rows can be transposed with the t() command.

Similarly, much palaeontological data comes with some sort of spatial data about it's provenance integrated with the occurrence data. As such, the locality data is often duplicated for each unique species at a certain site. In order to simplify plotting georeferenced data, a function called `create.lats()` can be used to extract the site coordinates from a list, eliminating duplicate entries.

```
> data(fdata.list)
> create.lats(fdata.list, loc = "locality", long = "longitude",
+     lat = "latitude")
```

2.6 Distance/Similarity/Beta Diversity Indices

Measuring the ecologic distance between sets of samples is often a necessary first step in many multivariate analyses (Green 1980; Shi 1993). As such, it also is often a contentious one, with different researchers advocating different measures, at times with multiple correct arguments. Although I do not wish to provide a full explanation here of every single measure, I will provide a brief overview of those included in the `fossil` package. Some of these measures are best described as indices of beta diversity, although they are grouped here with other similarity measures for convenience, as they are typically used in a similar fashion.

All of the similarity functions can be used in the same way. The functions need two arguments, representing the two samples. It is important that the species occurrences are arranged in the same way for each site, and that any absent species are represented by a zero.

```
> sampleA <- c(1, 1, 0, 1, 1, 1, 1)
> sampleB <- c(0, 1, 1, 0, 0, 1, 1)
> sorenson(sampleA, sampleB)
[1] 0.6
```

The species estimator functions included can be broadly grouped into two categories, those that use occurrence data and those which use abundance data. As abundance data is not always available, especially in palaeontology, more measures that use occurrence data are included in the package. Occurrence based measures can also be used with abundance data, but the abundance matrix is converted to an occurrence matrix by the function.

One of the oldest and best known occurrence measures is the Jaccard measure, also known as the Coefficient of Community (Table 1, Jaccard 1901; Shi 1993). The measure has seen extensive use, largely due to its simplicity and intuitiveness (Shi 1993; Magurran 2004). A similar measure also in common use is the Sorenson measure (also known as Dice, Czekanowski or Coincidence Index), which places more emphasis on the shared species present rather than the unshared, as can be seen in the difference in values for the example data set. Again, the calculation is relatively simple and intuitive, and both indices have been shown to provide useful results (Wolda 1981; Hubálek 1982). Two other similar indices that are occasionally used are the Ochiai and Kulczynski measures. While Hubálek (1982) lists the Ochiai and Kulczynski indices as providing good results, the Jaccard or

Sorenson are typically more recommended if only because they are more commonly used.

One of the most common problems in palaeontology, and indeed in many ecological studies, is that of differing sample sizes. Comparing two sites of very unequal sampling intensities can give a biased view of the actual species overlap. For example, a subsample of a site could be considered identical to the original site, as all the species in the subsample will be within the original. However, all the previous measures would show less than complete similarity due to their mathematical properties. With this in mind, Simpson (1960) developed a measure which can account for variability of sample sizes. His formula scales the value by the number of species from the least sampled site, so that the subsample in this case would have full similarity with the original. The Simpson measure is often used with data that is highly variable in sampling intensity, such as fossil datasets, for this very reason.

Although the `fossil` package contains a number of occurrence based similarity indices, by no means are all included. For example, Shi (1993) lists 39 and Hubalek (1982) lists 43 different variations of similarity index, many of which are little used outside their original papers.

Though not as common in palaeontological data sets, abundance values can provide valuable information about a community that is not possible with occurrence data. Analyses of community structure are very limited without abundance data, and abundance data can provide more subtle distinctions between communities. As well, species abundances can provide some measure of sampling intensity.

Possibly the most widely used abundance based measure is the Bray-Curtis measure, due to its strong relationship to ecological distance under varying conditions (Bray and Curtis 1957; Minchin 1987; Faith et al. 1987; Clarke 1993). The measure is equivalent to the Sorenson coefficient when used as a similarity measure with occurrence data. The Morisita-Horn index, while not as common as the Bray-Curtis, is also a highly recommended measure due to its relative independence from sample size and diversity (Wolda 1981; Magurran 2004). While there are several variations of the measure, I have used the version found within Magurran (2004).

Luckily, though the diversity of indices may seem somewhat overwhelming, the package provides an easy way to use them with large data sets. An included function called `dino.dist()` will take a matrix of species occurrences versus locality (or any analogous groupings) and return a full pairwise distance matrix as output. This function is written such that any other similarity index, including those defined by other packages or by the user, can be specified and used to calculate the matrix.

2.7 Non-parametric Species Estimators and Rarefaction

An obvious problem in palaeontology is the incompleteness of the record, and therefore our incomplete knowledge of the number of species present, whether it be locally or globally. Modern ecologists suffer from the same problem, whereby it is impractical to sample every single member of even relatively small communities of organisms (Chazdon et al. 1998). However, smaller samples still contain important information about the community, and can be extrapolated from to provide estimates of the true richness of the total community. Of course, such extrapolations must account for sampling intensity and area (Gleason 1922; Preston 1948).

One of the most commonly used methods for dealing with unequal sampling intensity is rarefaction, or interpolation of the data (Sanders 1968). Rarefaction provides a method of comparison between different communities, whereby each community is "rarefied" back to an equal number of sampled specimens (Heck et al. 1975; Foote 1992; Colwell and Coddington 1994). Within the `fossil` package is a method for rarefaction known as a Coleman Curve (Coleman 1981; Coleman et al. 1982). This type of rarefaction is carried out through a resampling method rather than a rarefaction formula; resampling is computationally much simpler and faster, and provides indistinguishable results from the formula based method (Coleman 1981; Coleman et al. 1982; Colwell and Coddington 1994; Magurran 2004). The Coleman Curve is an empirical measure of the rarefied number of individuals, while the rarefaction function is a theoretical model of what the empirical curve would look like. Although rarefaction can be useful, it is very sensitive to the underlying pattern of species abundance, such that collections with much lower species evenness will often give lower estimates of species diversity than those with very even abundances, regardless if species diversities in reality are equal (See Gotelli and Colwell 2001, for an in depth treatment of the issue.).

Although rarefaction interpolates data back, non-parametric species estimators extrapolate from the data to find what the “true” number of species may have been (Colwell and Coddington 1994). The typical way these estimators operate is by using the number of rare species that are found in a sample as a way of calculating how likely it is there are more undiscovered species. As an example, the Chao 1 estimator (Chao 1984; Colwell and Coddington 1994) calculates the estimated true species diversity of a sample by the equation:

$$S_1^* = S_{obs} + \frac{F_1^2}{2F_2}$$

where S_{obs} is the number of species in the sample, F_1 is the number of singletons (i.e. the number of species with only a single occurrence in the sample) and F_2 is the number of doubletons (the number of species with exactly two occurrences in the sample). The idea behind the estimator is that if a community is being sampled, and rare species (singletons) are still being discovered, there is likely still more rare species not found; as soon as all species have been recovered at least twice (doubletons), there is likely no more species to be found. Tests of the estimator have shown that it does provide reasonable estimates, at least for modern data sets (Chao 1984; Colwell and Coddington 1994; Chazdon et al. 1998). Of course, as the value is an estimate there is a degree of uncertainty, and a method to calculate the variance for the estimators has been provided by Chao (1987) in the form of

$$var(S_1^*) = F_2 \left[\left(\frac{F_1/F_2}{4} \right)^4 + (F_1/F_2)^3 + \left(\frac{F_1/F_2}{2} \right)^2 \right]$$

Although the Chao 1 estimator works for abundance data, often only occurrence data are available. There is another estimator, named conveniently Chao 2 (Chao 1987; Colwell and Coddington 1994), which uses occurrence data from multiple samples in aggregate to estimate the species diversity of the whole. This estimator is (Fig. 2–2) defined as:

$$S_2^* = S_{obs} + \frac{Q_1^2}{2Q_2}$$

which is virtually identical to the Chao 1 estimator, with singletons (Q_1) being species occurring in only one sample and doubletons (Q_2) occurring in two samples. This estimator can also make use of the Chao 1 variance formula provided above, with the substitution of F_1 and F_2 for Q_1 and Q_2 respectively.

Chao and colleagues (Chao and Lee 1992; Chao et al. 1993; Lee and Chao 1994) have also published another pair of estimators, called the Abundance Coverage Estimator and the Incidence Coverage Estimator, which use abundance and occurrence based data sets respectively. These estimators are much more complex; the Abundance-based Coverage Estimator takes the form

$$S_{ace} = S_{common} + \frac{S_{rare}}{C_{ace}} + \frac{F_1}{C_{ace}}\gamma_{ace}^2$$

where S_{common} are the species which occur more than 10 times in the sampling, S_{rare} are those species which occur 10 times or less, C_{ace} is the sample abundance coverage estimator, and finally γ_{ace} is the estimated coefficient of variation for F_1 for rare species (See Chazdon et al. 1998, for a full explanation and definition of the estimator). In simpler terms, the formula uses the number of rare species (<= 10) and the number of singletons (F_1) to estimate how many more undiscovered species there might be. Although this formula is for the abundance estimator, virtually the same holds true for the incidence based estimator, except that instead of the species abundance, it uses the number of samples each species occurs in. Both of the coverage estimators have been found to give good results, and are highly recommended (Chazdon et al. 1998; Hortal et al. 2006)

Another estimator provided is the Jackknife estimator, developed by Burnham and Overton (1978, 1979) originally for use with capture/recapture studies. The formula

$$S_{jack1} = S_{obs} + Q_1 \left(\frac{m-1}{m} \right)$$

represents the first order version of the estimator; the variable m represents the total number of samples. Smith and van Belle (1984) also provided a second order variation, with the formula

$$S_{jack2} = S_{obs} + \left[\frac{Q_1(2m-3)}{m} - \frac{Q_2(m-2)^2}{m(m-1)} \right]$$

The second order Jackknife has shown to be one of the most effective estimators, and may be the best estimator at the moment for highly sparse palaeontological collections, as it is the least susceptible to sampling bias (Chazdon et al. 1998; Hortal et al. 2006).

Finally, for completeness I also provide the bootstrap estimator

$$S_{boot} = S_{obs} + \sum_{k=1}^{S_{obs}} (1-p_k)^2$$

developed by Smith and van Belle (1984). The bootstrap richness estimator has been generally regarded as one of the poorer species estimators, and Chazdon et al. (1998) in fact recommend against using it.

Though the various estimators vary greatly in their formulae, the functions within `fossil` take care of most of the nuances, and generally require only one argument, that being a species occurrence matrix or species abundance vector or matrix.

```
> data(fdata.mat)
> chao1(fdata.mat)
[1] 12.25
> jack1(fdata.mat)
[1] 12.98980
```

It is often best to use a number of these estimators in concert, as concurrence between their individual values can lend support to their results. Colwell (2009) has released a program for Windows called EstimateS which does exactly this; it can calculate multiple species estimators for a data set, along with their variances and a species accumulation curve. As Colwell's program is so useful, it was used as a template to create the function `spp.est()`. The function has several important options, namely the number of randomisations and whether or not to use abundance data. The `spp.est()` function calculates a rarefaction curve, the Chao, Coverage Estimators and Jacknife, as well as standard deviations for all the estimates. As a default the function will run 10 randomizations of the data, however for more accurate estimates a much larger number of randomizations should be run. It should be noted though that with a large data set and a large number of randomizations that the function may take a long time to complete. At this time, work has been undertaken to parallelize this function, enabling a large speed up in processing time when using a multicore or multiprocessor system.

2.8 Minimum Spanning Trees

Minimum Spanning Trees (MST) and the associated Minimum Spanning Networks/Forests (MSN) are a useful method of visually displaying relationships between samples, whether those samples are biogeographic or taxonomic in nature (Fig. 2–1, Gower and Ross 1969). The MST is closely related to the final product of a Single Linkage Cluster Analysis (SLCA Sneath 1957; Gower and Ross 1969) and connects all the points in a sample with the minimum number of connections ($n - 1$). The method used to find the tree - also the most common method - is to begin with a single point at random, and begin connecting to the closest point not already in the tree. When there are more than one equally close point, one will be chosen at random. The randomness aspect of the connections can be disabled in the options for the function, if so desired, such that the first listed point will be used as the start for the tree and if more than one point is equally close, the first listed will be chosen. Although there are other MST functions available for R (Oksanen et al. 2010), those other methods did not allow for a random start or random selection of equally minimal branches. The MSN is closely related to the MST; the MSN is a combination of all the possible MSTs. This could mean that if there was only one shortest MST that the MSN would be identical.

2.9 Biogeography and GIS

Biogeography is concerned with locations of organisms in space. The `fossil` package implements a number of functions to assist in converting georeferenced datasets into formats useful for both graphing within R and exporting to GIS programs. R was originally created as a statistical language, but its ability to use and display geographic data is quite advanced for a non-GIS system. The `sp` package (Pebesma and Bivand 2005) along with a number of geographic libraries allows a user to put in data in a number of projections and change projection and datum. For a thorough treatment of spatial data analysis with R, I highly recommend Bivand et al. (2008); here I provide only a cursory description of the topic.

The simplest geographic function to use is likely `create.lats()`, which as mentioned previously can extract the locality data from a list of taxa occurrences. With the output from this function, a number of further analyses can be done. For example, it is often useful to have the distances between two points in space; this can be easily accomplished with the `earth.dist()` function, which returns a matrix of pairwise distances in kilometres (Fig. 2–2). One note however is that the original matrix of locations must be in decimal degrees. Of course, the `sp` package provides functions to convert between coordinate systems if necessary.

Biogeography is concerned with species locations in space, and the sampling distributions of those species can cause some interesting effects in diversity calculations, namely the well researched species/area effect (Arrhenius 1921; Gleason 1922; Preston 1960; Connor and McCoy 1979; Rosenzweig 1995). Although palaeontology often pays little attention to this effect, Carrasco et al. (2005) have shown that it does hold true in fossil data sets. As a way to observe these effects efficiently, we have created the function `sac()` that can create a summary species area curve for a data set (Fig. 2–3). As it's arguments, it takes a table of longitude/latitude and a species occurrence matrix. It makes use of another function called `earth.poly()` which can take a table of locations and calculate which points create the vertices for a minimum spanning polygon/convex hull, as well as then calculate the true geographic area of the polygon.

Though the R environment is powerful when analysing GIS data, it lacks a large amount of visual interactivity with the data. Often, it is simply easier to use a GIS program to view geographic data, and as such I have tried to make it as simple as possible to move geographic data out of R. Currently the package provides helper functions for exporting both geographic points (`lats2Shape()`) and MSTs/MSNs (`msn2Shape`) to shapefile format using the package `shapefiles` (Stabler 2006). To use the functions, you need the `shapefile` package available on your system; the package can be downloaded using the `install.packages(shapefiles)` command. Once the shapefiles have been created, they can be saved using the `write.shapefile()` command. The shapefiles can then be loaded in any GIS program (Fig. 2–4).

```
> data(fdata.lats)
> shape.lats <- lats2Shape(fdata.lats)
> fdata.dist <- dino.dist(fdata.mat)
```

```
> fdata.mst <- dino.mst(fdata.dist)
> shape.mst <- msn2Shape(fdata.mst, fdata.lats)
```

2.10 Conclusions

I optimistically envisage the `fossil` package growing larger and larger in both function and use. As the project is Open Source, I encourage others to help aid in it's development both by simply using it in various and novel situations, as well as suggesting new possible methods, indices and functions that may be useful.
As well, I readily encourage others to use the original source code for their own purposes, with the only caveat that attribution is given where appropriate. I hope that encouraging the recopying and reuse of this code will save others time while developing their methods and allow more time for the actual data analysis.

Table 2–1: Names, formulas and alternate names for included similarity coefficients. Variables in the formulae are: a = number of shared species, b = number of species found only in the first sample, and c = the number of species found only in the second sample.

Coefficient Name	Formulae	Alternate Name	Function Call
Jaccard	$a/(a+b+c)$	Coefficient of Community	`jaccard()`
Sorenson	$2a/(2a+b+c)$	Dice, Czekanowski, Coincidence Index	`sorenson()`
Simpson	$a/(a+min(b,c))$	-	`simpson()`
Braun-Blanquet	$a/(a+max(b,c))$	-	`braun.blanquet()`
Ochiai	$a/\sqrt{(a+b)(a+c)}$	Coefficient of Closeness	`ochiai()`
Kulczynski	$[a/(a+b)+a/(a+c)]/2$	-	`kulczynski()`

Figure 2–1: Minimum Spanning Tree for the `fdata` example data set from the `fossil` package, overlain over a map of the USA. Letters correspond to locality name.

```
> data(fdata.mat)
> fdata.dist <- dino.dist(fdata.mat)
> fdata.mst <- dino.mst(fdata.dist)
> data(fdata.lats)
> library(maps)
> map("state")
> mstlines(fdata.mst, coordinates(fdata.lats))
> points(coordinates(fdata.lats), pch = 16, col = "white", cex = 3)
> points(coordinates(fdata.lats), pch = 1, cex = 3)
> text(coordinates(fdata.lats), labels = LETTERS[1:12])
```

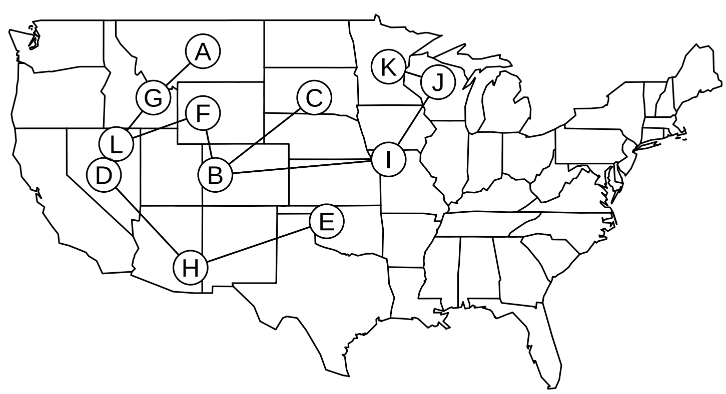


Figure 2–2: Distances between three selected locations from the `fdata` sample data. Distances given between points are in km.

```
> data(fdata.lats)
> fd.subset <- coordinates(fdata.lats)[1:3, ]
> earth.dist(fdata.lats[1:3, ])
         locA     locB
locB 893.4992
locC 776.3101 867.2648
> map("state")
> polygon(fd.subset)
> text(c(-110, -101, -106), c(42, 42, 47), labels = round(earth.dist(fd.subset)[c(1
+     3, 2)]))
> points(fd.subset, pch = 16, col = "white", cex = 3)
> points(fd.subset, pch = 1, cex = 3)
> text(fd.subset, label = LETTERS[1:3])
```

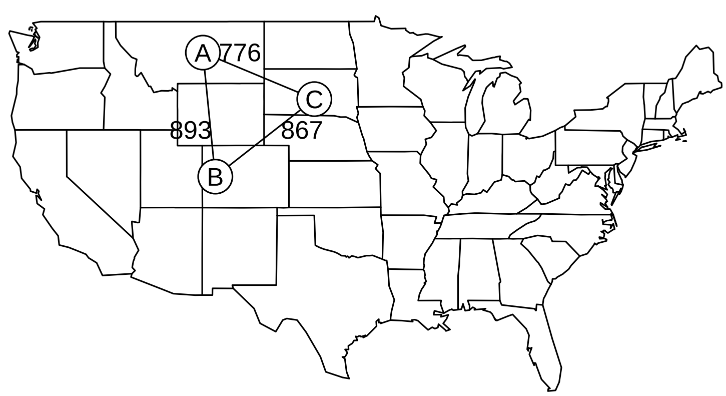


Figure 2–3: Species area curve for the `fdata` sample data.

```
> plot(log(sac(fdata.lats, fdata.mat)[[1]]), ylab = "log species richness",
+     xlab = "log area (km^2)")
```

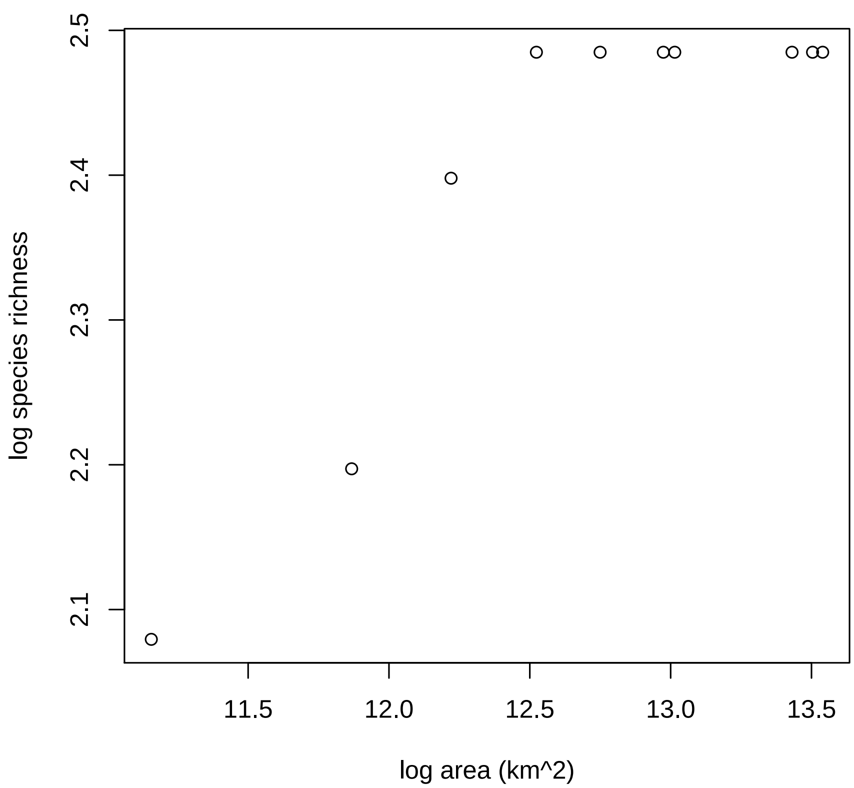
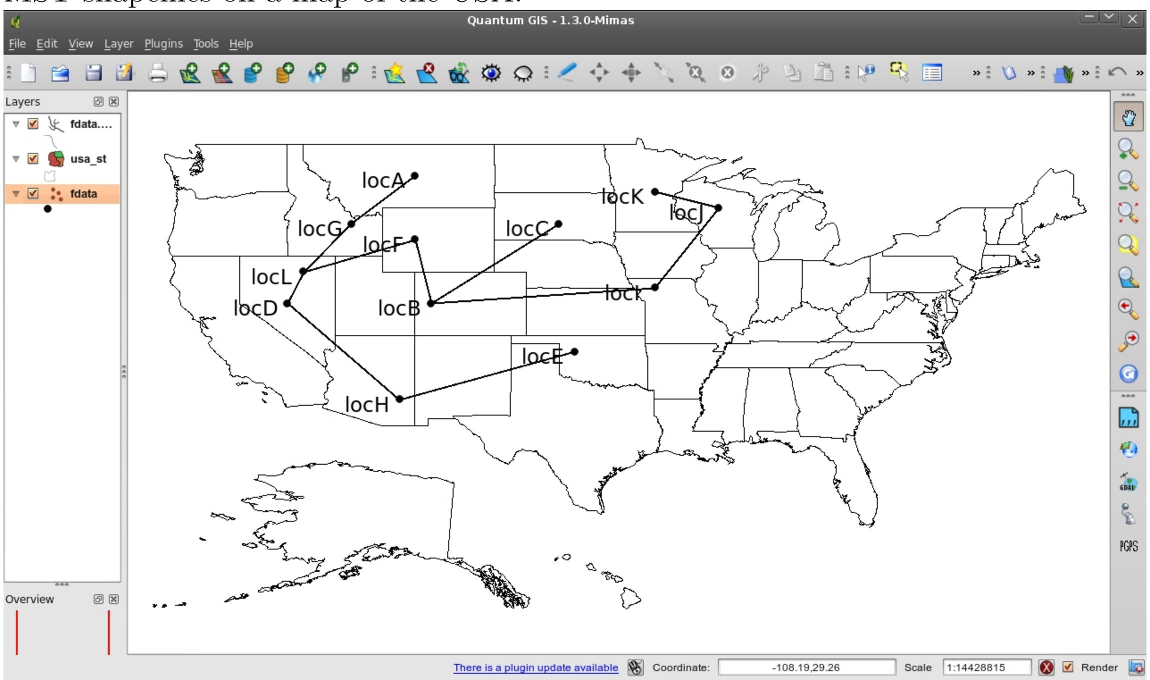


Figure 2–4: A screen shot from Quantum GIS, showing the exported latitude and MST shapefiles on a map of the USA.



Bridging Text

The next section in the methods is a manuscript that is being prepared for submission to a statistical software journal, describing a novel method of cluster analysis. The method is non-Euclidean, non-hierarchical and uses relational data, three traits which no other clustering method explicity combines. The method was originally created for biogeographic data, however it could be useful for a number of other type of studies, specifically any which would use NMDS ordination to plot the data. The method provides strict clusters optimized for within group similarity. Developing this novel statistical method will help with my further work on biogeographic provinces and patterns, as it is able to effectively cluster data sets which other methods cannot.

CHAPTER 3
A Non-Euclidean and Non-Hierarchical Clustering Method for Relational Data

3.1 Introduction

With many ecological datasets now containing thousands of records, semi-automated methods such as clustering are necessary to find underlying patterns in these highly complex collections. Clustering, defined as “a classificatory method which optimises intra-group homogeneity” (Lance and Williams 1967), is one of the most used types of multivariate analysis in ecology (James and McCulloch 1990; Hammer et al. 2001). Despite several tyes of cluster analyses being available, the underlying assumptions and limitations are very similar amongst the most common methods (James and McCulloch 1990).

A traditional cluster analysis will create a dendrogram that will typically be hierarchical in nature (James and McCulloch 1990; Shi 1993). These methods are best suited to scenarios such as phylogenetics, where a single (real or hypothetical) ancestor gives rise to multiple daughter descendants. However, this one-to-many structure may not be the most valid approach in all cases. In biogeography, a one-to-many relationship does not hold as species originate in multiple regions and move to multiple new sites, creating a many-to-many relationship. As such, it is best to use a method which allows greater flexibility in the way relationships are visualised.

The types of data sets to be clustered can be broadly divided into two groups: object and relational data. Object data is made up of an array of values, where each value describes a certain feature of a single sample (Hathaway and Bezdek 1994). Phylogenetic analyses begin with this type of data, with large tables composed of rows of homologous features, with each column describing a distinct specimen. In this way, each specimen is defined by its own characteristics, and no matter the number of other specimens present, the position of that specimen in n-dimensional space remains the same. Alternatively, relational data matrices are those where each value is a comparison between two individual samples. The comparison made can be one of similarity or dissimilarity. In this case, the characteristics of each sample is entirely dependent on the relationship of every other specimen relative to it, and the addition or deletion of samples can radically affect the character of the others.

One could question the use of relational data, as most relational data is derived from some sort of object data to begin with. However, there are many scenarios where a relational representation makes more sense than retaining object data, such as in biogeography. For example, when measuring individual characters of an organism, a zero means that the character has a zero measurement. However, in biogeography a zero measurement may mean either that there was truly no specimens of that species present or it could simply mean that the sample size was not large enough to capture the presence of all the species. In field studies, where time and money are often a concern, these two situations may be impossible to distinguish. Instead, using relational indices whereby each sample is compared to another can remove some of this bias; the use of similarity and dissimilarity measures as a way of assessing the relatedness of locations in biogeography is a well established practice.

While relational methods are not uncommon (i.e. UPGMA), many suffer from the aforementioned problems of forcing a hierarchical structure on the data. There are solutions such as c-means and k-medioid clustering, which are non-hierarchical and can use relational data. However, these methods make the assumption that the data has a Euclidean structure, while many of the measures used to calculate pairwise similarities are non-monotonic thereby creating non-Euclidean data matrices. Although methods allowing the use of relational data with more traditional c-means clustering exist (Hathaway and Bezdek 1994), such methods generally require that the data can be coerced into a Euclidean relational matrix. This coercion often fails with large and highly patchy data sets. Other options that do not require Euclidean data, such as Non-Metric Multidimensional Scaling (NMDS), can ordinate the data and are a useful way to present biogeographic information, but the actual division of the data into distinct groups still falls to the subjective eyes of the researcher.

As an alternative to traditional clustering and a complement to ordination techniques, I present here a clustering method called Non-Euclidean Relational Clustering (NERC). This method clusters relational data by minimising within group distances (dissimilarities) through an iterative process. While initially developed as a way to cluster non-Euclidean data sets, it can be equally useful for Euclidean data. It uses a branch-and-bound style heuristic approach (Lance and Williams 1967; Jain et al. 1999) as exhaustive searches for the ultimate optimal clustering matrix are far too time intensive for even moderately sized datasets. NERC allows for flexibility of relationships, as a hierarchical structure is not enforced, but still provides information on relationships between clusters. In order to demonstrate its function and utility, the NERC function is used here to cluster a small data set; several methods for assessing the "naturalness" of the final clusters are also introduced.

3.2 Methods

The NERC was written using the R Statistical Language (R Development Core Team 2010). The R Language was used as it is cross platform, Open Source and free to use, is widely used in statistical research, and is easy to extend with new functions and packages. A package called `fossil` (Vavrek 2010) with all of the functions discussed in this paper is available through the Comprehensive R Archive Network (CRAN) at `http://cran.r-project.org/web/packages/fossil/`. All data analysis and figure creation was done using R v2.9.2 on an Ubuntu 9.10 (Karmic) system. For a full copy of the data set and the R code used in calculations and figures, please consult the Supplemental Materials TK.

The NERC function has one required and three optional arguments, and takes the form `rclust(dist, clusters = 2, rand = 1000, counter = FALSE)`. The only required argument is a distance or dissimilarity matrix (the `dist` argument), either as a full matrix or lower triangle. The first optional argument (`clusters`) is the number of groups to be created. The number of groups used must be at least 2 but no greater than 1/2 the total number of samples, and must be a positive integer. The minimum value represents the smallest number of clusters without placing all samples within one group and the maximum value prevents clusters of one. The default value for the number of clusters is set to 2. The second optional argument gives the number of times the clustering process should be run. The last optional argument (`counter`) specifies whether to print the current run.

The way the function proceeds can be broken down into three distinct steps: the initialisation of clusters; the allocation of new elements to a cluster; and finally a reallocation process whereby the clusters are optimised (Lance and Williams 1967). The first step, initialisation of the clusters, begins by sampling a number of elements equal to the requested number of final clusters. Each of these selected samples is assigned to a different initial cluster. In the second step, the function searches for the greatest similarity (smallest value in a dissimilarity matrix) between any unassigned sample and any assigned sample. The unassigned sample with the highest similarity is assigned to the same group as that which it shares the greatest similarity, similar to Single Linkage Clustering Analysis (Gower and Ross 1969). This process then repeats, until all samples are assigned to a cluster. If at the second step any group has only one member the process restarts from the first step. After these two steps, a final optimisation is performed, whereby each sample is assessed as to their similarity to every group. If a sample has a greater similarity to another group rather than the one it is in, the optimisation routine will reassign the locality to another group. This is done one at a time, after which the similarity for each sample is recalculated, and the process keeps repeating itself until all samples have a greater average similarity to the other members within their groups than any other group. The process allows us to find local, but not necessarily ultimate, optima by minimising the overall dissimilarity within groups.

As with any cluster analysis, the purpose of NERC is to divide a set into multiple clusters, regardless of whether there are any actual divisions present. To test if the clustering method is actually picking up true divisions in the data, several other functions in the `fossil` package can be used. The first, and simplest is the calculation of a distance matrix of average within and between group distances using the `rclust.dist` function. As well as observing if the average distances within a cluster are much less than those between groups, the between group distances can be used to show relationships between different clusters, likely with some clusters being more similar than others. As well, this distance table shows which clusters are the most dense, with more tightly packed groups having a lower relative average within group distance.

A measure called the Cohesiveness Index (CoI) can also be used to evaluate the effectiveness of the clustering. The index is calculated by counting the number of within group connections which are part of the Minimum Spanning Tree (MST) divided by the total possible connections (where the total possible connections equals $n - 1$). It returns a value between zero and one for each cluster, where a value of one represents an exact congruence between the MST and the clusters. The index can also be considered to provide an abstraction of how close a Single Linkage Clustering Analysis would match the given clustering arrangement. The CoI is a semi-independent gauge of the integrity of the clusters apart from the average within group distance; if the CoI for all the groups is high, as well as their within group average distances relatively low, these together suggest the presence of distinct clusters.

The CoI function can also be used independently of the actual clustering function, so that clusters created by other methods can also be assessed. The function has two required arguments: a MST (binary) table (as is returned from any MST function in R) and a vector of group assignments (as is returned from the `rclust()` function).

To test if the clustering results are better than an entirely random grouping, a null model function is also provided. The function `rclust.null` uses a vector of group assignments and a distance matrix to calculate a null distribution for the clustering. It randomly assigns the samples to different groups, maintaining the same group sizes as in the initial group assignment vector, and recalculates the average within group distances with standard deviations. The means and standard deviations of the null model can be used to see if the actual clustering provides a significant improvement over a random sampling method.

Finally, it may be useful to compare results from two different cluster methods. The Rand Index (Rand 1971; Hubert and Arabie 1985) is a way of comparing two clustering outcomes, which provides an overall index between 0 and 1 of how well the two outcomes match, with 1 being a perfect match. The function is called `rand.index()`, and takes exactly two arguments; the first and second cluster identity vectors respectively. By using this index, one can compare the outcome from NERC to other more traditional clustering methods, as well as using it to observe the effectiveness of a clustering technique with a dataset that has a known clustering arrangement.

For a simple example using empirical data, I used a dataset from Gower and Ross (1969, originally from Delany and Healy (1966)) of white-toothed shrews from the Scilly and Channel Islands. A distance matrix was derived from skull measurements of shrews in 10 different English Channel locations, using canonical variate means as the centroid for each point (Gower and Ross 1969). The clustering was done using three groups, with 100 runs of clustering. The data was ordinated using the NMDS function provided by the ecodist (Goslee and Urban 2007) package.

3.3 Results

For the first run, I divided the original samples into two clusters. As a way to observe how the function operates on a data set, Fig. 3–1 shows how one hypothetical run might proceed. Initially, the Tresco and St. Martin's localities were selected by random sample in step one. Then, the second step proceeded to assign every other locality to one of the initial clusters, the order of which is represented by the diamonds on the connecting lines between the points (localities). The first sample to be assigned to one of the clusters was the St. Mary's sample, which had the smallest distance (greatest similarity) of any of the unassigned samples ($dist = 1.74$), as denoted by the first diamond. The function then proceeded in order to assign all the other samples (diamonds 2 to 8) until all samples were assigned to a cluster. Finally, the third optimisation step would occur, where the average within group and between group distances are compared for each sample. At this point, the St. Martin's locality would be shifted between clusters, as its average between group distance ($\overline{dist_b} = 3.18$) is less than its average within group distance ($\overline{dist_w} = 8.82$). After this, all the samples would have a smaller average within group distance than their average between group distance.

In the final configuration with two clusters (Fig. 3–2), one cluster consisting of Alderney, Guernsey and Cap Griz Nez (abbreviated as AlGC) was recovered, with the rest of the samples forming the other cluster (abbreviated as T+, for the Tresco locality). For this arrangement, the average within group distance was 3.35 for AlGC and 3.09 for T+. The CoI for both groups was 1, meaning in this case that the final clusters could be recreated by severing only one link in the MST.

When the function was run again 3 clusters (Fig. 3–3), the larger group in the first clustering (T+) was subsequently divided into two, with Jersey and Sark (abbreviated JS) forming their own cluster. The average within group distance for the T+ group became much smaller ($\overline{dist}_w = 2.32$) while increasing in the JS cluster ($\overline{dist}_w = 3.37$). The average between group distances clearly show the JS group as much more similar to the the T+ group ($\overline{dist}_b = 3.82$) than either group is to the AlGC cluster ($\overline{dist}_w = 9.53$ and 9.45, respectively). Also, the CoI for the JS cluster is 0 (i.e. no within group connections that correspond to the MST out of only one possibility), while for the other two groups the CoI is still 1.

It should be noted that the function was run with 4 clusters, however there was no stable solution found and the clustering failed.

3.4 Discussion

This small data set serves to show a few key points about NERC and ways of assessing the adequacy of its results. The two group clustering gave high values for the CoI for both clusters, while within group distances were relatively different (although not always significantly) from a random model. The 3 cluster arrangement on the other hand did little in the way of reducing the average within group distances versus the two cluster arrangement, and at the expense of reducing the CoI. As well, the JS group is well within the bounds of the null model. Based on the within group distances, CoI and null model, 3 clusters appears to be oversplitting.

A two cluster arrangement also makes the most sense based on the species used for the data set; *C. russula* is found only on the islands in the AlGC cluster, while *C. suaveolens* is found only on the islands inthe T+ cluster (Delany and Healy 1966). While further work must be done to refine the method, NERC is able to show the presence of major divsions within multivariate data sets. As well, though the three cluster arrangement combined the Jersey and Sark populations, the various measures of cluster strength were able to show that the JS cluster was not a "true" cluster but likely an artifact of oversplitting. In regards to the actual biologic data, Jersey and Sark show the most difference from the others in the T+ group, but their differences are in opposite directions to one another (Delany and Healy 1966).

While this dataset is highly simplistic, it serves as a useful example of how assessing results with multiple methods can give clusters which are more distinctive, with less over splitting of clusters. With larger datasets and more complex relationships this ability to visually assess results becomes even more difficult and so these assessment methods become even more useful.

3.5 Conclusions

Biogeographic data sets are often relational in nature, and relational clustering methods such as NERC can provide more approprite results than similar non-relational methods. Relational clustering methods allow a greater focus on how samples relate to one another rather than their absolute position in space. NERC compares samples directly to one another, while downplaying a large number of aspects that may group sites incorrectly in Euclidean space. This method can be used as an ideal complement to ordination methods (i.e. NMDS, PCoA) that do not provide explicit groupings. In combination, an ordination provides the backdrop upon which the samples can be plotted, while relational clustering can aid in the typically subjective division of the samples into distinct groupings. NERC is able to correctly divide data sets based on major divsions. Although it may at times create clusters that are not “natural”, using the different cluster strength indicators can show when clusters have strong support from the actual data. For biogeographic data sets, NERC is a useful tool which requires no assumptions to be violated in order to be used.

Table 3–1: Average within and between group pairwise distances for a two group clustering arrangement and comparison to a null model

Cluster	Average Pairwise Distance		Null Model		Cohesiveness
Name	T+	AlGC	$\overline{dist}_w$	σ	Index
T+	3.09	9.47	6.05	0.72	1
AlGC	9.47	3.35	6.22	1.98	1

Table 3–2: Average within and between group pairwise distances for a three group clustering arrangement and comparison to a null model

Cluster	Average Pairwise Distance			Null Model		Cohesiveness
Name	T+	AlGC	JS	$\overline{dist}_w$	σ	Index
T+	2.32	9.45	3.82	6.08	1.16	1
AlGC	9.45	3.35	9.53	6.01	2.07	1
JS	3.82	9.53	3.37	6.04	3.33	0

Figure 3–1: Flow diagram showing the order of progression of the NERC function; Tresco and St Martin's are the randomly selected points to begin the clusters. Numbers within the diamonds represent the order in which the sites were assigned to clusters. Abbreviations follow Gower and Ross (1969), originally from Delany and Healy (1966): Ag, St Agnes; Al, Aldemey; B, Bryher; C, Cap Griz Nez; G, Guernsey; J, Jersey; Mn, St Martin's; My, St Mary's; S, Sark; T, Tresco.

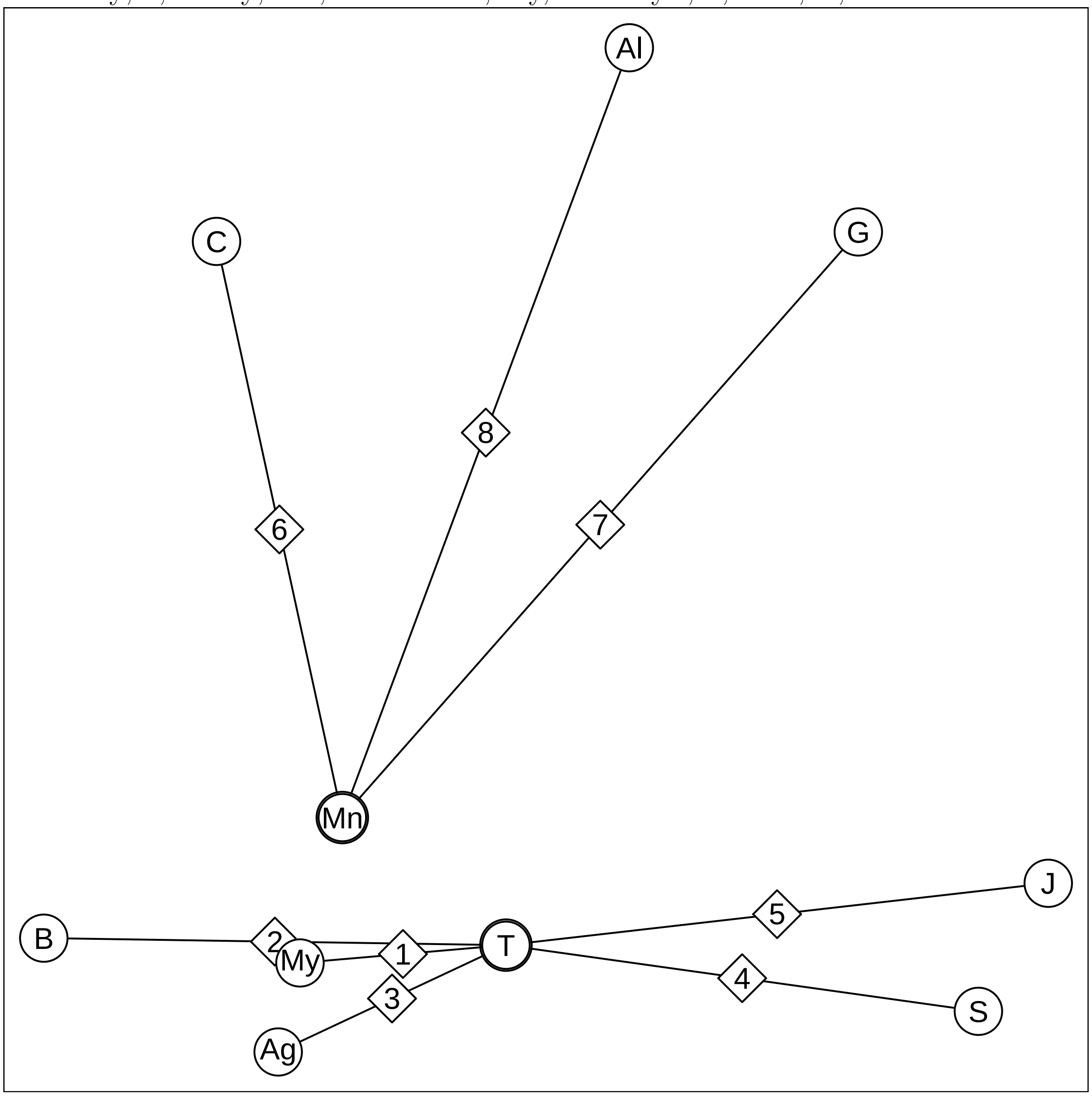


Figure 3–2: Non-Metric Multidimensional Scaling ordination with an Minimum Spanning Tree overlain and point shapes representing cluster assignments. Please note that while it may at first appear different than Gower and Ross (1969), the relative positions remain the same and the MST is identical. Abbreviations as in Figure 1.

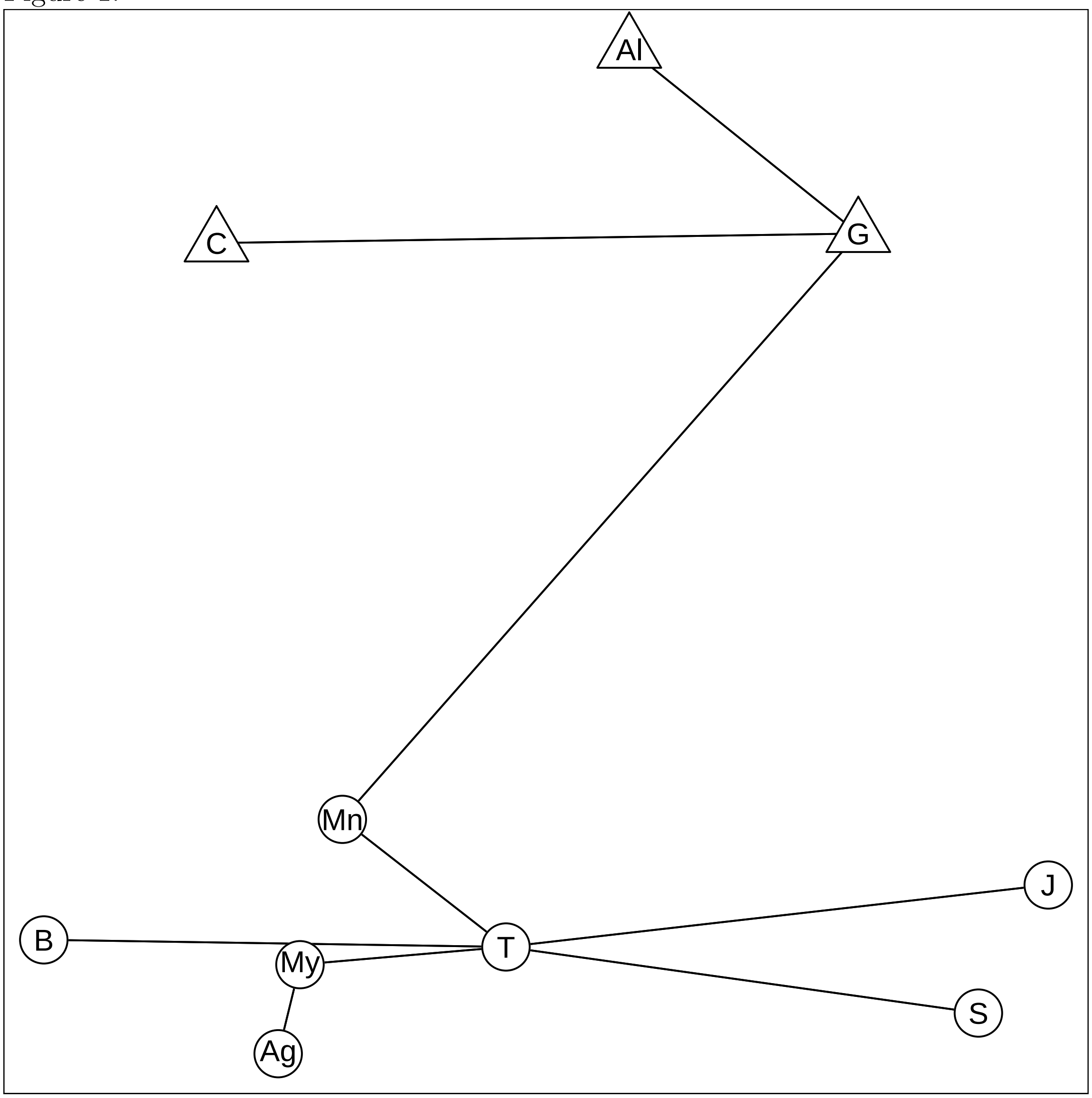
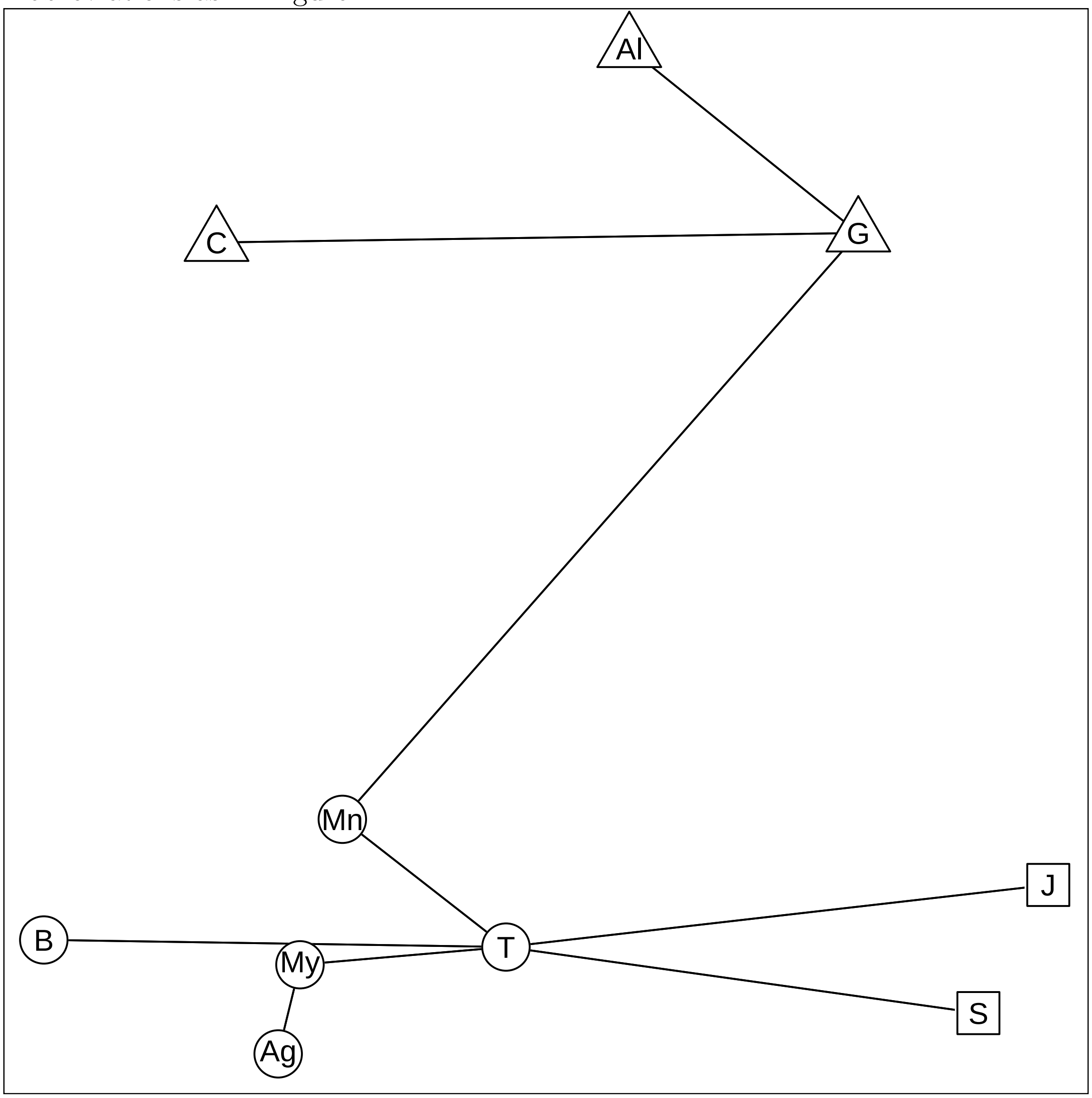


Figure 3–3: Same figure as previous, but with localities clustered into 3 groups. Abbreviations as in Figure 1.



Bridging Text

Chapter 4 describes a Triassic-aged flora from Axel Heiberg Island in the High Arctic. The paper was originally published in 2007 in the Canadian Journal of Earth Sciences (issue 44, pages 1653–1659). The flora is important because the Arctic is so inaccessible, meaning that there is so little material generally available and making any new material found of interest. In regards to palaeomacroecology, the best method for understanding patterns of large scale diversity in the fossil record is by collecting information on local assemblages (Johnson 2003). Without researchers on the ground, uncovering and describing new localities, there would be no data for palaeoecologists to study.

This fossil flora is from the Late Triassic, shortly before the Triassic/Jurassic boundary based on the presence of a species of peltasperm (a type of seed fern) by the name of *Lepidopteris*. This flora is the most diverse from the region for this time, and may represent a shift in climate in the area to a much drier environment.

Full citation:

Vavrek, M.J., Larsson, H.C.E. and Rybczynski, N. 2007. A Late Triassic flora from east-central Axel Heiberg Island, Nunavut, Canada. Canadian Journal of Earth Sciences, 44: 1653–1659.

Author Contributions: M.J.V. identified and photographed all the specimens; M.J.V. made all the figures and plates; M.J.V., H.C.E.L. and N.R. wrote the paper.

CHAPTER 4
A Late Triassic Flora From East-Central Axel Heiberg Island, Nunavut, Canada

4.1 Abstract

A new floral assemblage is described from the Fosheim Member of the Heiberg Formation on southern Axel Heiberg Island. The flora is relatively diverse, consisting of at least ten different collected species and three field identified species from a single small locality. The flora has some similarities to other European-Sinian floras, such as those previously found on Axel Heiberg and Ellesmere islands and in Scoresby Sound of eastern Greenland. The largest difference from previously described floras is the dominance of the bennettitalean *Pterophyllum astartense*, suggesting a drier climate in contrast to the more humid one previously proposed. The presence of *Lepidopteris ottonis* places the flora within the Upper Triassic *Lepidopteris* zone, suggesting a Norian age.

4.2 Introduction

The first recorded plant fossils in the Canadian High Arctic were found by members of a party sent in search of the Franklin Expedition in 1853 (Osborn (1855), reviewed in Tozer (1963) and Ash and Basinger (1991)). However, it was over a century before plant fossils were reported again from the region, when petrified wood and leaf impressions were collected in 1955 by the Geological Survey of Canada (GSC Operation Franklin; Fortier et al. 1963). Since then, Triassic plant macrofossils have been found scattered across the Arctic (Glenister 1963; Souther 1963; Tozer 1963; McLaren 1963; Ash and Basinger 1991).

The floral collections of the GSC from the Heiberg Formation, including specimens from Axel Heiberg, Ellesmere, Cornwall and Cameron islands were documented in detail in Ash and Basinger (1991). The flora described herein is from a new locality on Axel Heiberg Island. Although it shares many floral taxa with those described by Ash and Basinger (1991), it differs significantly in composition.

4.3 Geology

The Heiberg Formation was introduced by Tozer (1961), and defined by Souther (1963). Souther (1963) originally identified upper and lower members within the formation, but the formation has since been divided into three members and thoroughly redefined (Embry 1982, 1983, Fig. 4–1). The Romulus Member is the lowest unit and is of delta front and prodelta origin. It consists of coarsening upwards cycles of very fine- to fine-grained sandstone, siltstone and shale, and can range in thickness from 50 to 400 m (Embry 1983, 1991). The Fosheim Member, which is the middle unit and of delta plain origin, consists of mainly fine- to medium-grained sandstone with thin interbeds of carbonaceous siltstone, shale and coal and can range in thickness from 10 to 800 m (Embry 1983, 1991). The Remus Member is the uppermost unit and represents a strand plain and nearshore deposit, almost entirely composed of fine-grained sandstone, and ranges in thickness from 5 to 220 m (Embry 1983, 1991). The entire formation is representative of a large north-westward prograding delta complex which once emptied into the Sverdrup basin (Embry 1982). The formation boundaries are conformable with the underlying Barrow Formation and the overlying Jameson Bay Formation.

Dating of the formation comes from several marine interbeds containing invertebrates and from extensive palynological studies (Embry 1982, 1983; Suneby and Hills 1988). The Romulus Member is considered to be Norian in age based on marine macrofossils (Norford et al. 1973; Embry 1983) and Norian to Early Rhaetian based on palynomorphs (Suneby and Hills 1988). The Fosheim Member is considered as Norian to Pleinsbachian based on a pelecypod shell (Souther 1963; Embry 1983) and palynology (Suneby and Hills 1988), while the uppermost Remus Member is identified as Pleinsbachian to Late Toarcian based on marine fossils and palynomorphs (Embry 1983; Suneby and Hills 1988). Palynological data have shown that the upper and lower age limits of each member are variable across the formation (Suneby and Hills 1988). In total, the Heiberg Formation ranges from the Late Triassic (Early Norian) to the Early Jurassic (Early Pleinsbachian/Early Toarcian) (Embry 1983; Suneby and Hills 1988).

4.4 Materials and methods

The flora described in this paper comes from the Fosheim Member, at 79° 15.242' N 89° 21.033' W on Strand Fjord, east-central Axel Heiberg Island (Fig. 4–2). This locality is near several localities of Ash and Basinger (1991). All specimens were collected and/or recorded from an area of 100 m^2, from a single bed of approximately 10 cm thickness.

Fossils are preserved as carbonaceous compressions/impressions, with no cuticle present, in a dark grey to black argillaceous shale to siltstone matrix. Specimens were collected during summer 2004, and are accessioned in the Nunavut fossil collections at the Canadian Museum of Nature (CMN).

4.5 Systematic palaeontology

Order Osmundales Bromhead, 1838

Family Osmundaceae Berchtold et Presl, 1820

Genus *Todites* Brongniart, 1828

Todites sp. Brongniart, 1828

(Fig. 4–3A)

Description. A single specimen was recovered consisting of a partial pinna that measures 125 mm in length, with a pinna rachis 2 mm wide near the base narrowing to 1 mm wide near the apex (Fig. 4–3A). Pinnules are slightly falcate, about 12 mm long and 5 mm wide. The pinnule midvein is prominent, with lateral veins arising at approximately 30 degrees, with most veins forked once, but with some near the base of the pinnule forked twice. No teeth can be observed on the pinnule, however, this could be due to poor preservation of pinnule margins.

Comparisons. The Osmundaceae is one of the oldest known filicalean fern families, containing over 150 fossil and approximately 20 modern species (Hewitson 1962; Arnold 1964; Tidwell and Ash 1994). *Todites* is the generic name given to fossil ferns exhibiting similarities to the modern monospecific genus *Todea* (Seward 1910). The specimen described here compares closely with those described by Harris (1931) and Ash and Basinger (1991) in gross morphology, however, it shows little detail and is only tentatively assigned to *Todites*.

Order Peltaspermales Thomas 1933

Family Peltaspermaceae Thomas 1933

Genus *Lepidopteris* Schimper emend Townrow, 1956

Lepidopteris ottonis (Göppert) Schimper, 1869

(Fig. 4–3B)

Description. Of the two collected specimens, the largest and best preserved is a portion of a bipinnate leaf measuring 60 mm in length and 70 mm wide (Fig. 4–3B). The main rachis is 4 mm wide and the pinna rachides are 2 mm wide. The rachides are covered in distinctive blister-like swellings. Pinnae branch off at 50 degree angles. Pinnules are 2 to 3 mm wide and 4 to 6 mm long, and each has a prominent midvein. Lateral veins are not visible in the specimen. There are typically one or two intercalary pinnules borne on the main rachis in this taxon between the primary pinnae, however, these are not visible in our specimens.

Comparisons. *Lepidopteris* is the principal foliage type of the Peltaspermaceae during the Triassic, and all species of *Lepidopteris* have intercalary pinnules, pinnules set directly on the rachis between adjacent pinnae (Harris 1932*a*; Townrow 1960; Kerp and Haubold 1988). Another common characteristic of the genus are blister-like swellings found along the rachis and pinnules, and these swellings in fact enabled reconstruction of the plant from isolated organs (Harris 1932*a*; Thomas 1933; Townrow 1960). Little is known about the relationship of peltasperms to other gymnosperms, as they do not closely resemble any other known group (Townrow 1960). *L. ottonis* has previously been reported from northern hemisphere localities, including Greenland, China, Sweden and Germany, although this is the first report of it from the Canadian Arctic (Harris 1932*a*; Townrow 1960).

Order Bennettitales Engler 1892

Genus *Pterophyllum* Brongniart, 1828

Pterophyllum astartense Harris, 1932

(Fig. 4–3C)

Description. 19 specimens of this species have been recovered, ranging from fragments of pinnules to entire leaves (Fig. 4–3C). The leaf as a whole is lanceolate. The rachides are from 100 mm to greater than 170 mm in length, and 4 mm wide. Longitudinal ridges on the rachis can be seen, but transverse wrinkles are not present. Pinnules are borne alternately to oppositely, and curve slightly towards the apex of the rachis. Pinnules are 3 mm wide and 15 to 40 mm long. The bases of the pinnules are typically parallel, with some slightly expanded, and venation in the pinnules is parallel.

Comparisons. *Pterophyllum* is a form genus made up of the similar leaves of several generically different types of bennettitaleans (Harris 1932*b*). The different species contained within *Pterophyllum* show a mosaic of features, and often there are intermediate forms between the different species, making specific determination difficult. Previously, specimens of *Pterophyllum* were recovered from the Heiberg Formation, but these were referred to *P. subaequale* rather than *P.*

astartense (Ash and Basinger 1991). The specimens described here are definitely not *P. subaequale* because they lack the characteristic transverse wrinkles on the rachis found in that species, although their characteristics in virtually all other respects overlap.

Genus *Anomozamites* Schimper, 1870

Anomozamites sp. Schimper, 1870

(Fig. 4–3D)

Description. A single, nearly complete pinnate leaf was recovered (Fig. 4–3D). The overall shape of the leaf is long-lanceolate. The leaf is 180 mm long and up to 16 mm wide. The rachis is 2 mm wide, and shows some longitudinal striations, but no transverse wrinkles. The pinnules are up to 10 mm long and 7 mm wide, appear slightly falcate in outline, and are borne alternately to oppositely. There is no visible venation in the pinnules, possibly due to poor preservation.

Comparisons. *Anomozamites* is a form-genus sharing many similarities with *Pterophyllum* (Harris 1969). One of the main distinguishing characters between the two genera is that the length of the pinnules in *Anomozamites* tends to be less than double their width, while in *Pterophyllum* the opposite holds true, although exceptions within both genera can easily be found (Harris 1969). The specimen described here compares most closely to *Anomozamites nitida* and *A. minor* in its general morphology, but due to the lack of preservation of the cuticle no definite identification can be given.

Genus *Vardekloeftia* Harris, 1932

Vardekloeftia sp. Harris, 1932

(Fig. 4–3E)

Description. A single specimen of a round ovulate reproductive structure was recovered (Fig. 4–3E). The main body is 20 mm in diameter, and is covered in small bumps, 0.3 to 0.8 mm in diameter, which are the heads of the interseminal scales. The reproductive structure is attached to a fragmentary stem.

Comparisons. While the leaves of Triassic Bennettitales are rather common, the reproductive organs are poorly known (Harris 1932*b*). This species may represent the reproductive organs of *Pterophyllum astartense*, the most common bennittitalean leaf at the site, and small fragments of the leaf are in close association with the reproductive structure.

Order Czekanowskiales Pant, 1957

Genus *Czekanowskia* Heer, 1876

cf. *Czekanowskia* Heer, 1876

(Fig. 4–3F)

Description. Four specimens, each with multiple leaf bundles, were recovered. Branches are 4 mm in diameter, and leaf clusters occur every 15 to 25 mm along the branch (Fig. 4–3F). Clusters arise in two linear ranks along opposite sides of the branch. The leaves are borne in clusters of about 8, and each leaf is about 1 mm wide, but probably much greater than 30 mm long, although there are no complete leaves. Each leaf possesses a single median vein. Leaves may be forked, but no specimens in our possession exhibit it.

Comparisons. Leaves of this type are all fragmentary, and none show the necessary features for positive identification. However, the leaves generally conform to the size range of known Czekanowskiales and *Czekanowskia* is known from nearby contemporary localities (Ash and Basinger 1991).

Order Coniferales Engler and Prantl, 1889

Family Palissyaceae Florin, 1958

Genus *Stachyotaxus* Nathorst, 1908

Stachyotaxus elegans Nathorst, 1908

(Fig. 4–3G)

Description. Two specimens of leafy shoots of this species were collected from the site. Each shoot is up to 12 mm wide, and the main axis of the shoots are about 1 mm wide (Fig. 4–3G). Leaves arise from the stem oppositely at a high angle. The leaves are linear-lanceolate and single veined. The largest leaves are up to 2 mm wide and 10 mm long.

Comparisons. The Palissyaceae is an extinct family from the Triassic and Jurassic comprised of three genera, *Palissya*, *Stachyotaxus* and *Metridiostrobus* (Florin 1958; Delevoryas and Hope 1981). While Palissyaceae share some similarities in the female cones to the modern Cephalotaxaceae, the other organs are significantly different to warrant a separate family (Florin 1958). The genus *Stachyotaxus* is known only from the Late Triassic (Florin 1958), and has been reported from Sweden (Nathorst 1908), Greenland (Harris 1935) and elsewhere on Axel Heiberg Island (Ash and Basinger 1991).

Family incertae sedis

Genus *Podozamites* Braun, 1843

Podozamites cf. *P. mucronatus* Harris, 1935

(Fig. 4–4A)

Description. Eight different specimens of this species were collected. The main axis of the leafy shoot is about 3 mm wide (Fig. 4–4A). Leaves branch distichously at an angle of 50 degrees. Leaves are up to 15 mm wide at their widest point, and over 90 mm long at their longest. Leaves show numerous, dichotomizing parallel veins, about 8 to 12 veins near the base, increasing to about 30 after 80 mm. Leaf bases are constricted, and leaf apex is acute.

Comparisons. *Podozamites* was originally believed to be a Mesozoic cycad and regarded as a pinnate leaf, but later discoveries showed that several species displayed spiral phyllotaxy, and so is now classified as a conifer (Harris 1935). Later studies of the cuticle also affirmed this view (Stewart and Rothwell 1993). The genus as it stands is likely a mosaic of several genera, but due to the lack of readily identifiable and non-variable characters on the leaves, any division of the genus would be very difficult (Harris 1935). This first type of *Podozamites* has been referred to *P. mucronatus*, due to the similarity in leaf shape and size, although some of the specimens found are much longer than this species usually displays.

Podozamites cf. *P. schenki* Heer, 1876

(Fig. 4–4B)

Description. Three short leafy shoots were collected from the site. The main axes of the shoots are up to 2 mm wide (Fig. 4–4B). Leaves are spirally arranged, and branch off at 60 degrees. Leaf bases are constricted and tips are acutely pointed. Leaves are between 1.5 to 3 mm wide, and 20 to 45 mm long. Veins are parallel with 6 to 8 present near the base.

Comparisons. The specimens of *Podozamites* recovered fall into two distinct groups, based on leaf size and phyllotaxy. The first, referred to *P. mucronatus*, is much larger and shows distichous leaf arrangement, while *Podozamites* cf. *P. schenki* has much smaller leaves that are borne spirally. The difference in how the leaves branch eliminates the possibility that the smaller of these two species is simply an immature shoot of the larger.

Order incertae sedis

Unidentified cone

(Fig. 4–4C)

Description. Two specimens of an unidentified cone were recovered. The cone is at least 40 mm long, and approximately 10 mm wide (Fig. 4–4C). Each scale of the cone has several distinct ridges along its underside, and the scales are arranged helically. Scales are up to 4 mm wide and 6 mm long, and have a rounded end.

Comparisons. There are very few details that can be seen in these cones, making any identification uncertain. They do not appear to resemble the reproductive structures of any of the other species identified, however several of these species have poorly understood reproductive organs.

4.6 Discussion

This site has yielded a rather rich flora of at least 10 different species, as well as three more probable species that were not collected due to time and weather constraints. This site is currently one of the most diverse plant localities from Late Triassic rocks in the High Arctic of Canada, although due to the paucity of localities and research in this region this site may simply be of a typical richness.

Due to the location of the site and lack of appropriate tools and time, some fossils were only identified and photographed in the field but not collected. These uncollected specimens appear to represent three additional taxa: *Neocalamites* Halle; *Dictyophyllum exile* (Brauns) Nathorst; and *Ginkgoites* Seward. These taxa have previously been recovered from nearby localities, however, because no specimens were brought back for examination, a proper final diagnosis cannot be made.

While there are no lithologic markers to determine whether this site lies within the Triassic or the Jurassic in this area of the Fosheim Formation, the flora has some distinctive elements. The representative floras of the latest Triassic and the earliest Jurassic in Europe and Greenland are the *Lepidopteris* flora and the *Thaumatopteris* flora respectively (Harris 1937). *Lepidopteris ottonis* is considered to be the most important zone fossil of the *Lepidopteris* flora in Europe and Greenland, and its presence here provides strong evidence to place this flora in the *Lepidopteris* zone, and subsequently the Late Triassic (Harris 1937).

The flora described here shares many similarities to other Late Triassic European-Sinian area floras, especially those from Greenland (Harris 1926, 1931, 1932*a*,*b*, 1935) and Ellesmere and Axel Heiberg islands (Ash and Basinger 1991). However, the dominance of the bennettitalean *Pterophyllum astartense* differs from nearby localities. As bennettitaleans live in generally drier areas, their dominance could suggest a local region of well-drained soils, while nearby localities suggest an overall humid climate (Ash and Basinger 1991). However, there is also the possibility that this flora is indicative of a period of aridity in the region (Clemmensen et al. 1998). Eastern Greenland, at a similar latitude as Axel Heiberg, was near the transition between a southern dry steppe environment and a northern warm moist temperate climate during the Late Triassic (Clemmensen et al. 1998). The differences observed in the flora of the Canadian Arctic could be due to larger climatic shifts during this period. This flora may represent the earlier phase of the transition, with its more arid climate, while that of Ash and Basinger (1991) may represent a slightly later, more humid climate.

4.7 Conclusions

The new floral assemblage from Axel Heiberg is Late Triassic in age, as revealed by the presence of several index taxa. It most closely compares to other European-Sinian floras rather than Siberian floras, and contains similar species as the other Heiberg Formation floras and the nearby Late Triassic Scoresby Sound flora of east Greenland. The site is distinct from other Heiberg Formation floras in the abundance of Bennettitales, especially *Pterophyllum astartense*, and conifers. This floral composition may indicate a local region of well-drained soils, or may be the result of shifting climates in the region.

4.8 Acknowledgments

Funding for this work was provided by the Polar Continental Shelf Project to N.R. and H.C.E.L., the National Geographic Society to H.C.E.L. and N.R., Canada Research Chairs to H.C.E.L., a National Sciences and Engineering Research Council Discovery Grant to H.C.E.L., the Canadian Museum of Nature to N.R. and a National Sciences and Engineering Research Council Canada Graduate Scholarship (Masters) to M.J.V. Fieldwork was greatly aided by Erin Maxwell, Matthew Vickaryous, and Alex van Nievelt. Reviewers D.R. Braman, K.B. Pigg, J. Utting, and Associate Editors B.D.E. Chatterton and J. Jin are thanked for comments which greatly improved the manuscript.

Figure 4–1: Summary stratigraphic column for the Late Triassic-Early Jurassic Heiberg Formation, showing division of formation into members. Approximate position of this locality is denoted by an '**X**'. Adapted from Suneby and Hills (1988) and after Embry (1982, 1983).

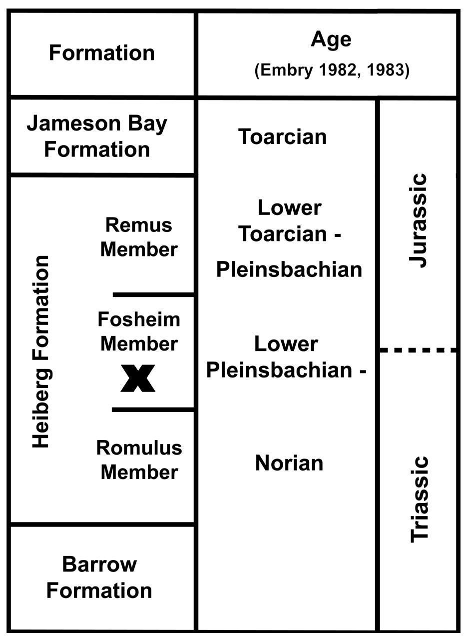


Figure 4–2: Close up of southern Axel Heiberg, showing the location of Axel Heiberg within Nunavut, and Nunavut within Canada. The locality described herein is indicated by an '**X**', while nearby localities described are identified by numbers 5-9, using the locality numbers of Ash and Basinger (1991).

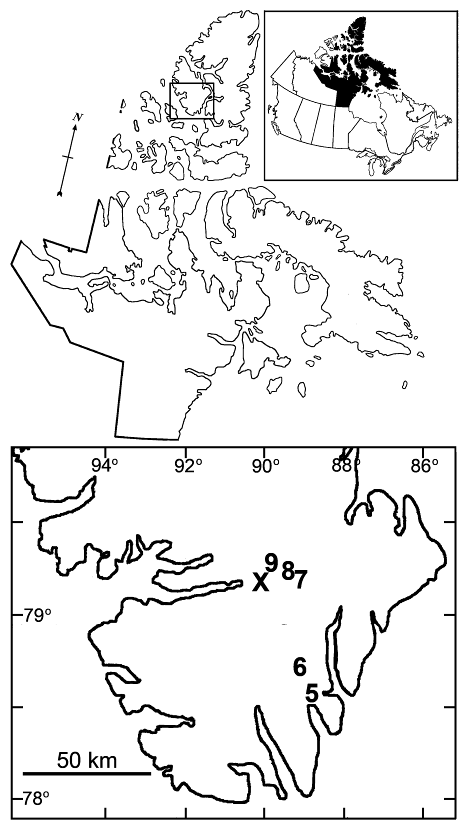


Figure 4–3: (A) cf. *Todites* sp., NUFM 026, partial pinna. (B) *Lepidopteris ottonis*, NUFM 046, pinna, with blister like swellings on pinna rachis. (C) *Pterophyllum astartense*, NUFM 029, leaf. (D) *Anomozamites* sp., NUFM 032, leaf. (E) *Vardekloeftia* sp., NUFM 024, reproductive structure. (F) cf. *Czekanowskia*, NUFM 035, leaf bundle. (G) *Stachyotaxus elegans*, NUFM 040, leafy shoot. Scale bars equal 1 cm.

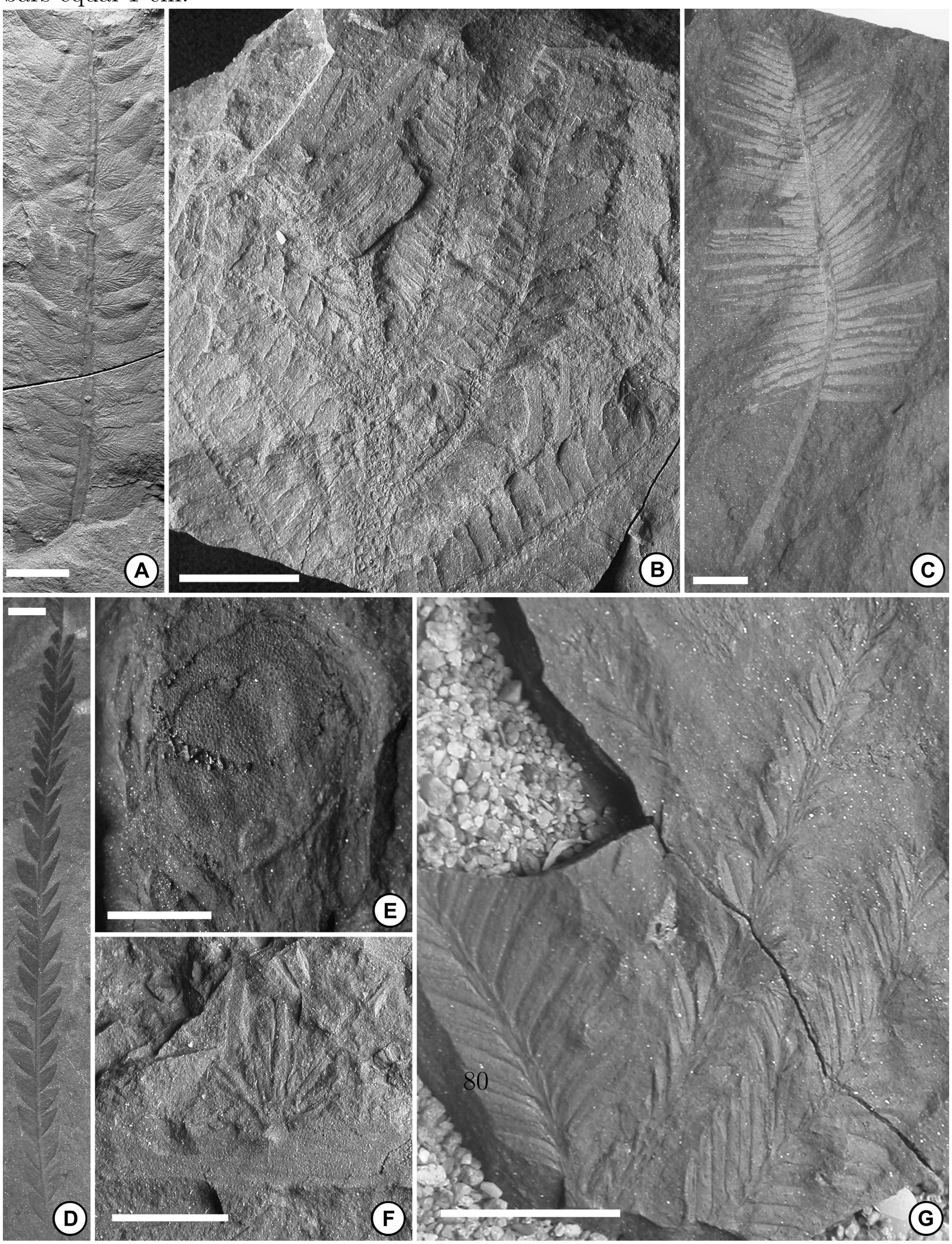
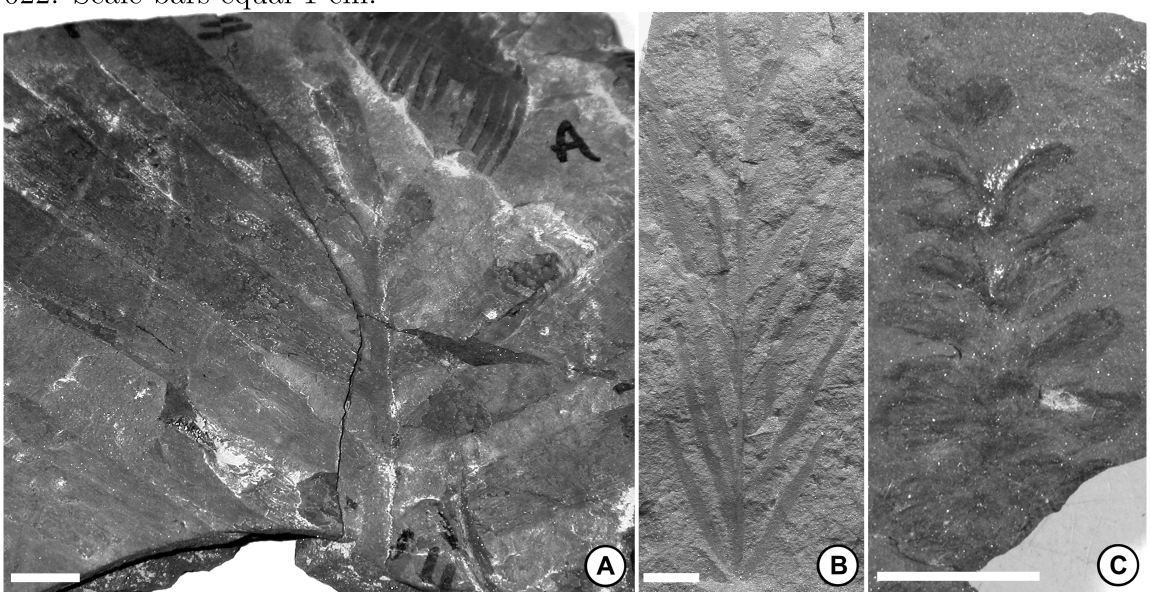


Figure 4–4: (A) *Podozamites* cf. *P. mucronatus*, NUFM 046, leafy shoot. (B) *Podozamites* cf. *P. schenki*, NUFM 034, leafy shoot. (C) Unidentified cone, NUFM 022. Scale bars equal 1 cm.



Bridging Text

The preceding three chapters have provided a foundation for the remainder of my thesis, in creating a set of analytical methods that can be used in a large scale analysis, as well as the process of collecting and adding more data. With the following chapter, I begin to apply some of these methods to the data. This chapter takes several types of methods that are often used in neoecological analyses, and applies them to the fossil record. In this way, I use our knowledge of modern patterns of species diversity to better understand the patterns in the fossil record, using the present to reveal the past.

This paper was originally published in the Proceedings of the National Academy of Sciences, in 2010, issue 107, pages 8265–8268. The paper used a number of modern ecological methods to quantify beta diversity in Maastrichtian dinosaur assemblages of North America. The main goal of the paper was to see if quantitative methods could recover the same endemic faunal provinces as had been previously reported. By using a common method of evaluating modern beta diversity, we were able to quantitatively test if there were large scale regions of dinosaurian endemism and provinciality in the Western Interior region of North America.

Full citation:

Vavrek, M.J. and Larsson, H.C.E. 2010. Low beta diversity of Maastrichtian dinosaurs of North America. Proceedings of the National Academy of Sciences, 107: 8265-8268.

Author Contributions: M.J.V. designed research; M.J.V. performed research; M.J.V. contributed new analytic tools; M.J.V. and H.C.E.L. analyzed data; and M.J.V. and H.C.E.L. wrote the paper.

CHAPTER 5
Low Beta Diversity of Maastrichtian Dinosaurs of North America

5.1 Abstract

Beta diversity is an important component of large scale patterns of biodiversity, but its explicit examination is more difficult than that of alpha diversity. Only recently have data sets large enough been presented to begin assessing global patterns of species turnover, especially in the fossil record. We present here the first analysis of beta diversity of a Maastrichtian (71-65 million years old) assemblage of dinosaurs from the Western Interior of North America, a region which covers approximately 1.5×10^6 km^2, borders an epicontinental sea and spans approximately 20 degrees of latitude. Previous qualitative analyses have suggested regional groupings of these dinosaurs and generally concluded that there were multiple distinct faunal regions. However, these studies did not directly account for sampling bias, which may artificially decrease similarity and increase turnover between regions. Our analysis used abundance-based data to account for sampling intensity and was unable to support any hypothesis of multiple distinct faunas; earlier hypothesized faunal delineations were likely a sampling artifact. Our results indicate a low beta diversity and support a single dinosaur community within the entire Western Interior region of latest Cretaceous North America. Homogeneous environments are a known driver of low modern beta diversities, and the warm equable climate of the late Cretaceous modulated by the epicontenental seaway is inferred to be an underlying influence on the low beta diversity of this ancient ecosystem.

5.2 Introduction

Alpha (α) diversity is defined by Whittaker (1972, 1977) as species richness on the local or habitat scale, and beta (β) diversity is defined as the difference in the types of species found in different areas of alpha diversity. Alpha and beta diversity together make up species richness at the landscape scale, called gamma (γ) diversity. Because beta diversity measures turnover across an area, it is closely related to the numbers of endemic species within each community, that in turn can be used to assess biotic provinciality. Quantitative estimates of modern beta diversity recover surprisingly low beta diversity values in spite of broad taxonomic and geographic sampling, regardless of the motility of the group in question (Harrison et al. 1992; Condit et al. 2002; Novotny et al. 2007). Instead, beta diversity appears most correlated with climate evenness (Pitman et al. 1999; Condit et al. 2002; Novotny et al. 2007).

We present the first beta diversity estimates for an ancient terrestrial ecosystem. The Western Interior of North America is perhaps the most intensely sampled dinosaur-bearing region in the world (Lehman 1987). The terrestrially deposited rock sequence is comprised of sparsely exposed Aptian (121 to 112 Ma) formations to extensively exposed Maastrichtian (71 to 65 Ma) formations. These Maastrichian rocks are largely floodplain deposits along the western shores of the epicontinental Western Interior Seaway (Fig. 5–1). The rich Maastrichtian deposits have been a focus of research on patterns of dinosaur distributions and biogeography on a sub-continental scale (Lehman 1987, 2001; Sloan 1969, 1976; Russell

1967, 1977; Sampson et al. 2004). We used this dinosaur assemblage for our analysis of beta diversity in the fossil record because it is currently the only data set large enough and the only region with extensive previous work on endemism and provinciality. Previous studies have concentrated on dinosaur faunal provinciality at local scales based on presence/absence data. Specifically, these analyses have suggested high levels of endemism at local, formational scales (Sampson et al. 2004). The most widely accepted community hypothesis divides the dinosaur fauna into three zones: a northern *Leptoceratops* zone; a southern *Alamosaurus* zone; and an interior *Triceratops* zone (Lehman 1987, 2001). The boundaries of these three regions have been slightly modified as more fossils have been found, and later studies incorporated these new dinosaur and as well as previous pterosaur (*Quetzalcoatlus*) finds (Lehman 2001).

The hypothesized provincial delineations lead to the prediction that beta diversity for this region as a whole should be relatively high. However, previous work was based on the occurrence of a few well known individual species that have well defined geographic boundaries within relatively small spatial scales. Such high levels of endemism would be unprecedented for any modern large bodied terrestrial fauna.

We use modern approaches to search for statistical support for areas of high endemism. These approaches included rarefaction and species estimators to calculate beta diversity and Minimum Spanning Trees (MST) and Non-Metric Multidimensional Scaling (NMDS) to search for readily apparent provinces (See Methods for full explanation). If there are strongly delineated provinces present, we predict that beta diversity for the Maastrichtian Western Interior dinosaur fauna is high and the NMDS/MST ordination will be fragmented into discrete clusters of sites.

5.3 Materials

This paper has made use of recent developments in the extensive cataloguing of dinosaur remains at the online and open Paleobiology Database (PaleoDB.org). All location and abundance data were downloaded from the Paleobiology Database on 14 January 2009, using the taxon name "Dinosauria" and a time span = "Campanian" to "Maastrichtian", with the following parameters: Continent = "North America"; Abundance Value = TRUE; State = TRUE; and Formations = TRUE. The majority of this data set originated from the work of Carrano (2000) to collect and collate the record of Dinosauria throughout the Mesozoic. The downloaded data was further filtered manually in OpenOffice.org Calc to exclude any taxon unidentifiable to genus. All avian taxa were excluded and Mexican and Alaskan faunas were removed to keep our data more comparable to previous work. Generic level identifications were used for the same reason. Formations were divided up by state and province, approximating the divisions of Lehman (1987). However, we found similar results when formations were not divided by state. In determining values of species richness for the entire data set, all formations where at least one genus with available abundance data was present were used. Dinosaur fossil records for many individual formations were too fragmentary to be used in an analysis of this type. In order to have adequate sample sizes for the rarefaction and species estimation, we used only sets which had greater than 100 specimens, limiting us to four assemblages. As well, while there were two cases of bonebeds included, these were eliminated due to their confounding effects on the various statistical methods. While records include both instances of specimens and individuals, representing isolated elements and relatively complete skeletons respectively, we did not differentiate so as to retain as much data as possible. Absolute values for species diversity were lower when analyses were run with only specimen records although beta diversity values were nearly unchanged (Table 5–1).

Diversity can be partitioned into different components at different levels of scale. Whittaker (1972, 1977) defined alpha (α) diversity as species richness on the local or habitat scale, with beta (β) diversity as the differences between areas of alpha diversity. Together, alpha and beta make up gamma (γ) diversity, or species richness at the landscape scale.

$$\gamma = \alpha\beta$$

By rearranging this relationship, Whittaker (1960) originally defined beta diversity as

$$\beta_W = \frac{\gamma}{\alpha}$$

While the measure is useful, its interpretation at times can be confusing. Because of the way it is calculated, the minimum value (complete similarity) is 1, while the maximum value (complete dissimilarity) is equal to the number of regions of alpha diversity used. Instead, Harrison et al. (1992) suggested the modification

$$\beta_{H1} = \frac{\frac{\gamma}{\alpha} - 1}{(N-1) \cdot 100}$$

which then gives a more intuitive value between 0 and 100, with 0 being complete similarity and 100 complete dissimilarity between sites. While many other methods of calculating beta diversity have been suggested, Whittaker's measure has remained one of the simplest and most commonly used (Magurran 2004).

Data analysis and all figures were done using the R Statistical package (R Development Core Team 2010), with the packages ecodist (Goslee and Urban 2007), fossil (Vavrek 2010), PBSmapping (Schnute et al. 2004), proj4 (Urbanek 2008) and shapefiles (Stabler 2006). For each formation, randomizations were run 1000 times. Rather than true rarefaction, Coleman curves were calculated, using Coleman's "random placement" method, which provides results virtually indistinguishable from rarefaction, but is computationally much simpler (Coleman 1981; Coleman et al. 1982; Colwell and Coddington 1994; Magurran 2004). Values reported for species estimators are the statistical average, calculated from 1000 randomizations. For the species estimators, values are from estimates calculated at n=100. The estimates were calculated at this level to compensate for any bias from sampling intensity. For the full data tables and R code used for the analysis and figure generation, please consult the appendices at the end of this thesis.

5.4 Results

Using observed values of generic richness, with alpha diversity calculated from the average richness of all formations ($n = 24, \overline{\alpha} = 5.46$), beta diversity is relatively high at $\beta_{H1} = 8.24$ (Table 5–2). This corroborates the high endemism found by previous workers, but takes no account of sampling effects. There is an obvious and unsurprising correlation between generic richness and sample size for this dataset ($R^2 = 0.79$, $p << 0.001$, Fig. 5–2), illustrating the bias of differential sampling intensity. Only 11 of the 24 formations have greater than 10 specimens recorded and thus low alpha diversities are due to insufficient sample sizes rather than actual conditions. This range of sampling intensity causes beta diversity estimates to be higher than those obtained from evenly, well sampled data (Novotny et al. 2007).

To reduce this bias when calculating the average alpha diversity, we eliminated all formations with fewer than 100 specimens. This eliminated all but four formations, namely Hell Creek (Montana), Hell Creek (North Dakota), Horseshoe Canyon (Alberta) and Lance (Wyoming). Formations with greater than 100 specimens are expected to give the most robust values of actual alpha diversity and are therefore more informative than keeping formations with fewer specimens. We felt that using only formations with more than 100 specimens was an appropriate compromise between maintaining high numbers of specimens while retaining enough localities for reasonable estimates. Average alpha diversity uncorrected for sampling bias (S_{obs}) for these four localities was 13.5. However, for the estimates of gamma diversity, we retained all formations for the calculation. This method would likely overestimate gamma diversity, as there were many species known for this region that have not been found within these four formations.

When beta diversity was recalculated with an average alpha diversity calculated from only the four formations mentioned above, beta diversity drops greatly to $\beta_{H1} = 3.33$. However, the sampling intensities for each formation and the region in total vary greatly, from n = 111 (Lance Wyoming) to n = 268 (Hell

Creek North Dakota). Rarefaction methods were used to compensate for these large differences (Fig. 5–3). After all samples were rarefied, generic richness for each formation dropped comparatively little ($\overline{\alpha} = 11.95$), while gamma diversity decreased by nearly one half. Consequently, beta diversity also showed a large decrease to $\beta_{H1} = 1.84$.

One criticism of rarefaction is that while accounting for sample size it does not reflect different rates of increase in species richness due to differences in underlying species evenness (Magurran 2004; Fager 1972). For example, an area with high species evenness will show a faster rate of increase than an area with a low evenness, due to very rare species taking much longer to be found. A way to compensate for this problem is by extrapolating to an estimate of total species richness, using species estimators.

Non-parametric estimators, though designed to be independent of sample size, are still affected by sampling to some extent. In order to correct for this, we ran estimators 1000 times on a randomized subsample from each locality with greater than 100 specimens. While the estimated richness from the three methods differed considerably, the overall result was a higher estimated alpha diversity ($\overline{\alpha} = 15.64$) and a lower β_{H1} value. The Chao 1, ACE, and Jacknife 1 estimations yielded beta diversities of 2.26, 2.03, and 2.08, respectively.

Finally, NMDS networks offer a test of similarity between all sites. If there were three distinct provinces as hypothesized, there would be at least three discrete clusters of sites within the network. If each individual site has large numbers of endemic taxa, the NMDS would show overdispersal, with all sites plotting in a ring to maximize site to site dissimilarities. The NMDS network plot is significantly different than either of these possibilities (Fig. 5–4B). Sites are scattered evenly throughout the plot space and no clusters are present, with little patterning in regards to endemism. The same network plotted on geographic space (Fig. 5–4A) emphasizes the lack of regional clusters. In general, dinosaur faunal similarity between sites is poorly associated with geographical distance (Fig. 5–4C), as pairwise values of generic similarity show little decay over distance.

5.5 Discussion

We find no evidence to support distinct faunal regions of dinosaurs during the Maastrichtian of the Western Interior of North America. While our estimates of beta diversity are much lower than what might be expected from direct observation of the fossil record, the effect of uneven sampling can be very large. When sampling the richness of a region, one should not expect to find all the species present, with rare species often not detected even in large samples (Siemann et al. 1996).

Another possible explanation for low beta diversity may be the time scale used. Studies on modern community associations are limited to relatively brief periods of sampling time. Moreover, the Maastrichtian represents approximately six million years and time averaging effects are undoubtedly confounding the data. Work on shorter time scales during the Pleistocene has shown that mammal species move independently, reorganizing community compositions in time scales of only 350 000 years (Potts and Deino 1995; Jablonski and Sepkoski Jr. 1996). Even the large faunal exchange between North and South American mammals during the Great American Interchange occurred over approximately 3 million years, with continental scale migrations of mammals from shrews to mastodonts (Marshall et al. 1979). While the climate of the Late Cretaceous was more stable than that of the Holocene, the suggestion that large, motile animals might maintain cohesive units seems unlikely.

Harrison et al. (1992) obtained beta diversity levels in British birds between $\beta_{H1} = 3.3$ and 5.7, higher levels than we calculated here ($\overline{\beta}_{H1} = 2.14$) and over a much smaller latitudinal range. Our results indicate less turnover across a greater distance and adds further evidence for a lack of dinosaurian endemism and provinciality during the Maastrichtian along the Western Interior Seaway. Novotny et al. (2007) present a lower beta diversity for tropical insects than we found for dinosaurs, noting that their study sampled a relatively climatically homogeneous landscape. One possible implication is that the low levels of beta diversity within Maastrichtian dinosaurs are due to climactic factors. During the latest Cretaceous the global climate was much warmer, with more equable temperatures and a greatly reduced latitudinal temperature gradient (Barron 1983; Amiot et al. 2004). The average yearly temperature at the equator for the Maastrichtian has been suggested to be only slightly warmer than it is today while polar regions were estimated to have been 15 to 25°C higher than they are today, with mean temperatures for the coldest months likely above freezing (Amiot et al. 2004).

Although this analysis does not support distinct communities of dinosaurs, some dinosaurs probably were restricted in their ranges. For example, the sauropod *Alamosaurus* likely did not live in the northern-most regions of the Western

Interior region. This taxon has not been recovered north of Utah despite its large size and conspicuousness. However, it seems counter-intuitive that an animal as large as *Alamosaurus* would not have dispersed at low levels to regions other than the restricted area it is found over its several million year existence. Modern species ranges often have one or more high density peaks, where the species is highly abundant, and tails, where the species can still be found but at very low densities (McGill and Collins 2003). The fossil record of *Alamosaurus* likely shows these high density regions, and due to random chance, individuals which may have lived in the low density regions of their ranges are simply not preserved. Therefore, while we may be seeing signatures of generic and species ranges within the fossil record, we are by no means seeing the entire region in which they lived. Rather than faunal provinces per se, previous research has likely recovered these high density areas for specific taxa.

Our results suggest that any one region of the Western North American Seaway would have had much higher dinosaur species richness than that observed, with about 16 species of dinosaur on average. At the continental scale, levels of beta diversity among dinosaur assemblages are comparable to modern terrestrial faunas with low endemicity. These results suggest that dinosaurs were not as restricted in their ranges as once thought, and that the fauna as a whole was largely homogeneous.

Figure 5–1: Approximate locations of Maastrichtian aged dinosaur-bearing formations from the Western Interior of North America used in this analysis. A)Abbreviations of formation names are as follows: Aguja (A); Denver (D); Ferris (Fe); Frenchman (Fn); Hell Creek Montana (HM); Hell Creek - North Dakota (HN); Hell Creek - South Dakota (HS); Horsehshoe Canyon (HC); Javelina (J); Kaiparowits (Kp); Kirtland (Kd); Lance - Montana (LM); Lance - South Dakota (LS); Lance - Utah (LU); Lance - Wyoming (LW); Laramie - Colorado (LaC); Laramie - Wyoming (LaW); McRae (M); North Horn (N); Pinyon Conglomerate (P); Scollard (Sc); St. Mary's River - Alberta (SA); St. Mary's River - Montana (SM); Tornillo - Texas (T). B) Locations of occurrences of the three indicator taxa, *Alamosaurus* (A), *Leptoceratops* (L) and *Triceratops* (T). Greyed area to the east is the approximate location of the inland seaway.

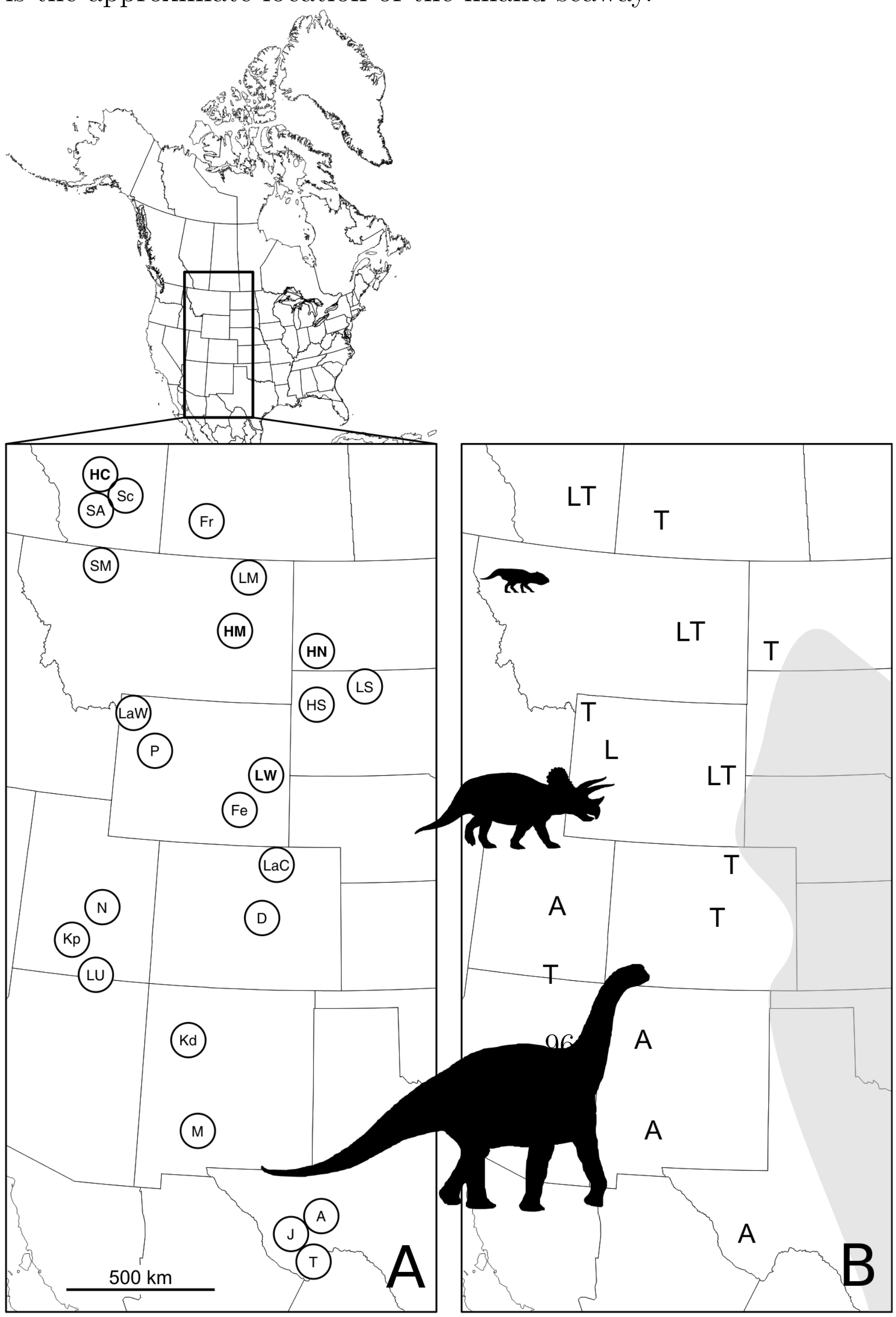


Figure 5–2: Comparison of sampling intensity (number of occurrences) and species richness for the Maastrichtian Western Interior

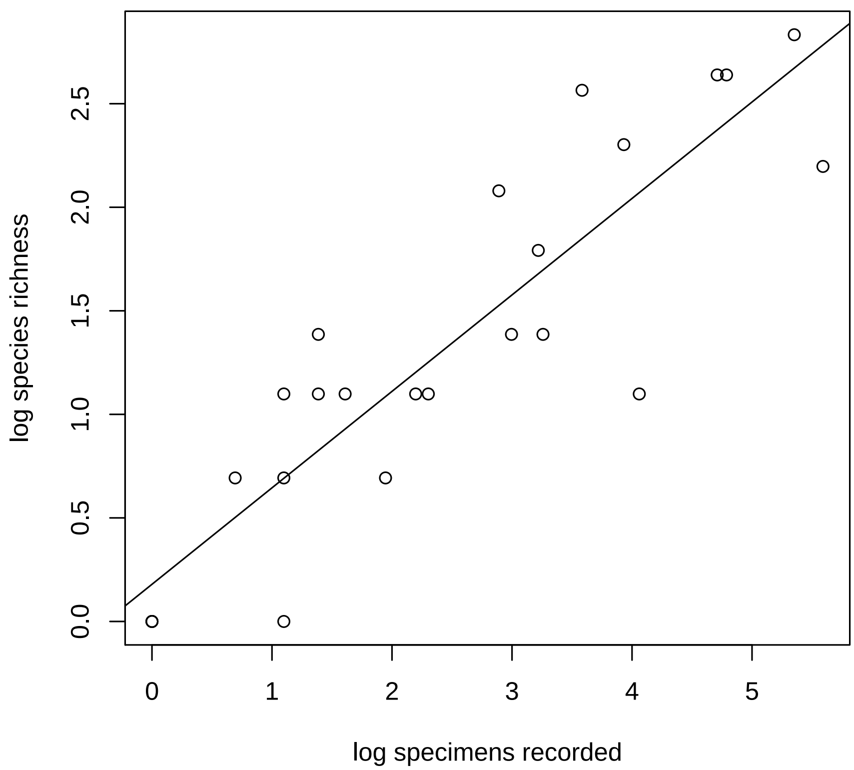


Figure 5–3: Comparison of methods of sample size correction. A) Smoothed curves for rarefaction using Coleman's random placement method, up to N = 100. B) Smoothed curves for Chao 1 values up to N = 100. C) ACE values. D) Jacknife 1 values.

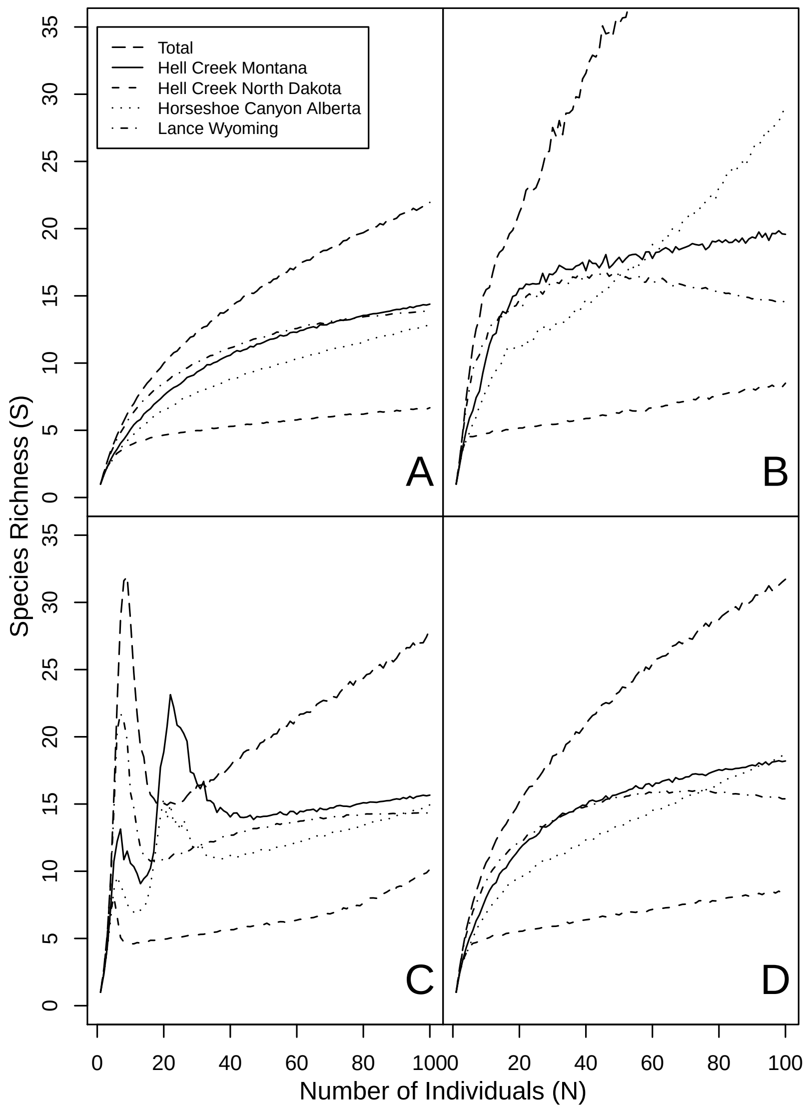


Figure 5–4: Relative pairwise similarities between localities. A) Map of the localities used in the analysis with an MST added to indicate the relative similarities between localities. The more similar two localities are to one another based on the species present, the darker the line connecting them will be. Note that no clusters are formed and many localities are more similar to far ranging ones than neighboring localities. B) A plot of the relative positions of an NMDS of the localities, showing no apparent clusters. C) Pairwise Dissimilarity of sites in comparison to their geographic distance from one another. As dissimilarity increases, the sites are interpreted to be less similar to one another. The line represents a Locally Weighted Sum of Squares function. The sites show a slight decay over distance although a linear regression is not significantly different from 0. These graphs indicate low beta diversities for this assemblage.

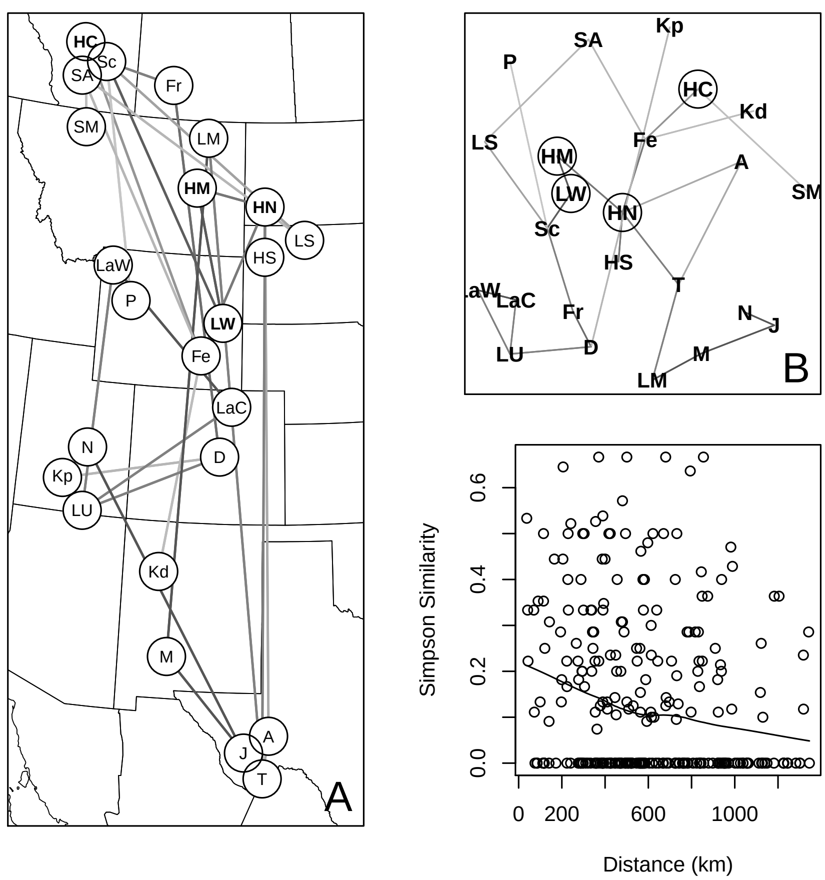


Table 5–1: Results for rarefaction and alternate methods of species estimation for the highlighted formations and full region. This table gives the results for the analysis when run with only the specimen data (removal of individual occurrences)

Locality	Number of Specimens	Species Richness (observed)	Rarefaction	Chao 1	ACE	Jackknife 1	Average Estimated Richness
Hell Creek Montana	180	16	14.06	18.63	15.46	17.8	17.3
Hell Creek North Dakota	263	9	6.38	7.9	9.26	8.1	8.42
Horseshoe Canyon Alberta	102	10	9.94	14.19	10.98	12.87	12.68
Lance Wyoming	82	12	12	12.17	12.24	12.99	12.46
Average α	156.75	11.75	10.59	13.22	11.99	12.94	12.71
Western Interior	865	38	20.03	38.38	26.2	29.02	31.2
β_{H1}	-	3.23	1.89	2.9	2.19	2.24	2.45

Table 5–2: Results for rarefaction and alternate methods of species estimation for the highlighted formations and full region

Locality	Number of Specimens	Species Richness (observed)	Rarefaction	Chao 1	ACE	Jackknife 1	Average Estimated Richness[1]
Hell Creek Montana	211	17	14.38	19.58	15.66	18.2	17.81
Hell Creek North Dakota	268	9	6.68	8.5	10.15	8.58	9.08
Horseshoe Canyon Alberta	120	14	12.85	28.99	14.94	18.73	20.88
Lance Wyoming	111	14	13.88	14.58	14.35	15.39	14.77
Average α[2]	177.5	13.5	11.95	17.91	13.77	15.22	15.64
Western Interior[3]	997	45	21.95	40.55	28.02	31.72	33.43
β_{H1}	-	3.33	1.84	2.26	2.03	2.08	2.14

[1] Average Estimated Richness is the mean value of Chao 1, ACE and Jacknife 1.

[2] Values for the average α were calculated using only the four listed formations.

[3] Values for the entire Western Interior region were calculated using all 24 localities from the dataset.

Bridging Text

Whereas the previous chapter on beta diversity was a way of using the present to inform the past, so too can we use the past to better understand the present. The following chapter describes changing latitudinal diversity gradients through deep time. Although the gradient has been quantified hundreds of times, our understanding of its origin and history remains poor. The fossil record provides us with a vast temporal scale on which to observe changes in the intensity of the latitudinal gradient, something unavailable from modern data sets. Though there is a growing literature of fossil latitudinal species gradients, much of the work has been done on isolated time periods and did not correlate the gradient with any proposed driver. This analysis tracks the diversity gradient in plant macrofloras through time and can then test the correlation of these shifts in the gradient to driving factors such climactic conditions. There is a relationship, though not significant, between the intensity of the latitudinal gradient and palaeotemperature estimates. By studying the gradient over a time series, we can begin to differentiate between the dozens of proposed drivers of the gradient, and come to a better understanding of how and why it came to be.

CHAPTER 6
Evolution of Macrofloral Latitudinal Diversity Gradients From the Late Cretaceous to Tertiary of North America

6.1 Abstract

The latitudinal gradient of species diversity is virtually ubiquitous across ecosystems, yet our understanding of the processes behind the pattern are poorly tested. Fossil data provide an opportunity to examine latitudinal gradients through time in an attempt to observe their changes in correlation with varying climactic, ecologic and geographic conditions. Using plant macrofossil data, we find that there is some correlation with average global temperatures though time, with periods of increased temperature showing a decreased latitudinal gradient. Although there remain many other processes and groups to test for a correlation, this paper serves as a starting point for a quantitative assessment of the Latitudinal Diversity Gradient using fossil data.

6.2 Introduction

One of the nearly universal patterns within ecology is the Latitudinal Diversity Gradient (LDG). As one travels away from equatorial regions, the numbers of species decreases, whether in the southern or northern hemispheres, in terrestrial or marine environments, amongst plants, invertebrates or vertebrates (Willig et al. 2003; Hillebrand 2004; Mittelbach et al. 2007). The gradient is the oldest diversity pattern to be recognized by ecologists, first noted by the geographer von Humbolt over 200 years ago (Hawkins 2001).

Although latitude itself is not the cause of the gradient, it can be correlated with several interrelated factors such as temperature, insolation, and area (Gaston 2000). Despite this, the underlying mechanism by which the LDG was created and is maintained remains ambiguous. In fact, over 25 different hypotheses have been suggested to explain the LDG (Gaston 2000), ranging from species origination/extinction rates to recolonization after glacial conditions (Pianka 1966; Hawkins et al. 2003). Rather than narrowing the focus, the number of explanatory hypotheses for the gradient has only been increasing during the last several decades, with few attempts to reject or determine relative influences of each (Platt 1964; Willig et al. 2003). Though the gradient has been quantified hundreds of times (Hillebrand 2004), it remains the major, unexplained pattern of natural history [Ricklefs in Lewin (1989)]; as Gaston (2000) stressed, a "predictive theory of species richness" is still distant.

6.3 Proposed drivers of the LDG

Though not exhaustive, we provide here an overview of the more common hypotheses for the existence of the LDG; for a more in depth review of the proposed mechanisms, we refer the reader to Pianka (1966), Willig et al. (2003) and Mittelbach et al. (2007). Broadly, the hypotheses put forth to explain the gradient can be divided into three categories: mechanisms of species coexistence and maintenance of diversity through ecological processes; evolutionary hypotheses with a focus on rates of speciation and extinction; and historical hypotheses concerned with the duration and extent of tropical environments through deep time. These groupings are used simply for convenience, and some of the hypotheses could be placed in more than one category.

6.3.1 Ecological Processes

Ecological mechanisms to explain the LDG have received a large amount of the research over the past several decades. Ecological processes encompass a variety of climactic factors, including insolation, precipitation, temperature and seasonality, and various combinations thereof. Generally, these types of hypotheses presume that the latitudinal gradient exists due to unequal carrying capacities between different regions (Fig. 6–1A; Mittelbach et al. 2007)

Insolation is perhaps one of the easiest hypotheses to address. The idea holds that the LDG is driven by the difference in the intensity of sunlight across the globe. Equatorial regions receive the most solar energy and the poles receive the least, and this fundamental difference in energy availability leads to the LDG. However, there are several sources that throw this theory into doubt. First, deep-sea benthic communities which are virtually unconnected to the solar energy cycle also demonstrate a latitudinal gradient, necessitating a different mechanism for their existence (Thomas and Gooday 1996). As well, insolation has changed relatively little over geologic time, meaning that previous studies which found a less intense gradient also do not support insolation as a valid hypothesis.

Although precipitation does affect species diversity over smaller scales, it also is unable to explain the LDG at a global scale. Precipitation does not vary monotonically with latitude, making it alone unlikely the cause of the LDG. There is some evidence in modern ecosystems that precipitation in combination with temperature, also calculated as evapotranspiration, does correlate well with the current LDG (Currie and Paquin 1987; Currie 1991).

In regards to temperature, there have been many more studies examining its effect and the closely related effects of factors such as climatic stability, environmental stability, environmental predictability, seasonality, and harshness (Willig et al. 2003). Often more broadly referred to as the ambient energy hypothesis, much of this research likely has to do simply with the availability of vast amounts of global temperature data than any other reason. The hypothesis holds that high-latitude regions generally have conditions that are colder and more environmentally unpredictable, making high latitudes more physiologically costly to live in (Brown 1988; Willig et al. 2003). From the palaeontological perspective, the ambient energy hypothesis does make some sense. Previous time periods with reduced diversity gradients may be due to higher temperatures or reduced seasonality, two conditions which have been demonstrated through multiple studies (Barron and Washington 1982; Huber et al. 2002; Sluijs et al. 2006; Wolfe and Upchurch 1987; Spicer and Parrish 1990).

6.3.2 Evolutionary Processes

Evolutionary hypotheses for the existence of the LDG, though not as extensively studied as ecological hypotheses, nonetheless have a much longer history of study (Darwin 1862; Wallace 1878; Mittelbach et al. 2007). Traditionally, arguments for evolutionary hypotheses invoke the idea that the tropics operate as either a "cradle", with elevated speciation rates, or a "museum", with decreased rates of extinction (Stebbins 1974). High rates of speciation can be caused by a number of means, including: higher rates of mutation due to higher temperatures (Rohde 1992; Allen et al. 2002); quicker genetic drift from smaller population sizes (Fedorov 1966); greater specialization possibilities due to lowered climactic variation (Haffer 1969; Dynesius and Jansson 2000); or stronger biotic interactions forcing faster rates of evolution (Dobzhansky 1950, see Table 1 for full summary of evolutionary mechanisms).

Despite the diversity of evolutionary hypotheses, a common difficulty with many of them is their circularity. Many of the proposed causes of the LDG rely on stronger biotic interactions taking place in the tropics, yet this often requires a higher diversity of species to be present. When viewed in the context of deep time, this means that the LDG has always existed, which we know to be untrue. While many of these mechanisms may be able to sustain the LDG, they are unable to initiate it.

Although biotic interactions may not be able to explain the origin of the Recent LDG, many other evolutionary hypotheses invoke underlying abiotic causes. One group of abiotic-driven hypotheses are those related to geography. Geographic factors as causes of the LDG were first proposed by Terborgh (1973) and later expanded upon by Rosenzweig (1995). These hypotheses are predicated on the fact that the earth is spherical, and due to this property there is simply more surface area at the equator than the poles. This increased surface area could lead to larger population and range sizes and thus lead to decreasing rates of extinction (Terborgh 1973; Rosenzweig 1995). Although the idea is intuitive, the majority of equatorial species have smaller populations and ranges than their more temperate or polar counterparts (Chown and Gaston 2000).

Abiotic environmental factors, such as temperature and seasonality, have also been suggested as underlying mechanisms of evolutionary hypotheses for the LDG. For example, Rohde (1992) and Allen et al. (2002) proposed that the high diversity in the tropics was due to increased rates of mutation (and subsequent speciation) as a result of higher temperatures (Mittelbach et al. 2007). There is support both for and against this hypothesis, though. Martin and Palumbi (1993) found that endothermic animals had much higher mutation rates than ectotherms. However, Weir and Schluter (2007) find that in New World birds and mammals speciation is actually more recent at the poles. Also, environmentally-driven evolutionary hypotheses can at times be difficult to distinguish from some ecological hypotheses. A number of authors (i.e. Darwin 1862; Wallace 1878; Fischer 1960) proposed that the tropics had a lower extinction rate because of the relative stability of the climate. This evolutionary hypothesis is very similar to the ambient energy hypothesis, with the main difference being a focus on rates of extinction rather than physiological effects. Arguably, these hypotheses could simply be slightly different symptoms of the same underlying cause.

6.3.3 Historical Processes

Historical hypotheses are the final category of explanations for the origin of the Recent LDG. This set of hypotheses makes the general assumption that older communities are more diverse. The size and extent of communities and biomes has varied greatly over time, in response to global warming and cooling, continental drift and sea-level change. Of these, only global warming and cooling have been studied in detail as possible explanations for the LDG. The Pleistocene glaciations have long been proposed as a possible cause for the LDG (Wallace 1878; Fischer 1960). However, the ample evidence for a pre-glacial LDG rejects the hypothesis that these glaciations were the only driver of the LDG. There is mounting evidence though that the LDG was less pronounced in the past, and the Pleistocene glaciations may have caused a steepening of the modern slope (Crame 2001; Mittelbach et al. 2007). For example, studies of both European and North American pollen samples demonstrate that the gradients became steeper over time (Silvertown 1985; Willig et al. 2003). However, the timing and duration of the changes in the slope do not coincide between the regions, and a separate analysis by Haskell (2001) found no change in the slope of the gradient in the last 10 000 years using familial pollen richness.

A final explanation for the LDG involves global cooling over a much longer time period. Fine and Ree (2006) found that while modern biome area was not significantly correlated with the LDG, when area was integrated with time the correlation became significant. To obtain a time-integrated area value, they combined the extent of biomes from past time periods since the Eocene; because the climate has been generally cooling since that time, the tropics are much smaller today than the past (Fine and Ree 2006). Since the tropics were so much larger in the past, there should be more species than other areas based simply on the species area effect. Unfortunately, they did not explore any of the presumed mechanisms (increased speciation/decreased extinction) by which area affects diversification (Mittelbach et al. 2007). Nonetheless, while there is some support for historical factors in creating the modern LDG, and the effects of time need to be considered in the study of the LDG.

There is an overwhelming amount of overlap between these different categories, and the true drivers of the LDG are likely some combination of factors from each group. For example, the tropics have smaller bodied species, possibly because of higher competition for resources and more complex food webs, leading to higher mutation rates and higher rates of speciation, all possibly driven by a higher rate of evapotranspiration. Species are being constantly affected by both the biotic and abiotic portions of the environment, which may combine to create non-intuitive results.

6.4 Evidence of the LDG in the Fossil Record

Although studies of latitudinal gradients in modern environments have the benefit of large amounts of data, the data covers only a relatively short amount of geologic time. Mechanistic studies on the cause of the gradient involving experimental manipulation of environments are often difficult to perform, as the processes of speciation and extinction are typically unobservable over short time scales. Controlled experimentation is not possible for fossil data either, however it has the advantage of large temporal scales over which great changes have occurred in both the position and shape of the continents (additional or less landmass at tropical latitudes), global temperatures and temperature gradients (environmental drivers of species richness) and opportunities for phylogenetic diversification and contraction (increased or decreased rates of speciation and extinction). Though the LDG has not received nearly as much attention by palaeoecologists as it has from neoecologists, there have been a handful of notable studies to date.

6.4.1 Marine Studies

Modern studies of the LDG have slight bias towards terrestrial ecosystems (Willig et al. 2003), yet most palaeoecological studies of the LDG are biased towards the marine record. This is due to the abundance of hard-shelled marine organisms in the fossil record (Crame 2002; Cecca et al. 2005; Leighton 2005) compared to relatively more depauperate terrestrial remains (Crane and Lidgard 1989; Ziegler et al. 1993).

Stehli et al. (1969) were the first to study gradients in the fossil record and their study covered several groups, including Cretaceous foraminifera and Permian brachiopod assemblages from the Northern Hemisphere. They demonstrated a distinct negative (or normal) gradient for all the groups they examined, however the methods they used make it difficult to compare the intensity of the LDG to any other study. They used a ratio of families instead of the more typical species richness in their calculations, and did not discuss how the gradient relates to modern gradients in intensity. They did provide some interesting correlations between taxon ages and latitude, with the lower latitudes containing a higher proportion of recently evolved taxa. To connect this mechanistically to the gradient, they suggest solar energy (ambient energy hypothesis) as the underlying driver.

While more concerned with biogeography and continental positions, Belasky (1994) did briefly discuss patterns of large scale species diversity. He found that rugose corals exhibit the highest diversity in the tropics, and progressively decrease towards the poles. He attributed this gradient to latitude-related thermal gradients of the time; Permian climate at the time showed a similar latitudinal thermal gradient to today (Stehli et al. 1969; Ziegler 1990; Kutzbach and Ziegler 1993).

In their study of benthic foraminiferans, Thomas and Gooday (1996) found that high latitude faunas showed a decrease in species diversity starting in the Eocene. They hypothesize that this decrease was due to the change from "greenhouse" to "icehouse" conditions that occurred around that time, and that richness was most affected by seasonality rather than temperature directly. However, these hypotheses must be taken at their word, as their figures are difficult to interpret and they provide no quantitative values of the gradient.

Crame (2002) found that Tithonian (Late Jurassic) bivalves show a general gradient that is similar to modern trends, however it is only significant if some low diversity samples near the equator are excluded. The author argued that these samples should be excluded because they likely represent undersampled localities, but the possibility of an ecological explanation also exists. As well, a situation where high latitude sites are undersampled due to accessibility and logistics is also possible, and could explain some modal gradients found where high diversity seems to correspond to the continental US and Europe (Raup 1976). Other studies (Shen and Shi 2004) have also shown a large deflection in the gradient, with a peak in diversity at 40°N, consequently leading to more questions than answers as to the mechanisms driving macro-scale diversity.

6.4.2 Terrestrial Studies

While the marine realm has seen the majority of work on ancient LDGs, there have been some studies of note in the terrestrial realm, largely dealing with pollen and plant remains. Crane and Lidgard (1989) studied the latitudinal diversification in angiosperms, tracing the dispersal and subsequent dominance of pollen floras through the Cretaceous. They found that the centre of angiosperm diversification was near the equator, with species then expanding to higher latitudes, with angiosperm dominance decreasing towards the poles. Although there was a distinct dominance gradient at the time, they did not address the actual LDG. Their study used relative percentages of angiosperm species pollen within different samples and they did not report the species diversities of any of the sites. It is difficult to state what the actual gradient looked like at the time, although their analysis does have implications of historical models of taxon radiation in creating the LDG.

Although their paper was more concerned with climate and phytogeography, Ziegler et al. (1993) did briefly discuss the presence of a LDG in macrofloras during the Triassic and Jurassic. However, the peak for this gradient was at about 40°N and may have been caused by sampling bias; the most diverse areas appear to be the most heavily sampled. This problem was likely exacerbated by binning the data into zones defined by latitudinal and longitudinal areas. Previous work (Anderson and Marcus 1993; Lyons and Willig 1999) has discussed how degree-delimited quadrats vary substantially in their absolute size between latitudes due to the effects of the curvature of the earth. At higher latitudes, areas defined by geographic coordinates are much smaller in actual areal extent than those at the equator, and even if there was no gradient one would be found due to the species area effect (Rosenzweig 1995; Anderson and Marcus 1993). In fact many studies of the LDG in the fossil record have involved binning data. In a study by Raup and Jablonski (1993) the gradient found was hump-shaped and may also be due to more intensive sampling at mid-latitudes, such as in the US and Europe. Likewise, Rees (2002) also provided a latitude-delimited analysis of generic diversity of Permian and Triassic plants, which showed a bimodal distribution. However, as the study was not directly concerned with the LDG, there was no compensation for sampling bias, and the areas with the highest diversities appear to correspond with those most intensively surveyed (e.g., Europe, North America). One of the few other papers to discuss terrestrial diversity gradients was that of Anderson et al. (1999), who suggest that the hot-house conditions of the Triassic may have led to a reversed LDG. However, they conclude this based on only two regions and no quantitative analysis, so yet again their data is difficult to assess.

Overall, our knowledge of the history of the LDG is patchy and at times contradictory. Although a flat, hump-shaped and positive (reversed) gradient have all been found, the majority show a negative (normal) gradient, albeit one with a reduced intensity (Crame 2001). Unfortunately, few studies to date have made any attempt to study the gradient in any sort of temporally continuous and extensive manner.

No matter the mechanism that has created the latitudinal gradient we see today, the origin of the Recent LDG must occur at some time in the geologic past, and the fossil record offers an exclusive view of the deep history that would have led to the emergence of the gradient. The purpose of this study is to create a testing framework for possible drivers of the LDG. We first derive the LDG through a series of consecutive time periods from the mid-Cretaceous to the present, and then correlate the intensity of the LDG to past global temperatures.

We predict that periods of "greenhouse" conditions (elevated global temperatures) will be correlated with a decrease in the intensity of the LDG.

6.5 Material and Methods

In order to study the LDG both through time and space, a large database of fossil occurrences needed to be built. We chose to examine the LDG from the Aptian (c. 120 mya) to the Miocene (c. 35 mya) with fossil data and additional data for the Recent. This timeframe included large temperature shifts, continental migration, and even a mass extinction. The database used for this study was compiled through both entry of primary sources and entries downloaded from the Paleobiology Database (PaleoDB). All occurrences that were added from primary sources will be merged with the PaleoDB as a permanent record of the data. Occurrence data from the PaleoDB were downloaded using the taxon name "Plantae" with the parameter Abundance Value = TRUE, Formations = TRUE, Latitude/Longitude = TRUE and Country = TRUE. All the data was further sorted to differentiate between all mega- and microfossils, so that only leaf impressions were used. The database was initially compiled using OpenOffice.org Calc, and later loaded into a PostgreSQL/PostGIS database. The tables for the database and code to convert it to PostgreSQL can be found in the supplementary materials. Data analysis and figure creation was done with the R Statistical Language (R Development Core Team 2010), using the 'fossil' (Vavrek 2010) and 'RODBC' (Ripley and Lapsley 2009) packages. For all R code used to connect to the database and create the figures, please refer to Appendix G.

From the database, plant occurrences were selected only if they met several certain criteria. Only macrofloral occurrences identified to at least the generic level were used in order to eliminate poorly known materials and to standardize at a certain taxonomic level. The latitudinal gradient has been consistently demonstrated with modern generic level data (Qian 1998; Willig et al. 2003). Localities without specific provenance were excluded. Although the database did contain global data, we only used sites from the US and Canada, as these two countries contained most of the localities to begin with, rather than including rare and possibly confounding data points from other continents. Although we collected both abundance and occurrence data, for the final analysis we used only the occurrence data. At present, abundance data for fossil floras is relatively sparse. After preliminary work with both types of data, we did not find significantly different results. In order to retain a greater latitudinal range and larger dataset we based the study on occurrence values.

To determine the gradient for any time period, we calculated a generalized linear model using locality diversity versus adjusted palaeolatitude. The palaeolatitude was calculated using a modified version of the Earth System History Geographic Information System (ESH-GIS) v. 02b (Scotese 2001). Locality diversity is based on the number of genera identified at any locality. We did not combine localities, as the uneven sampling meant that binning localities would have lead to even more misleading results (see Chapter 1 for a discussion on effects of binning data). Not all time periods were used in the final analysis due to insufficient numbers of localities. The final time periods/stages that were used were the Miocene, Eocene, Paleocene, Maastrichtian, Campanian, Cenomanian, Albian and Aptian. The Oligocene data was not used due to an insufficient spread in the position of the localities.

For a reference gradient, we have used the data from Gentry (1988) as a modern baseline. Previous actualistic studies on modern leaf litter accumulation found that samples similar in size to those from fossil studies closely reflect tree species within 10–25 m from the sample (about 300–2000 m^2, or 0.03–0.2 ha; Burnham et al. 1992). The study by Gentry (1988) sampled 0.1 ha plots, making it a good analogue to fossil datasets (Wing and Dimichele 1995). Gentry (1988) used species level data, and only provided absolute species counts for each locality.

6.6 Results and Discussion

Beginning with the Recent, data from Gentry (1988) reveals a steep negative slope for the LDG (slope $= -2.938$, $p << 0.001$; Figure 6–2, Table 6–2). There is high variance at lower latitudes but the regression is still significant. When reduced to only localities above 30°N in order to be more comparable to the fossil data, the slope is less steep although no longer significant (Figure 6–3). Miocene localities are spread throughout the western margin of North America, with one locality in the east, between palaeolatitudes 35 and 50°N (Figure 6–4). The LDG regression slope is -1.07 and is less intense compared to the Recent slope.

The Oligocene localities are clustered in three regions; the central western Rockies, Gulf of Mexico, and St. Lawrence Valley between palaeolatitudes 30 and 45°N (Figure 6–5). The southern locality data is sparse and there is high variance in the northern localities, the LDG slope is positive and approximately

1.51, however this very unusual value is likely due to the constrained latitudinal variation in all but one of the localities.

There is a reversed LDG again in the Eocene, however there is more data for this time period and the value is significantly different from zero. These localities are relatively evenly distributed across North America between palaeolatitudes 30 and 55°N (Figure 6–6). There appears to be some sampling bias near palaeolatitudes 40-45. In spite of the mid latitude oversampling, the linear regression is still positive, at 0.303 ($p = 0.033$).

Paleocene data are densely distributed along the western cordillera, with one locality on the eastern shore (Figure 6–7). These span between palaeolatitudes 34 and 60°N. There does not appear to be a bias in oversampling mid latitudes but the LDG slope is nearly flat (slope = 0.035) and not significant ($p = 0.394$).

Maastrichtian localities are spread along the western plateau from palaeolatitudes 37 to 54°N (Figure 6–8). There is a high sampling bias around palaeolatitude 46°N, due principally to the collections of Kirk Johnson. In spite of the hump-shaped collecting bias, the LDG regression is negative (slope = -0.4) and significant ($p = 0.007$).

Campanian localities are also confined to the western cordillera and span palaeolatitudes 38 to 54°N (Figure 6–9). There is again a sampling bias between palaeolatitudes 45 and 50°N. The LDG regression is negative at -0.992 ($p = 0.002$).

Cenomanian data is widespread throughout North America ranging from palaeolatitudes 37 to 59°N (Figure 6–10). There is no apparent sampling bias and the LDG slope is nearly flat at 0.06 but significant ($p = 0.027$).

Albian localities are the most extensive ranging from the southeastern corner of North America to northern Alaska (Figure 6–11). These span palaeolatitudes 36 to °N. The LDG slope is nearly horizontal at -0.08 and significant ($p = 0.283$).

Aptian localities are the most sparse and confined to a handful of sites related to the western cordillera and a single locality in the eastern USA (Figure 6–12). The localities span palaeolatitudes 37 to 56°N. The LDG slope is positive at 0.49 and significant ($p = 0.033$).

Throughout the time period we examined, the LDG was less intense compared to its modern state (Fig 6–13). Some time periods such as the Paleocene and Eocene appear to have no gradient whatsoever. The gradient became more intense towards the end of the Cretaceous, reduced in intensity from the K/T boundary to the Eocene, and then again increased in intensity until today.

The evolving LDG over the past 120 million years likely has roots in speculated drivers of the current LDG. The easiest to assess in the palaeontological record are environmental factors known to influence extant biodiversity patterns. Although many of these are unknown in the palaeontological record, general environmental factors such as the latitudinal temperature gradient (LTG) can be assessed. The LTG is steeply negative today. Historically, the LTG has changed dramatically with lush, temperate climates at the poles during greenhouse Earth phases. Specific latitudinal temperatures are unknown for most of the planet in the geologic record, but the LTG is closely related to the mean annual temperature, a value that is well known from the geologic record. High mean annual temperatures are associated with a shallow LTG, and high polar temperatures. Lower mean annual temperatures are associated with a steep LTG, such as today.

We can compare the LDG and LTG through a time series to observe any possible correlation (Figure 6–13). In general, global mean annual temperatures above 14°C are associated with positive LDG slopes and lower temperatures with negative LDG slopes. The two periods of global cooling are associated with negatively trending LDGs. A Spearman Rank Correlation which included all the slopes yields a non-significant but still correlated value of $\rho = 0.64 (p \leq 0.076)$. Fluctuations in the LDG slope generally parallel changes in the mean annual global temperature. Only during Maastrichtian does a significant LDG and global temperature conflict. This may present interesting exceptions or highlight the need for better constrained fossil or LTG data during and immediately bounding the Maastrichtian. The current steeply negative LDG appears to have its origin during the past 35 million years. During this time, the Earth's annual mean temperatures have been steadily declining toward current icehouse conditions.

The LDG does show some correlation with average global temperatures over time, however this correlation is not significant, possibly due to several confounding factors that we have not been able to fully quantify as yet. One possibility for the increasing intensity of the gradient is the radiation and expansion of the angiosperms during the same time as our study (Crame 2001). Crane and Lidgard (1989) found a steadily increasing latitudinal gradient in angiosperms during the Late Cretaceous as the angiosperms were rapidly radiating. Magallón and Sanderson (2001) found that the most recently evolved angiosperm groups were also the most speciose. Based on this, the increase in slope of the latitudinal gradient we found could be due to the radiation of angiosperms at low latitudes while other plant groups remained static. Other groups, such as gastropods and bivalves, also show a similarly recent radiation, with younger species more concentrated in the tropics (Crame 2000; Jablonski et al. 2006). There is still some question of how ubiquitous these recent radiations are across other groups as an all encompassing explanation. There is some evidence from birds that this "cradle" (Stebbins 1974) explanation for the LDG may not suit all groups (Weir and Schluter 2007).

Another possible confounding factor for our study may simply be sampling. In assessing the LDG for several fossil groups of marine bivalves, Crame (2001) suggested that a number of relatively low diversity tropical faunas may have occurred due to poor sampling. Though this seems likely, it could be equally possible that high latitude locations were also undersampled, thus leading to the hump-shaped gradient like that found by Raup and Jablonski (1993).

Another difficulty with this type of analysis is that this type of data often generates polygon-shaped patterns (Blackburn et al. 1992; Blackburn and Gaston 1998). These polygonal relationships are difficult to handle with conventional statistics, and may violate the assumptions of some methods (Blackburn and Gaston 1998). Some of the time periods, such as the Eocene, appear to exhibit this type of distribution, and future work using statistical methods that compensate for these types of distributions must be tested.

Finally, another problem is the type of measure of temperature that we use. Global average temperatures during the Cretaceous and Tertiary were typically higher than today, but not all regions experienced this warming equally (Frakes 1979; Hallam 1985; Frakes 1999). For example, Amiot et al. (2004) found that during the Maastrichtian polar regions were much warmer than today while equatorial areas were similar to modern temperatures. Ideally then we would compare the latitudinal diversity gradient to the latitudinal temperature gradient, as the temperature gradient is a possible mechanism which could be driving the diversity gradient. However, as we do not have a full data for the latitudinal temperature gradient through time, we must rely on the more coarse value of global temperature.

6.7 Conclusion

There is a correlated trend between global temperature and latitudinal diversity gradient intensity. However, further work needs to be done to more fully refine latitudinal gradients through more constrained time periods and latitudinal temperature gradients. It is likely that the diversity gradient is correlated with the latitudinal temperature gradient rather than simply global temperatures, however until further latitudinal temperature gradient data becomes available closer examination is not possible. As well, a wider latitudinal range of localities would help to refine estimates of the LDG.

Although the presence of latitudinal gradients in the geologic past has been more firmly established for the marine realm, the slope and magnitude of the gradient is less clear on the land. The steeply negative LDG we see today must have formed or begun forming at some point in the past, likely on the timescale of millions of years ago, but the catalyst of the gradient is still poorly understood.

What is needed to gain a better perspective on the gradient are explicit tests of varying hypotheses, beginning with the data available at present and continuing as more of the palaeobiodiversity of the world is catalogued into large databases. Establishing the existence of a pattern in the fossil record is but a first step, and for palaeontology to better contribute to our knowledge of ecosystems, we must look to discovering the processes that drive these patterns, especially in the light that fossil data sets may be the only ones that can reasonably provide an answer. Johnson (2003) in reference to the fossil record states that "the best hope for understanding global diversity patterns is to collect information on local assemblages".

A final, all encompassing answer to the gradient is unlikely to be found after integrating our knowledge of fossil data sets into the analysis. This may be due in some respects to the limitations of fossils, but has more to do with the fact that the gradient likely has no single answer. When we look out on the world today, what we see is the product of many powerful, complex and highly variable forces acting both antagonistically and in concert. Through geologic time, these forces have likely changed in their importance to the gradient: biotic forces, excepting periods of mass extinctions, are likely to have steadily increased in importance trough time, while a factor such as temperature is more important towards the extremes of its variability. Rates of species extinction and origination take time to manifest, and may lag behind their causes, creating an even more complex situation. Nonetheless, the species diversity gradient we see in the world today has many simple answers to its origins, unfortunately no single one of them is adequate.

Table 6–1: Mechanisms of action for increased speciation rates and decreased extinction rates as a cause for the latitudinal gradient. Table after Mittelbach et al. (2007).

Increased Speciation

Mechanism	Type
1. Genetic drift in small populations accelerates evolutionary rates (Fedorov 1966)	biotic
2. Stronger biotic interactions lead to greater specialization (Dobzhansky 1950) and faster speciation (Fischer 1960; Schemske 2002)	biotic
3. Higher likelihood of parapatric (Moritz et al. 2000) and sympatric speciation (Gentry 1989) in the tropics	biotic
4. Larger area of the tropics provides more opportunities for isolation (Terborgh 1973; Rosenzweig 1995)	geographic
5. Reduced climatic variation results in higher speciation at lower latitudes (Haffer 1969; Dynesius and Jansson 2000)	environmental
6. Narrower physiological tolerances in tropical organisms reduce dispersal across unfavourable environments (Janzen 1967)	environmental
7. Higher temperatures result in increased evolutionary speed (Rohde 1992; Allen et al. 2002)	environmental

Decreased Extinction

Mechanism	Type
1. Larger tropical area leads to higher population numbers, larger species ranges, and lower chance of extinction (Terborgh 1973; Rosenzweig 1995)	geographic
2. Stability of tropical climates reduces the chance of extinction (Darwin 1862; Wallace 1878; Fischer 1960)	environmental

Table 6–2: Slopes and probabilities for the analysed time periods. Slopes are given for both the full Gentry (1988) dataset as well as a reduced dataset of only sites above 30°N. Significant slopes are highlighted in bold.

Age	Slope	*p*-value
Recent	**-2.938**	<< 0.001
Recent above 30°N	-0.39	0.065
Miocene	-1.07	0.254
Oligocene	1.51	0.299
Eocene	**0.303**	0.033
Paleocene	0.035	0.396
Maastrichtian	**-0.412**	0.007
Campanian	**-0.992**	0.002
Cenomanian	**0.065**	0.027
Albian	-0.084	0.283
Aptian	**0.488**	0.033

Figure 6–1: The accumulation of species richness over time under three general scenarios. A) The tropics and temperate regions do not differ in rate of accumulation but have different carrying capacities. B) The tropics have a higher diversification rate than temperate regions. C) The tropics have had more time for diversification due to their antiquity. Figure originally from Mittelbach et al. (2007)

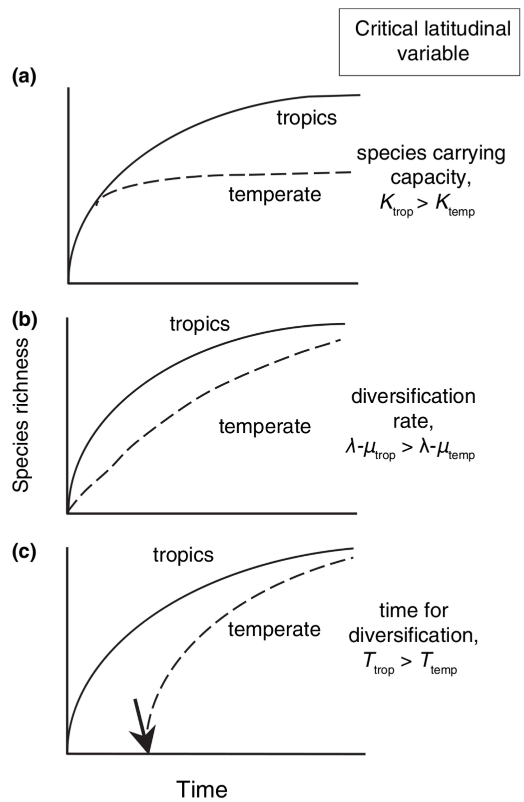


Figure 6–2: Gentry (1988) data

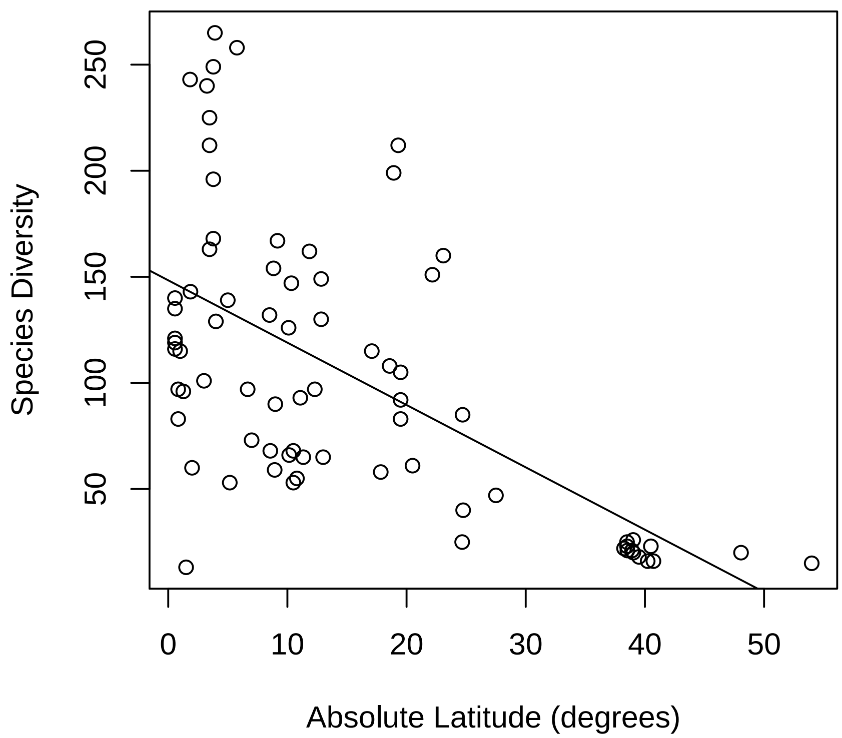


Figure 6–3: Gentry (1988) data above 30°N

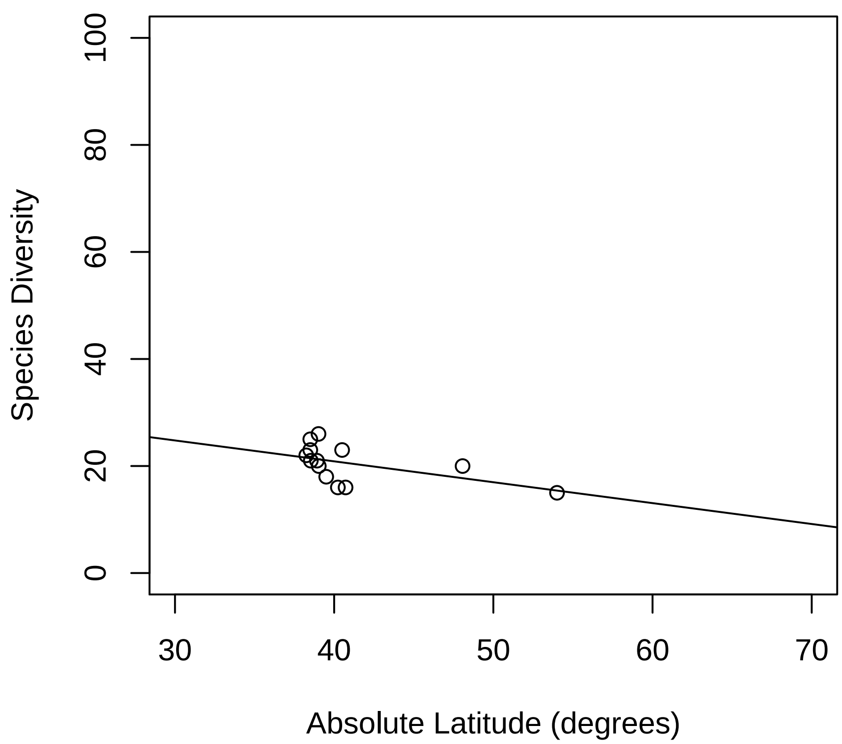


Figure 6–4: (A) Map of Miocene localities for which there is occurrence data and (B) latitude as compared to species diversity for the time period. Positions have been rotated to their palaeopositions using the rotations from Scotese (2002) and palaeogeographic maps are from the ESH-GIS package (Scotese 2001).

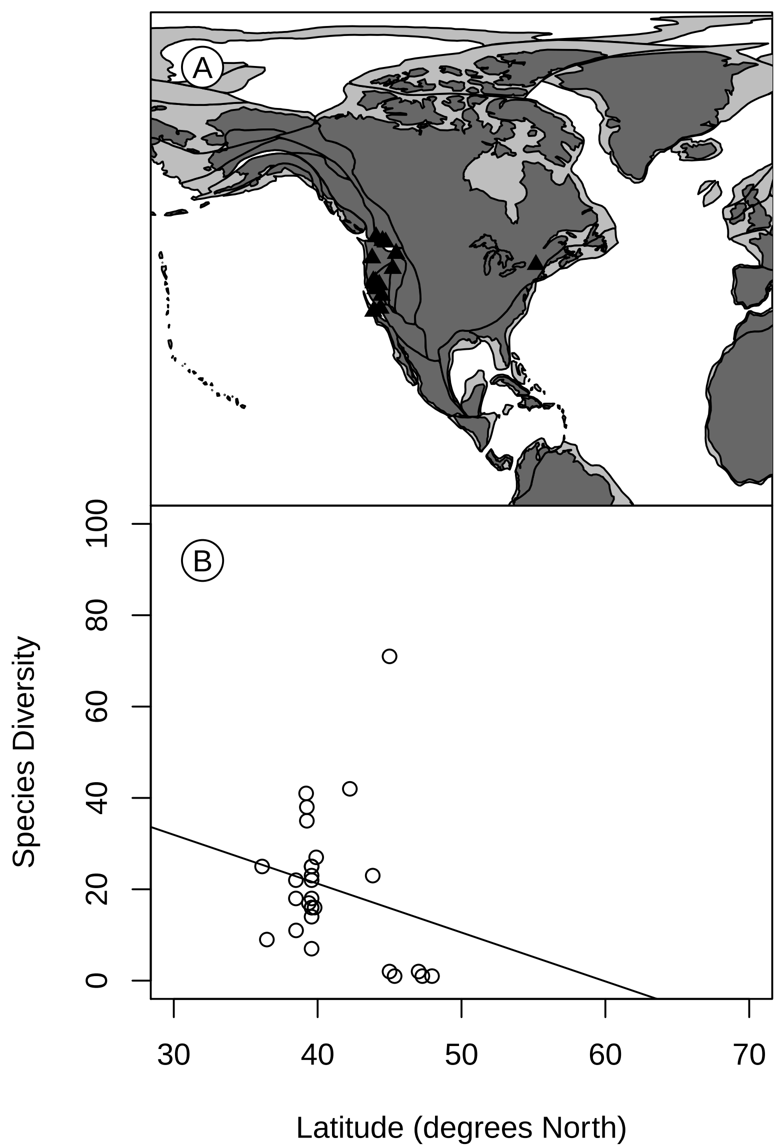


Figure 6–5: (A) Map of Oligocene localities for which there is occurrence data and (B) latitude as compared to species diversity for the time period. Positions have been rotated to their palaeopositions using the rotations from Scotese (2002) and palaeogeographic maps are from the ESH-GIS package (Scotese 2001).

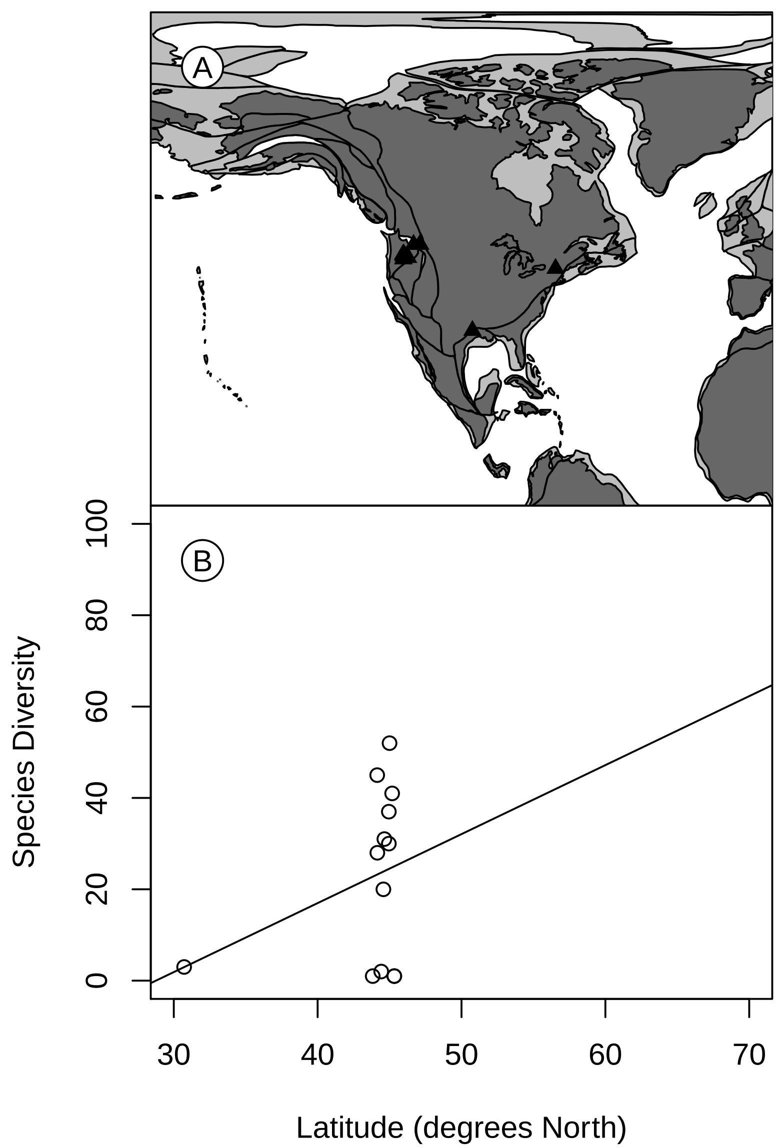


Figure 6–6: (A) Map of Eocene localities for which there is occurrence data and (B) latitude as compared to species diversity for the time period. Positions have been rotated to their palaeopositions using the rotations from Scotese (2002) and palaeogeographic maps are from the ESH-GIS package (Scotese 2001).

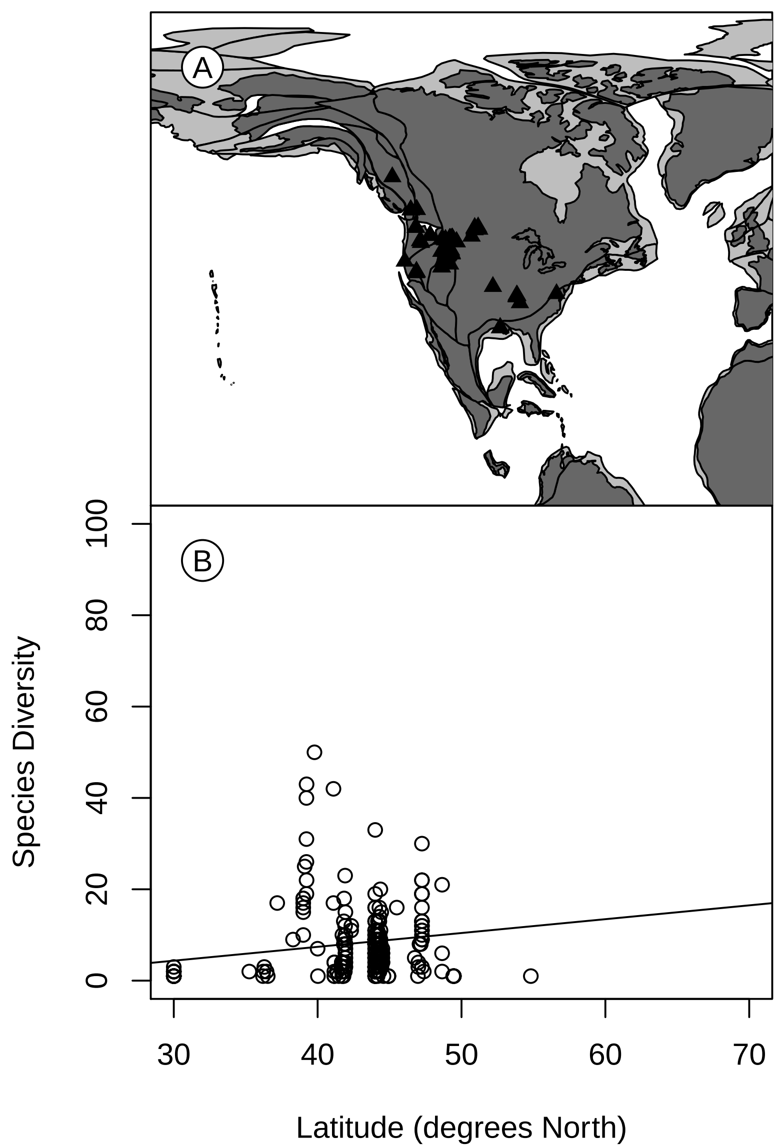


Figure 6–7: (A) Map of Paleocene localities for which there is occurrence data and (B) latitude as compared to species diversity for the time period. Positions have been rotated to their palaeopositions using the rotations from Scotese (2002) and palaeogeographic maps are from the ESH-GIS package (Scotese 2001).

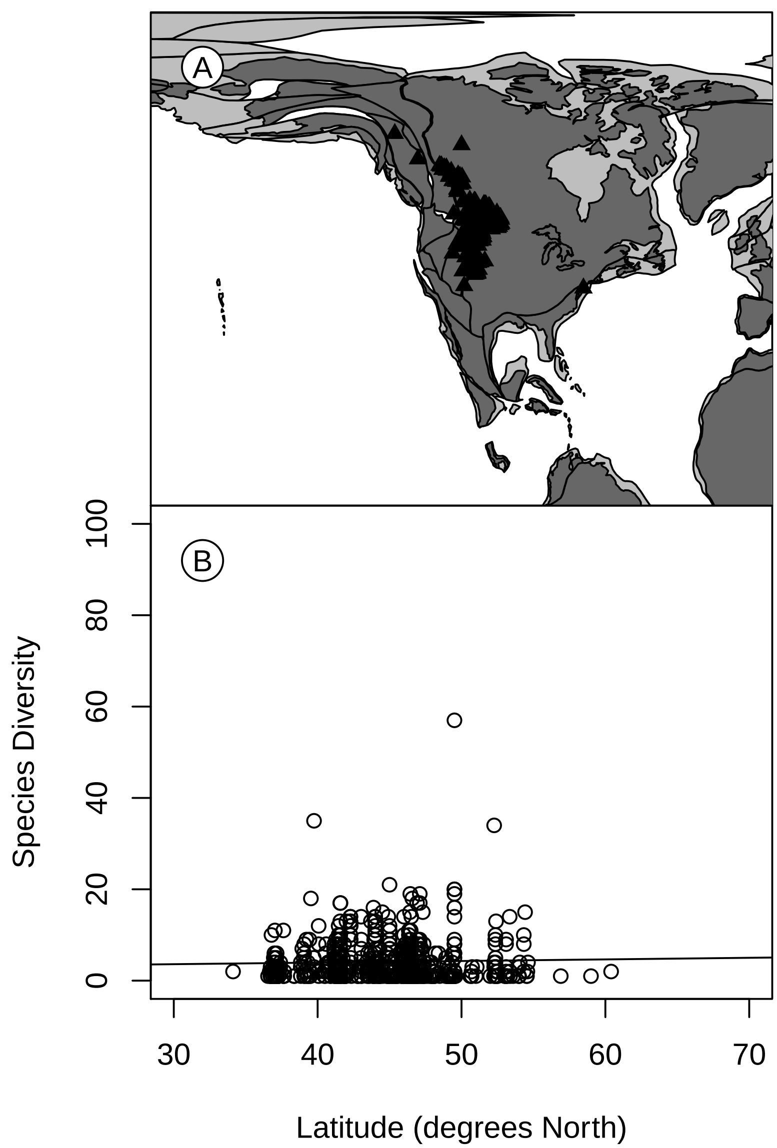


Figure 6–8: (A) Map of Maastrichtian localities for which there is occurrence data and (B) latitude as compared to species diversity for the time period. Positions have been rotated to their palaeopositions using the rotations from Scotese (2002) and palaeogeographic maps are from the ESH-GIS package (Scotese 2001).

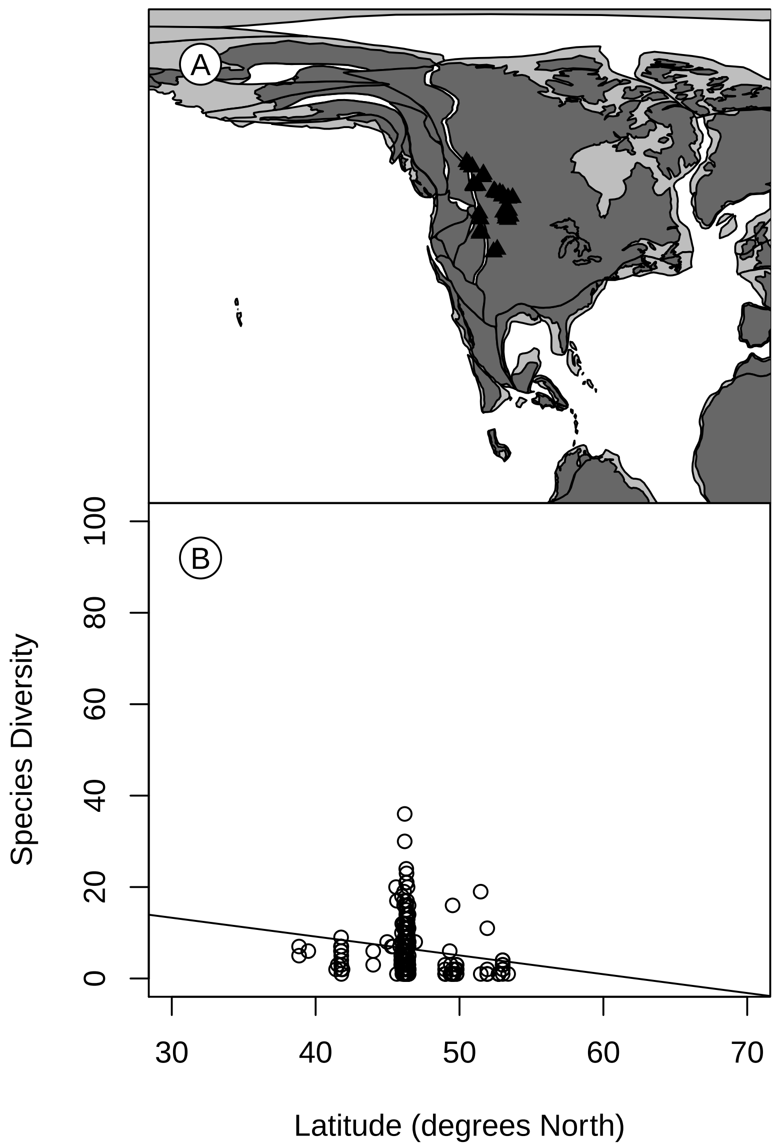


Figure 6–9: (A) Map of Campanian localities for which there is occurrence data and (B) latitude as compared to species diversity for the time period. Positions have been rotated to their palaeopositions using the rotations from Scotese (2002) and palaeogeographic maps are from the ESH-GIS package (Scotese 2001).

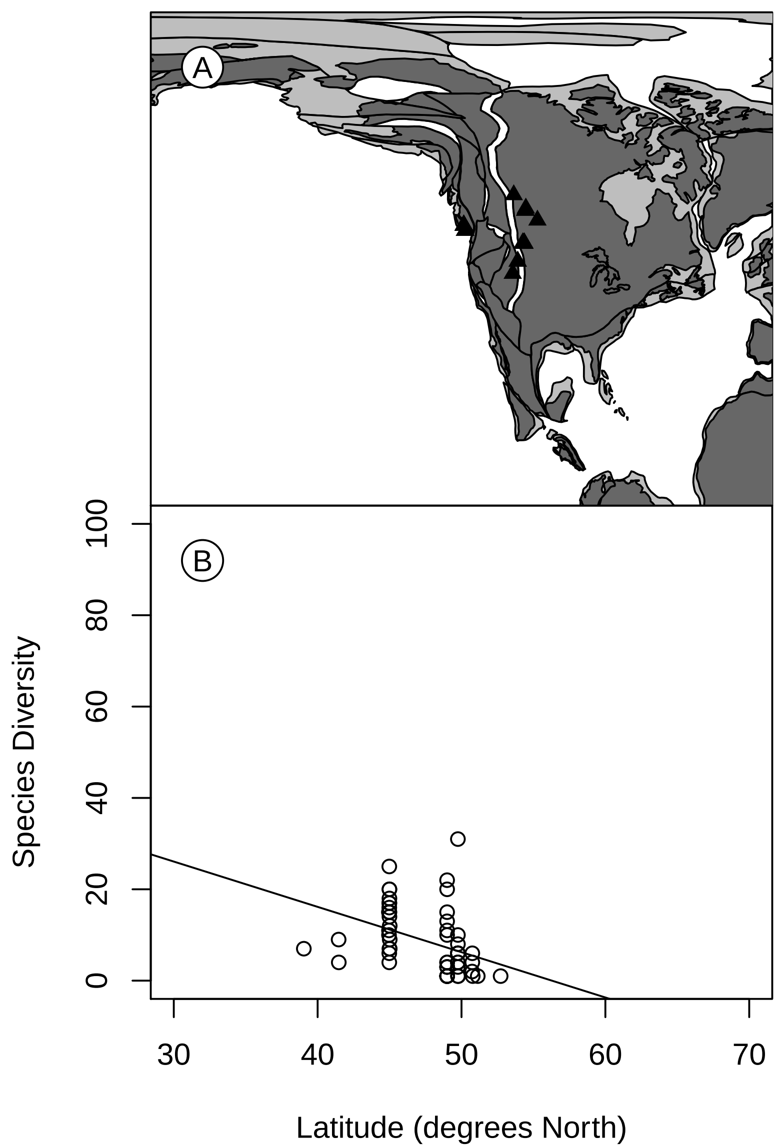


Figure 6–10: (A) Map of Cenomanian localities for which there is occurrence data and (B) latitude as compared to species diversity for the time period. Positions have been rotated to their palaeopositions using the rotations from Scotese (2002) and palaeogeographic maps are from the ESH-GIS package (Scotese 2001).

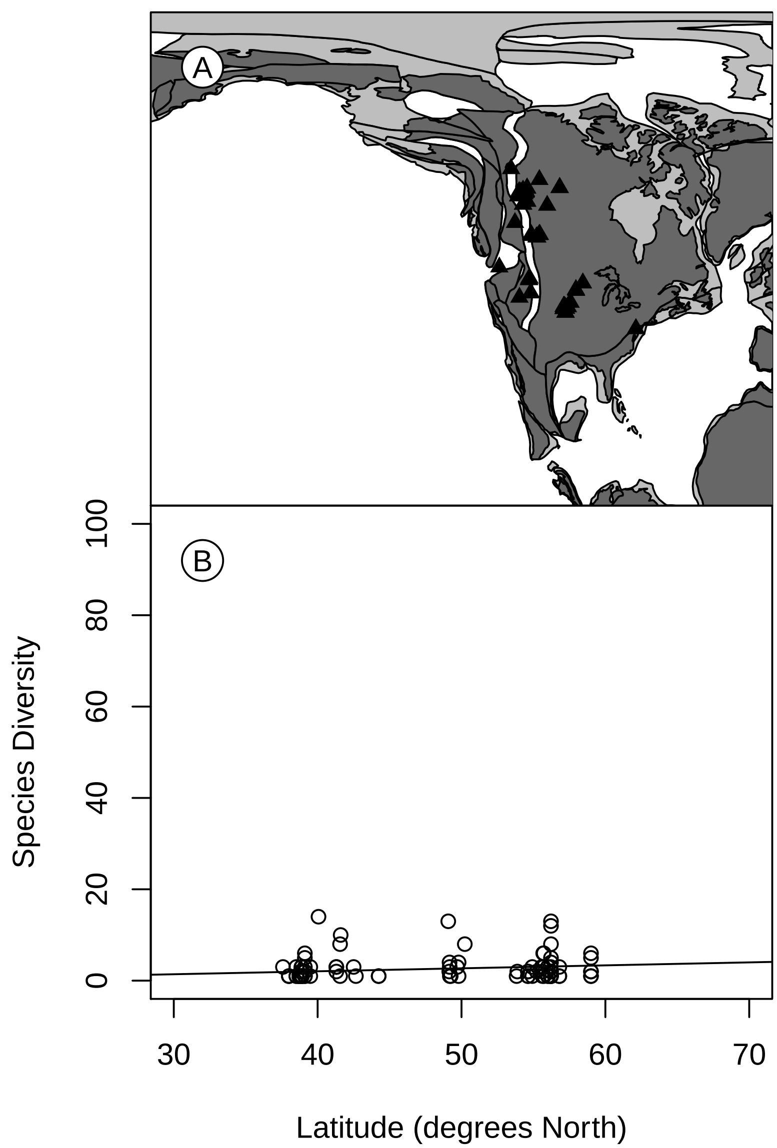


Figure 6–11: (A) Map of Albian localities for which there is occurrence data and (B) latitude as compared to species diversity for the time period. Positions have been rotated to their palaeopositions using the rotations from Scotese (2002) and palaeogeographic maps are from the ESH-GIS package (Scotese 2001).

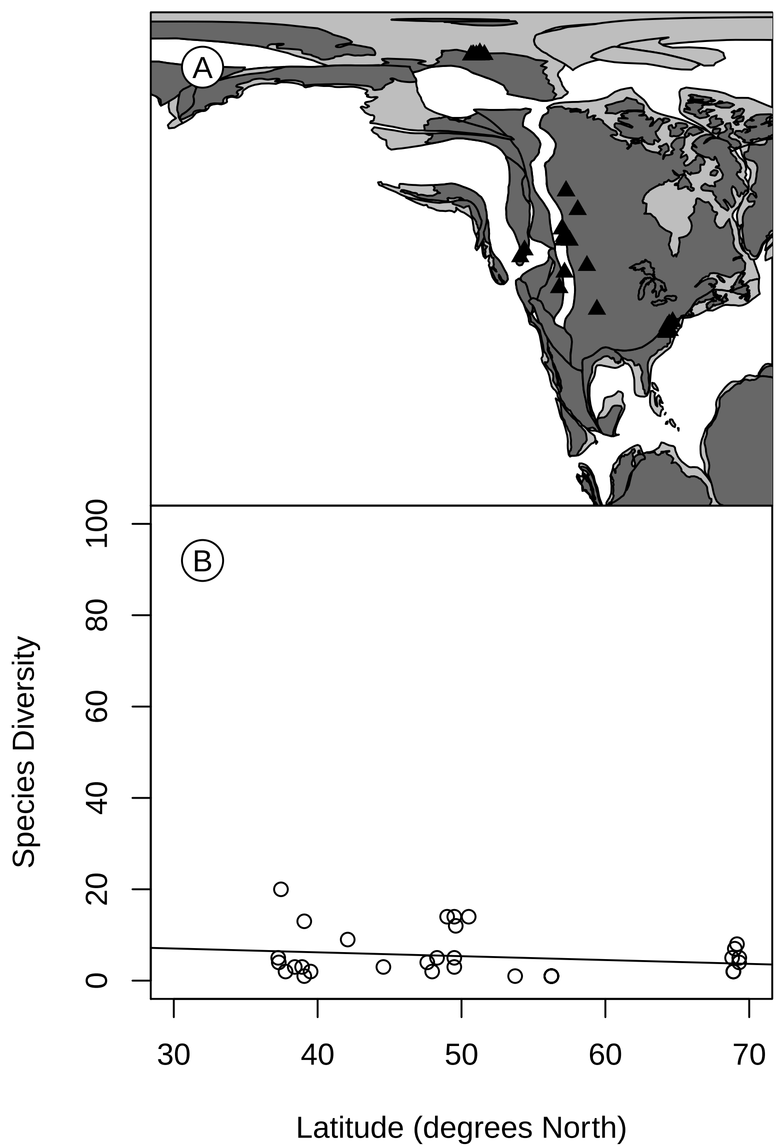


Figure 6–12: (A) Map of Aptian localities for which there is occurrence data and (B) latitude as compared to species diversity for the time period. Positions have been rotated to their palaeopositions using the rotations from Scotese (2002) and palaeogeographic maps are from the ESH-GIS package (Scotese 2001).

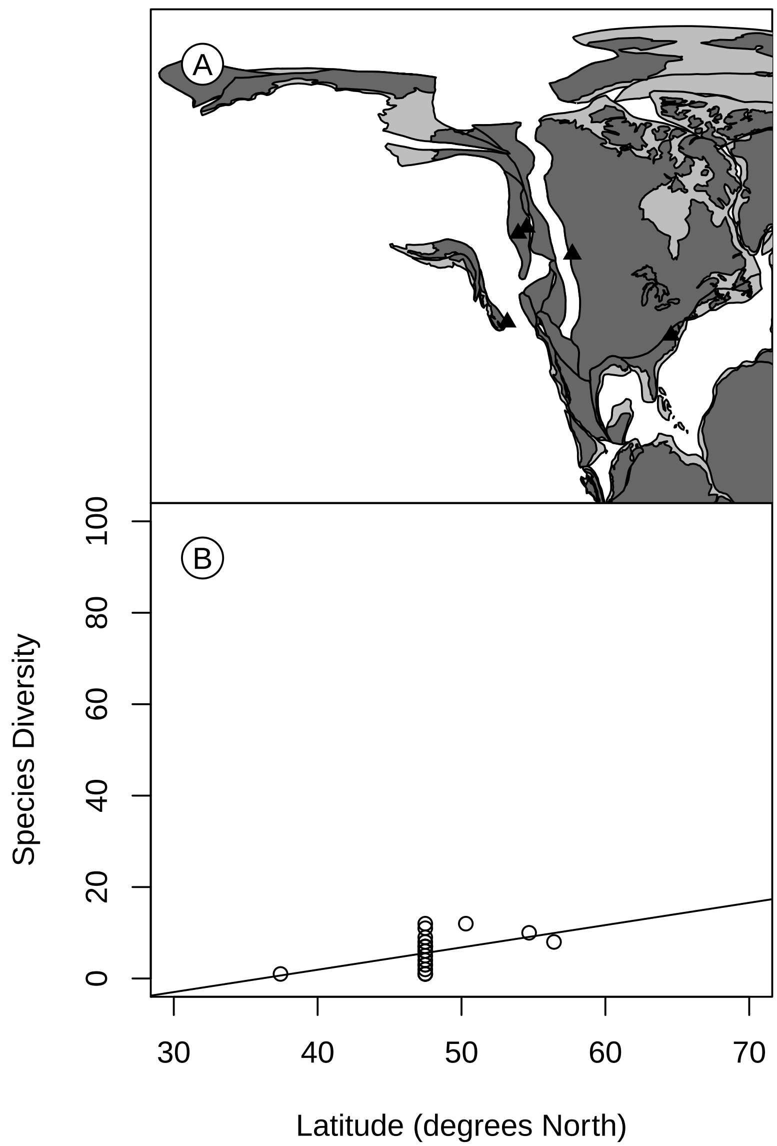
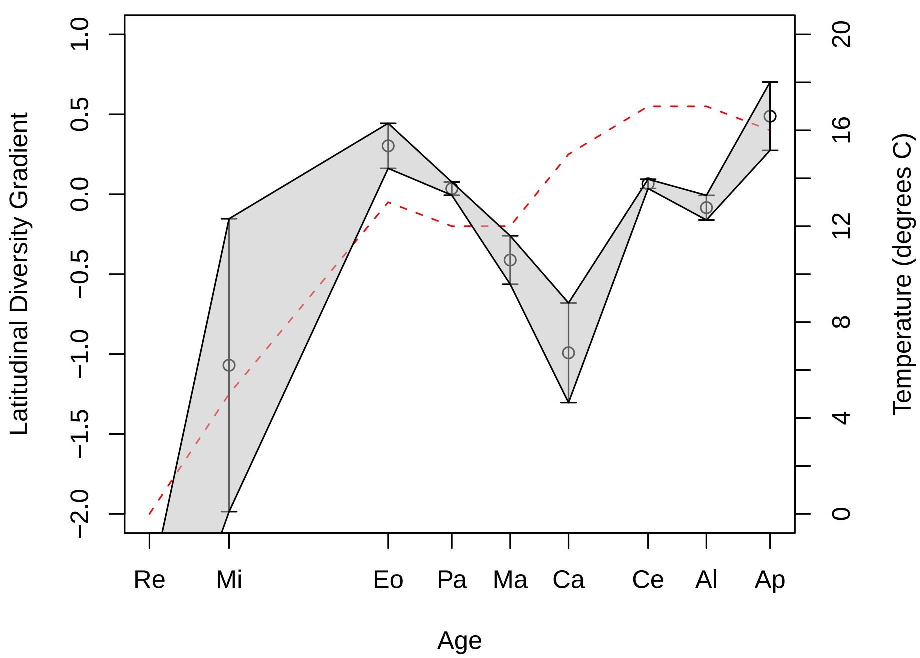


Figure 6–13: Slopes of the latitudinal diversity gradient through time. Points are the value of the slope for a given time period, with tails representing standard error. Grey region is error around slopes. Dashed line represents average global temperatures above present. More negative values represent a more pronounced/intense gradient. Abbreviations for ages are: Re = Recent; Mi = Miocene; Eo = Eocene; Pa = Paleocene; Ma = Maastrichtian; Ca = Campanian; Ce = Cenomanian; Al = Albian; Ap = Aptian.



CHAPTER 7
Advancing the Science of Palaeomacroecology

7.1 Introduction

Though palaeoecology has long grappled with questions of biodiversity over large scales, it has not always done so with an eye to its neoecological counterpart. Macroecological studies of the fossil record often use different terminology than those using modern ecosystem data, and so palaeoecological studies may fail to be noticed within the larger neoecological community.

Macroecology has often been criticized because of its difficulty of experimentally testing hypotheses. When trying to explain the distribution of species across continents, there is no simple way of experimentally manipulating variables in an ethical or practical manner (Blackburn and Gaston 1998). Some have argued that the side effects of modern civilizations will create an unintended pseudo-experiment in macroecology (Kerr et al. 2007), yet nonetheless neoecology is still lacking in this regard. Though palaeoecology has the same criticism, it does have the advantage of deep time. "Deep time" here refers to geologic time scales in the millions of years. Deep time is long enough for species to evolve, colonize, and expand over the landscape, and eventually go extinct (although humans have empirically shown that extinction events can occur over relatively short time spans). Rates of speciation and extinction are the primary drivers of species diversity, no matter their underlying causes (Rosenzweig 1995).

Compared to neoecologists, palaeoecologists have always had too much time on their hands. The fossil record encompasses enormous time ranges and our studies have been correspondingly phrased as such. To provide an idea of the differences in time scales, the longest running modern ecological experiment is the Park Grass Experiment, which has been running continuously since 1865 (Silvertown et al. 2006). This represents a mere fraction of the millions of years of fossil record available to palaeoecologists.

Species durations are typically on the range of millions of years, with many marine groups on the order of tens of millions of years (Stanley 1985). To try and capture the natural ebb and flow of species diversity, studies must look much further back than just the few decades that neoecology allows. The idea that macroecology is non-experimental is perhaps a red herring; even if we could conduct large scale experiments over long time scales, we would not solve the pressing questions in regards to species conservation and extinction that need answers now. An important consideration which is often overlooked is that nature has gone through dramatic changes that have been recorded in the fossil record. This puts palaeontology in a beneficial position with regards to the study of macroecology.

The idea of using the fossil record as a pseudo-experimental test case for macroecological studies is not novel (e.g. Jablonski et al. 2004, 2006; Alroy et al. 2008; Carrasco et al. 2009), and many have addressed problems and patterns of biodiversity over long time spans. The key comes in realizing the importance of palaeontological studies to macroecology, and that we are in fact studying macroecology using the fossil record. Palaeontologists need to engage their neoecological colleagues, by using terminology and testing concepts that are applicable to them. Leighton (2004) has discussed the importance of making palaeontology relevant, using the fossil record to try and answer the big questions of ecology like what factors control biodiversity. I wholeheartedly agree, and while he discusses refining our knowledge in order to study the finer scale patterns, I would add that the large scale patterns are also extremely important. A number of palaeoecologists have published studies with this goal in mind, and I follow here with a few examples of initial studies in areas of research with great potential.

7.2 The Species-area Curve

Raup (1976) was the first to give evidence that greater rock outcrop area led directly to higher species diversity. Essentially, what he demonstrated was the relationship between species and area (or the Species-Area Relationship, SAR), a large field within the macroecological literature (Arrhenius 1921; Gleason 1922; Preston 1960; Rosenzweig 1995). However, because time scales in ecological data sets are so short, modern ecologists rarely deal with a semi-temporal construct such as a rock unit. Recent papers on the SAR in the fossil record have begun to discuss the concept using neoecological terminology (Barnosky et al. 2005). Barnosky et al. (2005) have shown the existence of a species-area curve in the fossil record, and it must be accounted for when calculating the species richness of an area. It is important in our understanding of species distributions to understand how the shape of the species-area curve has changed (or remained constant) through the vast climactic and environmental changes of geologic time.

7.3 The Global Biodiversity Crisis

There is mounting evidence of a modern, sixth mass-extinction event in progress (Wilson 2003). Without palaeontology we would not have known about those five previous events, but even more important is how to use our knowledge of past events to better understand what is happening today. Neoecologists have only more recent records to rely on the extinction of organisms, but we need to establish proper baselines, both in general and for different groups. For example, we know that the average duration of a vascular plant species may be up to an order of magnitude longer than a mammal species. Average underlying rate differences gleaned from fossil evidence should help to inform us of what groups have shown the most susceptibility to human induced extinction when we compare the changes in extinction rates over time. Pragmatically, our money is limited in conservation efforts, and so it may be most proper to focus monies on certain groups that can be shown to have particularly accelerated rates of extinction due to anthropogenic effects.

7.4 The Latitudinal Gradient

The gradient of species diversity across latitudes is one of the few ubiquitous patterns in ecology and was discussed by scientists even before Darwin and Wallace (Mittelbach et al. 2007; Hawkins 2001). Though we can observe the present state of the gradient, our understanding of its full history is somewhat limited. In fact, there have been estimated to be over 25 different mechanisms that have created the gradient, with no easy way of disproving any of them (Gaston 2000). However, by examining the changes in the gradient through time in relation to changing climactic and geographic conditions, as well as evolutionary histories, we can begin to disprove at least some of these hypotheses. For example, it has been suggested that the latitudinal gradient was reduced in the past (Crame 2001), but the level of reduction is difficult to assess, and has not adequately been correlated to various environmental variables. In my study of the changing latitudinal gradient through time (Chapter 6), I found that there was a correlation (though non-significant) between global temperatures and the intensity of the gradient.

As well as evaluating changes in the gradient through time, there needs to be control as to how the gradient is quantified so that we can make it at least somewhat comparable to modern studies. For example, most modern studies use species richness, not percentages, and avoid binning by latitudinal band, so as to avoid species-area effects. However, many studies of the latitudinal gradient in the fossil record have used percentages or binning, making them difficult to integrate into our broader understanding. What is needed is a concerted effort to synthesize the already available data, no simple feat in itself, and begin analyzing it using techniques to those in neoecology.

7.5 Palaeomacroecology: A Summary

This thesis is intended to set a broad foundation for the development of Palaeomacroecology. To be able to conduct a successful study of large scale ecology in the fossil record, several variables come in to play. Palaeomacroecology often uses large, multivariate datasets that require efficient yet intricate statistical tools in order for researchers to understand the patterns within the dataset. By developing a large and growing set of statistical functions that anyone is free to use or modify, new research in palaeomacroecology can avoid replicating the work needed to create these tools. This allows new research to focus more on the actual data than the mechanics of the computer programs themselves. As well, in order to have large datasets to analyze we must continue to collect and describe new fossils, such as the flora I described from the Triassic rocks of Axel Heiberg. With these basic building blocks I have shown how we can use modern statistical techniques to aid in our understanding of fossil organisms such as dinosaurs. This research has an important impact on how we view dinosaurs ecologically: as motile, dynamic animals that colonized and dominated their environment. Finally, I used these same building blocks to show how the fossil record can benefit our understanding of Recent patterns of species diversity (i.e., the latitudinal diversity gradient). This also serves as an example of how we can use the vast temporal scales of the fossil record to study changes in the patterns of species diversity through time in relation to changing temperature, climate, and geographic positions.

There is a large area of potential research in macroecology in which palaeontology could play a role. Fully integrating the fossil record into macroecological analyses is not easy task, with many pitfalls to be aware of (see Chapter 1). However, we should not be deterred by the difficulty, as the potential benefits are worth it. As Markwick and Lupia (2002) noted, "[t]he fossil record is the only direct evidence about the biological evolution of life on Earth." To understand biological histories, we must embrace the fossil record, realizing that while it may not be able to answer every question, it is in many ways our best resource for understanding the history of biodiversity.

REFERENCES

Adrain, J. M., and S. R. Westrop. 2000. An empirical assessment of taxic paleobiology. Science **289**:110–112.

Allen, A. P., J. H. Brown, and J. F. Gillooly. 2002. Global biodiversity, biochemical kinetics, and the energetic-equivalence rule. Science **297**:1545–1548.

Alroy, J., M. Aberhan, D. J. Bottjer, M. Foote, F. T. Fursich, P. J. Harries, A. J. W. Hendy, S. M. Holland, L. C. Ivany, W. Kiessling, M. A. Kosnik, C. R. Marshall, A. J. McGowan, A. I. Miller, T. D. Olszewski, M. E. Patzkowsky, S. E. Peters, L. Villier, P. J. Wagner, N. Bonuso, P. S. Borkow, B. Brenneis, M. E. Clapham, L. M. Fall, C. A. Ferguson, V. L. Hanson, A. Z. Krug, K. M. Layou, E. H. Leckey, S. Nurnberg, C. M. Powers, J. A. Sessa, C. Simpson, A. Tomasovych, and C. C. Visaggi. 2008. Phanerozoic Trends in the Global Diversity of Marine Invertebrates. Science **321**:97–100.

Alroy, J., C. R. Marshall, R. K. Bambach, K. Bezusko, M. Foote, F. T. Fursich, T. A. Hansen, S. M. Holland, L. C. Ivany, D. Jablonski, D. K. Jacobs, D. C. Jones, M. A. Kosnik, S. Lidgard, S. Low, A. I. Miller, P. M. Novack-Gottshall, T. D. Olszewski, M. E. Patzkowsky, D. M. Raup, K. Roy, J. Sepkoski, J. J., M. G. Sommers, P. J. Wagner, and A. Webber. 2001. Effects of sampling standardization on estimates of Phanerozoic marine diversification. Proceedings of the National Academy of Sciences **98**:6261–6266.

Amiot, R., C. Lècuyer, E. Buffetaut, F. Fluteau, S. Legendre, and F. Martineau. 2004. Latitudinal temperature gradient during the Cretaceous Upper Campanian-Middle Maastrichtian: $\delta^{18}O$ record of continental vertebrates. Earth and Planetary Science Letters **226**:255–272.

Anderson, J., H. Anderson, S. Archangelsky, M. Bamford, S. Chandra, M. Dettmann, R. Hill, S. McLoughlin, and O. Rösler. 1999. Patterns of Gondwana plant colonisation and diversification. Journal of African Earth Sciences **28**:145–167.

Anderson, S., and L. F. Marcus. 1993. Effect of quadrat size on measurements of species density. Journal of Biogeography **20**:421–428.

Arnold, C. 1964. Mesozoic and Tertiary fern evolution and distribution. Memoirs of the Torrey Botanical Club **21**:58–66.

Arrhenius, O. 1921. Species and area. The Journal of Ecology **9**:95–99.

Ash, S., and J. Basinger. 1991. High latitude Upper Triassic Flora from the Heiberg Formation, Sverdrup Basin, Arctic Archipelago. Geological Survey of Canada Bulletin **4112**:101–131.

Barnosky, A. D., M. A. Carrasco, and E. B. Davis. 2005. The impact of the species-area relationship on estimates of paleodiversity. PLoS Biology **3**:1–6.

Barron, E. J. 1983. A warm, equable Cretaceous: The nature of the problem. Earth-Science Reviews **19**:305–338.

Barron, E. J., and W. M. Washington. 1982. Atmospheric circulation during warm geologic periods: Is the equator-to-pole surface-temperature gradient the controlling factor? Geology **10**:633–636.

Becker, R. A., J. M. Chambers, and A. R. Wilks. 1988. The New S Language. Wadsworth, Pacific Grove, CA.

Behrensmeyer, A., and R. Hook, 1992. Paleoenvironmental contexts and taphonomic modes. Pages 15–136 *in* A. Behrensmeyer, J. Damuth, W. DiMichele, R. Potts, H.-D. Sues, and S. Wing, editors. Terrestrial Ecosystems Through Time. Univ. of Chicago Press, Chicago.

Behrensmeyer, A. K., and R. E. Chapman, 1993. Models and simulations of time-averaging in terrestrial vertebrate accumulations. Pages 125–149 *in* Taphonomic Approaches to Time Resolution in Fossil Assemblages, volume 6. University of Tennessee, Knoxville, Short Courses in Paleontology.

Belasky, P. 1994. Biogeography of Permian corals and the determination of longitude in tectonic reconstructions of the Paleopacific region. Canadian Society of Petroleum Geologists, Memoir 17 pages 621–646.

Benton, M. J., M. A. Wills, and R. Hitchin. 2000. Quality of the fossil record through time. Nature **403**:534–537.

Bivand, R. S., E. J. Pebesma, and V. Gómez-Rubio. 2008. Applied Spatial Data Analysis with R. Springer, New York.

Blackburn, T. M., and K. J. Gaston. 1998. Some methodological issues in macroecology. The American Naturalist **151**:68–83.

Blackburn, T. M., and K. J. Gaston. 2002. Scale in macroecology. Global Ecology and Biogeography **11**:185–189.

Blackburn, T. M., J. H. Lawton, and J. N. Perry. 1992. A method of estimating the slope of upper bounds of plots of body size and abundance in natural animal assemblages. Oikos **65**:107–112.

Braun, W. J., and D. J. Murdoch. 2008. A First Course in Statistical Programming with R. Cambridge University Press, New York, NY, USA.

Bray, J. R., and J. T. Curtis. 1957. An ordination of the upland forest communities of southern Wisconsin. Ecological Monographs **27**:326–349.

Brown, J., 1988. Species diversity. Pages 57–89 *in* A. Myers and P. Giller, editors. Analytical Biogeography: An Integrated Approach to the Study of Animal and Plant Distributions. Chapman & Hall.

Brown, J. H. 1995. Macroecology. University of Chicago Press, Chicago.

Brown, J. H., and B. A. Maurer. 1989. Macroecology: the division of food and space among species on continents. Science **243**:1145–1150.

Burnham, K. P., and W. S. Overton. 1978. Estimation of the size of a closed population when capture probabilities vary among animals. Biometrika **65**:625–633.

Burnham, K. P., and W. S. Overton. 1979. Robust estimation of population size when capture probabilities vary among animals. Ecology **60**:927–936.

Burnham, R. J., 1993. Time resolution in terrestrial macrofloras: guidelines from modern accumulations. Pages 57–78 *in* S. Kidwell and A. Behrensmeyer, editors. Taphonomic Approaches to Time Resolution in Fossil Assemblages (Short Courses in Paleontology). Paleontological Society.

Burnham, R. J., S. L. Wing, and G. G. Parker. 1992. The reflection of deciduous forest communities in leaf litter: implications for autochthonous litter assemblages from the fossil record. Paleobiology **18**:30–49.

Carrano, M., 2000. Taxonomy and classification of non-avian Dinosauria: Online Systematics Archive 4: The Paleobiology Database. http://paleodb.org. URL `http://paleodb.org/cgi-bin/bridge.pl?user=Guest&action=displayPage&page=OSA_4_Dinosauri`.

Carrasco, M., B. Kraatz, E. Davis, and A. Barnosky., 2005. Miocene Mammal Mapping Project (MIOMAP). Technical report, University of California Museum of Paleontology. URL `http://www.ucmp.berkeley.edu/miomap/`.

Carrasco, M. A., A. D. Barnosky, and R. W. Graham. 2009. Quantifying the extent of North American mammal extinction relative to the pre-anthropogenic baseline. PLoS One **4**:e8331.

Cecca, F., B. Vrielynck, T. Lavoyer, and H. Gaget. 2005. Changes in the ammonite taxonomical diversity gradient during the late Jurassic and early Cretaceous. Journal of Biogeography **32**:535–547.

Chaney, R. 1959. Miocene floras of the Columbia Plateau. Part I. Composition and interpretation. Carnegie Institute of Washington Contributions to Paleontology **617**:1–134.

Chao, A. 1984. Non-parametric estimation of the number of classes in a population. Scandinavian Journal of Statistics **11**:265–270.

Chao, A. 1987. Estimating the population size for capture-recapture data with unequal catchability. Biometrics **43**:783–791.

Chao, A., and S.-M. Lee. 1992. Estimating the number of classes via sample coverage. Journal of the American Statistical Association **87**:210–217.

Chao, A., M.-C. Ma, and M. C. K. Yang. 1993. Stopping rules and estimation for recapture debugging with unequal failure rates. Biometrika **80**:193–201.

Chazdon, R., R. Colwell, J. Denslow, and M. Guariguata, 1998. Statistical methods for estimating species richness of woody regeneration in primary and secondary rain forests of NE Costa Rica. Pages 285–309 *in* F. Dallmeier and J. Comiskey, editors. Forest Biodiversity Research, Monitoring and Modeling: Conceptual Background and Old World Case Studies. Parthenon Publishing, Paris.

Chown, S. L., and K. J. Gaston. 2000. Areas, cradles and museums: the latitudinal gradient in species richness. Trends in Ecology & Evolution **15**:311–315.

Clarke, K. R. 1993. Non-parametric multivariate analyses of changes in community structure. Austral Ecology **18**:117–143.

Clemmensen, L. B., D. V. Kent, and F. A. J. Jr. 1998. A Late Triassic lake system in East Greenland: facies, depositional cycles and palaeoclimate. Palaeogeography, Palaeoclimatology, Palaeoecology **140**:135–159.

Coleman, B. D. 1981. On random placement and species-area relations. Mathematical Biosciences **54**:191–215.

Coleman, B. D., M. A. Mares, M. R. Willig, and Y.-H. Hsieh. 1982. Randomness, area, and species richness. Ecology **63**:1121–1133.

Colwell, R., 2009. EstimateS: Statistical estimation of species richness and shared species from samples. Version 8.2. URL `http://purl.oclc.org/estimates`.

Colwell, R. K., and J. A. Coddington. 1994. Estimating terrestrial biodiversity through extrapolation. Philosophical Transactions of the Royal Society B:

Biological Sciences **345**:101–118.

Condit, R., N. Pitman, J. Leigh, Egbert G., J. Chave, J. Terborgh, R. B. Foster, P. Nunez, S. Aguilar, R. Valencia, G. Villa, H. C. Muller-Landau, E. Losos, and S. P. Hubbell. 2002. Beta-diversity in tropical forest trees. Science **295**:666–669.

Connor, E. F., and E. D. McCoy. 1979. The statistics and biology of the species-area relationship. The American Naturalist **113**:791–833.

Crame, J. A. 2000. Evolution of taxonomic diversity gradients in the marine realm: evidence from the composition of Recent bivalve faunas. Paleobiology **26**:188–214.

Crame, J. A. 2001. Taxonomic diversity gradients through geological time. Diversity and Distributions **7**:175–189.

Crame, J. A. 2002. Evolution of taxonomic diversity gradients in the marine realm: a comparison of Late Jurassic and Recent bivalve faunas. Paleobiology **28**:184–207.

Crane, P. R., and S. Lidgard. 1989. Angiosperm diversification and paleolatitudinal gradients in Cretaceous floristic diversity. Science **246**:675–678.

Crawley, M. J., and J. E. Harral. 2001. Scale dependence in plant biodiversity. Science **291**:864–868.

Currie, D. J. 1991. Energy and large-scale patterns of animal- and plant-species richness. The American Naturalist **137**:27–49.

Currie, D. J., and V. Paquin. 1987. Large-scale biogeographical patterns of species richness of trees. Nature **329**:326–327.

Darwin, C. R. 1859. The Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life. John Murray, London, UK.

Darwin, C. R. 1862. The Voyage of the Beagle. John Murray, London, UK.

Delany, M. J., and M. J. R. Healy. 1966. Variation in the White-Toothed Shrews (*Crocidura* spp.) in the British Isles. Proceedings of the Royal Society of London. Series B, Biological Sciences **164**:63–74.

Delevoryas, T., and R. C. Hope. 1981. More evidence for conifer diversity in the Upper Triassic of North Carolina. American Journal of Botany **68**:1003–1007.

Dobzhansky, T. 1950. Evolution in the tropics. American Scientist **38**:209–221.

Donovan, S., and C. Paul, editors. 1998. The Adequacy of the Fossil Record. John Wiley and Sons, West Sussex, England.

Dynesius, M., and R. Jansson. 2000. Evolutionary consequences of changes in species' geographical distributions driven by Milankovitch climate oscillations. Proceedings of the National Academy of Sciences of the United States of America **97**:9115–9120.

Embry, A., 1982. The Upper Triassic-Lower Jurassic Heiberg deltaic complex of the Sverdrup Basin, Canadian Arctic Archipelago. Pages 189–217 *in* A. Embry and H. Balkwill, editors. Arctic geology and geophysics. Canadian Society of Petroleum Geologists Memoir.

Embry, A., 1983. Stratigraphic subdivision of the Heiberg Formation, eastern and central Sverdrup Basin, Arctic Islands. Pages 205–213 *in* Current Research, Part B. Geological Survey of Canada, Paper 83-1B.

Embry, A., 1991. Mesozoic history of the Arctic Islands. Pages 369–433 *in* H. Trettin, editor. Geology of the Innuitian Orogen and Arctic Platform of Canada and Greenland, volume 3. Geological Survey of Canada, Geology of Canada.

Ezard, T. H., and A. Purvis. 2009. paleoPhylo: free software to draw paleobiological phylogenies. Paleobiology **35**:460–464.

Fager, E. W. 1972. Diversity: a sampling study. The American Naturalist **106**:293–310.

Faith, D., P. Minchin, and L. Belbin. 1987. Compositional dissimilarity as a robust measure of ecological distance. Plant Ecology **69**:57–68.

Fedorov, A. A. 1966. The structure of the tropical rain forest and speciation in the humid tropics. Journal of Ecology **54**:1–11.

Fine, P. V. A., and R. H. Ree. 2006. Evidence for a time-integrated species-area effect on the latitudinal gradient in tree diversity. The American Naturalist **168**:796–804.

Fischer, A. G. 1960. Latitudinal variations in organic diversity. Evolution **14**:64–81.

Florin, R. 1958. On Jurassic Taxads and Conifers from North-Western Europe and Eastern Greenland. Acta Horta Bergiani **17**:257–402.

Foote, M. 1992. Rarefaction analysis of morphological and taxonomic diversity. Paleobiology **18**:1–16.

Fortier, Y., R. Blackadar, B. Glenister, H. Greiner, D. McLaren, N. McMillan, A. Norris, E. Roots, J. Souther, R. Thorsteinsson, and E. Tozer, editors.

1963. Geology of the North-Central Part of the Arctic Archipelago, Northwest Territories (Operation Franklin).

Frakes, L. 1979. Climates Throughout Geologic Time. Elsevier, Amsterdam.

Frakes, L. 1999. Estimating the global thermal state from Cretaceous sea surface and continental temperature data. Evolution of the Cretaceous Ocean–Climate System. Geol. Soc. Am. Spec. Pap. 332 **332**:49–57.

Gaston, K. J. 2000. Global patterns in biodiversity. Nature **405**:220–227.

Gentry, A. H. 1988. Changes in plant community diversity and floristic composition on environmental and geographical gradients. Annals of the Missouri Botanical Garden **75**:1–34.

Gentry, A. H., 1989. Speciation in tropical forests. Pages 113–134 *in* L. Holm-Nielsen, I. Nielsen, and H. Balslev, editors. Tropical Forests: Botanical Dynamics, Speciation and Diversity. Academic Press, San Diego, California.

Gleason, H. A. 1922. On the relation between species and area. Ecology **3**:158–162.

Glenister, B., 1963. Localities on southern Axel Heiberg Island. Pages 472–480 *in* Y. Fortier, R. Blackadar, B. Glenister, H. Greiner, D. McLaren, N. McMillan, A. Norris, E. Roots, J. Souther, R. Thorsteinsson, and E. Tozer, editors. Geology of the North-Central Part of the Arctic Archipelago, Northwest Territories (Operation Franklin). Geological Survey of Canada Memoir, 320.

Goslee, S., and D. Urban. 2007. The ecodist Package for Dissimilarity-based Analysis of Ecological Data. Journal of Statistical Software **22**:1–19. URL `http://cran.r-project.org/web/packages/ecodist/`.

Gotelli, N. J., and R. K. Colwell. 2001. Quantifying biodiversity: procedures and pitfalls in the measurement and comparison of species richness. Ecology Letters **4**:379–391.

Gower, J. C., and G. J. S. Ross. 1969. Minimum Spanning Trees and Single Linkage Cluster Analysis. Journal of the Royal Statistical Society. Series C (Applied Statistics) **18**:54–64.

Green, P. J. 2003. Diversities of gifts, but the same spirit. Journal of the Royal Statistical Society. Series D (The Statistician) **52**:423–438.

Green, R. H. 1980. Multivariate approaches in ecology: the assessment of ecologic similarity. Annual Review of Ecology and Systematics **11**:1–14.

Haffer, J. 1969. Speciation in Amazonian forest birds. Science **165**:131–147.

Hallam, A. 1985. A review of Mesozoic climates. Journal of the Geological Society **142**:433–445.

Hammer, Ø., D. A. Harper, and P. D. Ryan. 2001. PAST: paleontological statistics software package for education and data analysis. Palaeontologia Electronica **4**:9.

Harris, T. 1926. The Rhaetic flora of Scoresby Sound East Greenland. Meddelelser om Grønland **68**:43–148.

Harris, T. 1931. The fossil flora of Scoresby Sound East Greenland part 1: Cryptogams. Meddelelser om Gronland **85**:1–104.

Harris, T. 1932*a*. The fossil flora of Scoresby Sound East Greenland part 2: description of seed plants incertae sedis together with a discussion of certain cycadophyte cuticles. Meddelelser om Gronland **85**:1–112.

Harris, T. 1932*b*. The fossil flora of Scoresby Sound East Greenland part 3: Caytoniales and Bennettitales. Meddelelser om Gronland **85**:1–133.

Harris, T. 1935. The fossil flora of Scoresby Sound East Greenland part 4: Ginkoales, Coniferales, Lycopodiales and isolated fructifications. Meddelelser om Gronland **112**:1–176.

Harris, T. 1937. The fossil flora of Scoresby Sound East Greenland part 5: Stratigraphic relations of the plant beds. Meddelelser om Gronland **112**:1–112.

Harris, T. 1969. The Yorkshire Jurassic flora, Part III: Bennetitales. British Museum (Natural History), London.

Harrison, L. B., and H. C. E. Larsson. 2008. Estimating evolution of temporal sequence changes: a practical approach to inferring ancestral developmental sequences and sequence heterochrony. Systematic Biology **57**:378–387.

Harrison, S., S. J. Ross, and J. H. Lawton. 1992. Beta diversity on geographic gradients in Britain. The Journal of Animal Ecology **61**:151–158.

Haskell, J. 2001. The latitudinal gradient of diversity through the Holocene as recorded by fossil pollen in Europe. Evolutionary Ecology Research **3**:345–360.

Hathaway, R. J., and J. C. Bezdek. 1994. NERF *c*-means: Non-Euclidean relational fuzzy clustering. Pattern Recognition **27**:429–437.

Hawkins, B. A. 2001. Ecology's oldest pattern? Trends in Ecology & Evolution **16**:470.

Hawkins, B. A., R. Field, H. V. Cornell, D. J. Currie, J.-F. Guégan, D. M. Kaufman, J. T. Kerr, G. G. Mittelbach, T. Oberdorff, E. M. O'Brien, E. E. Porter, and J. R. G. Turner. 2003. Energy, water and broad-scale geographic patterns of species richness. Ecology **84**:3105–3117.

Heck, J., Kenneth L., G. van Belle, and D. Simberloff. 1975. Explicit calculation of the rarefaction diversity measurement and the determination of sufficient sample size. Ecology **56**:1459–1461.

Hewitson, W. 1962. Comparative morphology of the Osmundaceae. Annals of the Missouri Botanical Garden **49**:57–93.

Hillebrand, H. 2004. On the generality of the latitudinal diversity gradient. The American Naturalist **163**:192–211.

Hortal, J., P. Borges, and C. Gaspar. 2006. Evaluating the performance of species richness estimators: sensitivity to sample grain size. Journal of Animal Ecology **75**:274–287.

Hubálek, Z. 1982. Coefficients of association and similarity, based on binary (presence-absence) data: an evaluation. Biological Reviews **57**:669–689.

Huber, B. T., R. D. Norris, and K. G. MacLeod. 2002. Deep-sea paleotemperature record of extreme warmth during the Cretaceous. Geology **30**:123–126.

Hubert, L., and P. Arabie. 1985. Comparing partitions. Journal of Classification **2**:193–218.

Hunt, G., 2008. paleoTS: Modeling evolution in paleontological time-series.

Ihaka, R., and R. Gentleman. 1996. R: a language for data analysis and graphics. Journal of Computational and Graphical Statistics **5**:299–314.

Jablonski, D., and D. M. Raup. 1995. Selectivity of End-Cretaceous marine bivalve extinctions. Science **268**:389–391.

Jablonski, D., K. Roy, and J. Valentine, 2004. Evolutionary macroecology and the fossil record. Pages 368–390 *in* T. M. Blackburn and K. J. Gaston, editors. Macroecology: Concepts and Consequences: The 43rd Annual Symposium of the British Ecological Society Held at the University of Birmingham. Cambridge University Press.

Jablonski, D., K. Roy, and J. W. Valentine. 2006. Out of the Tropics: evolutionary dynamics of the latitudinal diversity gradient. Science **314**:102–106.

Jablonski, D., and J. J. Sepkoski Jr. 1996. Paleobiology, community ecology, and scales of ecological pattern. Ecology **77**:1367–1378.

Jaccard, P. 1901. Étude comparative de la distribution florale dans une portion des Alpes et des Jura. Bulletin del la Société Vaudoise des Sciences Naturelles **37**:547–579.

Jackson, J., and K. Johnson. 2001. Measuring past biodiversity. Science **293**:2401–2404.

Jackson, J., J. Todd, H. Fortunato, and P. Jung. 1999. Diversity and assemblages of Neogene Caribbean Mollusca of lower Central America. Bulletins of American Paleontology **356**:193–230.

Jain, A. K., M. N. Murty, and P. J. Flynn. 1999. Data clustering: a review. ACM Computing Surveys **31**:264–323.

James, F. C., and C. E. McCulloch. 1990. Multivariate analysis in ecology and systematics: panacea or Pandora's Box? Annual Review of Ecology and Systematics **21**:129–166.

Janzen, D. H. 1967. Why mountain passes are higher in the tropics. The American Naturalist **101**:233–249.

Johnson, K. G. 2003. New data for old questions. Paleobiology **29**:19–21.

Kerp, J., and H. Haubold. 1988. Aspects of Permian palaeobotany and palynology. VIII: On the reclassification of the West- and Central European species of the form-genus *Callipteris* Brongniart 1849. Review of Palaeobotany and Palynology **54**:135–150.

Kerr, J. T., H. M. Kharouba, and D. J. Currie. 2007. The macroecological contribution to global change solutions. Science **316**:1581–1584.

Kidwell, S. M. 2001. Preservation of species abundance in marine death assemblages. Science **294**:1091–1094.

Kutzbach, J. E., and A. Ziegler. 1993. Simulation of Late Permian climate and biomes with an atmosphere-ocean model: comparisons with observations. Philosophical Transactions: Biological Sciences **341**:327–340.

Lance, G. N., and W. T. Williams. 1967. A general theory of classificatory sorting strategies: II. Clustering systems. The Computer Journal **10**:271–277.

Lee, S.-M., and A. Chao. 1994. Estimating population size via sample coverage for closed capture-recapture models. Biometrics **50**:88–97.

Lehman, T., 2001. Late Cretaceous dinosaur provinciality. Pages 310–328. *in* D. Tanke and K. Carpenter, editors. Mesozoic vertebrate life: New research inspired by the paleontology of Philip J. Currie. University of Indiana Press, Bloomington.

Lehman, T. M. 1987. Late Maastrichtian paleoenvironments and dinosaur biogeography in the western interior of North America. Palaeogeography, Palaeoclimatology, Palaeoecology **60**:189–217.

Leighton, L. 2005. The latitudinal diversity gradient through deep time: testing the “Age of the Tropics” hypothesis using carboniferous productidine brachiopods. Evolutionary Ecology **19**:563–581.

Leighton, L. R. 2004. Are we asking the right question? Palaios **19**:313–315.

Leisch, F., and A. J. Rossini. 2003. Reproducible statistical research. Chance **16**:46–50.

Lewin, R. 1989. Biologists disagree over bold signature of nature: adaptation to the different annual climatic fluctuations at low and high latitudes is suggested as an explanation for a famous, unexplained pattern of nature. Science **244**:527–528.

Lyons, S. K., and M. R. Willig. 1999. A hemispheric assessment of scale dependence in latitudinal gradients of species richness. Ecology **80**:2483–2491.

MacFadden, B. J. 2005. Vertebrate ecology and evolution in 'Deep Time'. Trends in Ecology & Evolution **20**:355–355.

Maddison, W. P., and D. Maddison, 2009. Mesquite: a modular system for evolutionary analysis. Version 2.72. URL `http://mesquiteproject.org`.

Magallón, S., and M. J. Sanderson. 2001. Absolute diversification rates in angiosperm clades. Evolution **55**:1762–1780.

Magurran, A. E. 2004. Measuring Biological Diversity. Blackwell, Oxford.

Markwick, P. J., and R. Lupia. 2002. Palaeontological databases for palaeobiogeography, palaeoecology and biodiversity: a question of scale. Geological Society, London, Special Publications **194**:169–178.

Marshall, L. G., R. F. Butler, R. E. Drake, G. H. Curtis, and R. H. Tedford. 1979. Calibration of the Great American Interchange. Science **204**:272–279.

Martin, A. P., and S. R. Palumbi. 1993. Body size, metabolic rate, generation time, and the molecular clock. Proceedings of the National Academy of Sciences of the United States of America **90**:4087–4091.

Martin, R. 1999. Taphonomy: A process approach. Cambridge University Press, Cambridge, England.

McGill, B., and C. Collins. 2003. A unified theory for macroecology based on spatial patterns of abundance. Evolutionary Ecology Research **5**:469–492.

McLaren, D., 1963. Mount Nicolay, Cornwall Island. Pages 529–532 *in* Y. Fortier, R. Blackadar, B. Glenister, H. Greiner, D. McLaren, N. McMillan, A. Norris, E. Roots, J. Souther, R. Thorsteinsson, and E. Tozer, editors. Geology of the North-Central Part of the Arctic Archipelago, Northwest Territories (Operation Franklin). Geological Survey of Canada Memoir, 320.

Minchin, P. R. 1987. An evaluation of the relative robustness of techniques for ecological ordination. Plant Ecology **69**:89–107.

Mittelbach, G. G., D. W. Schemske, H. V. Cornell, A. P. Allen, J. M. Brown, M. B. Bush, S. P. Harrison, A. H. Hurlbert, N. Knowlton, H. A. Lessios, C. M. McCain, A. R. McCune, L. A. McDade, M. A. McPeek, T. J. Near, T. D. Price, R. E. Ricklefs, K. Roy, D. F. Sax, D. Schluter, J. M. Sobel, and M. Turelli.

2007. Evolution and the latitudinal diversity gradient: speciation, extinction and biogeography. Ecology Letters **10**:315–331.

Moritz, C., J. Patton, C. Schneider, and T. B. Smith. 2000. Diversification of rainforest faunas: an integrated molecular approach. Annual Review of Ecology and Systematics **31**:533–563.

Nathorst, A. 1908. Palaobotanische Mitteilungen, 7, uber *Palissy*, *Stachyotaxus*, und *Paleotaxus*. Kungliga Svenska Vetenskapsakademiens Handligar **43**:1–20.

Norford, B., W. Brideaux, T. Chamney, M. Copeland, H. Frebold, J. Hopkins, WV.S., J. Jeletzky, B. Johnson, D. McGregor, A. Norris, A. Pedder, E. Tozer, and T. Uyeno. 1973. Biostratigraphic determinations of fossils from the subsurface of the Yukon Territory and the Districts of Franklin, Keewatin and MacKenzie. Geological Survey of Canada, Paper 72-38 .

Novotny, V., S. E. Miller, J. Hulcr, R. A. I. Drew, Y. Basset, M. Janda, G. P. Setliff, K. Darrow, A. J. A. Stewart, J. Auga, B. Isua, K. Molem, M. Manumbor, E. Tamtiai, M. Mogia, and G. D. Weiblen. 2007. Low beta diversity of herbivorous insects in tropical forests. Nature **448**:692–695.

Oksanen, J., F. G. Blanchet, R. Kindt, P. Legendre, R. G. O'Hara, G. L. Simpson, P. Solymos, M. H. H. Stevens, and H. Wagner, 2010. vegan: Community Ecology Package. URL `http://CRAN.R-project.org/package=vegan`.

Osborn, S., 1855. The journal of Her Majesty's Sledge "John Barrow", (Lieutenant Sherard Osborn commanding Her Majesty's Steam Vessel Pioneer in charge). Whilst forming a part of the Southwestern Division of Sledges under the Orders of Commander G.H. Richards. Pages 88–261 *in* Further Papers Relative to the

Search for Sir John Franklin and the Crews of H.M.S Erebus and Terror. British Parliamentary Papers, Eyre and Spottiswoode. Queens Printer, London.

Palmer, M. W., and P. S. White. 1994. Scale dependence and the species-area relationship. The American Naturalist **144**:717–740.

Paradis, E., J. Claude, and K. Strimmer. 2004. APE: analyses of phylogenetics and evolution in R language. Bioinformatics **20**:289–290.

Paul, C., 1998. Adequacy, completeness and the fossil record. Pages 1–22 *in* S. Donovan and C. Paul, editors. The Adequacy of the Fossil Record. John Wiley and Sons, West Sussex, England.

Pebesma, E. J., and R. S. Bivand. 2005. Classes and methods for spatial data in R. R News **5**:9–13. URL `http://CRAN.R-project.org/doc/Rnews/`.

Pianka, E. R. 1966. Latitudinal gradients in species diversity: a review of concepts. The American Naturalist **100**:33–46.

Pitman, N. C. A., J. Terborgh, M. R. Silman, and P. Nuñez V. 1999. Tree species distributions in an Upper Amazonian forest. Ecology **80**:2651–2661.

Platt, J. R. 1964. Strong inference. Science **146**:347–353.

Potts, R., and A. Deino. 1995. Mid-Pleistocene change in large mammal faunas of East Africa. Quaternary Research **43**:106–113.

Preston, F. W. 1948. The commonness, and rarity, of species. Ecology **29**:254–283.

Preston, F. W. 1960. Time and space and the variation of species. Ecology **41**:611–627.

Qian, H. 1998. Large-scale biogeographic patterns of vascular plant richness in North America: an analysis at the generic level. Journal of Biogeography

25:829–836.

R Development Core Team, 2010. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.r-project.org.

Rand, W. M. 1971. Objective criteria for the evaluation of clustering methods. Journal of the American Statistical Association **66**:846–850.

Raup, D. M. 1976. Species diversity in the Phanerozoic: an interpretation. Paleobiology **2**:289–297.

Raup, D. M., and R. E. Crick. 1979. Measurement of faunal similarity in paleontology. Journal of Paleontology **53**:1213–1227.

Raup, D. M., and D. Jablonski. 1993. Geography of End-Cretaceous marine bivalve extinctions. Science **260**:971–973.

Rees, P. 2002. Land-plant diversity and the end-Permian mass extinction. Geology **30**:827–830.

Rees, P. M., C. R. Noto, J. M. Parrish, and J. T. Parrish. 2004. Late Jurassic climates, vegetation, and dinosaur distributions. The Journal of Geology **112**:643–653.

Ripley, B., and M. Lapsley, 2009. RODBC: ODBC Database Access. URL http://CRAN.R-project.org/package=RODBC.

Rohde, K. 1992. Latitudinal gradients in species diversity: the search for the primary cause. Oikos **65**:514–527.

Rosenzweig, M. L. 1995. Species Diversity in Space and Time. Cambridge University Press, Cambridge, UK.

Russell, D. A. 1967. A census of dinosaur specimens collected in western Canada. National Museum of Canada Natural History Paper **36**:1–13.

Russell, D. A. 1977. A vanished world – the dinosaurs of western Canada. National Museum of Canada Natural History Series **4**:1–142.

Sampson, S. D., M. A. Loewen, E. M. Roberts, J. A. Smith, L. E. Zanno, and T. A. Gates. 2004. Provincialism in Late Cretaceous terrestrial faunas: new evidence from the Campanian Kaiparowits Formation of Utah. Journal of Vertebrate Paleontology **24**:108A.

Sanders, H. L. 1968. Marine benthic diversity: a comparative study. The American Naturalist **102**:243–282.

Scheihing, M. H., and H. W. Pfefferkorn. 1984. The taphonomy of land plants in the orinoco delta: A model for the incorporation of plant parts in clastic sediments of late carboniferous age of euramerica. Review of Palaeobotany and Palynology **41**:205–240.

Schemske, D., 2002. Tropical diversity: patterns and processes. Pages 163–173 *in* R. Chazdon and T. Whitmore, editors. Ecological and Evolutionary Perspectives on the Origins of Tropical Diversity: Key Papers and Commentaries. University of Chicago Press, Chicago, IL.

Schnute, J. T., N. Boers, and R. Haigh, 2004. PBS Mapping 2: user's guide. Technical Report 2549, Canadian Technical Report of Fisheries and Aquatic Sciences.

Scotese, C. R., 2001. Earth System History Geographic Information System v. 02b. PALEOMAP Project, Arlington, Texas.

Scotese, C. R., 2002. PointTracker for Windows. Arlington, TX.

Seward, A. 1910. Fossil plants. Cambridge University Press, Cambridge, U.K.

Shen, S.-Z., and G. R. Shi. 2004. Capitanian (Late Guadalupian, Permian) global brachiopod palaeobiogeography and latitudinal diversity pattern. Palaeogeography, Palaeoclimatology, Palaeoecology **208**:235–262.

Sheppard, C. R. C. 1998. Biodiversity patterns in Indian Ocean corals, and effects of taxonomic error in data. Biodiversity and Conservation **7**:847–868.

Shi, G. R. 1993. Multivariate data analysis in palaeoecology and palaeobiogeography – a review. Palaeogeography, Palaeoclimatology, Palaeoecology **105**:199–234.

Siemann, E., D. Tilman, and J. Haarstad. 1996. Insect species diversity, abundance and body size relationships. Nature **380**:704–706.

Silvertown, J. 1985. History of a latitudinal diversity dradient: woody plants in Europe 13,000-1000 Years B.P. Journal of Biogeography **12**:519–525.

Silvertown, J., P. Poulton, E. Johnston, G. Edwards, M. Heard, and P. Biss. 2006. The Park Grass Experiment 1856-2006: its contribution to ecology. Journal of Ecology **94**:801–814.

Simpson, G. 1960. Notes on the measurement of faunal resemblance. American Journal of Science **258**:300–311.

Sloan, R. E. 1969. Cretaceous and Paleocene terrestrial communities of western North America. Proceedings of the North American Paleontological Convention **Part E**:427–453.

Sloan, R. E., 1976. The ecology of dinosaur extinction. Pages 134–154 *in* C. S. Churcher, editor. Athlon, Essays on Palaeontology in Honour of L. S. Russell.

Royal Ontario Museum Life Sciences Miscellaneous Publication.

Sluijs, A., S. Schouten, M. Pagani, M. Woltering, H. Brinkhuis, J. S. S. Damsté, G. R. Dickens, M. Huber, G.-J. Reichart, R. Stein, J. Matthiessen, L. J. Lourens, N. Pedentchouk, J. Backman, K. Moran, and the Expedition 302 Scientists. 2006. Subtropical Arctic Ocean temperatures during the Palaeocene/Eocene thermal maximum. Nature **441**:610–613.

Smith, A. B. 2001. Large-scale heterogeneity of the fossil record: implications for Phanerozoic biodiversity studies. Philosophical Transactions of the Royal Society B: Biological Sciences **356**:351–367.

Smith, E. P., and G. van Belle. 1984. Nonparametric estimation of species richness. Biometrics **40**:119–129.

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

Sneath, P. H. A. 1957. The application of computers to taxonomy. Journal of General Microbiology **17**:201–226.

Souther, J., 1963. Geologic traverse across Axel Heiberg Island from Buchanan Lake to Strand Fiord. Pages 529–532 *in* Y. Fortier, R. Blackadar, B. Glenister, H. Greiner, D. McLaren, N. McMillan, A. Norris, E. Roots, J. Souther, R. Thorsteinsson, and E. Tozer, editors. Geology of the North-Central Part of the Arctic Archipelago, Northwest Territories (Operation Franklin). Geological Survey of Canada Memoir, 320.

Spicer, R. 1981. The sorting and deposition of allochthonous plant material in a modern environment at Silwood Lake, Silwood Park, Berkshire, England. Geological Survey, Professional Paper **1143**:1–77.

Spicer, R., and J. Wolfe. 1987. Plant taphonomy of late Holocene deposits in Trinity (Clair Engle) Lake, northern California. Paleobiology **13**:227–245.

Spicer, R. A., and J. T. Parrish. 1990. Late Cretaceous-early Tertiary palaeoclimates of northern high latitudes: a quantitative view. Journal of the Geological Society **147**:329–341.

Stabler, B., 2006. shapefiles: Read and Write ESRI Shapefiles. URL `http://cran.r-project.org/web/packages/shapefiles/`.

Stallman, R., 1999. The GNU Operating System and the Free Software Movement. Pages 53–70 *in* C. DiBona, S. Ockman, and M. Stone, editors. Open Sources: Voices from the Open Source Revolution. O'Reilly, Sebastopol, California.

Stanley, S. M. 1985. Rates of evolution. Paleobiology **11**:13–26.

Stebbins, G. L. 1974. Flowering Plants: Evolution above the Species Level. The Belknap Press of Harvard University Press, Cambridge, MA.

Stehli, F. G., R. G. Douglas, and N. D. Newell. 1969. Generation and maintenance of gradients in taxonomic diversity. Science **164**:947–949.

Stewart, W., and G. Rothwell. 1993. Paleobotany and the Evolution of Plants. Cambridge University Press, Cambridge, U.K.

Suneby, L., and L. Hills. 1988. Palynological zonation of the Heiberg Formation (Triassic-Jurassic) Eastern Sverdrup Basin, Arctic Canada. Bulletin of Canadian Petroleum Geology **36**:347–361.

Terborgh, J. 1973. On the notion of favorableness in plant ecology. The American Naturalist **107**:481–501.

Thomas, E., and A. J. Gooday. 1996. Cenozoic deep-sea benthic foraminifers: Tracers for changes in oceanic productivity? Geology **24**:355–358.

Thomas, H. 1933. On some pteridospermous plants from the Mesozoic rocks of South Africa. Philosophical Transactions of the Linnean Society of London B **222**:193–265.

Tidwell, W., and S. Ash. 1994. A review of selected Triassic to Early Cretaceous ferns. Journal of Plant Research **107**:417–442.

Townrow, J. 1960. The Peltaspermaceae, a pteridosperm family of Permian and Triassic age. Palaeontology **3**:333–361.

Tozer, E., 1961. Summary account of Mesozoic and Tertiary stratigraphy, Canadian Arctic Archipelago. Pages 381–402 *in* G. Raasch, editor. Geology of the Arctic, Proceedings of the First International Symposium on Arctic Geology. University of Toronto Press, Toronto.

Tozer, E., 1963. Northwestern Cameron Island. Pages 639–644 *in* Y. Fortier, R. Blackadar, B. Glenister, H. Greiner, D. McLaren, N. McMillan, A. Norris, E. Roots, J. Souther, R. Thorsteinsson, and E. Tozer, editors. Geology of the North-Central Part of the Arctic Archipelago, Northwest Territories (Operation Franklin)., volume 320. Geological Survey of Canada Memoir, 320.

Urbanek, S., 2008. proj4: A simple interface to the PROJ.4 cartographic projections library. URL `http://www.rforge.net/proj4/`.

Vavrek, M. J., 2010. fossil: Palaeoecological and Palaeogeographical Analysis Tools. URL `http://cran.r-project.org/web/packages/fossil/`.

Vermeij, G., and G. Herbert. 2004. Measuring relative abundance in fossil and living assemblages. Paleobiology **30**:1–4.

Wallace, A. R. 1878. Tropical nature and other essays. Macmillan New York.

Watson, H. 1859. Cybele Brittanica. Longman and Company, London.

Weir, J. T., and D. Schluter. 2007. The latitudinal gradient in recent speciation and extinction rates of birds and mammals. Science **315**:1574–1576.

Westrop, S. R., and J. M. Adrain. 2001. Sampling at the species level: Impact of spatial biases on diversity gradients. Geology **29**:903–a–906.

Whittaker, R., 1977. Evolution of species diversity in land communities. Pages 250–268 *in* M. Hecht, W. Steere, and B. Wallace, editors. Evolutionary Biology, volume 10. Plenum Press, New York.

Whittaker, R. H. 1960. Vegetation of the Siskiyou Mountains, Oregon and California. Ecological Monographs **30**:279–338.

Whittaker, R. H. 1972. Evolution and measurement of species diversity. Taxon **21**:213–251.

Willig, M. R., D. M. Kaufman, and R. D. Stevens. 2003. Latitudinal gradients of biodiversity: pattern, process, scale, and synthesis. Annual Review of Ecology, Evolution, and Systematics **34**:273–309.

Wilson, E. O. 2003. The Future of Life. Vintage Books, New York.

Wing, S. L., and W. A. Dimichele. 1995. Conflict between local and global changes in plant diversity through geological time. PALAIOS **10**:551–564.

Wolda, H. 1981. Similarity indices, sample size and diversity. Oecologia **50**:296–302.

Wolfe, J. A., and G. R. Upchurch. 1987. North American nonmarine climates and vegetation during the Late Cretaceous. Palaeogeography, Palaeoclimatology, Palaeoecology **61**:33–77.

Ziegler, A., 1990. Phytogeographic patterns and continental configurations during the Permian Period. Pages 363–379 *in* W. McKerrow and C. Scotese, editors. Palaeozoic palaeogeography and biogeography, volume 12. Geological Society [London] Memoir.

Ziegler, A. M., J. M. Parrish, Y. Jiping, E. D. Gyllenhaal, D. B. Rowley, J. T. Parrish, N. Shangyou, A. Bekker, and M. L. Hulver. 1993. Early Mesozoic phytogeography and climate. Philosophical Transactions: Biological Sciences **341**:297–305.

Appendix A: Copyright License

Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License

(c) "Distribute" means to make available to the public the original and copies of the Work or Adaptation, as appropriate, through sale or other transfer of ownership.
(d) "License Elements" means the following high-level license attributes as selected by Licensor and indicated in the title of this License: Attribution, Noncommercial, ShareAlike.
(e) "Licensor" means the individual, individuals, entity or entities that offer(s) the Work under the terms of this License.
(f) "Original Author" means, in the case of a literary or artistic work, the individual, individuals, entity or entities who created the Work or if no individual or entity can be identified, the publisher; and in addition (i) in the case of a performance the actors, singers, musicians, dancers, and other persons who act, sing, deliver, declaim, play in, interpret or otherwise perform literary or artistic works or expressions of folklore; (ii) in the case of a phonogram the producer being the person or legal entity who first fixes the sounds of a performance or other sounds; and, (iii) in the case of broadcasts, the organization that transmits the broadcast.
(g) "Work" means the literary and/or artistic work offered under the terms of this License including without limitation any production in the literary, scientific and artistic domain, whatever may be the mode or form of its expression including digital form, such as a book, pamphlet and other writing; a lecture, address, sermon or other work of the same nature; a dramatic or dramatico-musical work; a choreographic work or entertainment in dumb show; a musical composition with or without words; a cinematographic work to which are assimilated works expressed by a process analogous to cinematography; a work of drawing, painting, architecture, sculpture, engraving or lithography; a photographic work to which are assimilated works expressed by a process analogous to photography; a work of applied art; an illustration, map, plan, sketch or three-dimensional work relative to geography, topography, architecture or science; a performance; a broadcast; a phonogram; a compilation of data to the extent it is protected as a copyrightable work; or a work performed by a variety or circus performer to the extent it is not otherwise considered a literary or artistic work.
(h) "You" means an individual or entity exercising rights under this License who has not previously violated the terms of this License with respect

to the Work, or who has received express permission from the Licensor to exercise rights under this License despite a previous violation.

(i) "Publicly Perform" means to perform public recitations of the Work and to communicate to the public those public recitations, by any means or process, including by wire or wireless means or public digital performances; to make available to the public Works in such a way that members of the public may access these Works from a place and at a place individually chosen by them; to perform the Work to the public by any means or process and the communication to the public of the performances of the Work, including by public digital performance; to broadcast and rebroadcast the Work by any means including signs, sounds or images.

(j) "Reproduce" means to make copies of the Work by any means including without limitation by sound or visual recordings and the right of fixation and reproducing fixations of the Work, including storage of a protected performance or phonogram in digital form or other electronic medium.

2. **Fair Dealing Rights**
Nothing in this License is intended to reduce, limit, or restrict any uses free from copyright or rights arising from limitations or exceptions that are provided for in connection with the copyright protection under copyright law or other applicable laws.

3. **License Grant**
Subject to the terms and conditions of this License, Licensor hereby grants You a worldwide, royalty-free, non-exclusive, perpetual (for the duration of the applicable copyright) license to exercise the rights in the Work as stated below:
   (a) to Reproduce the Work, to incorporate the Work into one or more Collections, and to Reproduce the Work as incorporated in the Collections;
   (b) to create and Reproduce Adaptations provided that any such Adaptation, including any translation in any medium, takes reasonable steps to clearly label, demarcate or otherwise identify that changes were made to the original Work. For example, a translation could be marked "The original work was translated from English to Spanish," or a modification could indicate "The original work has been modified.";
   (c) to Distribute and Publicly Perform the Work including as incorporated in Collections; and,
   (d) to Distribute and Publicly Perform Adaptations.

The above rights may be exercised in all media and formats whether now known or hereafter devised. The above rights include the right to make such modifications as are technically necessary to exercise the rights in other media and formats. Subject to Section 8(f), all rights not expressly granted by Licensor are hereby reserved, including but not limited to the rights described in Section 4(e).

4. **Restrictions**
   The license granted in Section 3 above is expressly made subject to and limited by the following restrictions:
   (a) You may Distribute or Publicly Perform the Work only under the terms of this License. You must include a copy of, or the Uniform Resource Identifier (URI) for, this License with every copy of the Work You Distribute or Publicly Perform. You may not offer or impose any terms on the Work that restrict the terms of this License or the ability of the recipient of the Work to exercise the rights granted to that recipient under the terms of the License. You may not sublicense the Work. You must keep intact all notices that refer to this License and to the disclaimer of warranties with every copy of the Work You Distribute or Publicly Perform. When You Distribute or Publicly Perform the Work, You may not impose any effective technological measures on the Work that restrict the ability of a recipient of the Work from You to exercise the rights granted to that recipient under the terms of the License. This Section 4(a) applies to the Work as incorporated in a Collection, but this does not require the Collection apart from the Work itself to be made subject to the terms of this License. If You create a Collection, upon notice from any Licensor You must, to the extent practicable, remove from the Collection any credit as required by Section 4(d), as requested. If You create an Adaptation, upon notice from any Licensor You must, to the extent practicable, remove from the Adaptation any credit as required by Section 4(d), as requested.
   (b) You may Distribute or Publicly Perform an Adaptation only under: (i) the terms of this License; (ii) a later version of this License with the same License Elements as this License; (iii) a Creative Commons jurisdiction license (either this or a later license version) that contains the same License Elements as this License (e.g., Attribution-NonCommercial-ShareAlike 3.0 US) ("Applicable License"). You must include a copy of, or the URI, for Applicable License with every copy of each Adaptation You Distribute or Publicly Perform. You may not

offer or impose any terms on the Adaptation that restrict the terms of the Applicable License or the ability of the recipient of the Adaptation to exercise the rights granted to that recipient under the terms of the Applicable License. You must keep intact all notices that refer to the Applicable License and to the disclaimer of warranties with every copy of the Work as included in the Adaptation You Distribute or Publicly Perform. When You Distribute or Publicly Perform the Adaptation, You may not impose any effective technological measures on the Adaptation that restrict the ability of a recipient of the Adaptation from You to exercise the rights granted to that recipient under the terms of the Applicable License. This Section 4(b) applies to the Adaptation as incorporated in a Collection, but this does not require the Collection apart from the Adaptation itself to be made subject to the terms of the Applicable License.

(c) You may not exercise any of the rights granted to You in Section 3 above in any manner that is primarily intended for or directed toward commercial advantage or private monetary compensation. The exchange of the Work for other copyrighted works by means of digital file-sharing or otherwise shall not be considered to be intended for or directed toward commercial advantage or private monetary compensation, provided there is no payment of any monetary compensation in connection with the exchange of copyrighted works.

(d) If You Distribute, or Publicly Perform the Work or any Adaptations or Collections, You must, unless a request has been made pursuant to Section 4(a), keep intact all copyright notices for the Work and provide, reasonable to the medium or means You are utilizing: (i) the name of the Original Author (or pseudonym, if applicable) if supplied, and/or if the Original Author and/or Licensor designate another party or parties (e.g., a sponsor institute, publishing entity, journal) for attribution ("Attribution Parties") in Licensor's copyright notice, terms of service or by other reasonable means, the name of such party or parties; (ii) the title of the Work if supplied; (iii) to the extent reasonably practicable, the URI, if any, that Licensor specifies to be associated with the Work, unless such URI does not refer to the copyright notice or licensing information for the Work; and, (iv) consistent with Section 3(b), in the case of an Adaptation, a credit identifying the use of the Work in the Adaptation (e.g., "French translation of the Work by Original Author," or "Screenplay based on original Work by Original Author").

The credit required by this Section 4(d) may be implemented in any reasonable manner; provided, however, that in the case of a Adaptation or Collection, at a minimum such credit will appear, if a credit for all contributing authors of the Adaptation or Collection appears, then as part of these credits and in a manner at least as prominent as the credits for the other contributing authors. For the avoidance of doubt, You may only use the credit required by this Section for the purpose of attribution in the manner set out above and, by exercising Your rights under this License, You may not implicitly or explicitly assert or imply any connection with, sponsorship or endorsement by the Original Author, Licensor and/or Attribution Parties, as appropriate, of You or Your use of the Work, without the separate, express prior written permission of the Original Author, Licensor and/or Attribution Parties.

(e) For the avoidance of doubt:

   i. Non-waivable Compulsory License Schemes. In those jurisdictions in which the right to collect royalties through any statutory or compulsory licensing scheme cannot be waived, the Licensor reserves the exclusive right to collect such royalties for any exercise by You of the rights granted under this License;

   ii. Waivable Compulsory License Schemes. In those jurisdictions in which the right to collect royalties through any statutory or compulsory licensing scheme can be waived, the Licensor reserves the exclusive right to collect such royalties for any exercise by You of the rights granted under this License if Your exercise of such rights is for a purpose or use which is otherwise than noncommercial as permitted under Section 4(c) and otherwise waives the right to collect royalties through any statutory or compulsory licensing scheme; and,

   iii. Voluntary License Schemes. The Licensor reserves the right to collect royalties, whether individually or, in the event that the Licensor is a member of a collecting society that administers voluntary licensing schemes, via that society, from any exercise by You of the rights granted under this License that is for a purpose or use which is otherwise than noncommercial as permitted under Section 4(c).

(f) Except as otherwise agreed in writing by the Licensor or as may be otherwise permitted by applicable law, if You Reproduce, Distribute or Publicly Perform the Work either by itself or as part of any Adaptations or Collections, You must not distort, mutilate, modify or take other derogatory action in relation to the Work which would be prejudicial to the Original Author's honor or reputation. Licensor agrees that in those jurisdictions (e.g. Japan), in which any exercise of the right granted in Section 3(b) of this License (the right to make Adaptations) would be deemed to be a distortion, mutilation, modification or other derogatory action prejudicial to the Original Author's honor and reputation, the Licensor will waive or not assert, as appropriate, this Section, to the fullest extent permitted by the applicable national law, to enable You to reasonably exercise Your right under Section 3(b) of this License (right to make Adaptations) but not otherwise.

5. **Representations, Warranties and Disclaimer**
   UNLESS OTHERWISE MUTUALLY AGREED TO BY THE PARTIES IN WRITING AND TO THE FULLEST EXTENT PERMITTED BY APPLICABLE LAW, LICENSOR OFFERS THE WORK AS-IS AND MAKES NO REPRESENTATIONS OR WARRANTIES OF ANY KIND CONCERNING THE WORK, EXPRESS, IMPLIED, STATUTORY OR OTHERWISE, INCLUDING, WITHOUT LIMITATION, WARRANTIES OF TITLE, MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE, NONINFRINGEMENT, OR THE ABSENCE OF LATENT OR OTHER DEFECTS, ACCURACY, OR THE PRESENCE OF ABSENCE OF ERRORS, WHETHER OR NOT DISCOVERABLE. SOME JURISDICTIONS DO NOT ALLOW THE EXCLUSION OF IMPLIED WARRANTIES, SO THIS EXCLUSION MAY NOT APPLY TO YOU.
6. **Limitation on Liability**
   EXCEPT TO THE EXTENT REQUIRED BY APPLICABLE LAW, IN NO EVENT WILL LICENSOR BE LIABLE TO YOU ON ANY LEGAL THEORY FOR ANY SPECIAL, INCIDENTAL, CONSEQUENTIAL, PUNITIVE OR EXEMPLARY DAMAGES ARISING OUT OF THIS LICENSE OR THE USE OF THE WORK, EVEN IF LICENSOR HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES.
7. **Termination**
   (a) This License and the rights granted hereunder will terminate automatically upon any breach by You of the terms of this License. Individuals or entities who have received Adaptations or Collections from You under this License, however, will not have their licenses terminated provided such individuals or entities remain in full compliance with those licenses. Sections 1, 2, 5, 6, 7, and 8 will survive any termination of this License.

(b) Subject to the above terms and conditions, the license granted here is perpetual (for the duration of the applicable copyright in the Work). Notwithstanding the above, Licensor reserves the right to release the Work under different license terms or to stop distributing the Work at any time; provided, however that any such election will not serve to withdraw this License (or any other license that has been, or is required to be, granted under the terms of this License), and this License will continue in full force and effect unless terminated as stated above.

8. **Miscellaneous**
   (a) Each time You Distribute or Publicly Perform the Work or a Collection, the Licensor offers to the recipient a license to the Work on the same terms and conditions as the license granted to You under this License.
   (b) Each time You Distribute or Publicly Perform an Adaptation, Licensor offers to the recipient a license to the original Work on the same terms and conditions as the license granted to You under this License.
   (c) If any provision of this License is invalid or unenforceable under applicable law, it shall not affect the validity or enforceability of the remainder of the terms of this License, and without further action by the parties to this agreement, such provision shall be reformed to the minimum extent necessary to make such provision valid and enforceable.
   (d) No term or provision of this License shall be deemed waived and no breach consented to unless such waiver or consent shall be in writing and signed by the party to be charged with such waiver or consent.
   (e) This License constitutes the entire agreement between the parties with respect to the Work licensed here. There are no understandings, agreements or representations with respect to the Work not specified here. Licensor shall not be bound by any additional provisions that may appear in any communication from You. This License may not be modified without the mutual written agreement of the Licensor and You.
   (f) The rights granted under, and the subject matter referenced, in this License were drafted utilizing the terminology of the Berne Convention for the Protection of Literary and Artistic Works (as amended on September 28, 1979), the Rome Convention of 1961, the WIPO Copyright Treaty of 1996, the WIPO Performances and Phonograms Treaty of 1996 and the Universal Copyright Convention (as revised on July 24, 1971). These rights and subject matter take effect in the relevant

jurisdiction in which the License terms are sought to be enforced according to the corresponding provisions of the implementation of those treaty provisions in the applicable national law. If the standard suite of rights granted under applicable copyright law includes additional rights not granted under this License, such additional rights are deemed to be included in the License; this License is not intended to restrict the license of any rights under applicable law.

Appendix B: Supplemental R Code for Chapter 2

```
###################################################
### chunk number 1: data-in
###################################################
#lines starting with a # are comments, and are ignored by the R interpreter
#install.packages('fossil')
library(fossil)


###################################################
### chunk number 2: list-to-occ-mat
###################################################
data(fdata.list)
create.matrix(fdata.list, tax.name = 'species', locality = 'locality')


###################################################
### chunk number 3: list-to-abund-mat
###################################################
data(fdata.list)
create.matrix(fdata.list, tax.name = 'species', locality = 'locality', abund=TRUE,
    abund.col = 'abundance')


###################################################
### chunk number 4: list-to-lats
###################################################
data(fdata.list)
create.lats(fdata.list, loc = 'locality', long = 'longitude', lat = 'latitude')


###################################################
### chunk number 5: sim-measure
###################################################
sampleA <- c(1,1,0,1,1,1,1)
sampleB <- c(0,1,1,0,0,1,1)
sorenson(sampleA, sampleB)


###################################################
### chunk number 6: spp-ests
###################################################
data(fdata.mat)
chao1(fdata.mat)
jack1(fdata.mat)


###################################################
### chunk number 7: shapefiles
###################################################
data(fdata.lats)
shape.lats<-lats2Shape(fdata.lats)
fdata.dist<-dino.dist(fdata.mat)
```

```
fdata.mst<-dino.mst(fdata.dist)
shape.mst<-msn2Shape(fdata.mst, fdata.lats)


###################################################
### chunk number 8: mst-map
###################################################
data(fdata.mat)
fdata.dist<-dino.dist(fdata.mat)
fdata.mst<-dino.mst(fdata.dist)
data(fdata.lats)
library(maps)
map('state')
mstlines(fdata.mst, coordinates(fdata.lats))
points(coordinates(fdata.lats), pch=16, col='white', cex=3)
points(coordinates(fdata.lats), pch=1, cex=3)
text(coordinates(fdata.lats), labels=LETTERS[1:12])


###################################################
### chunk number 9: geo-dist-map
###################################################
data(fdata.lats)
fd.subset<-coordinates(fdata.lats)[1:3,]
earth.dist(fdata.lats[1:3,])
map('state')
polygon(fd.subset)
text(c(-110,-101,-106), c(42, 42, 47), labels=round(earth.dist(fd.subset)[c(1,3,2)])
    )
points(fd.subset, pch=16, col='white', cex=3)
points(fd.subset, pch=1, cex=3)
text(fd.subset, label=LETTERS[1:3])


###################################################
### chunk number 10: spp-area-fig
###################################################
plot(log(sac(fdata.lats, fdata.mat)[[1]]), ylab = 'log species richness', xlab = '
    log area (km^2)')
```

Appendix C: Supplemental R Code for Chapter 3

```
#################################################
#this code is the original used for all data analysis and figure creation in the
    paper.
#The following is released under a Creative Commons Attribution-Noncommercial-Share
    Alike 3.0 Unported license
#full terms of the license can be found at: http://creativecommons.org/licenses/by-
    nc-sa/3.0/
#Copyright Matthew Vavrek (2010)
#################################################

###################################################
### chunk number 1: data
###################################################
#relational clustering paper example using white toothed shrews
library(fossil)
library(ecodist)


#shrews
read.csv('shrews.csv', header = T, row.names=1)->shrews
nmds(as.dist(shrews), mindim=2, maxdim=2, nits=10)->shrews.nmds
shrews.msn<-dino.msn(as.dist(shrews))

#clustering into 2 groups
rcs2<-rclust(shrews, 2, 100)
rcs2.null <- rclust.null(rcs2, shrews)

#Cohesiveness Index calculations
coi(shrews.msn, rcs2)
#to shift group numbers so that T+ is #1 and AlGC is #2
gn2 <- c(rcs2[1], rcs2[8])
gp2 <- numeric(2)
for (i in 1:2) gp2[i] <- which(gn2==i)
grps2 <- gp2[rcs2] #this is the final shifted group assignment vector
#the group names
grp.names2 <- c('T+', 'AlGC')
#for table 1, the within and between group average distances
tab2<-round(rclust.dist(grps2, shrews),2)
null2 <- round(rclust.null(grps2, shrews),2)
coi2 <- coi(shrews.msn, grps2)

#clustering into 3 groups
rcs3<-rclust(shrews, 3, 100)

#to shift group numbers so that T+ is #1 and AlGC is #2 and JS is #3
gn3 <- c(rcs3[1], rcs3[8], rcs3[7])
gp3 <- numeric(3)
for (i in 1:3) {
  gp3[i] <- which(gn3==i)
}
grps3 <- gp3[rcs3] #this is the final shifted group assignment vector
#the group names
```

```
grp.names3 <- c('T+', 'AlGC', 'JS')
#for table 3, the between and within group average distances
tab3 <- round(rclust.dist(grps3, shrews),2)
null3 <- round(rclust.null(grps3, shrews),2)
coi3 <- coi(shrews.msn, grps3)


###################################################
### chunk number 2: fig1
###################################################
pts<-shrews.nmds$conf[[which.min(shrews.nmds$stress)]]
par(mar=c(0,0,0,0))
plot(pts, col=rcs2, pch='', xaxt = 'n', yaxt = 'n')

flowlines <- matrix(c(1,5,1,2,1,3,1,6,1,7,4,10,4,9,4,8), 8,2, byrow = TRUE)
for (i in 1:8) lines(x = pts[flowlines[i,],1], y = pts[flowlines[i,],2])
for (i in 1:8) points(x = mean(pts[flowlines[i,],1]), y = mean(pts[flowlines[i,],2])
    , col='white', pch=18, cex=3.5)
for (i in 1:8) points(x = mean(pts[flowlines[i,],1]), y = mean(pts[flowlines[i,],2])
    , pch=5, cex=2.5)
for (i in 1:8) text(x = mean(pts[flowlines[i,],1]), y = mean(pts[flowlines[i,],2]),
    labels = i)

points(pts, col='white', pch=19, cex=3.5)
points(pts, pch=1, cex=3.5)
points(pts[c(1,4),], pch=1, cex=3.8)
text(pts, labels=colnames(shrews))


###################################################
### chunk number 3: fig2
###################################################
pts<-shrews.nmds$conf[[which.min(shrews.nmds$stress)]]
#sets the point style for the graph
cs <- c(19,17)
ocs <- c(1,2)
par(mar=c(0,0,0,0))
plot(pts, pch='', xaxt = 'n', yaxt = 'n')
mstlines(shrews.msn, pts)
points(pts, col = 'white', pch = cs[grps2], cex=3.5)
points(pts, pch = ocs[grps2], cex=3.5)
text(pts, labels=colnames(shrews))


###################################################
### chunk number 4: fig3
###################################################
pts<-shrews.nmds$conf[[which.min(shrews.nmds$stress)]]
#sets the point style for the graph
cs <- c(19,17,15)
ocs <- c(1,2,22)
par(mar=c(0,0,0,0))
plot(pts, pch='', xaxt = 'n', yaxt = 'n')
mstlines(shrews.msn, pts)
points(pts, col='white', pch = cs[grps3], cex=3.5)
points(pts, pch = ocs[grps3], cex=3.5)
text(pts, labels=colnames(shrews))
```

Appendix D: R Code for Complete Data for Chapter 5

```
###################################################
### chunk number 1: bdiv
###################################################
#this code was originally integrated as part of the LaTeX file that created the paper.
#The following is released under a Creative Commons Attribution-Noncommercial-Share Alike 3.0 Unported license
#full terms of the license can be found at: http://creativecommons.org/licenses/by-nc-sa/3.0/

#This code was tested on an Ubuntu 9.10 distribution running R version 2.9.2

#To set up the environment, be sure that the following files are located in the same directory as your working directory:
#-si_dataset_occurrence.csv
#-si_dataset_location.csv
#-si_appendix_map.shp

#and that you have the following packages installed:
#-fossil
#-shapefiles
#-PBSmapping
#-proj4
#-ecodist
#included in the following in the .Rnw file is the R code for the dinosaur beta diversity paper
#loading the necessary environment and datasets
library(fossil)
library(shapefiles)
library(PBSmapping)
library(proj4)
library(ecodist)
#library(rimage)

#bdiv is the name of the file that is directly taken from the PaleoDB
bdiv<-read.csv("si_dataset_occurrence.csv")
# bdiv.mat is the species occurrence matrix (ie species X locality)
bdiv.mat<-create.matrix(bdiv,locality='formation', abund=T)

#reading in the locality lat/long
bdiv.lats<-read.csv('si_dataset_location.csv', header=TRUE, row.names=1)

#create the distance matrix, minimum spanning tree and nmds for later
bdiv.dist<-dino.dist(bdiv.mat)
bdiv.msn<-dino.msn(bdiv.dist)
bdiv.nmds<-nmds(bdiv.dist, mindim=2, maxdim=2)


##removing bonebed counts
bdiv.mat["Edmontosaurus","Hell Creek South Dakota"]<-1
bdiv.mat["Edmontosaurus","Lance Wyoming"]<-11
```

```
##############################################################
#don't forget to change those randomizations to 1000 (rand = randomizations)
#the randomizations can be set to a low number to speed up processing, but for nicer
     graphs a higher number should be used
rand<-1000

##calculating spp.est curves for all sites >100 specimens
hcm.spp.est<-spp.est(bdiv.mat[,"Hell␣Creek␣Montana"],rand, counter=F)
hcnd.spp.est<-spp.est(bdiv.mat[,"Hell␣Creek␣North␣Dakota"],rand, counter=F)
hca.spp.est<-spp.est(bdiv.mat[,"Horseshoe␣Canyon␣Alberta"],rand, counter=F)
lw.spp.est<-spp.est(bdiv.mat[,"Lance␣Wyoming"],rand, counter=F)
bdiv.spp.est<-spp.est(rowSums(bdiv.mat), rand, counter=F)
total.localities<-dim(bdiv.mat)[2]
total.specimens<-dim(bdiv.spp.est)[1]
total.species<-bdiv.spp.est[total.specimens,2]

tmp<-bdiv.mat
tmp[tmp>0]<-1
observed.beta<-round(total.species/mean(colSums(tmp)), 2)
total.average.alpha <- round(mean(colSums(tmp)), 2)

#number of specimens/individuals per sample site
csums<-colSums(bdiv.mat)

total.greater.than.10.specimens<-length(csums[csums>10])

localities<-c("Hell␣Creek␣Montana","Hell␣Creek␣North␣Dakota","Horseshoe␣Canyon␣
    Alberta","Lance␣Wyoming")

min.loc<-min(colSums(bdiv.mat[,localities]))
min.loc.name<-names(which.min(colSums(bdiv.mat[,localities])))

max.loc<-max(colSums(bdiv.mat[,localities]))
max.loc.name<-names(which.max(colSums(bdiv.mat[,localities])))

# the following few lines calculate the linear model (regression) for the number of
    species observed versus the number of
# specimens. A graph of this is also included in the supplementary material, but for
     this portion of the paper just the
# values from the regression are used.
a<-log(colSums(bdiv.mat))
b<-log(apply(bdiv.mat,2,function(x) length(x[x>0])))
lmfig2<-lm(b~a)
r2.value<-round(summary(lmfig2)$r.squared, 2)
prob<-round(pf(summary(lmfig2)$fstatistic[1], summary(lmfig2)$fstatistic[2],
summary(lmfig2)$fstatistic[3], lower.tail = FALSE), 3)
if (prob==0) prob<-0.001

plot(a, b, xlab = 'log␣specimens␣recorded', ylab = 'log␣species␣richness')
abline(lmfig2)


###################################################
### chunk number 2: tab1
###################################################
#the following creates an easier to address table of values
sptab<-matrix(,7,8)
cols<-c(2,5,8,11)
```

```
sptab[1,1]<-dim(hcm.spp.est)[1]
sptab[1,2]<-hcm.spp.est[dim(hcm.spp.est)[1],2]
sptab[1,3:6]<-hcm.spp.est[100,cols]

sptab[2,1]<-dim(hcnd.spp.est)[1]
sptab[2,2]<-hcnd.spp.est[dim(hcnd.spp.est)[1],2]
sptab[2,3:6]<-hcnd.spp.est[100,cols]

sptab[3,1]<-dim(hca.spp.est)[1]
sptab[3,2]<-hca.spp.est[dim(hca.spp.est)[1],2]
sptab[3,3:6]<-hca.spp.est[100,cols]

sptab[4,1]<-dim(lw.spp.est)[1]
sptab[4,2]<-lw.spp.est[dim(lw.spp.est)[1],2]
sptab[4,3:6]<-lw.spp.est[100,cols]

sptab[5,1]<-total.specimens
sptab[5,2]<-total.species
sptab[5,3:6]<-bdiv.spp.est[100,cols]

for (i in 1:5) sptab[i,7]<-mean(sptab[i,4:6])
for (i in 1:7) sptab[6,i]<-mean(sptab[1:4,i])
for (i in 2:7) sptab[7,i]<-sptab[5,i]/mean(sptab[1:4,i])

sptab<-round(sptab,2)


###################################################
### chunk number 3: fig1
###################################################
#the following sets up an environment to plot figure 1 from the bdiv_sci paper


read.shp('si_appendix_map.shp')->a
convert.to.simple(a)->b

b<-cbind(b[,1],0,b[,2:3])
for (i in 1:length(levels(as.factor(b[,1])))) b[b[,1]==i,2]<-1:length(b[b[,1]==i,2])

colnames(b)<-c('PID', 'POS', 'X', 'Y')
b<-as.data.frame(b)

as.PolySet(b)->d


nproj<-"+proj=lcc␣+lat_1=43.26666666666667␣+lat_2=42.06666666666667␣+lat_0=41.5␣+lon
    _0=-102.5␣+x_0=1500000␣+y_0=1000000␣+ellps=GRS80␣+datum=NAD83␣+units=m␣+no_defs"

xm<-d
xm[,3:4] <- project(d[,3:4], nproj)

lats.prj<-bdiv.lats
lats.prj[,1:2]<-project(bdiv.lats[,2:1], nproj)
locs<-c('A', 'D', 'Fe', 'Fr', 'HM', 'HN', 'HS', 'HC', 'J', 'Kp', 'Kd', 'LM', 'LS', '
    LU', 'LW', 'LaC', 'LaW', 'M', 'N', 'P', 'Sc', 'SA', 'SM', 'T')
```

```
##the following creates a figure showing all of N America and a closeup of the
    region of interest (Western Interior)
n<-max(xm[,4])
s<-min(xm[,4])+1700000
e<-max(xm[,3])
w<-min(xm[,3])
bs<-(-500000)
bn<-(2200000)
bw<-(450000)
be<-(1850000)
#the affects the layout of the figure
layout(matrix(c(1,2,2,4,3,3), 3, 2))
par(oma=c(2,2,2,2))
par(mar=c(0,0,0,0))
plot(1, type='n', yaxt='n', xaxt='n', ylim=c(s,n), xlim=c(w,e), xaxs='i', yaxs='i',
    bty='n')
for (i in 1:length(levels(as.factor(xm[,1])))) polygon(xm[xm[,1]==i,3:4], lwd=0.6)
polygon(c(bw,bw,be,be), c(bs, bn, bn, bs), lwd=1.5)
lines(c(w, bw), c(s, bs))
lines(c(e, be), c(s, bs))

par(mar=c(0,0,0,0))
plot(1, type='n', yaxt='n', xaxt='n', ylim=c(bs,bn), xlim=c(bw,be), xaxs='i', yaxs='
    i')
for (i in 1:length(levels(as.factor(xm[,1])))) polygon(xm[xm[,1]==i,3:4], lwd=0.6)
points(lats.prj, pch=16, cex=4, col='white')
points(lats.prj, pch=1, cex=4, col='black')
mainpts<-c(5,6,8,15)
text(lats.prj[-mainpts,], labels=locs[-mainpts], font=1, cex=0.8)
text(lats.prj[mainpts,], labels=locs[mainpts], font=2, cex=0.8)

par(mar=c(0,0,0,0))
plot(1, type='n', yaxt='n', xaxt='n', ylim=c(bs,bn), xlim=c(bw,be), xaxs='i', yaxs='
    i')
for (i in 1:length(levels(as.factor(xm[,1])))) polygon(xm[xm[,1]==i,3:4], lwd=0.6)
points(lats.prj[bdiv.mat["Alamosaurus",]>0,], pch='A', font = 2)
points(lats.prj[bdiv.mat["Leptoceratops",]>0,]-10000, pch='L', font = 2)
points(lats.prj[bdiv.mat["Triceratops",]>0,]+10000, pch='T', font = 2)


###################################################
### chunk number 4: fig2
###################################################
#################################################################
##plotting various curves
loc.names <- c("Total", "Hell␣Creek␣Montana", "Hell␣Creek␣North␣Dakota", "Horseshoe␣
    Canyon␣Alberta", "Lance␣Wyoming")
loc.lines <- c(5,1,2,3,4)

##raw spp
par(mfrow=c(2,2),oma=c(4,4,4,4))
par(mar=c(0,0,0,0))
plot(hcm.spp.est[1:100,2], lty = 1, type = "l", ylim = c(0,35), xaxt="n")
lines(hcnd.spp.est[1:100,2], lty = 2)
lines(hca.spp.est[1:100,2], lty = 3)
lines(lw.spp.est[1:100,2], lty = 4)
lines(bdiv.spp.est[1:100,2], lty = 5)
points(97, 2, pch ='A', cex = 2)
```

```
legend(0, 35, legend = loc.names, lty = loc.lines, cex=0.8)

##chao 1
par(mar=c(0,0,0,0))
plot(hcm.spp.est[1:100,5], lty = 1, type = "l", ylim = c(0,35),xaxt="n",yaxt="n")
lines(hcnd.spp.est[1:100,5], lty = 2)
lines(hca.spp.est[1:100,5], lty = 3)
lines(lw.spp.est[1:100,5], lty = 4)
lines(bdiv.spp.est[1:100,5], lty = 5)
points(97, 2, pch ='B', cex = 2)

##ACE
par(mar=c(0,0,0,0))
plot(hcm.spp.est[1:100,8], lty = 1, type = "l", ylim = c(0,35))
lines(hcnd.spp.est[1:100,8], lty = 2)
lines(hca.spp.est[1:100,8], lty = 3)
lines(lw.spp.est[1:100,8], lty = 4)
lines(bdiv.spp.est[1:100,8], lty = 5)
points(97, 2, pch ='C', cex = 2)

###Jack 1
par(mar=c(0,0,0,0))
plot(hcm.spp.est[1:100,11], lty = 1, type = "l", ylim = c(0,35),yaxt="n")
lines(hcnd.spp.est[1:100,11], lty = 2)
lines(hca.spp.est[1:100,11], lty = 3)
lines(lw.spp.est[1:100,11], lty = 4)
lines(bdiv.spp.est[1:100,11], lty = 5)
points(97, 2, pch ='D', cex = 2)

mtext("Species␣Richness␣(S)",side=2,outer=T, padj=-3)
mtext("Number␣of␣Individuals␣(N)",side=1,outer=T,padj=3)


###################################################
### chunk number 5: fig3
###################################################

#the affects the layout of the figure
layout(matrix(c(1,1,2,3), 2, 2))

#map with mst
par(mar=c(2,2,2,2))
plot(1, type='n', yaxt='n', xaxt='n', ylim=c(bs,bn), xlim=c(bw,be), xaxs='i', yaxs='
    i')
for (i in 1:length(levels(as.factor(xm[,1])))) polygon(xm[xm[,1]==i,3:4], lwd=0.6)
for (i in 1:ncol(bdiv.msn)) {
  for (j in 1:ncol(bdiv.msn)) {
    if (bdiv.msn[i, j] == 1) {
      lines(c(lats.prj[i,1], lats.prj[j,1]), c(lats.prj[i,2], lats.prj[j,2]), col=
          gray(as.matrix(bdiv.dist)[i,j]), lwd=1.5)
    }
  }
}

points(lats.prj, pch=16, cex=4, col='white')
points(lats.prj, pch=1, cex=4, col='black')
mainpts<-c(5,6,8,15)
```

```r
text(lats.prj[-mainpts,], labels=locs[-mainpts], font=1, cex=0.8)
text(lats.prj[mainpts,], labels=locs[mainpts], font=2, cex=0.8)
points(1750000, -400000, pch ='A', cex = 2)

#nmds with mst
par(mar=c(1,2,2,2))
nmds2plot<-bdiv.nmds$conf[[which.min(bdiv.nmds$stress)]]

plot(nmds2plot[, 1], nmds2plot[, 2], pch=locs, xlab='', ylab='', cex=2, type='n',
    xaxt = 'n', yaxt = 'n')
for (i in 1:ncol(bdiv.msn)) {
  for (j in 1:ncol(bdiv.msn)) {
    if (bdiv.msn[i, j] == 1) {
      lines(c(nmds2plot[i,1], nmds2plot[j,1]), c(nmds2plot[i,2], nmds2plot[j,2]),
          col=gray(as.matrix(bdiv.dist)[i,j]))
    }
  }
}
text(nmds2plot, labels=locs, font = 2)
mainpts<-c(5,6,8,15)
points(nmds2plot[mainpts,], pch=1, cex=4)
points(max(nmds2plot[,1])-0.05, min(nmds2plot[,2])+0.05, pch = 'B', cex = 2)

#distance vs similarity
par(mar=c(4,4,1,2))
plot(earth.dist(bdiv.lats), 1-bdiv.dist, ylab='Simpson␣Similarity', xlab='Distance␣(
    km)')
lines(lowess(earth.dist(bdiv.lats), 1-bdiv.dist))
#abline(glm(1-bdiv.dist~earth.dist(bdiv.lats)))
points(2400, 0.05, pch = 'C', cex = 2)
```

Appendix E: R Code for Specimen-only Data for Chapter 5

```
###################################################
### chunk number 1: bdiv
###################################################
#this code is a modification of the original code used in the paper, such that only
    specimen data is used
#The following is released under a Creative Commons Attribution-Noncommercial-Share
    Alike 3.0 Unported license
#full terms of the license can be found at: http://creativecommons.org/licenses/by-
    nc-sa/3.0/

#This code was tested on an Ubuntu 9.10 distribution running R version 2.9.2

#To set up the environment, be sure that the following files are located in the same
     directory as your working directory:
#-si_dataset_occurrence.csv
#-si_dataset_location.csv
#-si_appendix_map.shp

#and that you have the following packages installed:
#-fossil
#-shapefiles
#-PBSmapping
#-proj4
#-ecodist

#loading the necessary environment and datasets
library(fossil)
library(shapefiles)
library(PBSmapping)
library(proj4)
library(ecodist)

#bdiv is the name of the file that is directly taken from the dinosauria
bdiv<-read.csv("si_dataset_occurrence.csv")

################################################
#########IMPORTANT##############################
################################################
#This line makes it so only specimens are used
bdiv<-bdiv[bdiv[,4]=='specimens',]

# bdiv.mat is the species occurrence matrix (ie species X locality)
bdiv.mat<-create.matrix(bdiv,locality='formation', abund=T)

#reading in the locality lat/long
bdiv.lats<-read.csv('si_dataset_location.csv', header=TRUE, row.names=1)

#create the distance matrix, minimum spanning tree and nmds for later
bdiv.dist<-dino.dist(bdiv.mat)

##removing bonebed counts
bdiv.mat["Edmontosaurus","Hell Creek South Dakota"]<-1
```

```
bdiv.mat["Edmontosaurus","Lance␣Wyoming"]<-11

############################################################
#the randomizations can be set to a low number to speed up processing, but for nicer
     graphs a higher number should be used
rand<-1000

##calculating spp.est curves for all sites >100 specimens
hcm.spp.est<-spp.est(bdiv.mat[,"Hell␣Creek␣Montana"], rand, counter=F)
hcnd.spp.est<-spp.est(bdiv.mat[,"Hell␣Creek␣North␣Dakota"], rand, counter=F)
hca.spp.est<-spp.est(bdiv.mat[,"Horseshoe␣Canyon␣Alberta"], rand, counter=F)
lw.spp.est<-spp.est(bdiv.mat[,"Lance␣Wyoming"], rand, counter=F)
bdiv.spp.est<-spp.est(rowSums(bdiv.mat), rand, counter=F)
total.localities<-dim(bdiv.mat)[2]
total.specimens<-dim(bdiv.spp.est)[1]
total.species<-bdiv.spp.est[total.specimens,2]

tmp<-bdiv.mat
tmp[tmp>0]<-1
observed.beta<-round(total.species/mean(colSums(tmp)), 2)

csums<-colSums(bdiv.mat)

total.greater.than.10.specimens<-length(csums[csums>10])

localities<-c("Hell␣Creek␣Montana","Hell␣Creek␣North␣Dakota","Horseshoe␣Canyon␣
    Alberta","Lance␣Wyoming")

min.loc<-min(colSums(bdiv.mat[,localities]))
min.loc.name<-names(which.min(colSums(bdiv.mat[,localities])))

max.loc<-max(colSums(bdiv.mat[,localities]))
max.loc.name<-names(which.max(colSums(bdiv.mat[,localities])))

# the following few lines calculate the linear model (regression) for the number of
    species observed versus the number of
# specimens. A graph of this is also included in the supplementary material, but for
     this portion of the paper just the
# values from the regression are used.
a<-log(colSums(bdiv.mat))
b<-log(apply(bdiv.mat,2,function(x) length(x[x>0])))
lmfig2<-lm(b~a)
r2.value<-round(summary(lmfig2)$r.squared, 2)
prob<-round(pf(summary(lmfig2)$fstatistic[1], summary(lmfig2)$fstatistic[2],
summary(lmfig2)$fstatistic[3], lower.tail = FALSE), 3)
if (prob==0) prob<-0.001


###################################################
### chunk number 2: tab1
###################################################
sptab<-matrix(,7,8)
cols<-c(2,5,8,11)

sptab[1,1]<-dim(hcm.spp.est)[1]
sptab[1,2]<-hcm.spp.est[dim(hcm.spp.est)[1],2]
sptab[1,3:6]<-hcm.spp.est[100,cols]
```

```
sptab[2,1]<-dim(hcnd.spp.est)[1]
sptab[2,2]<-hcnd.spp.est[dim(hcnd.spp.est)[1],2]
sptab[2,3:6]<-hcnd.spp.est[100,cols]

sptab[3,1]<-dim(hca.spp.est)[1]
sptab[3,2]<-hca.spp.est[dim(hca.spp.est)[1],2]
sptab[3,3:6]<-hca.spp.est[100,cols]

sptab[4,1]<-dim(lw.spp.est)[1]
sptab[4,2]<-lw.spp.est[dim(lw.spp.est)[1],2]
sptab[4,3:6]<-lw.spp.est[82,cols]

sptab[5,1]<-total.specimens
sptab[5,2]<-total.species
sptab[5,3:6]<-bdiv.spp.est[100,cols]

for (i in 1:5) sptab[i,7]<-mean(sptab[i,4:6])
for (i in 1:7) sptab[6,i]<-mean(sptab[1:4,i])
for (i in 2:7) sptab[7,i]<-sptab[5,i]/mean(sptab[1:4,i])

sptab<-round(sptab,2)
```

Appendix F: Dinosaur Locality Data for Chapter 5

```
,lat,long
Aguja Texas,30.75,-102.75
Denver Colorado,39,-105
Ferris Wyoming,42,-106
Frenchman Saskatchewan,50,-108
Hell Creek Montana,47,-106.5
Hell Creek North Dakota,46.5,-103
Hell Creek South Dakota,45,-103
Horseshoe Canyon Alberta,51,-113
Javelina Texas,30.25,-103.75
Kaiparowits Utah,38,-112
Kirtland New Mexico,35.5,-107.5
Lance Montana,48.5,-106
Lance South Dakota,45.5,-101
Lance Utah,37.1,-111
Lance Wyoming,43,-105
Laramie Colorado,40.5,-104.5
Laramie Wyoming,44.5,-110.5
McRae New Mexico,33,-107
North Horn Utah,39,-111
Pinyon Conglomerate Wyoming,43.5,-109.5
Scollard Alberta,50.5,-111.75
St. Mary River Alberta,50,-113
St. Mary River Montana,48.5,-112.5
Tornillo Texas,29.5,-103
```

Appendix G: Dinosaur Occurrence Data for Chapter 5

```
"locality","genus","abundance","occurrences.abund_unit","collections.state","
    collections.stage","formation␣minus␣state","formation"
25073,"Chasmosaurus",2,"individuals","Texas","Campanian","Aguja","Aguja␣Texas"
22709,"Chasmosaurus",10,"specimens","Texas","Maastrichtian","Aguja","Aguja␣Texas"
68003,"Richardoestesia",1,"specimens","Texas","Campanian","Aguja","Aguja␣Texas"
68006,"Richardoestesia",1,"specimens","Texas","Campanian","Aguja","Aguja␣Texas"
68007,"Richardoestesia",1,"specimens","Texas","Maastrichtian","Aguja","Aguja␣Texas"
68001,"Richardoestesia",2,"specimens","Texas","Campanian","Aguja","Aguja␣Texas"
68008,"Richardoestesia",2,"specimens","Texas","Maastrichtian","Aguja","Aguja␣Texas"
68003,"Richardoestesia",3,"specimens","Texas","Campanian","Aguja","Aguja␣Texas"
68002,"Saurornitholestes",1,"specimens","Texas","Campanian","Aguja","Aguja␣Texas"
68003,"Saurornitholestes",1,"specimens","Texas","Campanian","Aguja","Aguja␣Texas"
68003,"Saurornitholestes",1,"specimens","Texas","Campanian","Aguja","Aguja␣Texas"
68006,"Saurornitholestes",1,"specimens","Texas","Campanian","Aguja","Aguja␣Texas"
68014,"Saurornitholestes",1,"specimens","Texas","Campanian","Aguja","Aguja␣Texas"
68004,"Saurornitholestes",2,"specimens","Texas","Campanian","Aguja","Aguja␣Texas"
68014,"Saurornitholestes",2,"specimens","Texas","Campanian","Aguja","Aguja␣Texas"
68007,"Saurornitholestes",3,"specimens","Texas","Maastrichtian","Aguja","Aguja␣Texas
    "
68012,"Saurornitholestes",4,"specimens","Texas","Campanian","Aguja","Aguja␣Texas"
68001,"Saurornitholestes",5,"specimens","Texas","Campanian","Aguja","Aguja␣Texas"
68007,"Saurornitholestes",6,"specimens","Texas","Maastrichtian","Aguja","Aguja␣Texas
    "
68008,"Saurornitholestes",9,"specimens","Texas","Maastrichtian","Aguja","Aguja␣Texas
    "
49000,"Ornithomimus",1,"individuals","Colorado","Maastrichtian","Denver","Denver␣
    Colorado"
46139,"Triceratops",1,"specimens","Colorado","Maastrichtian","Denver","Denver␣
    Colorado"
73927,"Triceratops",1,"individuals","Colorado","Maastrichtian","Denver","Denver␣
    Colorado"
49523,"Tyrannosaurus",1,"specimens","Colorado","Maastrichtian","Denver","Denver␣
    Colorado"
49538,"Tyrannosaurus",1,"individuals","Colorado","Maastrichtian","Denver","Denver␣
    Colorado"
14541,"Ankylosaurus",1,"specimens","Wyoming","Maastrichtian","Ferris","Ferris␣
    Wyoming"
14544,"Ankylosaurus",1,"specimens","Wyoming","Maastrichtian","Ferris","Ferris␣
    Wyoming"
75289,"Ankylosaurus",1,"specimens","Wyoming","Maastrichtian","Ferris","Ferris␣
    Wyoming"
75430,"Ankylosaurus",1,"specimens","Wyoming","Maastrichtian","Ferris","Ferris␣
    Wyoming"
75435,"Ankylosaurus",1,"specimens","Wyoming","Maastrichtian","Ferris","Ferris␣
    Wyoming"
14543,"Dromaeosaurus",1,"specimens","Wyoming","Maastrichtian","Ferris","Ferris␣
    Wyoming"
75280,"Dromaeosaurus",1,"specimens","Wyoming","Maastrichtian","Ferris","Ferris␣
    Wyoming"
75443,"Dromaeosaurus",1,"specimens","Wyoming","Maastrichtian","Ferris","Ferris␣
    Wyoming"
14542,"Edmontonia",1,"specimens","Wyoming","Maastrichtian","Ferris","Ferris␣Wyoming"
```

```
75287,"Edmontonia",1,"specimens","Wyoming","Maastrichtian","Ferris","Ferris␣Wyoming"
75298,"Edmontonia",1,"specimens","Wyoming","Maastrichtian","Ferris","Ferris␣Wyoming"
75443,"Edmontonia",1,"specimens","Wyoming","Maastrichtian","Ferris","Ferris␣Wyoming"
14796,"Ornithomimus",1,"specimens","Wyoming","Danian","Ferris","Ferris␣Wyoming"
14542,"Ornithomimus",2,"specimens","Wyoming","Maastrichtian","Ferris","Ferris␣
    Wyoming"
14543,"Richardoestesia",1,"specimens","Wyoming","Maastrichtian","Ferris","Ferris␣
    Wyoming"
70343,"Richardoestesia",1,"specimens","Wyoming","Maastrichtian","Ferris","Ferris␣
    Wyoming"
75432,"Richardoestesia",1,"specimens","Wyoming","Maastrichtian","Ferris","Ferris␣
    Wyoming"
75287,"Richardoestesia",2,"specimens","Wyoming","Maastrichtian","Ferris","Ferris␣
    Wyoming"
14541,"Saurornitholestes",1,"specimens","Wyoming","Maastrichtian","Ferris","Ferris␣
    Wyoming"
14543,"Saurornitholestes",1,"specimens","Wyoming","Maastrichtian","Ferris","Ferris␣
    Wyoming"
14544,"Saurornitholestes",1,"specimens","Wyoming","Maastrichtian","Ferris","Ferris␣
    Wyoming"
70343,"Saurornitholestes",1,"specimens","Wyoming","Maastrichtian","Ferris","Ferris␣
    Wyoming"
75280,"Saurornitholestes",1,"specimens","Wyoming","Maastrichtian","Ferris","Ferris␣
    Wyoming"
75296,"Saurornitholestes",1,"specimens","Wyoming","Maastrichtian","Ferris","Ferris␣
    Wyoming"
75297,"Saurornitholestes",1,"specimens","Wyoming","Maastrichtian","Ferris","Ferris␣
    Wyoming"
75432,"Saurornitholestes",1,"specimens","Wyoming","Maastrichtian","Ferris","Ferris␣
    Wyoming"
75287,"Saurornitholestes",2,"specimens","Wyoming","Maastrichtian","Ferris","Ferris␣
    Wyoming"
75429,"Saurornitholestes",2,"specimens","Wyoming","Maastrichtian","Ferris","Ferris␣
    Wyoming"
75431,"Saurornitholestes",2,"specimens","Wyoming","Maastrichtian","Ferris","Ferris␣
    Wyoming"
14542,"Saurornitholestes",3,"specimens","Wyoming","Maastrichtian","Ferris","Ferris␣
    Wyoming"
75447,"Saurornitholestes",3,"specimens","Wyoming","Maastrichtian","Ferris","Ferris␣
    Wyoming"
75297,"Struthiomimus",1,"specimens","Wyoming","Maastrichtian","Ferris","Ferris␣
    Wyoming"
14544,"Stygimoloch",1,"specimens","Wyoming","Maastrichtian","Ferris","Ferris␣Wyoming
    "
14542,"Troodon",1,"specimens","Wyoming","Maastrichtian","Ferris","Ferris␣Wyoming"
14543,"Troodon",1,"specimens","Wyoming","Maastrichtian","Ferris","Ferris␣Wyoming"
75298,"Troodon",1,"specimens","Wyoming","Maastrichtian","Ferris","Ferris␣Wyoming"
75435,"Troodon",1,"specimens","Wyoming","Maastrichtian","Ferris","Ferris␣Wyoming"
75442,"Troodon",1,"specimens","Wyoming","Maastrichtian","Ferris","Ferris␣Wyoming"
75284,"Tyrannosaurus",1,"specimens","Wyoming","Maastrichtian","Ferris","Ferris␣
    Wyoming"
75286,"Tyrannosaurus",1,"specimens","Wyoming","Maastrichtian","Ferris","Ferris␣
    Wyoming"
75432,"Tyrannosaurus",1,"specimens","Wyoming","Maastrichtian","Ferris","Ferris␣
    Wyoming"
75446,"Tyrannosaurus",1,"specimens","Wyoming","Maastrichtian","Ferris","Ferris␣
    Wyoming"
```

```
54105,"Dromaeosaurus",3,"specimens","Saskatchewan","Maastrichtian","Frenchman","
    Frenchman␣Saskatchewan"
48628,"Thescelosaurus",1,"individuals","Saskatchewan","Maastrichtian","Frenchman","
    Frenchman␣Saskatchewan"
54105,"Triceratops",1,"specimens","Saskatchewan","Maastrichtian","Frenchman","
    Frenchman␣Saskatchewan"
54105,"Triceratops",20,"specimens","Saskatchewan","Maastrichtian","Frenchman","
    Frenchman␣Saskatchewan"
54105,"Tyrannosaurus",1,"specimens","Saskatchewan","Maastrichtian","Frenchman","
    Frenchman␣Saskatchewan"
11931,"Ankylosaurus",1,"specimens","Montana","Maastrichtian","Hell␣Creek","Hell␣
    Creek␣Montana"
14534,"Ankylosaurus",1,"specimens","Montana","Maastrichtian","Hell␣Creek","Hell␣
    Creek␣Montana"
45879,"Ankylosaurus",1,"specimens","Montana","Maastrichtian","Hell␣Creek","Hell␣
    Creek␣Montana"
14515,"Avisaurus",1,"specimens","Montana","Maastrichtian","Hell␣Creek","Hell␣Creek␣
    Montana"
14507,"Chirostenotes",1,"specimens","Montana","Maastrichtian","Hell␣Creek","Hell␣
    Creek␣Montana"
14533,"Chirostenotes",1,"specimens","Montana","Maastrichtian","Hell␣Creek","Hell␣
    Creek␣Montana"
14534,"Chirostenotes",1,"specimens","Montana","Maastrichtian","Hell␣Creek","Hell␣
    Creek␣Montana"
14535,"Chirostenotes",1,"specimens","Montana","Maastrichtian","Hell␣Creek","Hell␣
    Creek␣Montana"
14536,"Chirostenotes",1,"specimens","Montana","Maastrichtian","Hell␣Creek","Hell␣
    Creek␣Montana"
14548,"Chirostenotes",1,"specimens","Montana","Maastrichtian","Hell␣Creek","Hell␣
    Creek␣Montana"
14560,"Chirostenotes",1,"specimens","Montana","Maastrichtian","Hell␣Creek","Hell␣
    Creek␣Montana"
14562,"Chirostenotes",1,"specimens","Montana","Maastrichtian","Hell␣Creek","Hell␣
    Creek␣Montana"
14574,"Chirostenotes",1,"specimens","Montana","Maastrichtian","Hell␣Creek","Hell␣
    Creek␣Montana"
14639,"Chirostenotes",1,"specimens","Montana","Maastrichtian","Hell␣Creek","Hell␣
    Creek␣Montana"
59374,"Chirostenotes",1,"specimens","Montana","Maastrichtian","Hell␣Creek","Hell␣
    Creek␣Montana"
14582,"Chirostenotes",2,"specimens","Montana","Maastrichtian","Hell␣Creek","Hell␣
    Creek␣Montana"
14533,"Coelurosauria",1,"specimens","Montana","Maastrichtian","Hell␣Creek","Hell␣
    Creek␣Montana"
14535,"Coelurosauria",1,"specimens","Montana","Maastrichtian","Hell␣Creek","Hell␣
    Creek␣Montana"
14536,"Coelurosauria",1,"specimens","Montana","Maastrichtian","Hell␣Creek","Hell␣
    Creek␣Montana"
14582,"Edmontonia",1,"specimens","Montana","Maastrichtian","Hell␣Creek","Hell␣Creek␣
    Montana"
14546,"Edmontosaurus",1,"specimens","Montana","Maastrichtian","Hell␣Creek","Hell␣
    Creek␣Montana"
14549,"Edmontosaurus",1,"specimens","Montana","Maastrichtian","Hell␣Creek","Hell␣
    Creek␣Montana"
48836,"Edmontosaurus",1,"individuals","Montana","Maastrichtian","Hell␣Creek","Hell␣
    Creek␣Montana"
50088,"Edmontosaurus",1,"specimens","Montana","Maastrichtian","Hell␣Creek","Hell␣
    Creek␣Montana"
```

```
55385,"Edmontosaurus",1,"specimens","Montana","Maastrichtian","Hell␣Creek","Hell␣
    Creek␣Montana"
57914,"Edmontosaurus",1,"individuals","Montana","Maastrichtian","Hell␣Creek","Hell␣
    Creek␣Montana"
67129,"Edmontosaurus",1,"individuals","Montana","Maastrichtian","Hell␣Creek","Hell␣
    Creek␣Montana"
74009,"Edmontosaurus",1,"individuals","Montana","Maastrichtian","Hell␣Creek","Hell␣
    Creek␣Montana"
14674,"Edmontosaurus",2,"individuals","Montana","Maastrichtian","Hell␣Creek","Hell␣
    Creek␣Montana"
73947,"Leptoceratops",1,"individuals","Montana","Maastrichtian","Hell␣Creek","Hell␣
    Creek␣Montana"
73948,"Leptoceratops",1,"specimens","Montana","Maastrichtian","Hell␣Creek","Hell␣
    Creek␣Montana"
54002,"Nanotyrannus",1,"individuals","Montana","Maastrichtian","Hell␣Creek","Hell␣
    Creek␣Montana"
14674,"Nanotyrannus",5,"specimens","Montana","Maastrichtian","Hell␣Creek","Hell␣
    Creek␣Montana"
14616,"Pachycephalosaurus",1,"specimens","Montana","Maastrichtian","Hell␣Creek","
    Hell␣Creek␣Montana"
47627,"Pachycephalosaurus",1,"individuals","Montana","Maastrichtian","Hell␣Creek","
    Hell␣Creek␣Montana"
52904,"Pachycephalosaurus",1,"specimens","Montana","Maastrichtian","Hell␣Creek","
    Hell␣Creek␣Montana"
14508,"Richardoestesia",1,"specimens","Montana","Maastrichtian","Hell␣Creek","Hell␣
    Creek␣Montana"
14521,"Richardoestesia",1,"specimens","Montana","Maastrichtian","Hell␣Creek","Hell␣
    Creek␣Montana"
14525,"Richardoestesia",1,"specimens","Montana","Maastrichtian","Hell␣Creek","Hell␣
    Creek␣Montana"
14546,"Richardoestesia",1,"specimens","Montana","Maastrichtian","Hell␣Creek","Hell␣
    Creek␣Montana"
14548,"Richardoestesia",1,"specimens","Montana","Maastrichtian","Hell␣Creek","Hell␣
    Creek␣Montana"
14580,"Richardoestesia",1,,"Montana","Maastrichtian","Hell␣Creek","Hell␣Creek␣
    Montana"
14582,"Richardoestesia",1,"specimens","Montana","Maastrichtian","Hell␣Creek","Hell␣
    Creek␣Montana"
14605,"Richardoestesia",1,"specimens","Montana","Maastrichtian","Hell␣Creek","Hell␣
    Creek␣Montana"
14616,"Richardoestesia",1,"specimens","Montana","Maastrichtian","Hell␣Creek","Hell␣
    Creek␣Montana"
14639,"Richardoestesia",1,"specimens","Montana","Maastrichtian","Hell␣Creek","Hell␣
    Creek␣Montana"
14650,"Richardoestesia",1,"specimens","Montana","Maastrichtian","Hell␣Creek","Hell␣
    Creek␣Montana"
14653,"Richardoestesia",1,"specimens","Montana","Maastrichtian","Hell␣Creek","Hell␣
    Creek␣Montana"
14647,"Richardoestesia",2,"specimens","Montana","Maastrichtian","Hell␣Creek","Hell␣
    Creek␣Montana"
14624,"Richardoestesia",4,"specimens","Montana","Maastrichtian","Hell␣Creek","Hell␣
    Creek␣Montana"
14535,"Richardoestesia",5,"specimens","Montana","Maastrichtian","Hell␣Creek","Hell␣
    Creek␣Montana"
14549,"Richardoestesia",5,"specimens","Montana","Maastrichtian","Hell␣Creek","Hell␣
    Creek␣Montana"
14568,"Richardoestesia",5,"specimens","Montana","Maastrichtian","Hell␣Creek","Hell␣
    Creek␣Montana"
```

```
14549,"Richardoestesia",74,"specimens","Montana","Maastrichtian","Hell␣Creek","Hell␣Creek␣Montana"
53516,"Sphaerotholus",1,,"Montana","Maastrichtian","Hell␣Creek","Hell␣Creek␣Montana"
52903,"Stegoceras",1,"specimens","Montana","Maastrichtian","Hell␣Creek","Hell␣Creek␣Montana"
72143,"Stegoceras",1,"specimens","Montana","Maastrichtian","Hell␣Creek","Hell␣Creek␣Montana"
12977,"Stygimoloch",1,"specimens","Montana","Maastrichtian","Hell␣Creek","Hell␣Creek␣Montana"
14550,"Stygimoloch",1,"specimens","Montana","Maastrichtian","Hell␣Creek","Hell␣Creek␣Montana"
47626,"Stygimoloch",1,"specimens","Montana","Maastrichtian","Hell␣Creek","Hell␣Creek␣Montana"
48844,"Stygimoloch",1,"specimens","Montana","Maastrichtian","Hell␣Creek","Hell␣Creek␣Montana"
48845,"Stygimoloch",1,"specimens","Montana","Maastrichtian","Hell␣Creek","Hell␣Creek␣Montana"
48846,"Stygimoloch",1,"specimens","Montana","Maastrichtian","Hell␣Creek","Hell␣Creek␣Montana"
48848,"Stygimoloch",1,"individuals","Montana","Maastrichtian","Hell␣Creek","Hell␣Creek␣Montana"
14535,"Thescelosaurus",1,"specimens","Montana","Maastrichtian","Hell␣Creek","Hell␣Creek␣Montana"
14549,"Thescelosaurus",1,"specimens","Montana","Maastrichtian","Hell␣Creek","Hell␣Creek␣Montana"
14574,"Thescelosaurus",1,"specimens","Montana","Maastrichtian","Hell␣Creek","Hell␣Creek␣Montana"
14582,"Thescelosaurus",1,"specimens","Montana","Maastrichtian","Hell␣Creek","Hell␣Creek␣Montana"
14654,"Thescelosaurus",1,"specimens","Montana","Maastrichtian","Hell␣Creek","Hell␣Creek␣Montana"
48622,"Thescelosaurus",1,"individuals","Montana","Maastrichtian","Hell␣Creek","Hell␣Creek␣Montana"
48622,"Thescelosaurus",1,"individuals","Montana","Maastrichtian","Hell␣Creek","Hell␣Creek␣Montana"
48629,"Thescelosaurus",1,"specimens","Montana","Maastrichtian","Hell␣Creek","Hell␣Creek␣Montana"
54080,"Thescelosaurus",1,"individuals","Montana","Maastrichtian","Hell␣Creek","Hell␣Creek␣Montana"
54081,"Thescelosaurus",1,"individuals","Montana","Maastrichtian","Hell␣Creek","Hell␣Creek␣Montana"
14639,"Thescelosaurus",2,"specimens","Montana","Maastrichtian","Hell␣Creek","Hell␣Creek␣Montana"
14674,"Thescelosaurus",4,"specimens","Montana","Maastrichtian","Hell␣Creek","Hell␣Creek␣Montana"
14615,"Triceratops",1,"specimens","Montana","Maastrichtian","Hell␣Creek","Hell␣Creek␣Montana"
14649,"Triceratops",1,"specimens","Montana","Maastrichtian","Hell␣Creek","Hell␣Creek␣Montana"
35510,"Triceratops",1,"individuals","Montana","Maastrichtian","Hell␣Creek","Hell␣Creek␣Montana"
55353,"Triceratops",1,"individuals","Montana","Maastrichtian","Hell␣Creek","Hell␣Creek␣Montana"
55354,"Triceratops",1,"individuals","Montana","Maastrichtian","Hell␣Creek","Hell␣Creek␣Montana"
55355,"Triceratops",1,"individuals","Montana","Maastrichtian","Hell␣Creek","Hell␣Creek␣Montana"
```

```
64427,"Triceratops",1,"individuals","Montana","Maastrichtian","Hell␣Creek","Hell␣
    Creek␣Montana"
64428,"Triceratops",1,"individuals","Montana","Maastrichtian","Hell␣Creek","Hell␣
    Creek␣Montana"
64429,"Triceratops",1,"individuals","Montana","Maastrichtian","Hell␣Creek","Hell␣
    Creek␣Montana"
14550,"Triceratops",2,"specimens","Montana","Maastrichtian","Hell␣Creek","Hell␣Creek
    ␣Montana"
67129,"Triceratops",2,"individuals","Montana","Maastrichtian","Hell␣Creek","Hell␣
    Creek␣Montana"
14616,"Triceratops",3,"specimens","Montana","Maastrichtian","Hell␣Creek","Hell␣Creek
    ␣Montana"
49016,"Triceratops",3,"individuals","Montana","Maastrichtian","Hell␣Creek","Hell␣
    Creek␣Montana"
14521,"Troodon",1,"specimens","Montana","Maastrichtian","Hell␣Creek","Hell␣Creek␣
    Montana"
14568,"Troodon",1,"specimens","Montana","Maastrichtian","Hell␣Creek","Hell␣Creek␣
    Montana"
14582,"Troodon",1,"specimens","Montana","Maastrichtian","Hell␣Creek","Hell␣Creek␣
    Montana"
14650,"Troodon",1,"specimens","Montana","Maastrichtian","Hell␣Creek","Hell␣Creek␣
    Montana"
14549,"Troodon",3,"specimens","Montana","Maastrichtian","Hell␣Creek","Hell␣Creek␣
    Montana"
38544,"Tyrannosaurus",1,"specimens","Montana","Maastrichtian","Hell␣Creek","Hell␣
    Creek␣Montana"
48847,"Tyrannosaurus",1,"individuals","Montana","Maastrichtian","Hell␣Creek","Hell␣
    Creek␣Montana"
48848,"Tyrannosaurus",1,"individuals","Montana","Maastrichtian","Hell␣Creek","Hell␣
    Creek␣Montana"
60766,"Tyrannosaurus",1,"individuals","Montana","Maastrichtian","Hell␣Creek","Hell␣
    Creek␣Montana"
60766,"Tyrannosaurus",1,"specimens","Montana","Maastrichtian","Hell␣Creek","Hell␣
    Creek␣Montana"
14674,"Tyrannosaurus",5,"specimens","Montana","Maastrichtian","Hell␣Creek","Hell␣
    Creek␣Montana"
14610,"Avisaurus",1,"specimens","North␣Dakota","Maastrichtian","Hell␣Creek","Hell␣
    Creek␣North␣Dakota"
45098,"Richardoestesia",1,"specimens","North␣Dakota","Maastrichtian","Hell␣Creek","
    Hell␣Creek␣North␣Dakota"
45112,"Richardoestesia",1,"specimens","North␣Dakota","Maastrichtian","Hell␣Creek","
    Hell␣Creek␣North␣Dakota"
45113,"Richardoestesia",1,"specimens","North␣Dakota","Maastrichtian","Hell␣Creek","
    Hell␣Creek␣North␣Dakota"
45122,"Richardoestesia",1,"specimens","North␣Dakota","Maastrichtian","Hell␣Creek","
    Hell␣Creek␣North␣Dakota"
45136,"Richardoestesia",1,"specimens","North␣Dakota","Maastrichtian","Hell␣Creek","
    Hell␣Creek␣North␣Dakota"
45110,"Richardoestesia",2,"specimens","North␣Dakota","Maastrichtian","Hell␣Creek","
    Hell␣Creek␣North␣Dakota"
45114,"Richardoestesia",2,"specimens","North␣Dakota","Maastrichtian","Hell␣Creek","
    Hell␣Creek␣North␣Dakota"
45102,"Richardoestesia",3,"specimens","North␣Dakota","Maastrichtian","Hell␣Creek","
    Hell␣Creek␣North␣Dakota"
45120,"Richardoestesia",3,"specimens","North␣Dakota","Maastrichtian","Hell␣Creek","
    Hell␣Creek␣North␣Dakota"
45130,"Richardoestesia",3,,"North␣Dakota","Maastrichtian","Hell␣Creek","Hell␣Creek␣
    North␣Dakota"
```

```
45299,"Richardoestesia",3,"specimens","North␣Dakota","Maastrichtian","Hell␣Creek","
    Hell␣Creek␣North␣Dakota"
45118,"Richardoestesia",6,"specimens","North␣Dakota","Maastrichtian","Hell␣Creek","
    Hell␣Creek␣North␣Dakota"
14610,"Richardoestesia",8,"specimens","North␣Dakota","Maastrichtian","Hell␣Creek","
    Hell␣Creek␣North␣Dakota"
45109,"Richardoestesia",9,"specimens","North␣Dakota","Maastrichtian","Hell␣Creek","
    Hell␣Creek␣North␣Dakota"
45111,"Richardoestesia",9,"specimens","North␣Dakota","Maastrichtian","Hell␣Creek","
    Hell␣Creek␣North␣Dakota"
45099,"Richardoestesia",10,"specimens","North␣Dakota","Maastrichtian","Hell␣Creek","
    Hell␣Creek␣North␣Dakota"
45144,"Richardoestesia",11,"specimens","North␣Dakota","Maastrichtian","Hell␣Creek","
    Hell␣Creek␣North␣Dakota"
45098,"Saurornitholestes",1,"specimens","North␣Dakota","Maastrichtian","Hell␣Creek",
    "Hell␣Creek␣North␣Dakota"
45102,"Saurornitholestes",1,"specimens","North␣Dakota","Maastrichtian","Hell␣Creek",
    "Hell␣Creek␣North␣Dakota"
45145,"Saurornitholestes",1,"specimens","North␣Dakota","Maastrichtian","Hell␣Creek",
    "Hell␣Creek␣North␣Dakota"
45111,"Saurornitholestes",2,"specimens","North␣Dakota","Maastrichtian","Hell␣Creek",
    "Hell␣Creek␣North␣Dakota"
45128,"Saurornitholestes",2,"specimens","North␣Dakota","Maastrichtian","Hell␣Creek",
    "Hell␣Creek␣North␣Dakota"
45118,"Saurornitholestes",3,"specimens","North␣Dakota","Maastrichtian","Hell␣Creek",
    "Hell␣Creek␣North␣Dakota"
45099,"Saurornitholestes",10,"specimens","North␣Dakota","Maastrichtian","Hell␣Creek"
    ,"Hell␣Creek␣North␣Dakota"
14610,"Saurornitholestes",14,"specimens","North␣Dakota","Maastrichtian","Hell␣Creek"
    ,"Hell␣Creek␣North␣Dakota"
48843,"Stygimoloch",1,"specimens","North␣Dakota","Maastrichtian","Hell␣Creek","Hell␣
    Creek␣North␣Dakota"
45098,"Thescelosaurus",1,"specimens","North␣Dakota","Maastrichtian","Hell␣Creek","
    Hell␣Creek␣North␣Dakota"
45102,"Thescelosaurus",1,"specimens","North␣Dakota","Maastrichtian","Hell␣Creek","
    Hell␣Creek␣North␣Dakota"
45136,"Thescelosaurus",1,"specimens","North␣Dakota","Maastrichtian","Hell␣Creek","
    Hell␣Creek␣North␣Dakota"
45128,"Thescelosaurus",2,"specimens","North␣Dakota","Maastrichtian","Hell␣Creek","
    Hell␣Creek␣North␣Dakota"
45144,"Thescelosaurus",2,"specimens","North␣Dakota","Maastrichtian","Hell␣Creek","
    Hell␣Creek␣North␣Dakota"
45099,"Thescelosaurus",12,"specimens","North␣Dakota","Maastrichtian","Hell␣Creek","
    Hell␣Creek␣North␣Dakota"
14610,"Thescelosaurus",28,"specimens","North␣Dakota","Maastrichtian","Hell␣Creek","
    Hell␣Creek␣North␣Dakota"
45103,"Torosaurus",1,"specimens","North␣Dakota","Maastrichtian","Hell␣Creek","Hell␣
    Creek␣North␣Dakota"
24851,"Triceratops",1,"individuals","North␣Dakota","Maastrichtian","Hell␣Creek","
    Hell␣Creek␣North␣Dakota"
45123,"Triceratops",1,"specimens","North␣Dakota","Maastrichtian","Hell␣Creek","Hell␣
    Creek␣North␣Dakota"
45099,"Troodon",2,"specimens","North␣Dakota","Maastrichtian","Hell␣Creek","Hell␣
    Creek␣North␣Dakota"
14610,"Troodon",3,"specimens","North␣Dakota","Maastrichtian","Hell␣Creek","Hell␣
    Creek␣North␣Dakota"
45098,"Tyrannosaurus",1,"specimens","North␣Dakota","Maastrichtian","Hell␣Creek","
    Hell␣Creek␣North␣Dakota"
```

```
45112,"Tyrannosaurus",1,"specimens","North␣Dakota","Maastrichtian","Hell␣Creek","
    Hell␣Creek␣North␣Dakota"
45114,"Tyrannosaurus",1,"specimens","North␣Dakota","Maastrichtian","Hell␣Creek","
    Hell␣Creek␣North␣Dakota"
45126,"Tyrannosaurus",1,"individuals","North␣Dakota","Maastrichtian","Hell␣Creek","
    Hell␣Creek␣North␣Dakota"
45138,"Tyrannosaurus",1,"specimens","North␣Dakota","Maastrichtian","Hell␣Creek","
    Hell␣Creek␣North␣Dakota"
45141,"Tyrannosaurus",1,"specimens","North␣Dakota","Maastrichtian","Hell␣Creek","
    Hell␣Creek␣North␣Dakota"
45146,"Tyrannosaurus",1,"specimens","North␣Dakota","Maastrichtian","Hell␣Creek","
    Hell␣Creek␣North␣Dakota"
45312,"Tyrannosaurus",1,"specimens","North␣Dakota","Maastrichtian","Hell␣Creek","
    Hell␣Creek␣North␣Dakota"
45381,"Tyrannosaurus",1,"specimens","North␣Dakota","Maastrichtian","Hell␣Creek","
    Hell␣Creek␣North␣Dakota"
45299,"Tyrannosaurus",2,"specimens","North␣Dakota","Maastrichtian","Hell␣Creek","
    Hell␣Creek␣North␣Dakota"
45118,"Tyrannosaurus",3,"specimens","North␣Dakota","Maastrichtian","Hell␣Creek","
    Hell␣Creek␣North␣Dakota"
45102,"Tyrannosaurus",4,"specimens","North␣Dakota","Maastrichtian","Hell␣Creek","
    Hell␣Creek␣North␣Dakota"
45128,"Tyrannosaurus",4,"specimens","North␣Dakota","Maastrichtian","Hell␣Creek","
    Hell␣Creek␣North␣Dakota"
45136,"Tyrannosaurus",4,"specimens","North␣Dakota","Maastrichtian","Hell␣Creek","
    Hell␣Creek␣North␣Dakota"
45111,"Tyrannosaurus",6,"specimens","North␣Dakota","Maastrichtian","Hell␣Creek","
    Hell␣Creek␣North␣Dakota"
45109,"Tyrannosaurus",7,"specimens","North␣Dakota","Maastrichtian","Hell␣Creek","
    Hell␣Creek␣North␣Dakota"
45145,"Tyrannosaurus",11,"specimens","North␣Dakota","Maastrichtian","Hell␣Creek","
    Hell␣Creek␣North␣Dakota"
14610,"Tyrannosaurus",15,"specimens","North␣Dakota","Maastrichtian","Hell␣Creek","
    Hell␣Creek␣North␣Dakota"
45099,"Tyrannosaurus",17,"specimens","North␣Dakota","Maastrichtian","Hell␣Creek","
    Hell␣Creek␣North␣Dakota"
45144,"Tyrannosaurus",21,"specimens","North␣Dakota","Maastrichtian","Hell␣Creek","
    Hell␣Creek␣North␣Dakota"
58813,"Dracorex",1,"individuals","South␣Dakota","Maastrichtian","Hell␣Creek","Hell␣
    Creek␣South␣Dakota"
24790,"Edmontosaurus",1,"specimens","South␣Dakota","Maastrichtian","Hell␣Creek","
    Hell␣Creek␣South␣Dakota"
47034,"Edmontosaurus",5500,"specimens","South␣Dakota","Maastrichtian","Hell␣Creek","
    Hell␣Creek␣South␣Dakota"
45390,"Richardoestesia",1,"specimens","South␣Dakota","Maastrichtian","Hell␣Creek","
    Hell␣Creek␣South␣Dakota"
45391,"Richardoestesia",1,"specimens","South␣Dakota","Maastrichtian","Hell␣Creek","
    Hell␣Creek␣South␣Dakota"
45390,"Thescelosaurus",3,"specimens","South␣Dakota","Maastrichtian","Hell␣Creek","
    Hell␣Creek␣South␣Dakota"
53128,"Torosaurus",1,"individuals","South␣Dakota","Maastrichtian","Hell␣Creek","Hell
    ␣Creek␣South␣Dakota"
45388,"Tyrannosaurus",1,"specimens","South␣Dakota","Maastrichtian","Hell␣Creek","
    Hell␣Creek␣South␣Dakota"
45389,"Tyrannosaurus",3,"specimens","South␣Dakota","Maastrichtian","Hell␣Creek","
    Hell␣Creek␣South␣Dakota"
45390,"Tyrannosaurus",13,"specimens","South␣Dakota","Maastrichtian","Hell␣Creek","
    Hell␣Creek␣South␣Dakota"
```

```
11899,"Albertosaurus",1,"specimens","Alberta",,"Horseshoe␣Canyon","Horseshoe␣Canyon␣Alberta"
11900,"Albertosaurus",1,"specimens","Alberta",,"Horseshoe␣Canyon","Horseshoe␣Canyon␣Alberta"
11901,"Albertosaurus",1,"specimens","Alberta",,"Horseshoe␣Canyon","Horseshoe␣Canyon␣Alberta"
11902,"Albertosaurus",1,"specimens","Alberta",,"Horseshoe␣Canyon","Horseshoe␣Canyon␣Alberta"
55363,"Albertosaurus",8,"individuals","Alberta",,"Horseshoe␣Canyon","Horseshoe␣Canyon␣Alberta"
67612,"Arrhinoceratops",1,"individuals","Alberta",,"Horseshoe␣Canyon","Horseshoe␣Canyon␣Alberta"
11916,"Daspletosaurus",1,"specimens","Alberta",,"Horseshoe␣Canyon","Horseshoe␣Canyon␣Alberta"
75475,"Eotriceratops",1,"individuals","Alberta","Maastrichtian","Horseshoe␣Canyon","Horseshoe␣Canyon␣Alberta"
51966,"Euoplocephalus",1,"individuals","Alberta","Maastrichtian","Horseshoe␣Canyon","Horseshoe␣Canyon␣Alberta"
60805,"Euoplocephalus",1,"individuals","Alberta","Maastrichtian","Horseshoe␣Canyon","Horseshoe␣Canyon␣Alberta"
61628,"Euoplocephalus",1,"specimens","Alberta","Maastrichtian","Horseshoe␣Canyon","Horseshoe␣Canyon␣Alberta"
61629,"Euoplocephalus",1,"specimens","Alberta","Maastrichtian","Horseshoe␣Canyon","Horseshoe␣Canyon␣Alberta"
49079,"Hypacrosaurus",1,"individuals","Alberta",,"Horseshoe␣Canyon","Horseshoe␣Canyon␣Alberta"
51965,"Montanoceratops",1,"specimens","Alberta","Maastrichtian","Horseshoe␣Canyon","Horseshoe␣Canyon␣Alberta"
45752,"Ornithomimus",4,"specimens","Alberta","Maastrichtian","Horseshoe␣Canyon","Horseshoe␣Canyon␣Alberta"
45732,"Ornithomimus",1,"individuals","Alberta","Maastrichtian","Horseshoe␣Canyon","Horseshoe␣Canyon␣Alberta"
46351,"Ornithomimus",1,"individuals","Alberta","Maastrichtian","Horseshoe␣Canyon","Horseshoe␣Canyon␣Alberta"
52479,"Ornithomimus",1,"individuals","Alberta","Maastrichtian","Horseshoe␣Canyon","Horseshoe␣Canyon␣Alberta"
52480,"Ornithomimus",1,"individuals","Alberta","Maastrichtian","Horseshoe␣Canyon","Horseshoe␣Canyon␣Alberta"
34744,"Pachyrhinosaurus",1,"specimens","Alberta","Maastrichtian","Horseshoe␣Canyon","Horseshoe␣Canyon␣Alberta"
64352,"Pachyrhinosaurus",2,"specimens","Alberta",,"Horseshoe␣Canyon","Horseshoe␣Canyon␣Alberta"
48624,"Parksosaurus",1,"specimens","Alberta","Maastrichtian","Horseshoe␣Canyon","Horseshoe␣Canyon␣Alberta"
60563,"Richardoestesia",11,"specimens","Alberta","Maastrichtian","Horseshoe␣Canyon","Horseshoe␣Canyon␣Alberta"
76062,"Saurornitholestes",1,"specimens","Alberta","Maastrichtian","Horseshoe␣Canyon","Horseshoe␣Canyon␣Alberta"
60563,"Saurornitholestes",8,"specimens","Alberta","Maastrichtian","Horseshoe␣Canyon","Horseshoe␣Canyon␣Alberta"
52469,"Struthiomimus",1,"individuals","Alberta","Maastrichtian","Horseshoe␣Canyon","Horseshoe␣Canyon␣Alberta"
64324,"Troodon",1,"specimens","Alberta","Maastrichtian","Horseshoe␣Canyon","Horseshoe␣Canyon␣Alberta"
60563,"Troodon",65,"specimens","Alberta","Maastrichtian","Horseshoe␣Canyon","Horseshoe␣Canyon␣Alberta"
61847,"Alamosaurus",3,"individuals","Texas","Maastrichtian","Javelina","Javelina␣Texas"
```

```
75467,"Gryposaurus",1,"specimens","Utah","Campanian","Kaiparowits","Kaiparowits␣Utah
    "
58520,"Hagryphus",1,"individuals","Utah","Campanian","Kaiparowits","Kaiparowits␣Utah
    "
70276,"Ornithomimus",1,"individuals","Utah","Campanian","Kaiparowits","Kaiparowits␣
    Utah"
52060,"Parasaurolophus",1,"specimens","Utah","Maastrichtian","Kaiparowits","
    Kaiparowits␣Utah"
47010,"Alamosaurus",1,"specimens","New␣Mexico","Maastrichtian","Kirtland","Kirtland␣
    New␣Mexico"
54103,"Alamosaurus",1,"specimens","New␣Mexico","Maastrichtian","Kirtland","Kirtland␣
    New␣Mexico"
70335,"Alamosaurus",1,"specimens","New␣Mexico","Maastrichtian","Kirtland","Kirtland␣
    New␣Mexico"
46992,"Alamosaurus",8,"specimens","New␣Mexico","Maastrichtian","Kirtland","Kirtland␣
    New␣Mexico"
62535,"Ankylosaurus",1,"specimens","New␣Mexico","Maastrichtian","Kirtland","Kirtland
    ␣New␣Mexico"
53981,"Chasmosaurus",1,,"New␣Mexico",,"Kirtland","Kirtland␣New␣Mexico"
53514,"Daspletosaurus",1,"specimens","New␣Mexico","Maastrichtian","Kirtland","
    Kirtland␣New␣Mexico"
62530,"Monoclonius",1,"specimens","New␣Mexico","Maastrichtian","Kirtland","Kirtland␣
    New␣Mexico"
54104,"Nodocephalosaurus",1,"specimens","New␣Mexico","Campanian","Kirtland","
    Kirtland␣New␣Mexico"
48839,"Parasaurolophus",1,"specimens","New␣Mexico","Campanian","Kirtland","Kirtland␣
    New␣Mexico"
53899,"Parasaurolophus",1,"specimens","New␣Mexico","Campanian","Kirtland","Kirtland␣
    New␣Mexico"
62317,"Parasaurolophus",1,"individuals","New␣Mexico","Campanian","Kirtland","
    Kirtland␣New␣Mexico"
34828,"Pentaceratops",1,"individuals","New␣Mexico","Maastrichtian","Kirtland","
    Kirtland␣New␣Mexico"
46991,"Pentaceratops",1,"individuals","New␣Mexico","Maastrichtian","Kirtland","
    Kirtland␣New␣Mexico"
62549,"Pentaceratops",1,"specimens","New␣Mexico","Maastrichtian","Kirtland","
    Kirtland␣New␣Mexico"
64094,"Pentaceratops",1,"specimens","New␣Mexico","Maastrichtian","Kirtland","
    Kirtland␣New␣Mexico"
53982,"Pentaceratops",2,"specimens","New␣Mexico",,"Kirtland","Kirtland␣New␣Mexico"
47821,"Prenocephale",1,"specimens","New␣Mexico","Campanian","Kirtland","Kirtland␣New
    ␣Mexico"
48838,"Saurolophus",1,"individuals","New␣Mexico","Maastrichtian","Kirtland","
    Kirtland␣New␣Mexico"
76105,"Saurornitholestes",1,"specimens","New␣Mexico","Campanian","Kirtland","
    Kirtland␣New␣Mexico"
53514,"Sphaerotholus",1,"specimens","New␣Mexico","Maastrichtian","Kirtland","
    Kirtland␣New␣Mexico"
53515,"Sphaerotholus",2,"specimens","New␣Mexico","Campanian","Kirtland","Kirtland␣
    New␣Mexico"
46995,"Struthiomimus",1,"specimens","New␣Mexico",,"Kirtland","Kirtland␣New␣Mexico"
47009,"Struthiomimus",1,"specimens","New␣Mexico",,"Kirtland","Kirtland␣New␣Mexico"
47013,"Struthiomimus",1,"specimens","New␣Mexico",,"Kirtland","Kirtland␣New␣Mexico"
14518,"Alamosaurus",2,"specimens","New␣Mexico","Maastrichtian","Kirtland","Kirtland␣
    New␣Mexico"
49005,"Tyrannosaurus",1,"individuals","Montana","Maastrichtian","Lance","Lance␣
    Montana"
```

```
47628,"Edmontonia",1,"individuals","South␣Dakota","Maastrichtian","Lance","Lance␣
    South␣Dakota"
47628,"Pachycephalosaurus",1,"individuals","South␣Dakota","Maastrichtian","Lance","
    Lance␣South␣Dakota"
47628,"Pachycephalosaurus",1,"individuals","South␣Dakota","Maastrichtian","Lance","
    Lance␣South␣Dakota"
50084,"Thescelosaurus",1,"individuals","South␣Dakota","Maastrichtian","Lance","Lance
    ␣South␣Dakota"
48549,"Triceratops",1,"specimens","Utah","Maastrichtian","Lance","Lance␣Utah"
14514,"Ankylosaurus",1,"specimens","Wyoming","Maastrichtian","Lance","Lance␣Wyoming"
14585,"Ankylosaurus",1,"specimens","Wyoming","Maastrichtian","Lance","Lance␣Wyoming"
45878,"Ankylosaurus",1,"specimens","Wyoming","Maastrichtian","Lance","Lance␣Wyoming"
14505,"Coelurosauria",1,"specimens","Wyoming","Maastrichtian","Lance","Lance␣Wyoming
    "
14514,"Coelurosauria",1,"specimens","Wyoming","Maastrichtian","Lance","Lance␣Wyoming
    "
14514,"Dromaeosaurus",2,"specimens","Wyoming","Maastrichtian","Lance","Lance␣Wyoming
    "
14505,"Edmontosaurus",1,"specimens","Wyoming","Maastrichtian","Lance","Lance␣Wyoming
    "
46128,"Edmontosaurus",1,"individuals","Wyoming","Maastrichtian","Lance","Lance␣
    Wyoming"
46209,"Edmontosaurus",1,"individuals","Wyoming","Maastrichtian","Lance","Lance␣
    Wyoming"
49238,"Edmontosaurus",1,"individuals","Wyoming","Maastrichtian","Lance","Lance␣
    Wyoming"
54176,"Edmontosaurus",1,"individuals","Wyoming","Maastrichtian","Lance","Lance␣
    Wyoming"
54178,"Edmontosaurus",1,"individuals","Wyoming","Maastrichtian","Lance","Lance␣
    Wyoming"
57169,"Edmontosaurus",1,"specimens","Wyoming","Maastrichtian","Lance","Lance␣Wyoming
    "
57170,"Edmontosaurus",1,"specimens","Wyoming","Maastrichtian","Lance","Lance␣Wyoming
    "
57764,"Edmontosaurus",1,"individuals","Wyoming","Maastrichtian","Lance","Lance␣
    Wyoming"
14585,"Edmontosaurus",2,"specimens","Wyoming","Maastrichtian","Lance","Lance␣Wyoming
    "
26760,"Edmontosaurus",300000,"specimens","Wyoming","Maastrichtian","Lance","Lance␣
    Wyoming"
34747,"Leptoceratops",1,"individuals","Wyoming","Maastrichtian","Lance","Lance␣
    Wyoming"
14585,"Ornithomimus",1,"specimens","Wyoming","Maastrichtian","Lance","Lance␣Wyoming"
14606,"Ornithomimus",1,"individuals","Wyoming","Maastrichtian","Lance","Lance␣
    Wyoming"
14609,"Ornithomimus",1,"specimens","Wyoming","Maastrichtian","Lance","Lance␣Wyoming"
14585,"Pachycephalosaurus",1,"specimens","Wyoming","Maastrichtian","Lance","Lance␣
    Wyoming"
52106,"Pachycephalosaurus",1,"specimens","Wyoming","Maastrichtian","Lance","Lance␣
    Wyoming"
64644,"Pachycephalosaurus",1,"specimens","Wyoming","Maastrichtian","Lance","Lance␣
    Wyoming"
64645,"Pachycephalosaurus",1,"specimens","Wyoming","Maastrichtian","Lance","Lance␣
    Wyoming"
72145,"Pachycephalosaurus",1,"specimens","Wyoming","Maastrichtian","Lance","Lance␣
    Wyoming"
14505,"Richardoestesia",1,"specimens","Wyoming","Maastrichtian","Lance","Lance␣
    Wyoming"
```

```
14566,"Richardoestesia",1,"specimens","Wyoming","Maastrichtian","Lance","Lance␣
    Wyoming"
14577,"Richardoestesia",1,"specimens","Wyoming","Maastrichtian","Lance","Lance␣
    Wyoming"
14585,"Richardoestesia",1,"specimens","Wyoming","Maastrichtian","Lance","Lance␣
    Wyoming"
14585,"Richardoestesia",5,"specimens","Wyoming","Maastrichtian","Lance","Lance␣
    Wyoming"
14596,"Richardoestesia",6,"specimens","Wyoming","Maastrichtian","Lance","Lance␣
    Wyoming"
14585,"Saurornithoides",5,"specimens","Wyoming","Maastrichtian","Lance","Lance␣
    Wyoming"
13144,"Thescelosaurus",1,"specimens","Wyoming","Maastrichtian","Lance","Lance␣
    Wyoming"
48626,"Thescelosaurus",1,"specimens","Wyoming","Maastrichtian","Lance","Lance␣
    Wyoming"
50085,"Thescelosaurus",1,"individuals","Wyoming","Maastrichtian","Lance","Lance␣
    Wyoming"
14566,"Thescelosaurus",2,"specimens","Wyoming","Maastrichtian","Lance","Lance␣
    Wyoming"
14585,"Thescelosaurus",6,"specimens","Wyoming","Maastrichtian","Lance","Lance␣
    Wyoming"
46134,"Torosaurus",1,"individuals","Wyoming","Maastrichtian","Lance","Lance␣Wyoming"
46135,"Torosaurus",1,"individuals","Wyoming","Maastrichtian","Lance","Lance␣Wyoming"
14505,"Triceratops",1,"specimens","Wyoming","Maastrichtian","Lance","Lance␣Wyoming"
14519,"Triceratops",1,"specimens","Wyoming","Maastrichtian","Lance","Lance␣Wyoming"
46131,"Triceratops",1,"individuals","Wyoming","Maastrichtian","Lance","Lance␣Wyoming
    "
46132,"Triceratops",1,"individuals","Wyoming","Maastrichtian","Lance","Lance␣Wyoming
    "
46136,"Triceratops",1,"individuals","Wyoming","Maastrichtian","Lance","Lance␣Wyoming
    "
50086,"Triceratops",1,"individuals","Wyoming","Maastrichtian","Lance","Lance␣Wyoming
    "
50087,"Triceratops",1,"individuals","Wyoming","Maastrichtian","Lance","Lance␣Wyoming
    "
50930,"Triceratops",1,"specimens","Wyoming","Maastrichtian","Lance","Lance␣Wyoming"
50931,"Triceratops",1,"specimens","Wyoming","Maastrichtian","Lance","Lance␣Wyoming"
52107,"Triceratops",1,"specimens","Wyoming","Maastrichtian","Lance","Lance␣Wyoming"
52681,"Triceratops",1,"individuals","Wyoming","Maastrichtian","Lance","Lance␣Wyoming
    "
52682,"Triceratops",1,"individuals","Wyoming","Maastrichtian","Lance","Lance␣Wyoming
    "
52683,"Triceratops",1,"individuals","Wyoming","Maastrichtian","Lance","Lance␣Wyoming
    "
52684,"Triceratops",1,"individuals","Wyoming","Maastrichtian","Lance","Lance␣Wyoming
    "
52684,"Triceratops",1,"individuals","Wyoming","Maastrichtian","Lance","Lance␣Wyoming
    "
52685,"Triceratops",1,"individuals","Wyoming","Maastrichtian","Lance","Lance␣Wyoming
    "
52686,"Triceratops",1,"individuals","Wyoming","Maastrichtian","Lance","Lance␣Wyoming
    "
52687,"Triceratops",1,"individuals","Wyoming","Maastrichtian","Lance","Lance␣Wyoming
    "
52688,"Triceratops",1,"individuals","Wyoming","Maastrichtian","Lance","Lance␣Wyoming
    "
```

```
52690,"Triceratops",1,"individuals","Wyoming","Maastrichtian","Lance","Lance␣Wyoming
    "
52691,"Triceratops",1,"individuals","Wyoming","Maastrichtian","Lance","Lance␣Wyoming
    "
52692,"Triceratops",1,"individuals","Wyoming","Maastrichtian","Lance","Lance␣Wyoming
    "
52693,"Triceratops",1,"individuals","Wyoming","Maastrichtian","Lance","Lance␣Wyoming
    "
52694,"Triceratops",1,"individuals","Wyoming","Maastrichtian","Lance","Lance␣Wyoming
    "
52695,"Triceratops",1,"individuals","Wyoming","Maastrichtian","Lance","Lance␣Wyoming
    "
54177,"Triceratops",1,"individuals","Wyoming","Maastrichtian","Lance","Lance␣Wyoming
    "
54179,"Triceratops",1,"individuals","Wyoming","Maastrichtian","Lance","Lance␣Wyoming
    "
14585,"Triceratops",2,"specimens","Wyoming","Maastrichtian","Lance","Lance␣Wyoming"
57161,"Triceratops",2,"specimens","Wyoming","Maastrichtian","Lance","Lance␣Wyoming"
14566,"Troodon",1,"specimens","Wyoming","Maastrichtian","Lance","Lance␣Wyoming"
13144,"Troodon",2,"specimens","Wyoming","Maastrichtian","Lance","Lance␣Wyoming"
13144,"Troodon",2,"specimens","Wyoming","Maastrichtian","Lance","Lance␣Wyoming"
14505,"Troodon",2,"specimens","Wyoming","Maastrichtian","Lance","Lance␣Wyoming"
14566,"Troodon",2,"specimens","Wyoming","Maastrichtian","Lance","Lance␣Wyoming"
14585,"Troodon",8,"specimens","Wyoming","Maastrichtian","Lance","Lance␣Wyoming"
45878,"Tyrannosaurus",1,"individuals","Wyoming","Maastrichtian","Lance","Lance␣
    Wyoming"
52871,"Tyrannosaurus",1,"specimens","Wyoming","Maastrichtian","Lance","Lance␣Wyoming
    "
61888,"Tyrannosaurus",1,"individuals","Wyoming","Maastrichtian","Lance","Lance␣
    Wyoming"
46210,"Edmontosaurus",1,"specimens","Colorado","Maastrichtian","Laramie","Laramie␣
    Colorado"
49536,"Torosaurus",1,"individuals","Colorado","Maastrichtian","Laramie","Laramie␣
    Colorado"
49529,"Triceratops",1,"individuals","Colorado","Maastrichtian","Laramie","Laramie␣
    Colorado"
46133,"Diceratops",1,"individuals","Wyoming","Maastrichtian","Laramie","Laramie␣
    Wyoming"
46120,"Edmontosaurus",1,"individuals","Wyoming","Maastrichtian","Laramie","Laramie␣
    Wyoming"
46119,"Triceratops",1,"individuals","Wyoming","Maastrichtian","Laramie","Laramie␣
    Wyoming"
46126,"Triceratops",1,"individuals","Wyoming","Maastrichtian","Laramie","Laramie␣
    Wyoming"
46127,"Triceratops",1,"individuals","Wyoming","Maastrichtian","Laramie","Laramie␣
    Wyoming"
46129,"Triceratops",1,"individuals","Wyoming","Maastrichtian","Laramie","Laramie␣
    Wyoming"
46138,"Triceratops",1,"individuals","Wyoming","Maastrichtian","Laramie","Laramie␣
    Wyoming"
46130,"Triceratops",2,"individuals","Wyoming","Maastrichtian","Laramie","Laramie␣
    Wyoming"
55289,"Alamosaurus",1,"specimens","New␣Mexico","Maastrichtian","McRae","McRae␣New␣
    Mexico"
55287,"Tyrannosaurus",1,"specimens","New␣Mexico","Maastrichtian","McRae","McRae␣New␣
    Mexico"
45718,"Alamosaurus",1,"individuals","Utah","Maastrichtian","North␣Horn","North␣Horn␣
    Utah"
```

```
45721,"Alamosaurus",1,"individuals","Utah","Maastrichtian","North␣Horn","North␣Horn␣Utah"
45722,"Alamosaurus",1,"specimens","Utah","Maastrichtian","North␣Horn","North␣Horn␣Utah"
45719,"Torosaurus",2,"individuals","Utah","Maastrichtian","North␣Horn","North␣Horn␣Utah"
48842,"Torosaurus",2,"specimens","Utah","Maastrichtian","North␣Horn","North␣Horn␣Utah"
53824,"Leptoceratops",1,"specimens","Wyoming","Maastrichtian","Pinyon␣Conglomerate","Pinyon␣Conglomerate␣Wyoming"
11898,"Albertosaurus",1,"specimens","Alberta","Maastrichtian","Scollard","Scollard␣Alberta"
45876,"Ankylosaurus",1,"individuals","Alberta","Maastrichtian","Scollard","Scollard␣Alberta"
45877,"Ankylosaurus",1,"individuals","Alberta","Maastrichtian","Scollard","Scollard␣Alberta"
64628,"Leptoceratops",1,"individuals","Alberta","Maastrichtian","Scollard","Scollard␣Alberta"
14569,"Leptoceratops",2,"specimens","Alberta","Maastrichtian","Scollard","Scollard␣Alberta"
64353,"Leptoceratops",3,"individuals","Alberta","Maastrichtian","Scollard","Scollard␣Alberta"
14569,"Pachycephalosaurus",1,"specimens","Alberta","Maastrichtian","Scollard","Scollard␣Alberta"
48627,"Thescelosaurus",1,"individuals","Alberta","Maastrichtian","Scollard","Scollard␣Alberta"
47130,"Torosaurus",1,"individuals","Alberta","Maastrichtian","Scollard","Scollard␣Alberta"
47128,"Triceratops",1,"specimens","Alberta","Maastrichtian","Scollard","Scollard␣Alberta"
47127,"Triceratops",3,"specimens","Alberta","Maastrichtian","Scollard","Scollard␣Alberta"
11917,"Tyrannosaurus",1,"specimens","Alberta","Maastrichtian","Scollard","Scollard␣Alberta"
11918,"Tyrannosaurus",1,"specimens","Alberta","Maastrichtian","Scollard","Scollard␣Alberta"
52072,"Anchiceratops",4,"specimens","Alberta",,"St.␣Mary␣River","St.␣Mary␣River␣Alberta"
14502,"Edmontonia",1,"specimens","Alberta","Maastrichtian","St.␣Mary␣River","St.␣Mary␣River␣Alberta"
52071,"Pachyrhinosaurus",1,"individuals","Alberta",,"St.␣Mary␣River","St.␣Mary␣River␣Alberta"
52249,"Pachyrhinosaurus",3,"specimens","Alberta","Maastrichtian","St.␣Mary␣River","St.␣Mary␣River␣Alberta"
52072,"Pachyrhinosaurus",10,"specimens","Alberta",,"St.␣Mary␣River","St.␣Mary␣River␣Alberta"
52072,"Troodon",1,"specimens","Alberta",,"St.␣Mary␣River","St.␣Mary␣River␣Alberta"
61631,"Euoplocephalus",1,"specimens","Montana","Maastrichtian","St.␣Mary␣River","St.␣Mary␣River␣Montana"
55266,"Montanoceratops",1,"individuals","Montana","Maastrichtian","St.␣Mary␣River","St.␣Mary␣River␣Montana"
55267,"Montanoceratops",1,"individuals","Montana","Maastrichtian","St.␣Mary␣River","St.␣Mary␣River␣Montana"
68010,"Saurornitholestes",1,"specimens","Texas","Maastrichtian","Tornillo","Tornillo␣Texas"
68011,"Saurornitholestes",1,"specimens","Texas","Maastrichtian","Tornillo","Tornillo␣Texas"
```

```
68011,"Saurornitholestes",1,"specimens","Texas","Maastrichtian","Tornillo","Tornillo
    ␣Texas"
68009,"Saurornitholestes",2,"specimens","Texas","Maastrichtian","Tornillo","Tornillo
    ␣Texas"
68009,"Saurornitholestes",3,"specimens","Texas","Maastrichtian","Tornillo","Tornillo
    ␣Texas"
69211,"Torosaurus",1,"specimens","Texas","Maastrichtian","Tornillo","Tornillo␣Texas"
69210,"Tyrannosaurus",1,"specimens","Texas","Maastrichtian","Tornillo","Tornillo␣
    Texas"
```

Appendix H: R Code for Chapter 6

```
###################################################
### chunk number 1: setup
###################################################
#set up environment with appropriate packages
library(fossil)
library(pointtracker)
library(RODBC)


#Hypothesis 1: latitude vs species diversity
############################################################################
##step 1: create all datasets for individual stages/epochs, both species matrices
    and lat/long tables
plantdb<-odbcConnect("plantdb", "postgres", "NbasscHs")


###################################################
### chunk number 2: newfuncs
###################################################

sel.data<-function(time, division) {
  dat<-sqlQuery(plantdb, paste("SELECT␣species_detail.locality,␣species_detail.genus
      ,␣species_detail.abundance,␣locality.lat,␣locality.long␣" ,
    "FROM␣␣species_detail␣LEFT␣OUTER␣JOIN␣locality␣ON␣(species_detail.locality␣=␣␣
        locality.locality)␣" ,
    "WHERE␣locality.", division, "␣=␣", time, "␣AND␣species_detail.abundance␣>␣0␣AND
        ␣species_detail.mega_micro␣=␣0␣",
    "AND␣(locality.country␣=␣'Canada'␣OR␣locality.country␣=␣'United␣States')", sep="
        "))
  mat<-create.matrix(dat, abund = T)
  return(mat)
}

sel.lats<-function(time, division) {
  lats<-sqlQuery(plantdb, paste("SELECT␣DISTINCT␣ON␣(locality.locality)␣locality.
      locality,␣locality.lat,␣locality.long" ,
    "FROM␣␣locality␣LEFT␣OUTER␣JOIN␣species_detail␣ON␣(locality.locality␣=␣species_
        detail.locality)" ,
    "WHERE␣locality.", division, "␣=␣", time, "␣AND␣(locality.country␣=␣'Canada'␣OR␣
        locality.country␣=␣'United␣States')␣AND␣species_detail.abundance␣>␣0␣AND␣
        species_detail.mega_micro␣=␣0"))
  return(lats)
}

map.plot<-function(time, lats) {
lats.rot<-cbind(lats[,1],point.tracker(lats[,2:3],time))
pmap<-read.csv(paste("/home/matthew/docs/paleomap/polygons/", time, "mya", sep=""))
plot(pmap[,3:4], type = "n", xaxs = "i", yaxs = "i", xlim = c(-180, 0), ylim = c
    (0,90), xlab = "Longitude", ylab = "Latitude")
levels(as.factor(pmap[,2]))->f
for (i in 1:length(f)) {
```

```
  if (is.na(charmatch("c", f[i]))==F) polygon(pmap[pmap[,2]==f[i],3:4], col = "#
      bebebe")
  else polygon(pmap[pmap[,2]==f[i],3:4], col = "#676767")
  }
points(lats.rot[,3:2], col=2, pch=17)
}

#### plotting specipal.oc.lats<-pal.oc.lats[order(pal.oc.lats[,1]),]es abundance vs
    latitude for 4 different estimates
quad.plot<-function(mat, lats) {
        x<-function(x) length(x[x>0])
        sobs.mio<-apply(mat,2,x)
        chao.mat<-apply(mat, 2, chao1)
        jack.mat<-apply(mat, 2, jack1)
        ace.mat<-apply(mat, 2, ACE)
        #yl<-max(c(sobs.mio, chao.mat, jack.mat, ace.mat))
        #xl<-min(lats[,2])

        par(mfrow=c(2,2),oma=c(4,4,4,4))
        par(mar=c(0,0,0,0))

        #first plot

        plot(lats[,2], sobs.mio, xlab="", ylab="",xaxt="n", ylim=c(0,100), xlim=c
            (30,65))
        lines(lowess(sobs.mio~lats[,2]))
        abline(glm(sobs.mio~lats[,2]))
        text(33, 90, "A", cex=2)

        #### plotting species abundance vs latitude for chao
        plot(lats[,2], chao.mat, xlab = "", ylab = "", xaxt="n", yaxt="n", ylim=c
            (0,100), xlim=c(30,65))
        lines(lowess(chao.mat~lats[,2]))
        abline(glm(chao.mat~lats[,2]))
        text(33, 90, "B", cex=2)

        #### plotting species abundance vs latitude for chao
        plot(lats[,2], jack.mat, xlab = "", ylab = "", ylim=c(0,100), xlim=c(30,65))
        lines(lowess(jack.mat~lats[,2]))
        abline(glm(jack.mat~lats[,2]))
        text(33, 90, "C", cex=2)

        #### plotting species abundance vs latitude for chao
                plot(lats[,2], ace.mat, xlab = "", ylab = "", yaxt="n", ylim=c
                    (0,100), xlim=c(30,65))
        lines(lowess(ace.mat~lats[,2]))
        abline(glm(ace.mat~lats[,2]))
        text(33, 90, "D", cex=2)
}

#
    ##################################################################################

#following functions are to create plots for occurrence data
sel.ocdata<-function(time, division) {
  mat<-sqlQuery(plantdb, paste("SELECT␣species_detail.locality,␣species_detail.genus
      ,␣species_detail.abundance,␣locality.lat,␣locality.long" ,
```

```
  "FROM␣␣species_detail␣LEFT␣OUTER␣JOIN␣locality␣ON␣(species_detail.locality␣=␣␣
      locality.locality)" ,
  "WHERE␣locality.", division, "␣=␣", time, "␣AND␣(locality.country␣=␣'Canada'␣OR␣
      locality.country␣=␣'United␣States')␣AND␣species_detail.mega_micro␣=␣0␣AND␣
      locality.lat␣<␣70"))
  mat<-mat[is.na(mat[,4])==FALSE,]
  mat<-create.matrix(mat, abund = F)
  return(mat)
}

sel.oclats<-function(time, division) {
  lats<-sqlQuery(plantdb, paste("SELECT␣species_detail.locality,␣species_detail.
      genus,␣species_detail.abundance,␣locality.lat,␣locality.long" ,
  "FROM␣␣species_detail␣LEFT␣OUTER␣JOIN␣locality␣ON␣(species_detail.locality␣=␣␣
      locality.locality)" ,
  "WHERE␣locality.", division, "␣=␣", time, "␣AND␣(locality.country␣=␣'Canada'␣OR␣
      locality.country␣=␣'United␣States')␣AND␣species_detail.mega_micro␣=␣0␣AND␣
      locality.lat␣<␣70"))
  lats<-lats[is.na(lats[,4])==FALSE,]
  lats<-lats[duplicated(lats[,1])==FALSE,c(1,4,5)]
  lats<-lats[order(lats[,1]),]
  return(lats)
}


oc.map.plot<-function(time, lats) {
  lats.rot<-cbind(lats[,1],point.tracker(lats[,2:3], time))
  pmap<-read.csv(paste("/home/matthew/docs/paleomap/polygons/", time, "mya", sep="")
      )
  plot(pmap[,3:4], type = "n", xaxs = "i", yaxs = "i", xlim = c(-180, 0), ylim = c
      (0,90), xlab = "Longitude", ylab = "Latitude")
  levels(as.factor(pmap[,2]))->f
  for (i in 1:length(f)) {
    if (is.na(charmatch("c", f[i]))==F) polygon(pmap[pmap[,2]==f[i],3:4], col = "#
        bebebe")
    else polygon(pmap[pmap[,2]==f[i],3:4], col = "#676767")
  }
  points(lats.rot[,3:2], col=2, pch=17)
}

oc.plot<-function(mat, lats) {
  x<-function(x) length(x[x>0])
  sobs.ocdata<-apply(mat,2,x)
  plot(lats[,2], sobs.ocdata, xlab="Latitude", ylab="Species␣Richness", ylim=c
      (0,100), xlim=c(30,65))
  lines(lowess(sobs.ocdata~lats[,2]))
  abline(glm(sobs.ocdata~lats[,2]))
  text(40,10, coefficients(glm(sobs.ocdata~lats[,2]))[2])
  return(glm(sobs.ocdata~lats[,2]))
}


###################################################
### chunk number 3: mio-oc
###################################################
###first for Miocene
mio.oc.mat<-sel.ocdata("'Miocene'", 'epoch')
mio.oc.lats<-sel.oclats("'Miocene'", 'epoch')
```

```
###################################################
### chunk number 4: mio-oc-map
###################################################
oc.map.plot(20, mio.oc.lats)


###################################################
### chunk number 5: mio-oc-plot
###################################################
mio.oc.plot<-oc.plot(mio.oc.mat, mio.oc.lats)


###################################################
### chunk number 6: mio
###################################################
###first for Miocene
mio.mat<-sel.data("'Miocene'", 'epoch')
mio.lats<-sel.lats("'Miocene'", 'epoch')


###################################################
### chunk number 7: mio-map
###################################################
map.plot(20, mio.lats)


###################################################
### chunk number 8: mio-quad-plot
###################################################
quad.plot(mio.mat, mio.lats)


###################################################
### chunk number 9: oligo-oc
###################################################
###first for Oligocene
oligo.oc.mat<-sel.ocdata("'Oligocene'", 'epoch')
oligo.oc.lats<-sel.oclats("'Oligocene'", 'epoch')


###################################################
### chunk number 10: oligo-oc-map
###################################################
oc.map.plot(20, oligo.oc.lats)


###################################################
### chunk number 11: oligo-oc-plot
###################################################
oligo.oc.plot<-oc.plot(oligo.oc.mat, oligo.oc.lats)


###################################################
### chunk number 12: eo-oc
###################################################
###first for Eocene
```

```
eo.oc.mat<-sel.ocdata("'Eocene'", 'epoch')
eo.oc.lats<-sel.oclats("'Eocene'", 'epoch')


###################################################
### chunk number 13: eo-oc-map
###################################################
oc.map.plot(50, eo.oc.lats)


###################################################
### chunk number 14: eo-oc-plot
###################################################
eo.oc.plot<-oc.plot(eo.oc.mat, eo.oc.lats)


###################################################
### chunk number 15: eo
###################################################
eo.mat<-sel.data("'Eocene'", 'epoch')
eo.lats<-sel.lats("'Eocene'", 'epoch')


###################################################
### chunk number 16: eo-map
###################################################
map.plot(50, eo.lats)


###################################################
### chunk number 17: eo-quad-plot
###################################################
quad.plot(eo.mat, eo.lats)


###################################################
### chunk number 18: pal-oc
###################################################
###first for Paleocene
pal.oc.mat<-sel.ocdata("'Paleocene'", 'epoch')
pal.oc.lats<-sel.oclats("'Paleocene'", 'epoch')


###################################################
### chunk number 19: pal-oc-map
###################################################
oc.map.plot(50, pal.oc.lats)


###################################################
### chunk number 20: pal-oc-plot
###################################################
pal.oc.plot<-oc.plot(pal.oc.mat, pal.oc.lats)


###################################################
### chunk number 21: pal
###################################################
```

```
pal.mat<-sel.data("'Paleocene'", 'epoch')
pal.lats<-sel.lats("'Paleocene'", 'epoch')


###################################################
### chunk number 22: pal-map
###################################################
map.plot(50, pal.lats)


###################################################
### chunk number 23: pal-quad-plot
###################################################
quad.plot(pal.mat, pal.lats)


###################################################
### chunk number 24: maas-oc
###################################################
###first for Maastrichtian
maas.oc.mat<-sel.ocdata("'Maastrichtian'", 'stage')
maas.oc.lats<-sel.oclats("'Maastrichtian'", 'stage')


###################################################
### chunk number 25: maas-oc-map
###################################################
oc.map.plot(50, maas.oc.lats)


###################################################
### chunk number 26: maas-oc-plot
###################################################
maas.oc.plot<-oc.plot(maas.oc.mat, maas.oc.lats)


###################################################
### chunk number 27: maas
###################################################
maas.mat<-sel.data("'Maastrichtian'", 'stage')
maas.lats<-sel.lats("'Maastrichtian'", 'stage')


###################################################
### chunk number 28: maas-map
###################################################
map.plot(50, maas.lats)


###################################################
### chunk number 29: maas-quad-plot
###################################################
quad.plot(maas.mat, maas.lats)


###################################################
### chunk number 30: camp-oc
###################################################
```

```
###first for Campanian
camp.oc.mat<-sel.ocdata("'Campanian'", 'stage')
camp.oc.lats<-sel.oclats("'Campanian'", 'stage')


###################################################
### chunk number 31: camp-oc-map
###################################################
oc.map.plot(50, camp.oc.lats)


###################################################
### chunk number 32: camp-oc-plot
###################################################
camp.oc.plot<-oc.plot(camp.oc.mat, camp.oc.lats)


###################################################
### chunk number 33: camp
###################################################
camp.mat<-sel.data("'Campanian'", 'stage')
camp.lats<-sel.lats("'Campanian'", 'stage')


###################################################
### chunk number 34: camp-map
###################################################
map.plot(50, camp.lats)


###################################################
### chunk number 35: camp-quad-plot
###################################################
quad.plot(camp.mat, camp.lats)


###################################################
### chunk number 36: cen-oc
###################################################
###first for Cenomanian
cen.oc.mat<-sel.ocdata("'Cenomanian'", 'stage')
cen.oc.lats<-sel.oclats("'Cenomanian'", 'stage')


###################################################
### chunk number 37: cen-oc-map
###################################################
oc.map.plot(50, cen.oc.lats)


###################################################
### chunk number 38: cen-oc-plot
###################################################
cen.oc.plot<-oc.plot(cen.oc.mat, cen.oc.lats)


###################################################
### chunk number 39: cen
```

```
###################################################
cen.mat<-sel.data("'Cenomanian'", 'stage')
cen.lats<-sel.lats("'Cenomanian'", 'stage')


###################################################
### chunk number 40: cen-map
###################################################
map.plot(50, cen.lats)


###################################################
### chunk number 41: cen-quad-plot
###################################################
quad.plot(cen.mat, cen.lats)


###################################################
### chunk number 42: alb-oc
###################################################
###first for Albian
alb.oc.mat<-sel.ocdata("'Albian'", 'stage')
alb.oc.lats<-sel.oclats("'Albian'", 'stage')


###################################################
### chunk number 43: alb-oc-map
###################################################
oc.map.plot(50, alb.oc.lats)


###################################################
### chunk number 44: alb-oc-plot
###################################################
alb.oc.plot<-oc.plot(alb.oc.mat, alb.oc.lats)


###################################################
### chunk number 45: alb
###################################################
alb.mat<-sel.data("'Albian'", 'stage')
alb.lats<-sel.lats("'Albian'", 'stage')


###################################################
### chunk number 46: alb-map
###################################################
map.plot(50, alb.lats)


###################################################
### chunk number 47: alb-quad-plot
###################################################
quad.plot(alb.mat, alb.lats)


###################################################
### chunk number 48: apt-oc
```

```
##################################################
###first for Aptian
apt.oc.mat<-sel.ocdata("'Aptian'", 'stage')
apt.oc.lats<-sel.oclats("'Aptian'", 'stage')


##################################################
### chunk number 49: apt-oc-map
##################################################
oc.map.plot(50, apt.oc.lats)


##################################################
### chunk number 50: summary-plot
##################################################
apt.oc.plot<-oc.plot(apt.oc.mat, apt.oc.lats)


##################################################
### chunk number 51: gentry
##################################################
gt<-read.csv('gentry_table.csv', header=F)

dms2dd<-function(x) {
dn<-regexpr('Âř', x)
if (dn>0) d<-as.numeric(substr(x,1,dn-1))
else d<-0
mn<-regexpr('\'',␣x)
if␣(mn>0)␣m<-as.numeric(substr(x,dn+1,mn-1))
else␣m<-0
sn<-regexpr('\"',␣x)
if␣(sn>0)␣s<-as.numeric(substr(x,mn+1,sn-1))
else␣s<-0
if␣(regexpr('S',␣x)>0␣|␣regexpr('W',␣x)>0)␣sw<-TRUE
else␣sw␣<-␣FALSE
dd<-d+(m/60)+(s/3600)
if␣(sw==TRUE)␣dd<-0-dd
return(dd)
}

dms2dd(gt[1,4])

a<-gt[,3]
b<-NULL
for␣(i␣in␣1:length(a))␣b<-c(b,dms2dd(a[i]))

z<-as.numeric(as.character(gt[-9,8]))
y<-abs(b[-9])

plot(y,z,xlab="",␣ylab="")
abline(glm(z~y))
summary(glm(z~y))
mod.oc.plot<-glm(z~y)


##################################################
###␣chunk␣number␣52:␣gentry-30
##################################################
```

```
p<-y[y>30]
q<-z[y>30]
plot(p,q,xlab="",␣ylab="",␣ylim=c(0,100),␣xlim=c(30,65))
abline(glm(q~p))


###################################################
###␣chunk␣number␣53:␣slope-plot-temp
###################################################
slopes<-c(coefficients(mod.oc.plot)[2],coefficients(mio.oc.plot)[2],coefficients(eo.
    oc.plot)[2],coefficients(pal.oc.plot)[2],coefficients(maas.oc.plot)[2],
    coefficients(camp.oc.plot)[2],coefficients(cen.oc.plot)[2],coefficients(alb.oc.
    plot)[2],coefficients(apt.oc.plot)[2])

dates<-c(0,15,45,␣60,␣68,␣77,␣96,␣105,␣117)

ci<-c(summary(mod.oc.plot)$coefficients[2,2],summary(mio.oc.plot)$coefficients[2,2],
    summary(eo.oc.plot)$coefficients[2,2],summary(pal.oc.plot)$coefficients[2,2],
    summary(maas.oc.plot)$coefficients[2,2],summary(camp.oc.plot)$coefficients[2,2],
    summary(cen.oc.plot)$coefficients[2,2],summary(alb.oc.plot)$coefficients[2,2],
    summary(apt.oc.plot)$coefficients[2,2])
plot(dates,␣slopes,␣ylim=c(-2.0,1.0),xaxt='n')
#$
ucl<-slopes+ci
lcl<-slopes-ci
arrows(dates,ucl,dates,lcl,length=.05,angle=90,code=3)
#lines(lowess(slopes~dates))
lines(dates,c(0,5,13,12,12,15,17,17,16)/20*3-2,␣col=2)
#abline(glm(slopes~dates))
polypts<-matrix(c(dates,␣rev(dates),␣ucl,␣rev(lcl)),␣ncol=2)
polygon(polypts,␣col=rgb(190/255,␣190/255,␣190/255,␣alpha=0.5))
axis(1,␣at=dates,␣labels=c('Mod','Mio','Eo','Pal','Maas','Camp','Cen','Alb','Apt'))
axis(4,␣at=(seq(0,20,2)/20*3-2),␣labels=seq(0,20,2))
mtext("Temperature (Celsius)",␣side=4,␣line=1)
```