Spring 5-11-2024

# PALEOECOLOGY OF THE LATE CRETACEOUS WESTERN INTERIOR SEAWAY AND GULF COASTAL PLAIN

Ceara Kathleen Quinn Purcell
*University of New Mexico*

Ceara Purcell

*Candidate*

Department of Earth and Planetary Sciences

*Department*

This dissertation is approved, and it is acceptable in quality and form for publication:

*Approved by the Dissertation Committee:*

Peter Fawcett , Chairperson

Corinne Myers

Louis Scuderi

Jason Moore

Tyler Mackey

# PALEOECOLOGY OF THE LATE CRETACEOUS WESTERN INTERIOR SEAWAY AND GULF COASTAL PLAIN

**by**

**CEARA KATHLEEN QUINN PURCELL**

B.S., Geological Sciences, Ohio University, Athens, 2021

DISSERTATION

Submitted in Partial Fulfillment of the
Requirements for the Degree of

**Doctor of Philosophy**
**Earth and Planetary Sciences**

The University of New Mexico
Albuquerque, New Mexico

**May 2024**

# DEDICATION

I dedicate this dissertation to my husband, Austin Pollard. He knows what he did.

## ACKNOWLADGEMENTS

I would like to acknowledge the numerous awesome people who helped me get this far. Firstly, I'd like to thank the academic influences I've have throughout the years. Specifically, my advisor, Dr. Cori Myers. She suffered through my unreasonable program pace and the neuroses that came with it with more than just good grace, so thank you. Thank you as well to Dr. Louis Scuderi and Dr. Carli Pietsch for helping with publications and for always being such good sports about reading my extremely long supplementary materials. I would also like to thank my other committee members, Dr. Jason Moore and Dr. Tyler Mackey, the entire UNM Earth and Planetary Sciences Department, and the amazing Paula Pascetti and Cecilia Arias for always being so helpful and kind. To all the researchers, collectors, and explorers who made the dataset I used possible, thank you from the bottom of my heart. Not only could I not have compiled so much alone, I would never have dreamed of trying.

For helping me get as far as a Ph.D., I'd like to thank my Master's advisor Dr. Alycia Stigall for her unrelenting enthusiasm and unfailing interest in my success. To all the professors who showed me the wonders of deserts, porphyry deposits, and caves, Todd, Russel, Hannah, and Greg, thank you so much. You guys are so cool, I hope I can be like you one day.

Additionally, without the financial support of the National Science Foundation through the Museum Research Traineeship and the UNM Department of Earth and Planetary Sciences this research would not be possible, and for that I am very grateful. Thank you especially to the hard work and dedication of Dr. Tom Turner, who loves all things science

and always made students feel welcome, and Loretta Esquibel, who constantly made the MRT a better and brighter group with her ingenuity and kindness. Thank you both.

Secondly, I would like to acknowledge the support and friendship of Rhiannon Nolan, Jone Naujokaityte, Alex Apgar, Joe Wislocki, Cole Cooper, Bri Gordon, Ian Forsythe, Sam Mignogna, Travis Heier, Stephanie Hunt, Alex Knoer, and Naveen Balaji. They helped me in so many ways, science and otherwise, but first and foremost they helped me have fun and stay sane.

To my family, Sheila, Kevin, Maura, Sean, Brenna, and Tara, thank you for always believing in me despite myself. I can finally say, in my professional opinion as a Doctor of Philosophy (which encompasses the study of truth): you guys are the inarguable best.

Finally, I'd like to thank my husband, Austin Pollard. You're the one I always want to talk to first, and last. If there's a better person than you, I've never met them, and I don't believe there's a better person out there for me. Clichés be damned — I couldn't have done any of this without you.

# PALEOECOLOGY OF THE LATE CRETACEOUS WESTERN INTERIOR SEAWAY AND GULF COASTAL PLAIN

**by**

**CEARA KATHLEEN QUINN PURCELL**

**B.S., Geological Sciences, Ohio University, Athens, 2021**

**Ph.D. Earth and Planetary Sciences, University of New Mexico, 2024**

## ABSTRACT

This dissertation presents research into the paleobiogeography and paleoecology of the Western Interior Seaway (WIS) and Gulf Coastal Plain (GCP) during the Late Cretaceous. The dataset utilized here includes an extensive record of marine invertebrates from over 200 years of fossil collecting and sedimentary data collated from over 45 different literature sources, representing approximately 17 million years of deposition. The high-resolution of this dataset, and its extensive geographic distribution and temporal duration, make it ideal for exploring various ecological questions. Using these data, I have attempted to reconstruct paleobiogeographic provinces as published in prior works, explore functional diversity patterns, and test some of the fundamental assumptions related to niche modeling in the fossil record. This work represents one of the most extensive and dynamic analyses of the WIS, focusing on elucidating primary assumptions regarding broad scale ecological relationships and the methods used to explore them.

# TABLE OF CONTENTS

# LIST OF TABLES

# LIST OF FIGURES

# PREFACE

This preface acknowledges the funding sources and outlines the roles that co-authors, and collaborators performed in each chapter pursuant to the requirements of the University of New Mexico and the Department of Earth and Planetary Sciences. Ceara K. Q. Purcell conducted greater than or equal to 51% of the work presented in each chapter of this dissertation and is therefore the primary author. This research was funded by NSF #2021744, #1924807, and #1601878.

Chapter 1 was published in the journal *Geology* with co-authors Dr. Corinne Myers and Dr. Louis Scuderi (Purcell et al., 2023). Dr. Myers and Dr. Scuderi contributed advice about methods and interpretations as well as providing textural edits for the manuscript. Ceara Purcell designed the project, conducted analyses, created tables and figures, wrote the manuscript, and interpreted the bulk of the results.

Chapter 2 was submitted (March 2024) to the journal *Paleobiology* and co-authored by Dr. Myers, Dr. Carlie Pietsch, and Mr. Naveen Balaji (Purcell et al., *in review*). Dr. Myers and Dr. Pietsch contributed advice about methods and interpretations as well as providing textural edits for the manuscript. Mr. Balaji assisted with a portion of R coding. Ceara Purcell designed the project, conducted analyses, created tables and figures, wrote the manuscript, and interpreted the bulk of the results.

Chapter 3 is unpublished work assessing fundamental biases of paleo-niche analysis that will be summited for publication in the journal *Paleobiology* following the dissertation defense.

# CHAPTER 1: INTRODUCTION

This analysis builds off the extensive research previously conducted on the Campanian and Maastrichtian paleontological record of the Western Interior Seaway (WIS) and adjacent Gulf Coastal Plain (GCP). It is intended to quantitatively test existing paleobiogeographic and ecological assumptions and build a foundational framework for future investigations into this interval and region. Specifically, my research utilizes relatively novel techniques applied to the fossil record that incorporate quantitative and qualitative methods to investigate fundamental biotic patterns and environmental associations that existed in the WIS and GCP during the upper Late Cretaceous, just prior to the end-Cretaceous mass extinction event (Figure 1).

The fundamental aspects of biotic associations present throughout Earth's history are not only informative in understanding the ancient world, but they may also provide us with the tools and base-line assumptions necessary to predict and mitigate detrimental taxonomic responses in the world today to projected environmental change. Sediments and fossils from the WIS and GCP represent an extremely well-documented, geographically extensive, and diverse assemblage that is well suited to investigating these issues. Modern, quantitative, biostatistical techniques have in recent history applied to fossil data to expand on previous findings (i.e., Dudie and Stigall, 2010; Keil, 2017; Dean et al. 2019). However, it is not enough to be able to apply these techniques; we also need to establish the basic principles, biases, and viable methods to be sure these techniques are providing accurate results and can be interpreted in useful ways. Prior to their application, however, it is important to establish the basic principles underlying the phenomena recorded by paleontological and

**Table of Acronyms**

| | |
|---|---|
| $\alpha_F$ | Functional richness |
| BC | Betweenness centrality |
| $BC_{ave}$ | Average betweenness centrality |
| CAM | Campanian |
| CC | Clustering coefficient |
| ENM | Ecological niche modeling |
| FE | Functional entity |
| FN | Fundamental niche |
| GCP | Gulf Coastal Plain |
| HFS | High faunal similarity |
| KPg | End Cretaceous mass extinction |
| kyrs | Thousands of years |
| LDG | Latitudinal diversity gradient |
| MAA | Maastrichtian |
| MST | Minimum spanning trees |
| My | Millions of years |
| NSF | National Science Foundation |
| PCS | Principal component analysis |
| PNC | Phylogenetic niche conservation |
| RN | Realized niche |
| SEI | Shannon's evenness index |
| SI | Supplementary information |
| SQS | Shareholder quorum subsampling |
| SME | Simpsons measure of evenness |
| TA | Transcontinental arch |
| TH | Threshold |
| WIS | Western Interior Seaway |

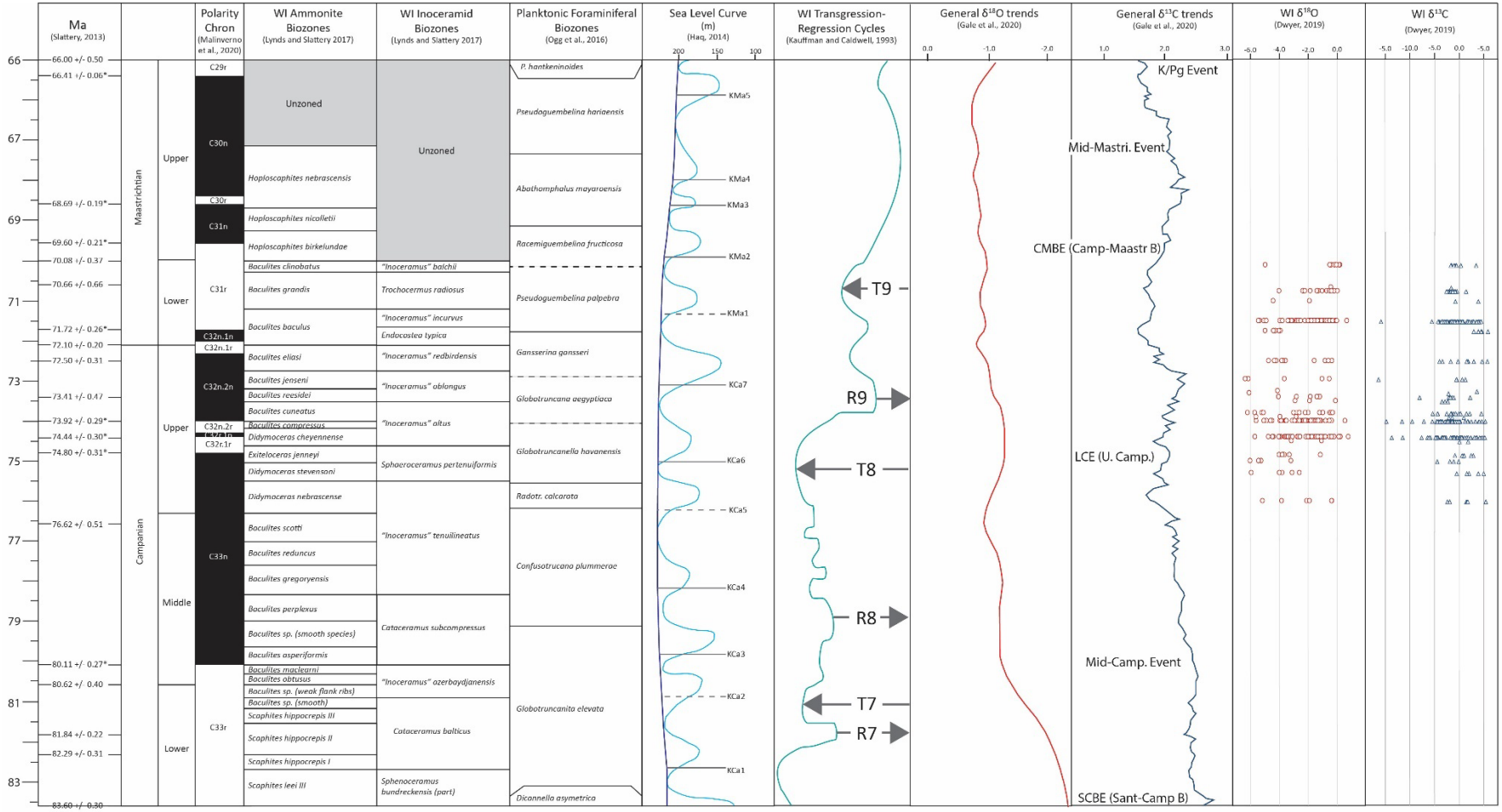



**Figure 1.** Campanian to Maastrichtian biostratigraphic chart showing polarity chorn intervals (Malinverno et al. 2020), biozones (Ogg et al. 2016; Lynds and Slattery 2017), sea level fluctuations (Haq 2014), global stable carbon and oxygen isotope trends (Gale et al. 2020), and stable oxygen and carbon isotope data collated for the Western Interior region (Dwyer, 2019). Ages marked with an asterisk have been added or updated based on Malinverno et al. (2020) from Slattery et al. (2013). WI isotope values modified from unpublished data by Dwyer (2019) which span the R9 regressive interval of the Bearpaw Cyclothem (*D. cheyennenses – B. baculus* ammonite zones).

environmental data, and the common biases which may influence our understanding of these phenomena. Questions that can potentially be investigated to better understand the Earth system in this way include how taxa and the functional entities distribute themselves across space and how species abiotic ecological niches, the primary ecological unit hypothetically underlying these biogeographical patterns, can be accurately and consistently assessed using fossil data. By collating one of the most extensive, high-resolution paleontological and sedimentological databases available to date, this research addresses some aspects of these issues and provides guidance for additional investigations, creating an essential framework from which even greater understanding of the Late Cretaceous and the modern world can be achieved.

**Geologic Setting: The Western Interior Seaway and Gulf Coastal Plain**

During the late Albian, rising sea levels flooded the North American continent, connecting the Artic Ocean in the north with the Tethys Sea in the south at the proto-Gulf of Mexico (Williams and Stelek 1975; Kauffman and Caldwell 1993; Roberts and Kirschbaum 1995; Ziegler and Rowley 1998). This epeiric seaway occupied the forearc basins formed by the tectonic movement of the North American Plate west relative to the European Plate, beginning in the Late Triassic (Monger 1993; DeCelles 2004; Miall et al. 2008). These basins were bound on the west by the Cordilleran Fold-Thrust Belt and on the east by the stable North American Craton (Kauffman and Caldwell 1993; DeCelles 2004; Miall et al. 2008). Sedimentary units from the WIS record tectonically influenced subsidence and associated deposition, primarily from erosion of the uplifted fold-thrust belt to the west. Deposits from the GCP, in contrast, represent deposition along the passive but subsiding margin of the North American Continent (Braile et al. 1986). Geologic units from both regions represent a

dense record of marine deposition from approximately 100 to 59.2 Ma (Caldwell, 1974; Slattery et al., 2013).

Across this interval, eustatic and tectonic, regional changes to sea level greatly influenced the paleoceanography and geometry of the WIS and its connection with the GCP, eventually resulting in the diminution and draining of the WIS itself at the end of the Cretaceous (Gill and Cobban 1966*a*; Kauffman 1984; Kauffman and Caldwell 1993; Slattery et al. 2013; Lowery et al. 2018). Oceanic differences between the WIS and GCP may also have influenced the macroevolution and macroecology of marine biota. The WIS had a unique and potentially dynamically shifting set of conditions, including distinct habitat shifts such as moving from siliciclastic to carbonate deposition eastward over its longitudinal extent (Slattery et al. 2013), the presence of a counterclockwise gyre and distinct cool and warm water masses (Fisher et al. 1994; Slingerland et al. 1996; Steel et al. 2012; Elderbak and Leckie 2016; Lowery et al. 2018), and potential periods of oceanic stratification (Cochran et al. 2003; He et al. 2005; Fricke et al. 2010; Petersen et al. 2016; Lowery et al. 2018). Distinct biotic patterns across both space and time have been noted throughout the Campanian and Maastrichtian of the WIS in particular, often associated with abiotic factors such as sea level (Sohl 1967, 1971, 1987; Jeletzky 1971; Kauffman 1973, 1984; Scott and Cobban 1986).

The dense fossil and sedimentological record of the WIS and GCP, their persistence across a critical interval just prior to a major mass extinction event, their broad latitudinal distribution, and the variation in environmental conditions they experienced make these two regions ideal for testing various hypotheses of ecological patterns and processes on geological timescales. Indeed, many previous workers have made great strides in

understanding the paleogeographic, biotic, and oceanic patterns that persisted throughout the WIS and GCP regions across this interval. Most of these studies are more than thirty years old, and the addition of novel methods and more extensive datasets make reevaluating these patterns using modern analytical methods both desirable and valuable. For example, it has long been assumed – based on investigations primarily conducted in the late 1980s – that the WIS was characterized by distinct biogeographical subprovinces whose distribution was closely associated with fluctuations in regional sea level (Kauffman 1984). No dedicated analysis has reassessed these claims using modern methods or updated taxonomic data.

Few palaeoecological studies, furthermore, focus on studying spatiotemporal patterns within a relatively "stable" interval of Earth's history, rather than an period of significant abiotic change such as a mass extinction interval (i.e., Dunhill et al., 2018; Foster and Twitchet, 2014). While these studies are relevant to determining how major abiotic perturbations influence biota, they rarely include long-term, in-depth analysis for background intervals before or after the mass extinction event itself. Understanding how ecological patterns present themselves during background intervals, is therefore highly relevant to establishing baseline assumptions. Since the end Cretaceous mass extinction occurs at the end of the Maastrichtian, these analyses illuminate the spatial and temporal ecological patterns just before this event, providing a point of comparison for shifts across the mass extinction interval.

Lastly, understanding how abiotic and biotic systems interact is useful for predicting the effects of modern environmental changes, such as warming climates and shifting sea levels. Unlike most modern analyses, paleontological analyses allow for long-term assessment of broad-scale patterns during intervals of environmental change. Deep time data

can therefore be used as analogues for current and future conditions, informing modern conservation efforts. These analyses provide information about various aspects of a marine system over approximately 17 million years, informing our understanding of our current ecological system and its potential responses to future environmental change.

**Research Chapters and Databases**

This dissertation consists of three research studies that focus on fundamental ecological aspects of the WIS and GCP region during the Campanian and Maastrichtian. They utilize an extensive fossil dataset including over 33,000 total occurrences of marine invertebrates, and composed of 1113 different species and 574 genera (Table 1, Figure 2). Major clades represented include bivalves, cephalopods, gastropods, echinoderms, corals, brachiopods, bryozoans, and crustaceans. The complete database characterizes a substantial portion of marine animal taxonomic and functional diversity preserved in the WIS present during the Late Cretaceous and their geographic extent. High taxonomic resolution make these data ideal for future studies into primary ecological characteristics such as paleobiogeographic patterns, functional assemblages, and realized/fundamental niche analyses (Peterson et al. 2011).

The first chapter focuses on reassessing the paleobiogeographic provinces present in the WIS during the Late Cretaceous and their abiotic associations (Purcell et al., 2023). Biotic provinces associated with latitudinal gradients that shift across space relative to major abiotic shifts such as sea level change were hypothesized to exist in the WIS and GCP across the interval (Kauffman 1984). However, only two major provinces were observed using quantitative network analysis, with greater provinciality resulting from sea level fall across

**Table 1**. Taxonomic database summary

| Class | Number of Unique Genera | Number of Unique Species |
|---|---|---|
| Gastropoda | 194 | 376 |
| Bivalvia | 192 | 395 |
| Cephalopoda | 82 | 220 |
| Malacostraca | 32 | 30 |
| Gymnolaemata | 18 | 15 |
| Echinoidea | 18 | 30 |
| Anthozoa | 7 | 10 |
| Polychaeta | 5 | 5 |
| Hexanauplia | 4 | 11 |
| Rhynchonellata | 4 | 2 |
| Scaphopoda | 3 | 7 |
| Xiphosura | 2 | 1 |
| Stenolaemata | 2 | 2 |
| Demospongea | 2 | 2 |
| Asteroidea | 1 | 1 |
| Crinoidea | 1 | 1 |
| Chromadorea | 1 | 0 |
| Homarus | 1 | 0 |
| Lingulata | 1 | 2 |
| Hydrozoa | 1 | 1 |
| Thecostraca | 1 | 0 |
| Maxillipoda | 1 | 1 |
| Bryozoa (undifferentiated) | 1 | 1 |
| **Total:** | **574** | **1113** |

the interval. Conclusions support a WIS biogeographically homogenized, that became increasingly distinct through time from the open-ocean-facing GCP.

The second chapter explores the distinction between functional and taxonomic diversity in the WIS and GCP. Within this analysis, the dataset was also tested for how spatial aggregation of occurrences might influenced results of studies of this ilk. Spatiotemporal patterns of functional diversity were compared with taxonomic diversity under the primary hypothesis that functional diversity would be both spatially and temporally stable relative to

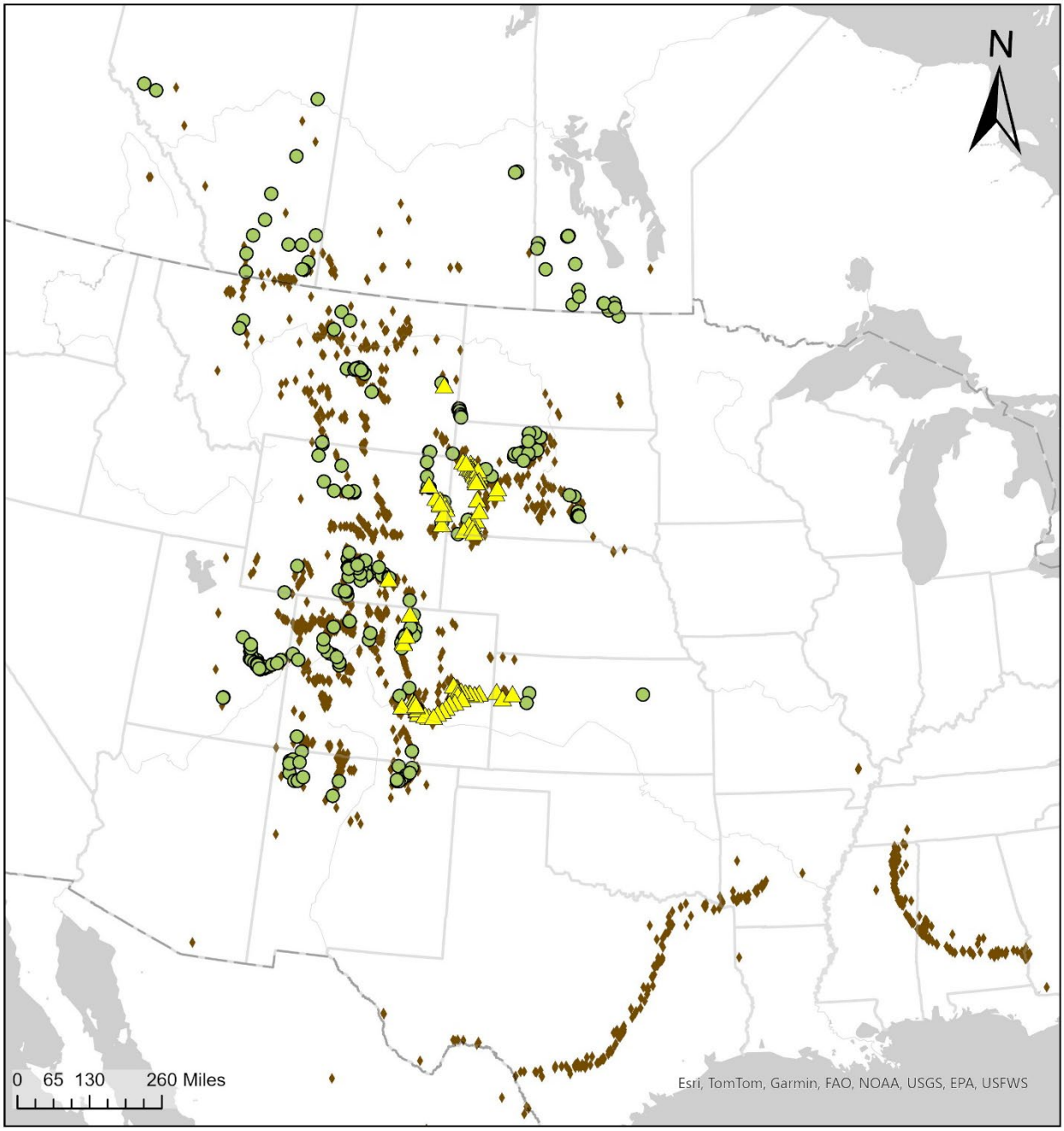



**Figure 2**. Distribution of taxonomic (brown diamonds), stratigraphic (green circles), and hydrothermal seep (yellow triangles) from the WIS across the study interval (Campanian – Maastrichtian). See Chapters 1 and 4 for details.

taxonomic diversity. This hypothesis was largely supported by the data, though differences in functional assemblages were observed in the WIS relative to the GCP. These patterns were found to reflect specific abiotic factors present in the WIS as compared to the GCP and to support ecological buffering against taxonomic loss during environmentally stressed intervals. diversity would be both spatially and temporally stable relative to taxonomic diversity. This hypothesis was largely supported by the data, though differences in functional assemblages were observed in the WIS relative to the GCP. These patterns were found to reflect specific abiotic factors present in the WIS as compared to the GCP and to support ecological buffering against taxonomic loss during environmentally stressed intervals.

The final chapter incorporates a high-resolution stratigraphic database from across the WIS to assess potential sources of bias in the application of ecological niche modeling using deep time paleontological data. This chapter specifically tests the influence of temporal resolution on ecological niche quantification and provides a case study test of phylogenetic niche conservation in fossil species as they relate to their genus and the potential utility of using generic-level data as a proxy for species characteristics. It was hypothesized that low temporal resolution negatively influences the fidelity of niche characterization, and that individual species within the same genus with present high niche conservation with each other and with the genus-level niche itself. Temporal resolution was found to negatively correlate with niche characterization, but genus-level data was not found to represent species-level niches well.

**Table 2**. The number of sedimentary data locality points at each level of biostratigraphic confidence in the twelve biozone intervals. See Chapter 4 for details.

| Interval | High Confidence | Medium Confidence | Low Confidence | Total Localities |
|---|---|---|---|---|
| *H. birkelundae-H. nebrascensis* | 5 | 10 | 5 | **20** |
| *B. clinolobatus* | 5 | 19 | 15 | **39** |
| *B. baculus-B. grandis* | 10 | 18 | 10 | **38** |
| *B. reesidei-B. eliasi* | 28 | 6 | 3 | **37** |
| *B. compressus-B. cuneatus* | 13 | 7 | 8 | **28** |
| *D. cheyennense* | 3 | 9 | 5 | **17** |
| *D. nebrascense-E. jenneyi* | 12 | 21 | 3 | **36** |
| *B. reduncus-B. scotti* | 13 | 7 | 11 | **31** |
| *B. perplexus-B. gregoryensis* | 38 | 10 | 19 | **67** |
| *B. maclearni-B. sp. (smooth)* | 24 | 17 | 30 | **71** |
| *B. obtusus* | 25 | 9 | 29 | **63** |
| *S. leei-B. sp. (weak flank ribs)* | 19 | 19 | 21 | **59** |
| Total | | | | 506 |

The sedimentological database included in the third research chapter will furthermore provide a foundation for exploring the more detailed localized aspects of ecology and abiotic conditions. Stratigraphic data incorporates sedimentary information relevant to environmental conditions present during sedimentation, including grain size, siliciclastic versus carbonate composition, sedimentary structures, bedding thickness, and bioturbation. These data were used to create WIS maps of continuous sedimentary characteristics constrained by paleo-shorelines.

# CHAPTER 2: FAUNAL PROVICIALITY IN THE LATE CRETACEOUS WESTERN INTERIOR SEAWAY USING NETWORK MODELING

**Ceara Purcell[1], Louis Scuderi[1], Corinne Myers[1]**

*[1]Department of Earth and Planetary Sciences, Northrop Hall, 221 Yale Blvd NE University of New Mexico, Albuquerque, New Mexico 87131*

## Abstract

The Western Interior Seaway (WIS) was historically divided into latitudinal faunal provinces that were taxonomically distinct from the adjacent Gulf Coastal Plain (GCP) and shift in space due to sea level changes. However, no rigorous quantitative analysis using recent taxonomic updates have reassessed these provinces and their associations. We used network modeling of macro-invertebrate WIS and GCP fauna to test whether biotic provinces existed and to examine their relationships with abiotic change. Results suggest a cohesive WIS unit existed across the Campanian, and distinct WIS and GCP provinces existed in the Maastrichtian. Sea level changes coincided with changes in network metrics. These results indicate that, while the WIS did not contain sub-provinces in the Late Cretaceous, environmental factors influenced faunal associations and their communication over time.

## Introduction

The Western Interior Seaway (WIS) and Gulf Coastal Plain (GCP) are characterized by a dense fossil record of marine invertebrates in the latest Cretaceous (~100–66 Ma; Caldwell, 1974; Slattery et al., 2013), spanning 45° latitude, that experienced a wide range of environmental shifts. Fluctuating sea levels (Figure 3C), for example, modified basin geometry and water mass distributions, impacting marine life (e.g., He et al., 2005; Kauffman, 1984; Lowery et al., 2018). A restricted connection between the WIS and the

open ocean affected oceanic conditions relative to the GCP and may have influenced biotic provinces.

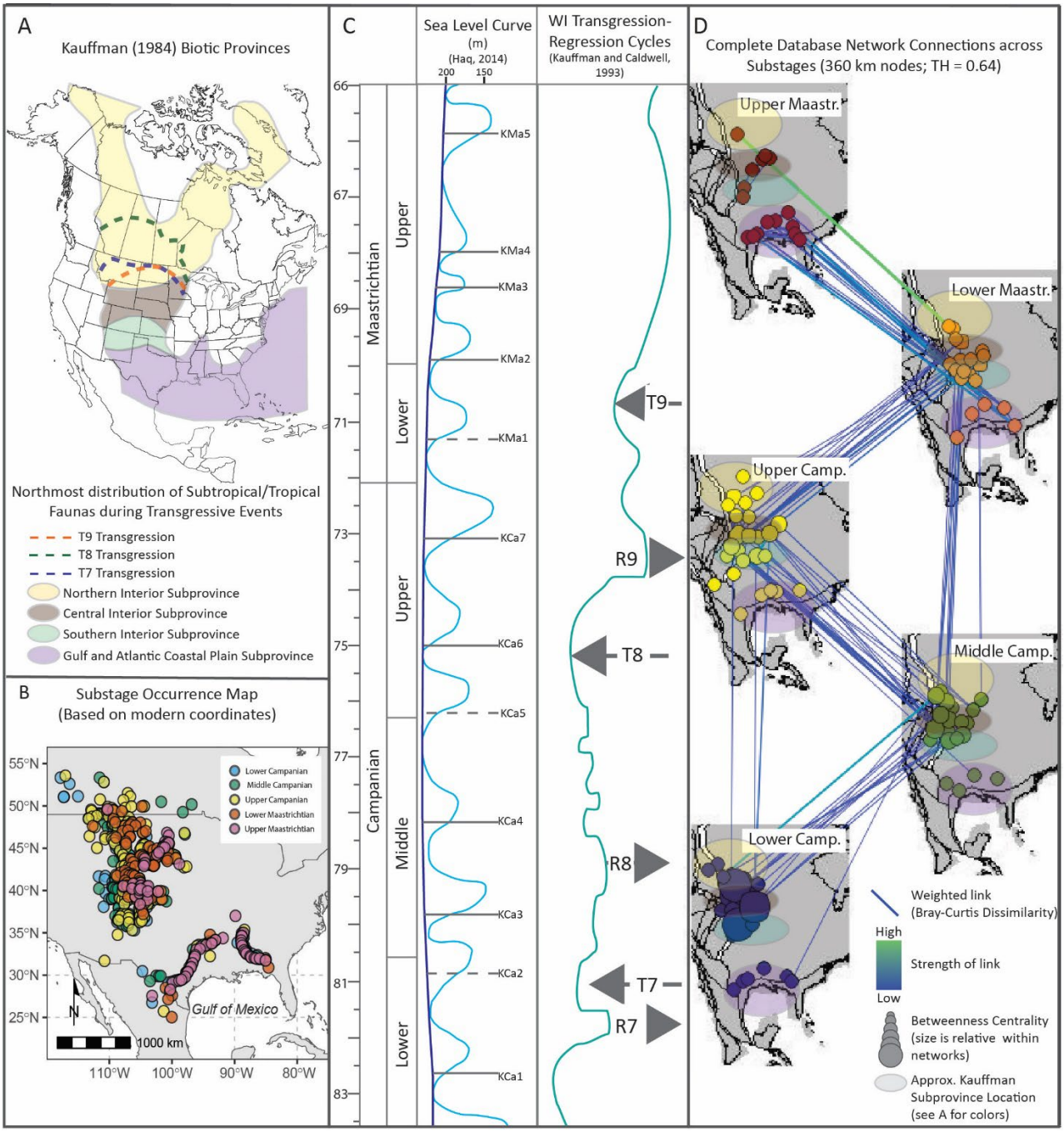



**Figure 3**. A) Previously defined subprovinces and northernmost extent of tropical/subtropical faunas during transgressions, modified from Kauffman (1984). B) Occurrence map of data from this study. C) Global and regional sea levels curves with major transgressive-regressive events. D) Complete database 360 km network with Kauffman's subprovinces indicated. A modified version of this figure with additional biostratigraphic and isotopic data is available as Appendix A, Figure S30.

Biotic provinces are geographic regions characterized by distinct ecological associations. Previous studies of biotic provinces using fossil materials have: attributed them to major climatic regions (e.g., Kocsis et al., 2021), associated shifting provinces with sea level fluctuations (e.g., Kauffman, 1984), compared spatiotemporal influences on taxonomic association patterns (e.g., Kiel, 2017), and observed changes in provinciality relative to taxonomic loss (e.g., Kocsis et al., 2018).

Quantifying biogeographic patterns can therefore shed light on macroecology over evolutionary time. Kauffman (1984) described three significant biotic incursions during transgressions based on changes to WIS subprovinces (Figure 3A). These subprovinces, determined using percent endemism of mollusk records from 1960-80s, are: (1) Northern Interior; (2) Southern Interior; and (3) Central Interior Subprovince (Kauffman 1984). Another identified faunal province was the Gulf and Atlantic Coastal Plain Subprovince. However, delineation of these paleobiogeographic provinces and their changes through time were based on qualitative assemblages limited by the available fossil data (Kauffman, 1984 and references therein). Analysis of WIS provinciality using current fossil data will improve the validity of these interpretations.

Network modeling analysis of faunal provinces is a novel approach (Kiel 2016, 2017) to quantifying faunal similarity across spatiotemporal units. Using a well-vetted database of over 33,000 fossil occurrences from the WIS and the GCP we used this approach to reevaluate WIS provinciality in the Campanian (CAM) and Maastrichtian (MAA) stages of the Late Cretaceous. While previous studies have used network analysis to explore provincialism in fossil taxa (Kiel, 2016, 2017; Rojas et al., 2021; Muscente et al., 2018; Kocsis et al., 2018, 2021), none have applied the technique to a geochemically unique,

restricted ocean system, characterized by over 100 years of dedicated sampling. This research may also inform general patterns of Earth-Life interaction over long timescales (e.g., the nature of faunal variation through time and space) and serve as a foundation for future WIS/GCP investigations.

## Methods

Records of marine invertebrates from the CAM and MAA of the WIS and the GCP were compiled from digital databases, including the Paleobiology Database (08/25/21 download) and iDigBio (08/30/21 download), and records from museum collections at the Black Hills Institute, USGS-Denver (Cobban Collection), and the Mackenzie (2007)thesis database (Table S1). Taxa were binned into the Early, Middle, and Late CAM and the Early and Late MAA stages (Figure 3B and S1). Localities were converted to paleocoordinates within a 60 km grid for analysis; nodes with fewer than three unique taxa were removed. The vetted database was analyzed prior to network modeling to determine fundamental sources of taxonomic and spatial bias which should be considered during network interpretations (Appendix A, Table S2, Figure S2-S9).

Faunal provinces were delimited for substages individually and for the complete database (combined substages) using threshold weighted networks (Keil, 2016) in the EDENetworks software (Appendix A, Table S3 & Figure S10-S20; Moalic et al., 2012; Kivelä et al., 2015; Kiel, 2016). Network components which disconnect at and below a network-specific threshold identified by EDENetworks known as the percolation point are interpreted as representing distinct “community” groups (Newman 2012) or faunal provinces. General patterns in network connections across all substages together were assessed using

coarser spatial aggregations of the data and minor network components were assessed for spatiotemporal consistency (Appendix A, Figure S21–S23).

Average network clustering coefficient (CC) values, indicative of network organization ranging from 0 (no cluster) to 1 (fully connected cluster), were compared with a null model of randomized networks to determine if the topology was more or less clustered than a random distribution (Table S4; Kiel, 2016). Link weights, or the degree of dissimilarity between nodes, and betweenness centrality (BC), a measure of the degree to which a node acts as a geographic connection between regions, were averaged by 5° paleolatitudinal bins for comparison (Appendix A, Figure S24-S26). Link weights were also binned by geographic distance to test for correlation between faunal dissimilarity and distance (Appendix A, Table S5 & Figure S27). Given latitudinal overlap between the WIS and GCP, around 35°N, we separated the data by major components and evaluated link similarity between the two regions independently (Appendix A, Table S6 & Figure S28). Sampling bias influence on network communication was assessed by binning average betweenness centrality ($BC_{ave}$), by generic richness as a proxy for sampling effort (Appendix A, Figure S25) and using minimum spanning trees (MST; Appendix A, Figure S29). Additional explanation of methods is provided in the supplementary materials.[2]

## Results and Discussion

### *WIS vs. GCP Provinces*

The presence of a single faunal province in the WIS is supported by all network permutations and subsequent analysis (Figure 4). This province is geographically consistent with the WIS (Appendix A, Figure S21), is maintained at all threshold levels (Appendix A,

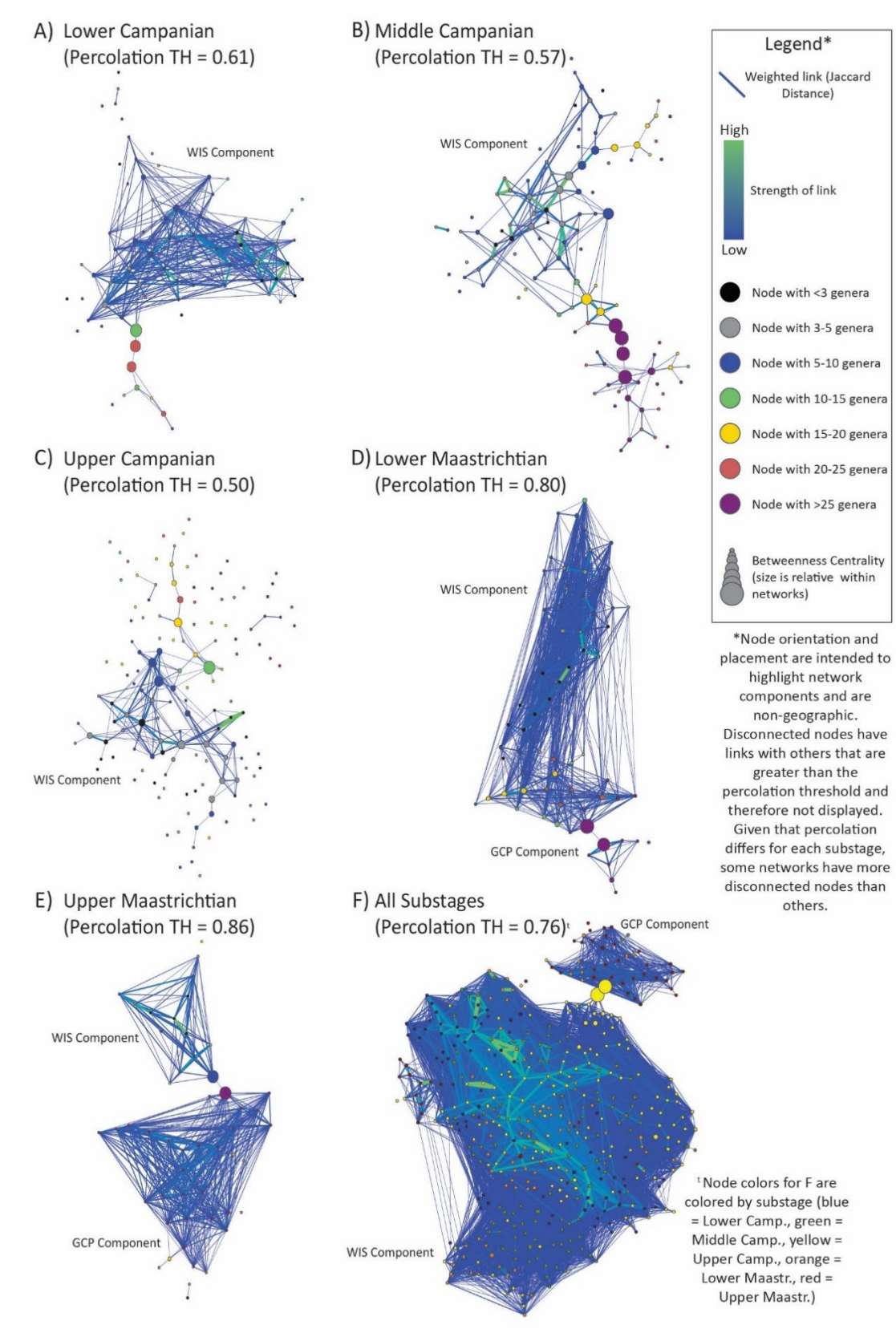



**Figure 4**. Substage and complete database networks at percolation points.

Figure S10-15), does not contain spatiotemporally consistent minor components (Appendix A, Figure S22-23) and is faunally distinct from GCP grid cells (Appendix A, Figure S21). For MAA substages, which contain more GCP fossil occurrences relative to the CAM, a well-supported GCP faunal province was observed (Figure 4, S16-S19). This distinct GCP province is supported by the full database network as well (Appendix A, Figure S20). Network randomization comparisons demonstrated that these results are non-random and likely reflect biogeographic patterns (Table S4; average CC >3 stand. dev. from mean). They do not result from sampling bias based on comparisons with spatial cluster analysis (Appendix A, Figure S5-S9), BC comparisons (Appendix A, Figure S25), and MST (Appendix A, Table S7-S12). Thus, our quantitative analysis does not support the existence of WIS biotic subprovinces but does support a distinct GCP province in the MAA (Kauffman 1984).

The WIS and GCP provinces may have resulted from geochemical and bathymetric changes across the transcontinental arch (TA), which may have acted as a bathymetric high between the regions (He et al., 2005; Lowery et al., 2018 and references therein), rather than by latitudinal factors (e.g., temperature). Geochemical studies have found evidence for non-normal marine conditions in the WIS, including low salinity or brackish conditions (Cochran et al., 2003; Dennis et al., 2013; Fricke et al., 2010) and lower $\delta^{18}O$ of seawater than the open ocean (Fricke et al. 2010; Petersen et al. 2016). There is also evidence for stratification within the WIS during the CAM and MAA, produced by mixing water masses (He et al. 2005; Lowery et al. 2018). These factors could have created a habitat barrier between the GCP and WIS, facilitating provincialism. The WIS fauna may have been more tolerant of non-normal conditions, supported by a lack of abundant reef-building and reef-associated

taxa (Gill and Cobban 1966b; Caldwell 1968; Kauffman 1984; Kauffman and Caldwell 1993).

The lack of a latitudinally-defined provinces within the WIS is unsurprising given a flattened latitudinal temperature gradient in the Cretaceous greenhouse (Mannion et al., 2014; Super et al., 2018). However, evidence for different water mass distributions and salinity/temperature gradients have long been associated with latitude and faunal gradients in the WIS (Fisher et al., 1994; Slingerland et al., 1996; Longman et al., 1998; Elderbak and Leckie, 2016; Lowery et al., 2018). During much of the study interval, a cool-water mass circulated south through the WIS from the northern connection with Greenland and northern Europe, interacting with the northward moving warm water mass from the Tethys, forming a counterclockwise gyre (Steel et al., 2012; Lowery et al., 2018). However, no evidence for provinces matching these water bodies was observed in our results. Instead, this gyre could have contributed to faunal homogenization despite ocean stratification or abiotic gradients. Further, the unique geochemical nature of the basin may have encouraged WIS incumbents and generalists to flourish over specialists or invaders. Dataset differences, including improved sampling and the lack of foraminifera in this study, may have hindered observation of Kauffman's (1984) subprovinces, though this requires further investigation. Indeed, the potential for along-seaway variation within specific WIS faunas, as observed by previous authors (i.e., Sohl, 1971; Jeletzky, 1971, etc.), was not tested by this analysis which tests for discrete clusters of faunal assemblages.

### *Decreasing Faunal Similarity and Sea Level Fall*

Network dissimilarity values increase through time (i.e., decreasing similarity), particularly in the WIS province, based on average link values per substage (Table 3); in

contrast, the GCP province shows increasing similarity through the MAA (Table S6). Bathymetric and geochemical changes coincide with these shifts, suggesting a potential relationship. The WIS gyre that may have promoted mixing of water masses, WIS dispersal, and mixing with the GCP (Fisher et al. 1994; Slingerland et al. 1996; Longman et al. 1998; Elderbak and Leckie 2016). However, falling sea levels likely impacted circulation patterns, water mass dynamics, and geochemical and environmental gradients (e.g., non-normal salinity, nutrient load) that would limit WIS migration and thereby insulate existing fauna from outside invasion (Cochran et al. 2003b; He et al. 2005; Fricke et al. 2010; Petersen et al. 2016). Shallowing along the TA may also have created a geographic barrier between the WIS and GCP as early as the Late CAM (Lehman, 1987; Lowery et al., 2018 and references therein).

**Table 3**. Average/median link weights within substages and between different substages. 95% confidence intervals (CI) of the mean indicated.

| | **Substages Link Weight Comparisons (Mean ± 95% CI/Median)** | | | | |
|---|---|---|---|---|---|
| **Substage Link Weights (Mean ± 95% CI/Median)** | Lower Camp. | Middle Camp. | Upper Camp. | Lower Maastr. | Upper Maastr. |
| Lower Camp. (0.78 ± 0.004/0.80) | - | 0.83 ± 0.002/0.84 | 0.85 ± 0.002/0.86 | 0.87 ± 0.003/0.88 | 0.96 ± 0.001/0.98 |
| Middle Camp. (0.79 ± 0.002/0.82) | - | - | 0.83 ± 0.001/0.85 | 0.86 ± 0.002/0.88 | 0.95 ± 0.001/0.96 |
| Upper Camp. (0.82 ± 0.002/0.84) | - | - | - | 0.86 ± 0.002/0.88 | 0.95 ± 0.001/0.96 |
| Lower Maastr. (0.83 ± 0.004/0.86) | - | - | - | - | 0.93 ± 0.003/0.95 |
| Upper Maastr. (0.85 ± 0.006/0.95) | - | - | - | - | - |

Dampened circulation and salinity gradients may have also caused declining faunal similarity within the WIS alone, as evidenced by lower average faunal similarity within each substage network across time (Table S6). Below ~1000 km faunal similarity comparisons show weak correlation between distance and link weight (Figure 5 and S28), indicating only slight decline in similarity over distance within a substage, despite decreasing similarity though time. Sedimentary evidence for tidal circulation influences through at least the Middle and Late CAM (Steel et al. 2012) suggest continued circulation and mixing that could promote homogenization. Distance comparisons for the WIS and GCP components individually produced similar patterns (Figure S28), indicating that these results are not basin specific. Additional study of WIS oceanography is needed to confidently assess the potential influence of Late Campanian oceanographic changes.

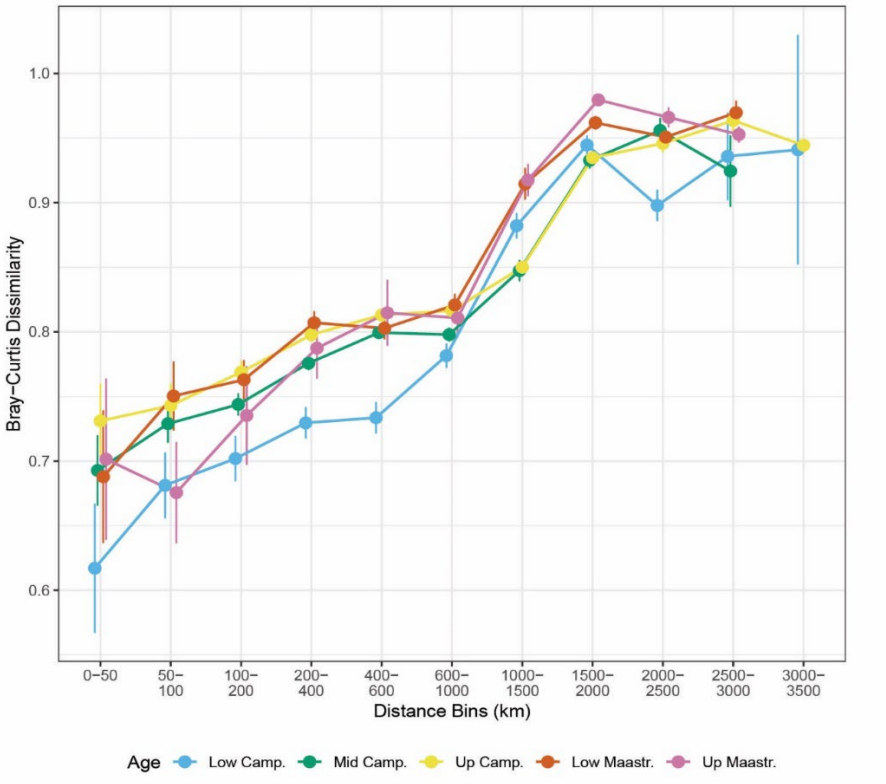



**Figure 5**. Plots of average link weights within geographic distance bins (WGS84 ellipsoid). Bars indicate 95% confidence intervals.

While faunal connectivity within the WIS and between the WIS and GCP decreased over time, MAA average link weights show that GCP faunal similarity increased (Appendix A, Table S6; Figure S21). As an longitudinally broad, open-ocean facing province, the GCP would have experienced normal marine conditions, less latitudinal variation, and the potential for long-distance dispersal, potentially supporting faunal similarity across the MAA by reducing the endemism. Within the WIS province, from the Early CAM to the Early MAA, network connections remained strong between substages across time, indicating weak faunal turnover in the region even as faunal similarity decreased, until the Late MAA when the province disconnects from previous iterations (Figure 2D). The Late MAA disconnect with previous iterations matches expectations of oceanic changes that decrease dispersal and habitat homogenization for WIS taxa (Elderbak & Leckie, 2016; Steel et al., 2012; Fisher et al., 1994; Longman et al., 1998; Slingerland et al., 1996).

*Latitudinal Patterns*

Results do not support latitude-based faunal provinces, despite changes in network metrics across latitudes (Figure 4). However, the region of highest average faunal similarity (HFS) shifts 5° north from the 40-45°N bin to the 45-50°N bin across the R9 regressive event at the end of the Middle CAM (Kauffman & Caldwell, 1993; Figure 2C). Prior to the R9 regression, sea levels were more stable and the HFS region was fixed (Appendix A, Figure S26). This HFS shift appears to reflect biogeographic patterns and is unlikely to be a product of data distributions given that faunal similarity compared to geodesic distance are relatively stable over <1,000 km (Figure 5). The geodesic distance covered by a 5° latitudinal bin in this region is approximately 555 km and the distance covered by two bins is approximately 1,110 km. Thus, similarity begins to strongly decrease over distances greater than 10°

latitude. This suggests regional control on network metrics. If faunal similarity only depended on distance, then similarity should show a uniform pattern across latitude rather than the observed peaks and dips (Appendix A, Figure S25). Therefore, a HFS region that shifts parallel to sea level likely represents a distinct biogeographical component influenced by oceanographic changes.

Similarly, although the region of highest $BC_{ave}$ (indicates highest faunal communication between regions) primarily occupies northern latitudes (40-60°N), it shifts south from the Middle to Late CAM (Appendix A, Figure S25). This region of highest communication may indicate intermediate habitat (Kiel 2016), uniformity of conditions (i.e., water depth), or currents which transport taxa long distances (Lowery et al., 2018). The latter would support larval migration of marine taxa, especially those with long planktonic larval stages (Nickols et al. 2015). However, more specific bathymetric, geochemical, and sedimentological evidence for habitat conditions is sorely needed, but outside the scope of this analysis. The region of highest $BC_{ave}$ may also correspond to Kauffman's (1984) mixing zone. Despite shifts in highest $BC_{ave}$, all networks show a minor or major peak in the central WIS (45-50°N; Appendix A, Figure S25) corresponding with a region of mixing water masses (Lowery et al. 2018). This supports oceanographic or habitat controls on fauna in the WIS.

**Conclusions**

Network analysis of the Late Cretaceous WIS and GCP supports a single biogeographic province throughout the CAM and an independent GCP province in the MAA with decreasing faunal connectivity through time. This contrasts with Kauffman's (1984) original division of the region into four "subprovinces." Decreasing faunal similarity over the

study interval is consistent with oceanographic and geochemical changes that restricted the WIS and exacerbated non-normal marine conditions. Though no overarching relationship between faunal associations and latitude are observed, regional movement of HFS and highest $BC_{ave}$ suggest environmental changes (i.e., falling sea levels and associated effects) were the primary control on biogeographic connections. We provide further evidence for the utility of network modeling to quantitatively characterize paleobiogeographic trends on evolutionary timescales relative to major environmental shifts, an important analytical tool in modern tests of marine biogeographic change under predicted global change.

**Acknowledgements**

Thanks to Timothy Nagle-McNaughton, Steffen Kiel, and Xi Gong for helpful discussions and assistance with methods and visualization. Thanks also to Steve Holland and two anonymous reviewers whose comments improved this manuscript. CP acknowledges support from NSF #2021744; CEM acknowledges support from NSF #1924807.

# CHAPTER 3: A NETWORK ANALYSIS APPROACH TO UNDERSTANDING TRENDS IN FUNCTIONAL DIVERSITY IN THE LATE CRETACEOUS OF NORTH AMERICA

**Ceara Purcell[1], Naveen Balaji[2], Carlie Pietsch[3], Corinne Myers[1]**

[1]*Department of Earth and Planetary Sciences, Northrop Hall, 221 Yale Blvd NE University of New Mexico, Albuquerque, New Mexico 87131 U.S.A. Email: cpurcell@unm.edu*

[2]*Data Engineer, BBInsight, 1750 East Golf Road, Suite 232, Schaumburg, Illinois 60173 U.S.A. Email: nvnbalji@gmail.com.*

[3]*Department of Geology, Duncan Hall, One Washington Square, San Jose University, San Jose, California 95192 U.S.A. Email: carlie.pietsch@sjsu.edu*

## Abstract.

Functional diversity is a complex aspect of macroecology that may respond to environmental changes differently than taxonomic diversity across space and time. To investigate spatiotemporal patterns of functional diversity over geologic time we utilize paleontological data from the environmentally diverse North American Western Interior Seaway and Gulf Coastal Plain through the Campanian and Maastrichtian geologic stages. Invertebrate taxa are classified into functional entities (FEs) based on their motility, attachment style, tiering level, and feeding strategy and analyzed across four levels of spatial aggregation over five geologic sub-stages. Network models of functional assemblages are also used to assess functional biogeographic patterns. Results indicate that regional functional diversity follows previously observed global trends of stability across geologic intervals and supports high functional redundancy (i.e., packing of genera) within specific FEs, buffing against loss. Unlike present-day taxa, however, the number of Late Cretaceous marine FEs do not decline across a latitudinal temperature gradient. Biogeographic patterns in functional diversity are instead associated with paleo-oceanographic conditions and the

resulting habitat differences between the Western Interior Seaway and Golf Coastal Plain. These analyses provide insight into the spatiotemporal consistency of functional diversity patterns relevant to understanding current, ancient, and future Earth-life dynamics.

## Introduction

The relationship between functional diversity and environmental change, while relatively poorly understood, is relevant to current, ancient, and future Earth systems. Given a relationship does exist between the two, how abiotic factors influence functional diversity under different circumstances has not been well constrained and functional diversity has previously been found to respond differently across space and time relative to environmental influences (Foster and Twitchett 2014; Dunhill et al. 2018; Edie et al. 2018; Schumm et al. 2019; Pimiento et al. 2020). For example, functional diversity seems to be little affected during mass extinction events despite severe loss of taxonomic diversity (Foster and Twitchett 2014; Dunhill et al. 2018; Edie et al. 2018). This is potentially due to functional redundancy, which is defined as the number of taxa within a given functional entity (FE) (Pimiento et al. 2020). In recent ecosystems functional diversity trends are negatively correlated with the latitudinal diversity gradient (LDG), that is, decreasing at greater latitudes, regardless of functional redundancy (Edie et al. 2018; Schumm et al. 2019; Floyd et al. 2020; Forsyth and Gilbert 2021). Identifying factors that may influence these patterns and how they differ during background intervals (versus mass extinction intervals) in the geologic record are relevant to a better understanding of macroecology on evolutionary timescales. Paleontological data is uniquely able to assess patterns of functional diversity across both space and time (Roopnarine and Angielczyk 2015; Muscente et al. 2018; Foster et al. 2020).

The latest Cretaceous (~ 85 – 66 Ma, Campanian-Maastrichtian) Western Interior Seaway (WIS) and Gulf Coastal Plain (GCP) of North America are characterized by a spatiotemporally dense fossil record of marine invertebrate assemblages living across a 45-degree latitudinal gradient (Caldwell 1974; Kauffman 1984; Slattery et al. 2013). Fluctuating sea level across the interval impacted marine geochemistry, water mass distributions, and biogeography (Figure 6) (e.g., He et al., 2005; Kauffman, 1984; Lowery et al., 2018; Purcell et al., 2023, etc.) (Figure 6). The WIS itself was characterized by an asymmetrical foreland basin geometry with a deep, siliciclastic foredeep to the west, a latitudinally intermittent central forebulge, and a broad, shallow carbonate shelf to the east (Slattery et al. 2013; Minor et al. 2022). Furthermore, variable restriction between the WIS and GCP promoted distinct habitat conditions between the regions, wherein the WIS experienced episodes of a stratified, potentially brackish, and geochemically distinct restricted seaway system and the GCP represented a normal-marine, unrestricted shoreline (Fisher et al. 1994; Slingerland et al. 1996; Longman et al. 1998; Elderbak and Leckie 2016; Lowery et al. 2018; Dwyer 2019). The Late Cretaceous WIS and GCP are therefore excellent regions in which to explore the spatiotemporal patterns of functional diversity and their relationship to macroecological patterns.

The Late Cretaceous faunal record in these regions is characterized by abiotic change and background levels of biotic turnover. Functional entities are globally stable across the end-Cretaceous mass extinction (KPg) itself (e.g., Edie et al., 2018), but none have yet analyzed regional changes in functional diversity during background intervals of the Late Cretaceous. Moreover, latitudinal changes in functional diversity have been poorly studied in the fossil record. Though present-day functional diversity declines as latitude increases (Edie et al.

2018; Schumm et al. 2019), the latest Cretaceous is characterized by a relatively flat latitudinal

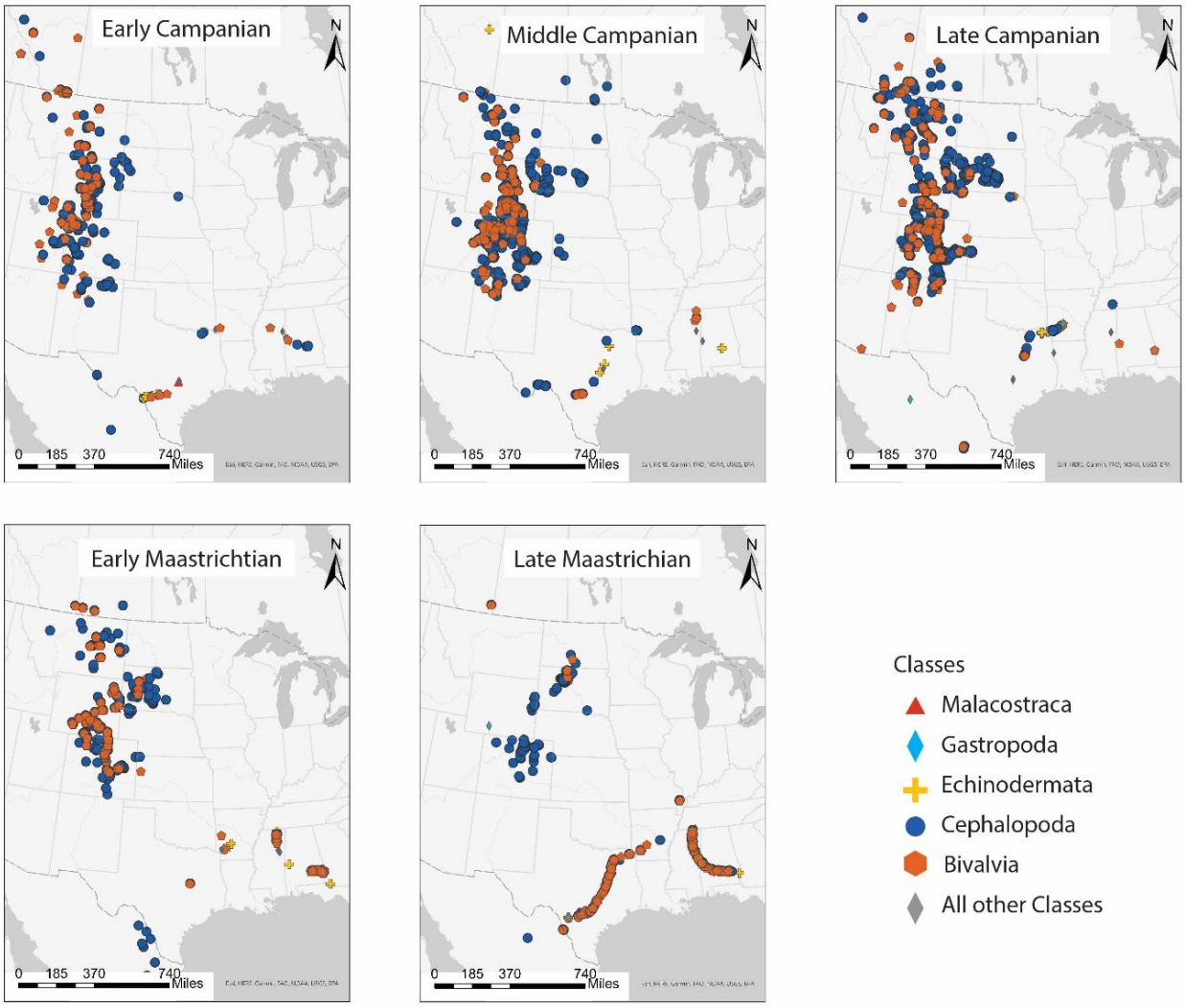



**Figure 6.** Campanian to Maastrichtian data locations.

temperature gradient that may present a more uniform distribution of FEs compared to today (Mannion et al. 2012, 2014; Super et al. 2018). Dramatic oceanographic differences between the GCP and the WIS have additionally been found to correspond with marine provinciality and altered biotic connectivity (Purcell et al. 2023). Regional differences in abiotic factors may therefore result in distinct Late Cretaceous patterns in the structure of functional

assemblages, or the types of FEs present in a specific time and place (Schumm et al. 2019). Global functional diversity analyses, which have previously been the focus of most paleo-FE studies, may also introduce longitudinal variations that overprint regional patterns (Mannion et al. 2014; Schumm et al. 2019). Generally, the influence of data aggregation across space on functional diversity analyses has not been rigorously tested. By focusing on a single, well-studied region such as the WIS/GCP, we can address these aspects of the biogeography FEs over time.

Thus, we use a well-vetted fossil occurrence database representing approximately 18 million years of an ecologically diverse suite of marine invertebrates, to observe how functional diversity relates to major environmental shifts in space and time within the WIS and adjacent GCP. We hypothesize that 1) functional richness and evenness are stable across space and within a single temporal interval, and across the minor sea level oscillations and environmental changes present during the Campanian and Maastrichtian; 2) the WIS and GCP lack notable biogeographic differences in specific FEs, despite previously observed taxonomic provinciality and habitat differences; 3) functional redundancy is associated with FE stability, both in terms of specific FEs presence and absence and in terms of FE richness and evenness, across space/time, and 4) these patterns are observable at several levels spatial aggregation. By documenting spatiotemporal FE stability we can contribute to a better understanding of how functional diversity connects with FE stability, habitat- and dispersal-based provinciality, and illustrate the influence of spatial resolution on functional diversity analyses. We demonstrate that packing of species and genera into FEs leads to stability of functional guild structure throughout the study interval. A major sea level regression and disconnect of the WIS and GCP leads to paleoenvironmental change and regional extirpation

of FEs. Thus, our research better illuminates the biogeography of functional diversity and its structure during a background interval that leads into a major taxonomic depletion event: the end-Cretaceous mass extinction (Edie et al. 2018). This work informs both modern and ancient understanding of functional diversity dynamics and its relationship to macroecology.

## Methods

Fossil occurrence data consists of marine invertebrates from WIS and GCP Campanian and Maastrichtian strata compiled from digital databases, including the Paleobiology Database (08/25/21 download) and iDigBio (08/30/21 download), and records from museum collections at the Black Hills Institute, USGS William A. Cobban Cretaceous Ammonite Collection, and a thesis database (Mackenzie 2007). All data were vetted to remove incomplete, erroneous, or poorly spatially and temporally resolved occurrences and to update taxonomic nomenclature. Stratigraphic ages were vetted to the substage level based on stratigraphic and biozone information or from literature sources (see Purcell et al. 2023). Spatial resolution that could not be confirmed to less than 30km of uncertainty were removed to avoid spatial ambiguity and records east of -80° longitude and south of 23° latitude were removed to avoid records from outside the WIS/GCP region.

The final dataset of 32,864 fossil occurrences includes a total of 552 genera of bivalves, cephalopods, gastropods, echinoderms, corals, brachiopods, bryozoans, and crustaceans (see Purcell et al. 2023 for additional details and references used in dataset vetting). These groups represent a wide variety of lifestyle modes both within clades and between distinct phylogenetic groups. They are furthermore important components in both modern and ancient shallow marine ecosystems and therefore excellent test subjects for ecological analysis. Taxa were binned into the Early, Middle, and Late Campanian and the

Early and Late Maastrichtian substages based on stratigraphic information. Tests confirming signal over noise in this dataset were previously published in Purcell et al. (2023).

Genera were classified into functional entities (FEs) based on 1) mobility and attachment style, 2) lifestyle mode or tiering level, and 3) feeding strategy following Bambach et al. (2007;see also Bush & Bambach, 2011) (Table 4). Characteristics were assigned based on extant relatives or higher order traits (i.e., Family or Order level traits, see excel file Table S36) (i.e., Family or Order level traits; e.g., Aberhan & Kiessling, 2015; Edie et al., 2018; Foster et al., 2020; Foster & Twitchett, 2014; Sessa et al., 2012). Taxa were grouped into FEs using the mFD package in R (Magneville et al. 2022). Functional diversity metrics were calculated using five levels of spatial aggregation, or units: regional (WIS+GCP, unit size 1), basinal (WIS and GCP, separately, unit size 2), 5° paleolatitudinal bins (unit size 3), 360km$^2$ grid cells (unit size 4), and 60km$^2$ grid cells (unit size 5). These units assess how functional diversity metrics change across space, test sensitivity to spatial binning, and allow assessment of patterns that may emerge at different spatiotemporal scales of analysis. Aggregations in 60km$^2$ and 360km$^2$ grid cells (referred to as "nodes" in Kiel 2017) are intended to capture local paleo-"community" structure, though they do not necessarily scale to present-day ecological communities or homogeneous environments. Unit locations were estimated based on the average location of occurrences they contain, converted to paleocoordinates using the chonosphere package in R (Kocsis and Raja 2020).

**Table 4.** Table of the FE present in the database and their code values.

| Mobility level | Code | Description |
|---|---|---|
| Mobile, unattached | MU | Includes actively mobile, freely mobile, slow moving, fast moving, and creeping unattached taxa. |
| Facultatively mobile, unattached | FU | Includes facultatively mobile and unattached taxa. |
| Facultatively mobile, attached | FA | Includes facultatively mobile and swimming taxa with some form of attachment (byssate, pedicle, or cementing) |
| Immobile, unattached | IU | Includes free-lying, boring taxa. |
| Immobile, attached | IA | Includes stationary and attached (cemented, epibiont, byssate) taxa. |
| **Feeding mechanism** | **Code** | **Description** |
| Suspension feeding | S | Includes all forms of micro-feeding from the water column (i.e., filter feeding, suspension feeding, etc.). Various diet types are included in this strategy (i.e., carnivore) but it is non-specific. |
| Deposit feeding | D | Includes deposit feeding in the subsurface or at the sediment surface and detritivore feeding. Some taxa have mixed deposit/suspension feeding but just classified as deposit feeding. |
| Herbivore / grazing | H | Includes all forms of herbivorous feeding, including grazing. |
| Carnivore / predatory | C | Includes all macro-carnivorous and predatory feeding strategies. Implies active feeding but includes carnivorous scavenging as well. |
| Omnivore | O | Includes feeding strategies that incorporate both carnivorous and herbivorous feeding. If multiple differing feeding modes were attributed to a taxon, including omnivore, omnivore listed. |
| Photosymbiotic | P | Includes photosymbiotic taxa. |
| Chemosymbiotic | CH | Includes chemosymbiotic taxa, including chemosymbiotic deposit feeders. |
| **Tiering Level/Lifestyle Mode** | **Code** | **Description** |
| Boring | B | Includes boring taxa and nestlers. |
| Epifaunal | E | Includes all taxa living above the sediment-water interface (non-burrowing) that do not have upward mobility in the water column (includes all epifaunal levels, i.e., intermediate epifaunal) |
| Semi-infaunal | PI | Includes all taxa which living partially beneath the sediment surface and taxa which both burrow and live epifaunally. |
| Infaunal | I | Includes all taxa living below the sediment surface, either shallow, deep, or at uncertain infaunal depths |
| Nektonic | N | Includes all taxa living freely in the water column, well above the sediment surface. |
| Nekto-benthic | NB | Includes all taxa that live in the water column close to the sediment surface or those with uncertainty as to which level of the water column they occupy. |

Generic richness, functional richness ($\alpha_F$; the number of unique FEs), and functional evenness were calculated for each spatiotemporal level (i.e., five time bins at five spatial scales). Functional evenness, or the relative abundance of genera within FEs, was calculated using Simpson's Measure of Evenness (SME). SME is calculated by normalizing the inverse Simpson's Diversity Index by the number of total FE present:

$$Simpson's\ Diversity\ Index\ (D) = \frac{1}{\sum {p_i}^2} \qquad (EQ\ 1)$$

$$Simpson's\ measure\ of\ evenness = \frac{D}{S} \qquad (EQ\ 2)$$

Where $p_i$ represents the proportion of genera in a given FE and S is the total number of FEs present. Values closer to 0 indicate lower evenness and higher values indicate higher evenness (SME has no upper limit); no value is assigned for units containing a single FE. Similarly, SME cannot be calculated if all FEs contain a single genus. In both cases, units were given SME values of 0. SME is weighted by the abundance or dominance of a specific FE and was chosen because it is less sensitive to taxon-richness and provides an estimate of the generic abundance distribution (Magurran, 2003; see SI for discussion of other metrics tested). Richness and evenness values were compared between temporally and spatially adjacent units by calculating the proportional change in $\alpha_F$ and SME values.

Fundamental sources of bias in the database were tested using several methods. The relationship between generic and functional richness was analyzed by correlating the two variables based on 60km unit aggregates. Bootstrapping of 60km unit $\alpha_F$ values based on the number of unique genera were performed to determine how $\alpha_F$ varies with generic richness ($\alpha_G$). Additionally, SQS subsampling (Alroy 2010) of the FEs in each substage was

performed to find the generic sample size estimates appropriate for capturing the majority of FEs (see SI page 2).

Functional entity assemblages within spatiotemporal units were compared below the regional level using network modeling in the EDENetworks software (Moalic et al. 2012; Kivelä et al. 2015) following methods outlined in Kiel (2017) and Purcell et al. (2023). EDENetworks for presence-absence data utilizes Bray-Curtis dissimilarity values to assign weights to the links between each unit in the network, allowing connections to be viewed at different threshold values of dissimilarity. Betweenness centrality, a measure of the degree to which a unit acts as a geographic connection between regions, was also calculated (Kivelä et al., 2015). Network topology patterns at the percolation point were used to determine if FE assemblages were provincial (i.e., geographically controlled) given the assumption that network components represent distinct biologically-based "community" groups (Newman 2012). The percolation point depicts the threshold at which all components in a network are connected without forming a single component (Kivelä et al. 2015). The specific FEs and associated PBDB paleoenvironmental characteristics of occurrences within components found at and just below percolation were qualitatively assessed (Appendix B, Table S35). To determine if the WIS and GCP differ in $\alpha_F$, functional diversity, and FEs present, the two basins (unit size 2) were analyzed separately using all previously described methods. Differences in the three most common clades (cephalopods, bivalves, and gastropods) were also compared. Other clades made up too few of the occurrences to form useful networks. See supplementary materials for details regarding methods and results (SI, page 23).

## Results

### *Relationship Between Functional and Generic Diversity*

Generic and functional richness have a strong positive correlation that loses strength at approximately more than 100 genera (Appendix B, Figure S1). Generic richness ($\alpha_G$), previously considered a proxy for sampling bias in this database (Purcell et al., 2023), is therefore not representative of potential $\alpha_F$ when more than ~100 genera are present. Furthermore, Bootstrap analysis of 60km units indicate standard error of $\alpha_F$ in units has a positive correlation with unit $\alpha_F$ that depreciates above approximately 20 FEs (Appendix B, Figure S3). This suggests that the distribution of genera in FEs becomes more variable (i.e., less even) when more than 20 FEs are present; that is, functional evenness should be more biased in units with < 20 FEs, while also poorly representing potential $\alpha_F$. However, SQS analysis of substages, which was used to determine what level of $\alpha_G$ best captures the majority of common FEs, indicates that $\alpha_G > 16$ in the Campanian and $\alpha_G > 18$ in the Maastrichtian are likely to capture the *most common FEs* (Appendix B, Table S1). Therefore, spatiotemporal units with fewer than 100 unique genera are under-sampled relative to $\alpha_F$ and functional evenness but units with more than 16 and 18 unique genera may reliably represent FE assemblages (paleo-"communities") for the Campanian and Maastrichtian, respectively.

All units containing less than 16 or 18 unique genera for the Campanian and Maastrichtian, respectively, were removed given that they are not considered to be representative of either FE assemblages or functional diversity metrics. This includes 67 out of 533 units at 60km (unit size 5), 80 out of 132 units at 360km (unit size 4) aggregated spatial resolution, and seven out of 32 paleolatitude bins (unit size 3). Removal of these

under-sampled units was not found to notably alter results at any level of aggregation. Figures and tables including all units regardless of generic richness are available in the SI.

### *Substage-level Diversity:*

There are a total of 38 FEs in the database (Appendix B, Table S2): 30 each in the Early Campanian and the Middle Campanian, 33 in the Late Campanian, 35 in the Early Maastrichtian, and all 38 in the Late Maastrichtian, representative of 18 taxonomic classes (Table 4 & 5). Functional entities are overwhelmingly assigned at taxonomic levels below class. Regional $\alpha_F$, in unit size 1, is highly stable, rising only slightly (~6% to 10%) across any two substages while generic richness is more variable, rising across the Campanian, before falling between the Late Campanian and Early Maastrichtian and rising through the Late Maastrichtian (Figure 6A, Appendix B, Table S3). Functional evenness (SME) decreases 50% between the Middle and Late Campanian and fluctuates slightly across the other substages (Figure 6B). SME change only corresponds positively with generic richness in the Early to Middle Campanian while $\alpha_F$ is stable (Appendix B, Table S3).

Generic and FE diversity are furthermore disconnected through time. Only three FEs are added between the Middle to Late Campanian, despite the sharp increase in generic richness (81%) (Figure 8). Simultaneously, mobile unattached epifaunal carnivores (MU-E-C in Figure 4; composed of gastropods and crustaceans) rose in dominance from ~8% to ~30% in proportional generic richness, while facultatively mobile unattached infaunal suspension feeders (FU-I-SF in Figure 8; composed of bivalves) fell from 17% to 10% (Appendix B, Table S4). No other FE experienced a change in proportional generic diversity of more than 5% across the interval. The few shifts in generic occupation of FEs caused functional evenness to decline sharply. Following the Late Campanian, MU-E-Cs remain dominant;

immobile attached epifaunal suspension feeders (IA-E-SF in Figure 8) and FU-I-SF, both previously the most diverse FEs, never again reach their former degree of dominance (Appendix B, Table S4). All other FEs have low proportional generic diversity throughout all five time bins (<9%). At any given substage no more than eight FEs are "gained"

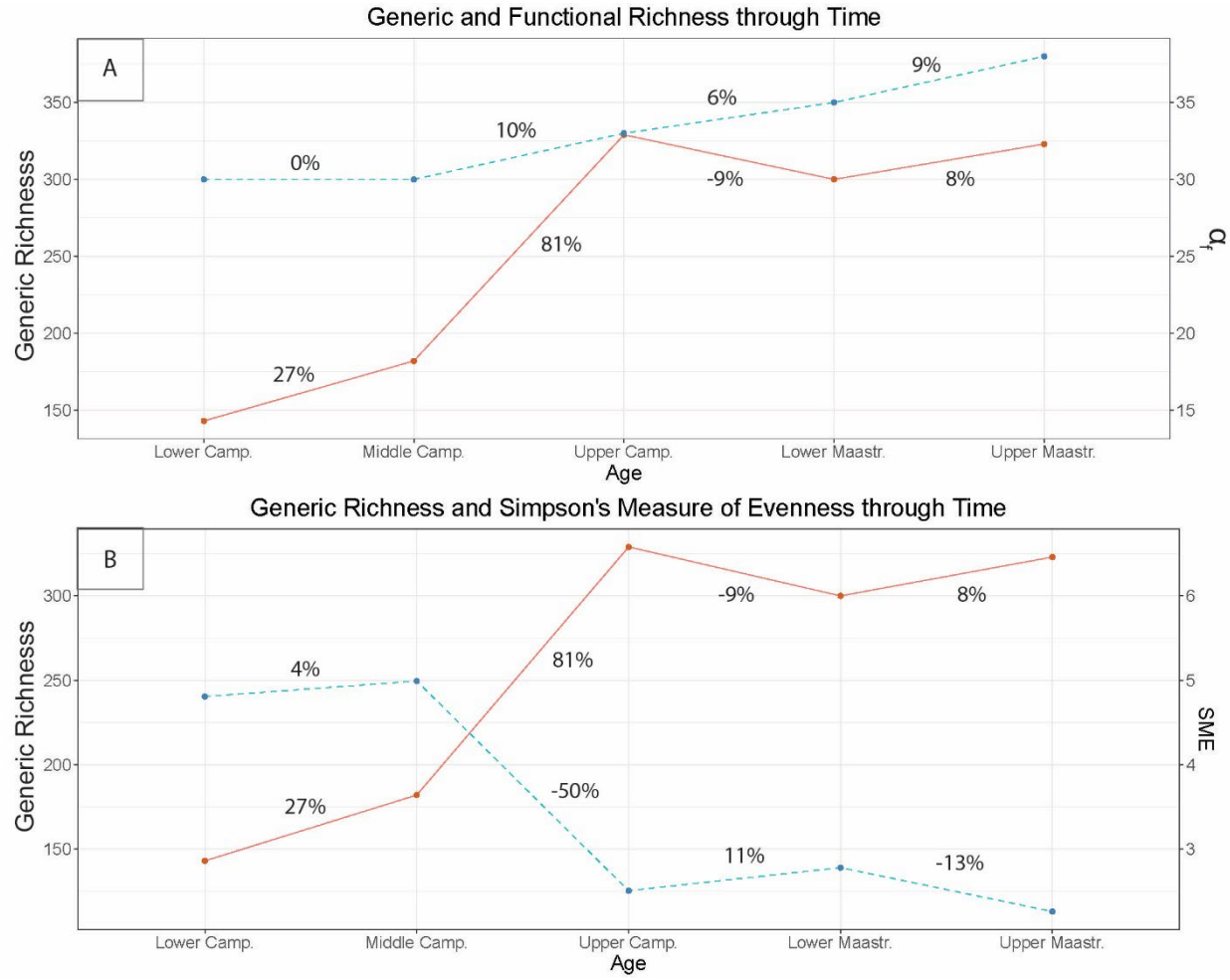



**Figure 7.** Line plots of substage values for A) generic richness (red, solid line) and functional richness (blue, dotted line) and B) generic richness (red, solid line) and SME (blue, dotted line). The left y-axis in all plots depicts generic richness values, while the right y-axis presents values for the other metric. Proportional change in values is denoted as percentages. Camp. = Campanian, Maastr. = Maastrichtian.

**Table 5**. List of taxonomic classes and the FE they represent. See Table S36 for details on genera. Asterisks denote FEs represented by a single class.

| FE | Class | FE | Class |
|---|---|---|---|
| IA-E-P | Anthozoa | IA-B-SF | Demospongea |
| IA-E-SF | Anthozoa | IA-E-SF | Demospongea |
| MU-PI-C | Asteroidea | MU-E-H | Echinoidea |
| MU-N-C | Bivalvia | MU-E-O | Echinoidea |
| *FU-I-C | Bivalvia | *MU-PI-D | Echinoidea |
| *FU-I-CH | Bivalvia | MU-I-D | Echinoidea |
| *MU-I-CH | Bivalvia | MU-E-C | Gastropoda |
| *FA-I-D | Bivalvia | MU-PI-C | Gastropoda |
| *FU-I-D | Bivalvia | MU-I-C | Gastropoda |
| MU-I-D | Bivalvia | *FU-E-H | Gastropoda |
| IA-E-P | Bivalvia | MU-E-H | Gastropoda |
| *IA-PI-P | Bivalvia | *FU-PI-H | Gastropoda |
| *IU-PI-P | Bivalvia | *MU-PI-H | Gastropoda |
| *FU-B-SF | Bivalvia | MU-E-O | Gastropoda |
| IA-B-SF | Bivalvia | FA-E-SF | Gastropoda |
| *IU-B-SF | Bivalvia | FU-E-SF | Gastropoda |
| FA-I-SF | Bivalvia | *MU-E-SF | Gastropoda |
| *FU-I-SF | Bivalvia | IA-E-SF | Gymnolaemata |
| *IU-I-SF | Bivalvia | IA-E-SF | Hexanauplia |
| FA-E-SF | Bivalvia | MU-E-O | Homarus |
| FU-E-SF | Bivalvia | IA-E-SF | Hydrozoa |
| IA-E-SF | Bivalvia | FA-I-SF | Lingulata |
| *IU-E-SF | Bivalvia | MU-E-C | Malacostraca |
| *FA-PI-SF | Bivalvia | MU-I-C | Malacostraca |
| *IA-PI-SF | Bivalvia | MU-I-D | Malacostraca |
| *MU-I-SF | Bivalvia | *MU-E-D | Malacostraca |
| *IA-I-SF | Bivalvia | MU-E-O | Malacostraca |
| IA-E-SF | Bryozoa | MU-E-C | Merostomata |
| MU-N-C | Cephalopoda | IA-E-SF | Rhynchonellata |
| *MU-NB-C | Cephalopoda | MU-I-D | Scaphopoda |
| MU-N-SF | Cephalopoda | IA-E-SF | Stenolaemata |
| *MU-NB-SF | Cephalopoda | IA-E-SF | Thecostraca |
| IA-E-SF | Crinoidea | | |

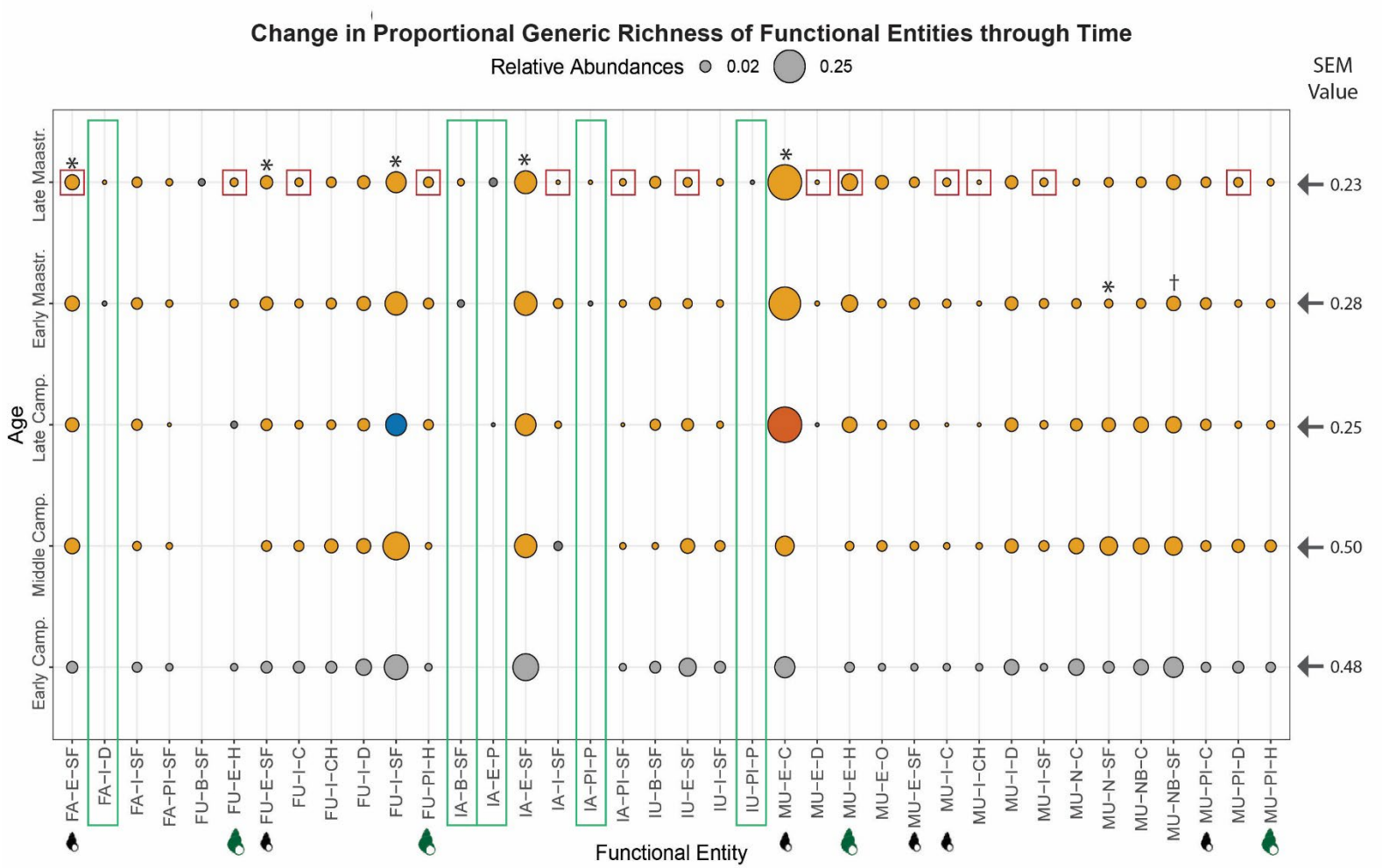



**Figure 8.** Plot of FE relative generic richness through time. Circle size indicates relative generic richness within each substage, color indicates a proportional increase (red) greater than 0.05, decrease (blue) greater than 0.05, change less than 0.05 (orange) from the previous substage. Grey indicates a value cannot be compared with a previous substage. Green rectangles denote FEs only found in the GCP and red squares denote FEs missing from the Late Maastrichtian WIS. Gastropod silhouettes denote gastropod-type FEs; larger green silhouettes indicate herbivorous FEs absent in the Late Maastrichtian 360km WIS network component. FEs that are present in every 60km unit just below percolation in each substage are denoted with an asterisk (GCP component) and a cross (WIS component). See Table 4 for FE naming key.

(i.e., not observed in the previous interval but currently present) and only two FEs are "lost" (i.e., previously observed but missing in the current interval). All FEs lost/gained have very low proportional generic diversity (Figure 8). The two FEs that are unobserved in a substage but present in the preceding and succeeding substages are facultatively unattached epifaunal herbivores (FU-E-H) and immobile attached epifaunal photosymbionts (IA-E-P), in the Middle Campanian and Late Maastrichtian, respectively (Figure 8, Appendix B, Table S4). While both FEs are among the least diverse FEs within the substage prior to their apparent extirpation (each represented by a single genus), in the succeeding substage they are not the least diverse and are represented by these same genera as well as either one (FU-E-H) or two (IA-E-P) additional genera.

*WIS vs GCP Influence*

The WIS, analyzed independently (unit size 2), has fairly stable $\alpha_F$ (4% and 7% increase) in the Campanian, but declines from the Late Campanian through the Maastrichtian (13-30% decline; Appendix B, Figure S4-S5, Table 6). WIS generic richness parallels this pattern with greater proportional change. In the GCP generic and functional richness increase throughout Appendix B, (Figure S6-S7), particularly between the Middle and Late Campanian, indicating that the GCP data strongly influences richness trends in the overall database. WIS and GCP SME change inversely to generic richness, even when $\alpha_F$ is stable or increases (Appendix B, Table S6 and S8).

Of the eight FEs "gained" across the study interval, five are only found in the GCP (Figure 8): three photosymbionts, one boring suspension feeder, and an attached infaunal deposit feeder. No photosymbionts and only one boring FE are present in the WIS at any

time bin. Sampling is not a reasonable driver of FE differences between the WIS and GCP. All WIS

**Table 6.** Table of regional generic richness, functional richness (αf), and functional evenness (SME) values in each substage. Values for the overall database are given first, followed by the values observed in the WIS and then GCP. Camp. = Campanian, Maastr. = Maastrichtian.

| | Generic Richness | Functional Richness ($\alpha_f$) | Functional Evenness (SME) |
|---|---|---|---|
| **Early Camp.** | 143 (108/47) | 30 (28/14) | 0.48 (0.53/0.61) |
| **Mid Camp.** | 182 (151/65) | 30 (29/22) | 0.50 (0.51/0.67) |
| **Late Camp.** | 329 (199/213) | 33 (31/29) | 0.25 (0.41/0.21) |
| **Early Maastr.** | 300 (119/248) | 35 (27/34) | 0.28 (0.51/0.26) |
| **Late Maastr.** | 323 (42/309) | 38 (19/38) | 0.23 (0.78/0.22) |
| **Mean** | **255.4 (123.8/176.4)** | **33.2 (26.8/27.4)** | **0.35 (0.55/0.39)** |
| **Median** | **300 (119/213)** | **33 (28/29)** | **0.28 (0.51/0.26)** |
| **S.D.** | **86.60 (57.79/115.3)** | **3.42 (4.60/9.58)** | **0.13 (0.14/0.23)** |

substages have more than adequate sampling to produce representative FE assemblages (i.e., > 16 and 18 unique genera for the Campanian and Maastrichtian, respectively), and all but the Late Maastrichtian WIS dataset have adequate generic richness to capture potential $\alpha_F$ (i.e., >100 unique genera). No FEs are completely lost between the Early Campanian and Late Maastrichtian in the GCP, however, several FEs are absent in the Campanian of the GCP where generic richness is frequently <100; this suggests that these absent FEs may be caused by poor sampling (Appendix B, Figure S7; Table 6 and S10). That said, generic richness in all GCP substages is adequate to capture the 20 most common FEs and the FEs with the highest overall proportional generic diversity providing a baseline community structure throughout the study interval.

*Class Influences.*

Cephalopods represent 85 unique genera and 13,170 fossil occurrences. They are characterized by only four FEs, all of them mobile and unattached, nektonic or nektobenthic, and carnivorous or suspension feeders (Table 5). All four FEs are present throughout the study interval and are fairly even (<16% SME change) despite changes to cephalopod generic richness (5-54% generic richness change; Appendix B, Figure S8-S9, Table S12). Bivalves make up 184 unique genera, 14,085 fossil occurrences, and represent 24 FEs in the database. Bivalve generic and functional richness generally increase across the study interval, but SME values are more stable, indicating even dispersal of additional genera to bivalve FEs (Appendix B, Figure S10-S11). Bivalve SME only changes notably between the Early and Middle Campanian, decreasing by ~23% while generic diversity increases by 33% and $\alpha_F$ increases by 6% (Appendix B, Table S14). Gastropods, which make up 184 unique genera, 4664 fossil occurrences, and represent 11 FEs, are less stable than bivalves or cephalopods (Appendix B, Figure S12-S13). The number of gastropod genera rises sharply between the Middle and Late Campanian (408%), resulting in a sharp increase in $\alpha_F$ (57%), and a clear decline in SME (66%; Appendix B, Figure S12, Table S16).

## *Functional versus Generic Diversity Across Distance and Spatial Scales:*

*Paleolatitudinal Bins (unit size 3).*

When the data are aggregated by paleolatitudinal bins and compared between substages, patterns of proportional change in all metrics are dissimilar from the regional analysis (unit size 1) except between the Middle and Late Campanian (Table 7 and 8; Appendix B, Figure S15). Proportional change across paleolatitudes bins (i.e., LDG comparison) indicates generally decreasing generic and functional richness and increasing

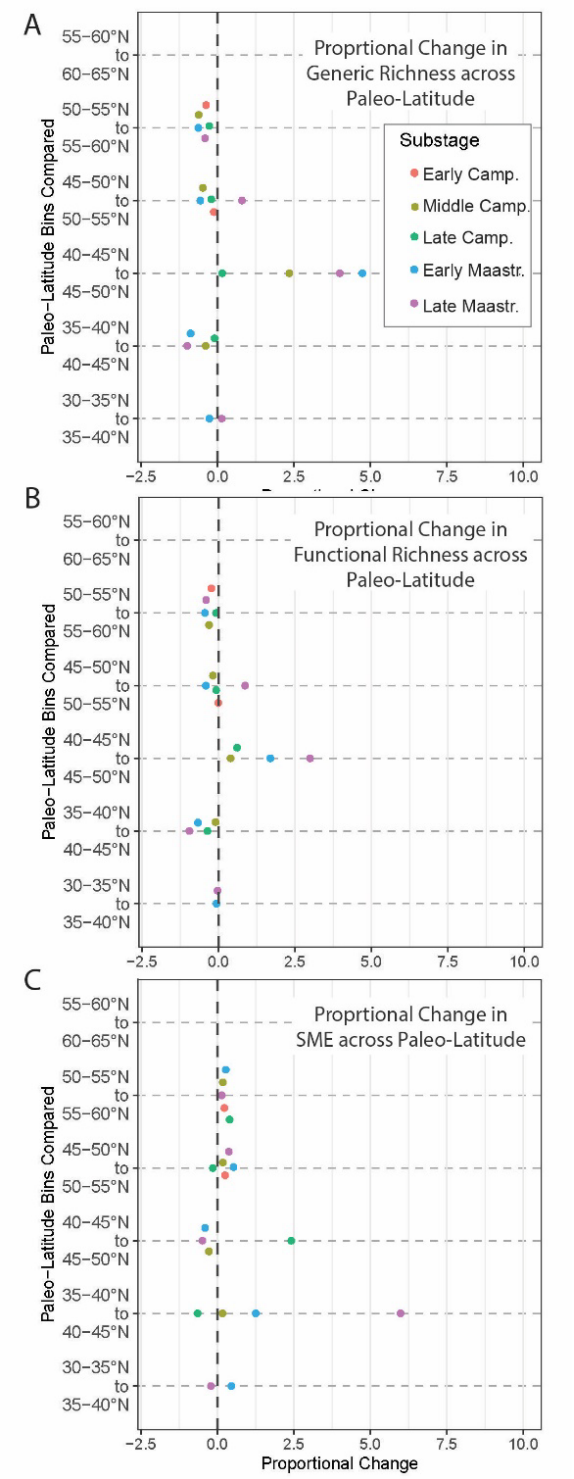



**Figure 9.** Beeswarm plot of proportional change in A) generic richness, B) functional richness (αf), and C) functional evenness (SME). Values reflect change in metric between adjacent paleolatitude for each substage. Camp. = Campanian, Maastr. = Maastrichtian.

**Table 7.** Table of paleolatitude bin generic richness, functional richness (αf), and functional evenness (SME) and mean, median, and standard deviation values for each substage. NA value indicate no value measured for that unit in that substage. LC = Early Campanian, MC = Middle Campanian, UM = Late Campanian, LM = Early Maastrichtian, UM = Late Maastrichtian.

| | **Generic Richness** | | | | | **Functional Richness ($\alpha_f$)** | | | | | **Functional Evenness (SME)** | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | EC | MC | LC | EM | LM | EC. | MC | LC | EM | LM | EC. | MC | LC | EM | LM |
| 30-35°N | 11 | NA | 14 | **210** | **238** | 4 | NA | 5 | **32** | **37** | 1.15 | NA | 0.83 | **0.26** | **0.27** |
| 35-40°N | 41 | 65 | **142** | **155** | **271** | 14 | 22 | **28** | **30** | **36** | 0.57 | 0.67 | **0.42** | **0.38** | **0.21** |
| 40-45°N | 14 | 40 | **129** | 19 | 3 | 7 | 20 | **18** | 10 | 2 | 1.00 | 0.78 | **0.15** | 0.86 | 1.50 |
| 45-50°N | 67 | **134** | **149** | **109** | 15 | 22 | **28** | **29** | **27** | 8 | 0.56 | **0.56** | **0.51** | 0.51 | 0.77 |
| 50-55°N | 59 | 70 | **119** | 48 | 27 | 22 | 23 | **27** | 16 | 15 | 0.71 | 0.66 | **0.43** | 0.78 | 1.06 |
| 55-60°N | 37 | 27 | 87 | 18 | 16 | 17 | 16 | 25 | 9 | 9 | 0.87 | 0.78 | 0.60 | 1.00 | 1.21 |
| 60-65°N | 5 | 1 | 7 | NA | NA | 5 | 1 | 4 | NA | NA | 0.00 | 0.00 | 0.88 | NA | NA |
| Mean | 33 | 56 | 92 | 93 | 95 | 13 | 18 | 19 | 21 | 18 | 0.69 | 0.58 | 0.54 | 0.63 | 0.84 |
| Median | 37 | 53 | 119 | 79 | 22 | 14 | 21 | 25 | 22 | 12 | 0.71 | 0.67 | 0.51 | 0.65 | 0.92 |
| S.D. | 24.3 | 45.8 | 59.4 | 78.6 | 124 | 7.7 | 9.4 | 10.8 | 10.3 | 15.0 | 0.37 | 0.29 | 0.25 | 0.29 | 0.52 |

**Table 8.** Summary statistics for proportional change in generic richness, functional richness (αf), and functional evenness (SME) for paleolatitude bins across adjacent substages. Camp. = Campanian, Maastr. = Maastrichtian. Values with less than the SQS sample estimate have been removed from these summary statistics.

| | | **Low Camp. - Mid Camp.** | **Mid. Camp. - Up. Camp.** | **Up. Camp. - Low Maastr.** | **Low Maastr. - Up. Maastr.** |
|---|---|---|---|---|---|
| **Generic Richness** | Mean | 0.38 | 1.29 | -0.48 | -0.23 |
| | Median | 0.39 | 1.18 | -0.60 | -0.27 |
| | SD | 0.54 | 0.93 | 0.39 | 0.62 |
| **Functional Richness ($\alpha_f$)** | Mean | 0.21 | 0.19 | -0.30 | -0.20 |
| | Median | 0.16 | 0.17 | -0.41 | -0.03 |
| | SD | 0.28 | 0.25 | 0.29 | 0.44 |
| **Functional Evenness (SME)** | Mean | 0.00 | -0.37 | 1.23 | 0.24 |
| | Median | -0.03 | -0.35 | 0.66 | 0.28 |
| | SD | 0.12 | 0.27 | 2.02 | 0.41 |

evenness across paleolatitude from low to high latitudes, except between the 40-45°N to 45-50°N bins where richness increases and evenness decreases in all substages (Figure 9). However, these patterns are considered to be strongly influenced by sampling bias, as will be discussed more below.

*360km and 60km Units (unit size 4).*

Temporal patterns apparent at the substage and paleolatitudinal bin level are present but dampened at higher spatial resolutions (360km and 60km units; Figure 10, Appendix B, Figures S18-19; Table S30 and S32). Proportional changes in generic and functional richness between 360km and 60km units are not well supported by linear regression but indicate that the two values are poorly correlated across distance in the Campanian and positively correlated (i.e., increasing with increased distance) in the Maastrichtian (Appendix B, Figure S20-S27). Proportional changes in 60km and 360km units SME values across distance are negatively correlated in all substages except in the Early Campanian 360km units where SME is slightly positively correlated with distance (Appendix B, Figure S22 and S26). Thus, SME decreases across distance in most substages and, regardless of spatial resolution, agrees with assumptions about spatial autocorrelation and previous clustering assessments of the database (Purcell et al., 2023). Functional richness only increases with distance in the Maastrichtian, most likely due to the addition of GCP units with high generic richness (Appendix B, Figure S1).

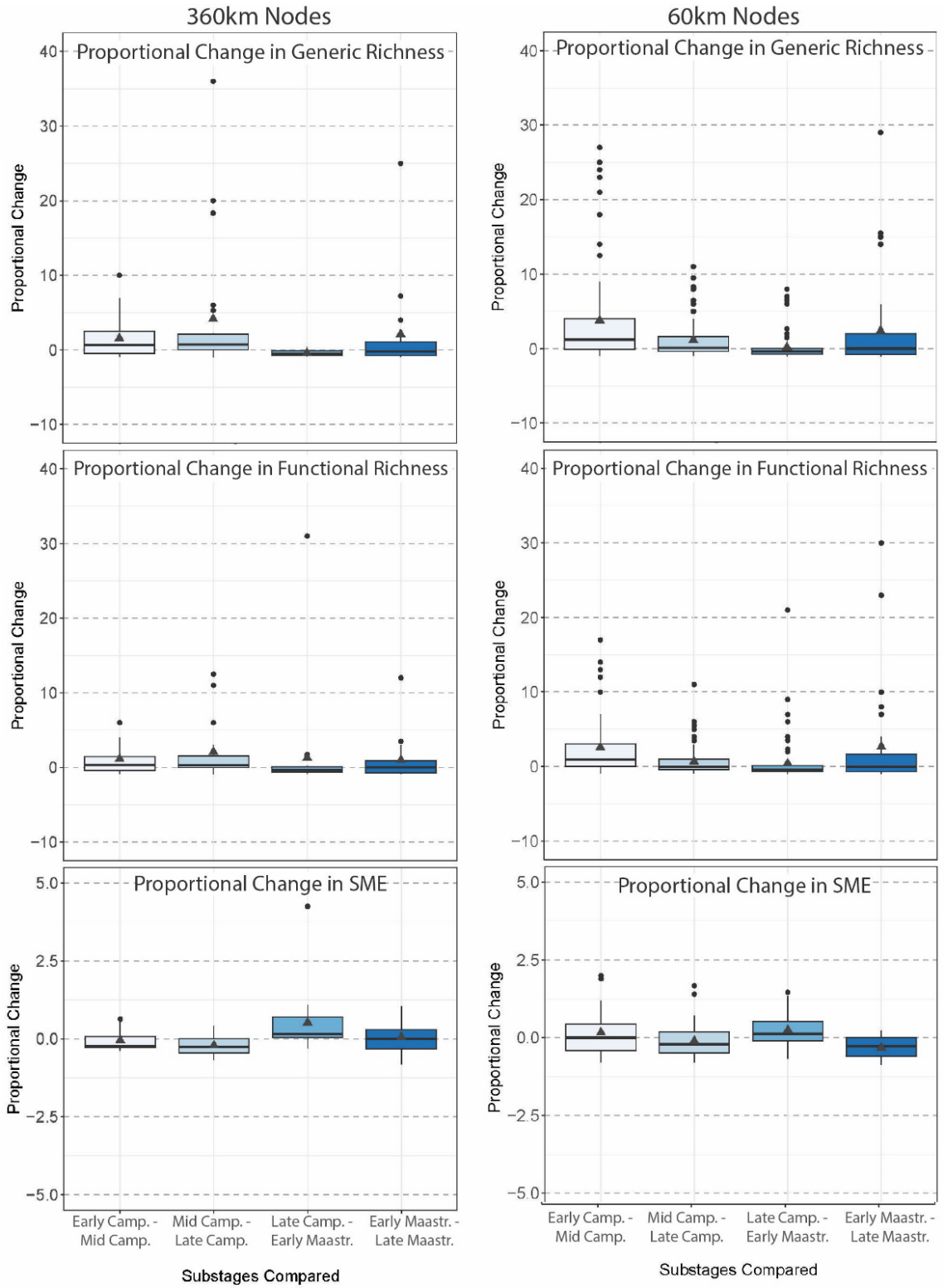



**Figure 10.** Box and whisker plots of proportional change between substages for generic richness, functional richness (αf), and functional evenness (SME) at the 360km and 60km unit scales. Triangles indicate mean values. Camp. = Campanian, Maastr. = Maastrichtian.

*Network models of Functional Assemblages:*

Network models indicate spatiotemporal units are well connected overall across the study region, though regional connectivity diminishes with higher spatial resolution (as expected) and is less robust in the Late Maastrichtian. Paleolatitude bin networks (unit size 3) show strong connections between both adjacent and nonadjacent paleolatitudes for all substages, except in bins with poor sampling (i.e., <SQS threshold; Figure 11). 360km and 60km unit networks (unit sizes 4 and 5, respectively) do not separate into distinct components in any substage except the Late Maastrichtian (Appendix B, Figure S28-S29, 12-14). The 360km Middle Campanian network has two geographically overlapping components, represented by a functionally diverse assemblage and a depauperate assemblage of only mobile unattached nektobenthic suspension feeders (MU-NB-SF in Figure 8; Appendix B, Table S33; Figure S28). However, these components are not observed in the 60km Middle Campanian unit network (Figure 12); this suggests that data aggregation influences network modeling results methodologically (e.g., via percolation points) and may not relate directly to biological associations. The 60km Early Maastrichtian network also has distinct geographically overlapping components representing functionally diverse and functionally depauperate assemblages, most likely due to sampling differences between units (Figure 12; Appendix B, Table S34).

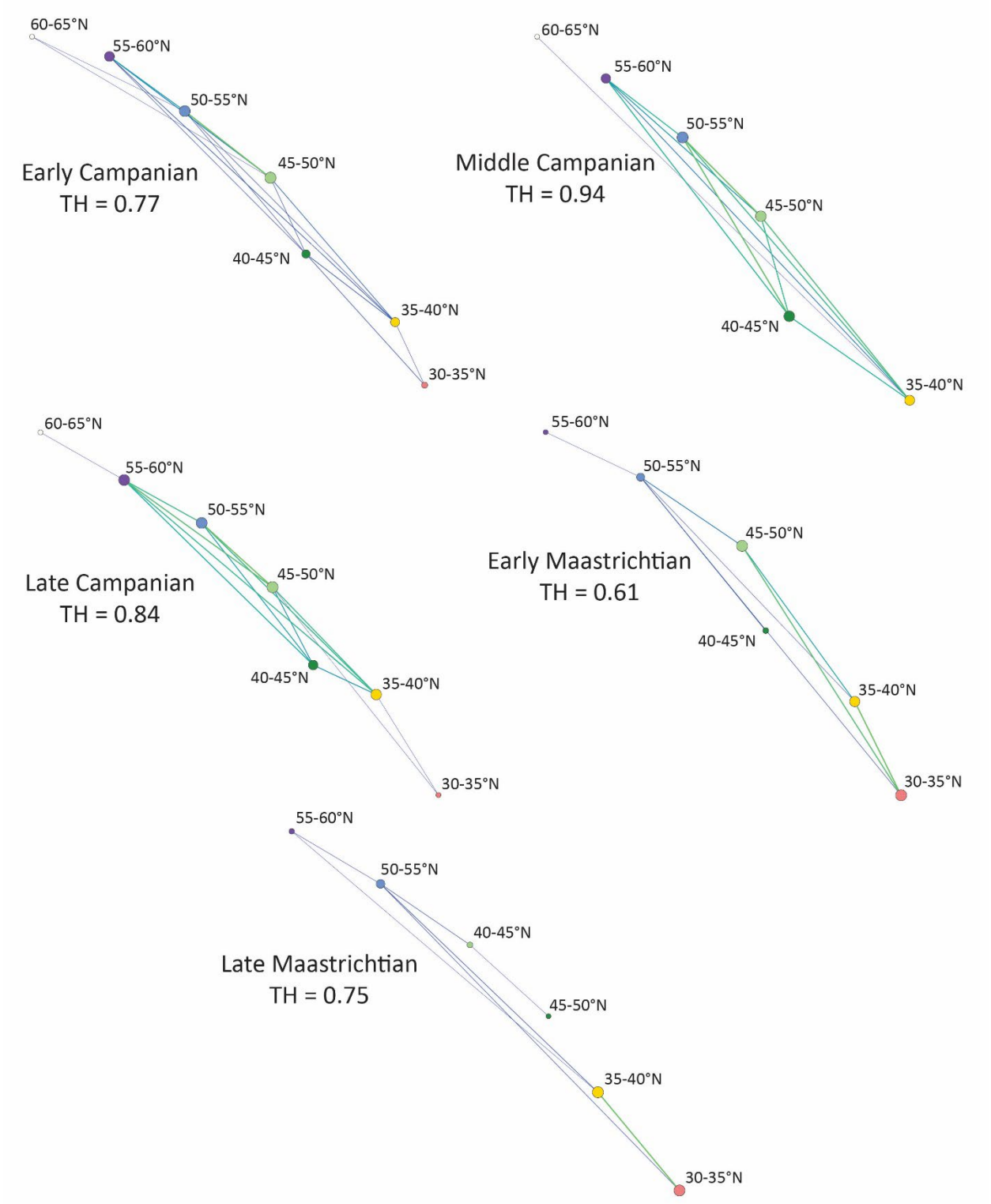



**Figure 11.** Network models across paleolatitude bins. Colored points represent the paleolatitude (cool = higher paleolatitude, warm = lower paleolatitude). Size indicates betweenness centrality. Line color and thickness indicate degree of similarity (thinner/thicker lines = greater similarity). TH =Threshold.

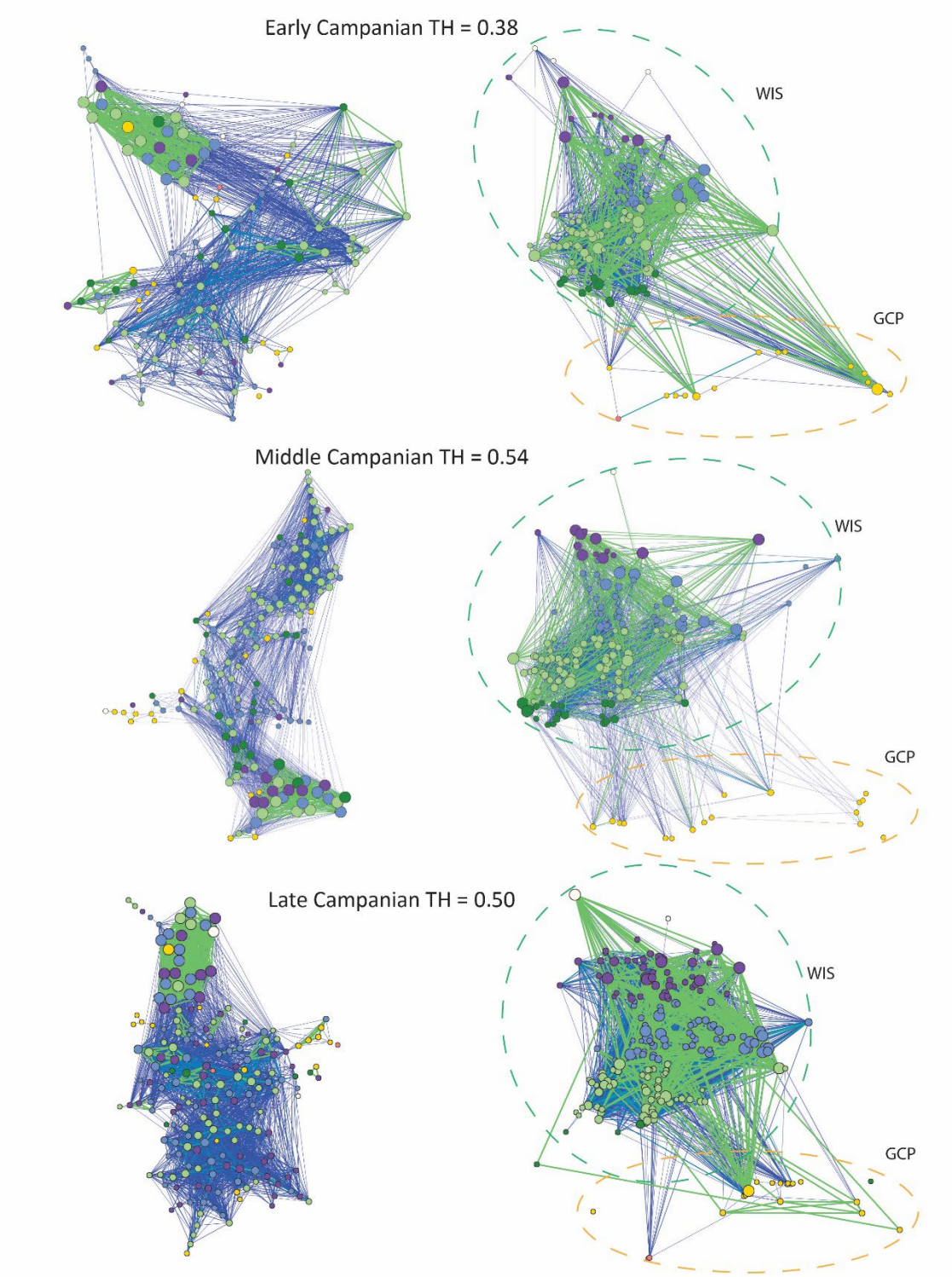



**Figure 12.** Campanian 60km unit networks of FE assemblages plotted without any geographic coordinates (left) and based on paleo-coordinates (right). Unit size represents Betweenness Centrality. Unit colors indicate paleolatitude as in Fig. 8 (warmer = low paleolatitudes and cooler = higher paleolatitudes). TH = Threshold.

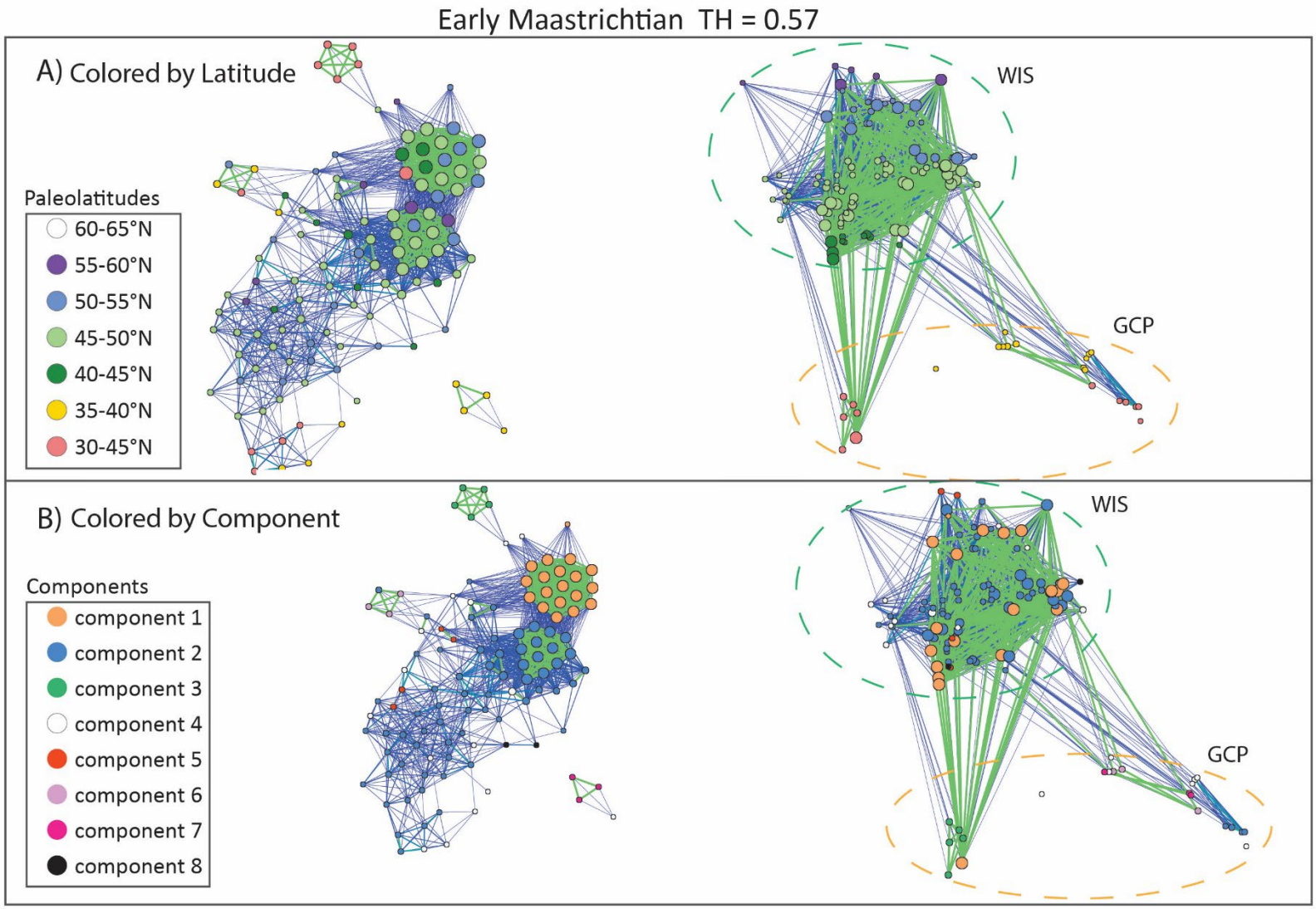



**Figure 13.** Early Maastrichtian 60km unit networks plotted by FE assemblages (left) and based on paleo-coordinates (right). Unit colors of upper networks indicate paleolatitude (warmer = lower paleolatitudes and cooler = higher paleolatitudes). Lower networks are colored by groups of highly similar units or components. TH = Threshold.

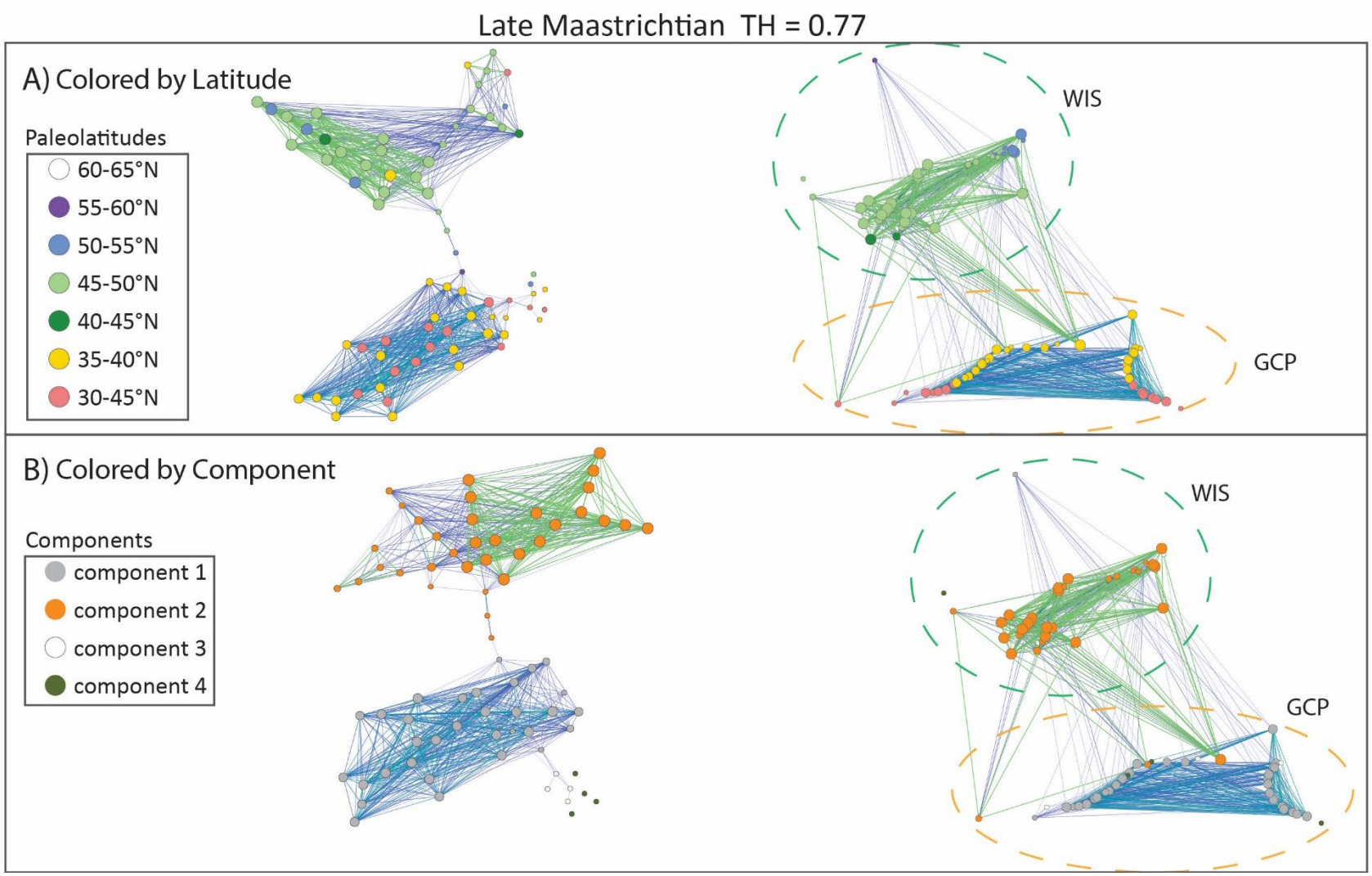



**Figure 14.** Late Maastrichtian 60km unit networks plotted without any geographic coordinates (left) and based on paleo-coordinates (right). Unit colors of upper networks indicate paleolatitude (warm = southern; cool = northern). Lower networks are colored by component. TH = Threshold**.**

In the Late Maastrichtian 360km and 60km unit networks, however, separate components at percolation occupy mainly either the WIS or GCP, indicating provincialism (Figure 14, S29; see Purcell et al. 2023 for more on provincialism). These components represent a functionally diverse GCP assemblage whose units all share five FEs and a functionally less diverse WIS assemblage (Appendix B, Table S33 and S34). When all 360km substage units are modeled together, strong links persist between all regions and substages regardless of distinct components observed in individual substage networks (Appendix B, Figure S30). WIS sampling is less extensive in the Late Maastrichtian than in the GCP, which may explain these differences.

Differentiation in network components based on dominantly siliciclastic and carbonaceous sediments is potentially present in 60km unit networks (Appendix B, Table S35), but a more thorough investigation beyond the scope of this paper would be necessary to determine its prevalence. Minimum Bray-Curtis dissimilarity values between units increase with distance in all substages except in the 360km Early Campanian network (Appendix B, Figure S31-S32), indicating spatial autocorrelation and diminishing FE similarity across distance, as expected. Class-level analysis of 60km networks indicates bivalves and cephalopods form strong, non-provincial network connections (Appendix B, Figure S33-S36) but gastropods form provincial components given their WIS-dominated Campanian distribution and GCP-dominated Maastrichtian distribution (Appendix B, Figure S37-S38). Three of the five FEs shared by all units in the Late Maastrichtian 60km unit network GCP component are represented by gastropods (Figure 8; Appendix B, Table S34). However, there is strong evidence that this is once again a product of sampling bias due to differential

preservation of taxonomically and ecologically informative features (Smith 2001; Smith et al. 2001; Dean et al. 2019).

## Discussion

Several previous studies looking at functional diversity across major environmental perturbations have suggested that functional ecological signals are more robust to disturbances such as mass extinctions (Foster and Twitchett 2014; Dunhill et al. 2018; Edie et al. 2018; Pimiento et al. 2020), but are not stable across latitudes (Edie et al. 2018; Schumm et al. 2019). Here we looked at the dynamics of functional entities (FEs) during a background (non-mass extinction) interval that experienced environmental changes in the form of sea level oscillations and their effects on ocean circulation in a restricted epicontinental seaway (WIS) as well as the adjacent normal-marine shelf (GCP). Our results indicate that functional diversity is indeed stable through time despite environmental shifts, and moreover, is stable across a 35° latitudinal gradient. Both patterns can be explained through functional redundancy found previously to produce stability despite environmental perturbations (Pimiento et al., 2020). The WIS and GCP are also functionally distinct, a characteristic of oceanic basin environmental differences previously observed by Edie et al. (2018). Our results suggest, overall, that paleontological data in deep time is relevant and useful for spatiotemporal comparisons of functional trends, but that aspects of data aggregation and type, such as class-level diversity should be taken into account.

### *Functional Diversity in Classes:*

Analyses of FD trends within specific molluscan classes showed some taxon-specificity, which should be noted for future research as a potential source of bias (e.g., especially among gastropods). In most substages, additional bivalve genera are distributed

into FEs evenly, resulting in <7% SME change despite more dramatic generic and functional richness fluctuations (Appendix B, Figure S11; Table S14). Cephalopods, on the other hand, are functionally stable but dramatic changes to cephalopod generic richness resulted in changes to functional evenness (<16%) that was not necessarily inverse to generic richness as would be expected (Appendix B, Figure S8, Table S12). Finally, gastropod richness and evenness values fluctuated strongly across all substages relative to other mollusks (Appendix B, Figure S13). Gastropod generic diversity increased sharply between the Middle and Late Campanian (408%), resulting in rising $\alpha_F$ (57%) and a clear decline in SME (66%). This parallels changes to database metrics not observed in either cephalopod or bivalve data alone, suggesting that gastropod patterns strongly influenced the overall results. These results show that trends in functional diversity observed within single taxonomic classes may poorly represent other clades or overall paleo-"community" structure. Thus, ideal datasets should include diverse taxonomic sampling to best observe broad paleo-"community" structure through space and time. However, it should be noted that ecological and sampling identification biases in specific clades, such as WIS gastropods, may also skew results. Close attention should therefore be given to the diversity of different higher level taxonomic groups in a dataset. Network models using the different mollusk classes also indicate that ecological differences can influence these results. Bivalves have high evenness, resulting in more consistent representation of FE within high-resolution spatial units that increases network connectivity while gastropods have higher FE diversity ($\alpha_F$) and low evenness, which could generate distinct network components. However, determining the specific causes are beyond the scope of this analyses.

*Functional Stability through Time:*

Our analyses overall indicate temporal stability in functional diversity across background intervals, despite changes to the environment, supporting hypothesis 1. Results show that $\alpha_F$ increases only slightly across the study interval, despite more rigorous (fluctuating) changes to patterns of generic richness. No FEs are lost, and only a few are "gained" between any two substages. FEs gained/lost between individual substages are rare, with low proportional generic richness, throughout their temporal range of existence. Given their rarity, timing of loss/gain may be influenced by taphonomic or sampling biases (e.g., the Signor-Lipps effect; (Signor and Lipps 1982; Nawrot et al. 2018; Dean et al. 2019). Moreover, previous analyses support that any loss of FEs likely reflects (brief) regional extirpation versus a global pattern of extinction, which is also observed here (Foster and Twitchett 2014; Dunhill et al. 2018; Edie et al. 2018; Pimiento et al. 2020). For example, immobile attached epifaunal photosymbionts (IA-E-P) is represented by only the genus *Titanosarcolites* in the Late Campanian, is unobserved in the Early Maastrichtian, potentially indicating extirpation, and is subsequently represented by *Titanosarcolites* as well as both *Dictyaraea* and *Gyropleura* in the Late Maastrichtian. Therefore, IA-E-P is most likely either present in the Early Maastrichtian, though unsampled, or only briefly absent due to the short-lived extirpation of *Titanosarcolites*. Paleolatitude bins and 60/360km units are less functionally stable, but this is likely due to spatial resolution and sampling: reduced geographic bin size often results in fewer than 100 genera/unit, which results in control of $\alpha_F$ and functional evenness by generic richness rather than ecology (Appendix B, Figure S1 and S3). Unsurprisingly, functional evenness is higher and more stable at higher spatial resolutions across time, which demonstrates a limit within spatial resolution to adequately

capture the full range of FEs present, or the distribution of genera within them. Thus, functional diversity trends are not necessarily consistent at all spatial scales, refuting hypothesis 4, though further discussion of the utility of observations made using different spatial resolutions is made below. This is an important finding in a field where higher resolution is typically considered more "real," with the implication being that there is a specific spatial resolution for plausibly reconstructing paleo-"communities." Determining whether this "best" resolution is generalizable or case-specific will take continued analyses that include tests of spatial sensitivity across different time periods and habitats.

### *A Flat Latitudinal Functional Diversity Gradient:*

Our analysis shows that the Late Cretaceous WIS + GCP region lack a latitudinal functional diversity gradient signal; this confirms hypothesis 1 in part, which predicted that functional diversity is stable across space, and contradicts previous studies of latitudinal functional diversity gradients in the Recent (Edie et al., 2018; Schumm et al., 2019). At face value, both generic and functional richness decline between adjacent paleolatitude bins from low to high latitude but few bins contain sufficient unique genera (i.e., >100) to adequately represent "real" functional diversity (Table 7). The influence of true diversity change cannot be confidently separated from sampling bias. Paleolatitude bin networks are consistently well-connected regardless of degree of paleolatitude bin separation spatially, and both 360km and 60km unit networks lack latitudinally distinct components, indicating similar functional assemblages (Figure 12 & 13). Even taken at face value, paleolatitudinal patterns would indicate diversity peaks in mid-latitudes of the study region, potentially due to sampling bias such as is present in current diversity studies (i.e., Menegotto & Rangel, 2018). Mid-latitude peaks in diversity, even when accounting for sampling bias, have been observed for

Cretaceous terrestrial fauna potentially resulting from landmass distributions (Mannion et al. 2012); these mid-latitude peaks may therefore be representative of shifting WIS basin geometries, as can be observed in the data distribution of this study (Figure 6).

There is, furthermore, poor evidence for a global Late Cretaceous LDG in the literature once sampling bias has been taken into account (Huang et al. 2014; Mannion et al. 2014; Brodie and Mannion 2022). In our WIS+GCP data, functional diversity metrics do not change consistently from south to north among bins with sufficient sampling (>100 GR) and in the Late Campanian, where paleolatitude bins have >84 genera from 35° to 60°N, $\alpha_F$ declines inconsistently by a total of ~12%, losing only three FEs. This contrasts sharply with the dramatic decline in $\alpha_F$ observed by Schumm and colleagues for present-day bivalves across the same 25° gradient (Schumm et al., 2019) (Figure 6), and also conflicts with LDG patterns observed in present-day birds showing decreased $\alpha_F$ at higher latitudes in parallel with generic richness (Edie et al., 2018; Schumm et al., 2019). Changes in functional diversity across latitude furthermore do not form consistent nested subsets of FEs (Table. S24-28), as is observed in the latitudinal gradient of today (Schumm et al., 2019). This suggests that FEs are not lost systematically from some maximum $\alpha_F$ assemblage, but instead the specific FE vary across latitudinal space via turnover.

This suggests that the lack of change in functional diversity across latitude is a distinct pattern of the Late Cretaceous compared to today, or at least one common to Greenhouse climate regimes in Earth history (Huang et al. 2014; Brodie and Mannion 2022). A flat Late Cretaceous latitudinal temperature gradient documented for the study area could have promoted reduced environmental disparity across latitude, thereby fostering functional stability (Huber et al. 1995; Mannion et al. 2012). A flat temperature gradient is supported by

the observation that terrestrial diversity in the Late Cretaceous correlates with landmass distributions rather than climate gradients (Baron, 1989; Poulsen, 2010; Mannion et al., 2012; Nicholson et al., 2016). The latitudinally extensive distribution of the WIS basin and its unique oceanographic conditions, including non-normal marine conditions such as lower salinity and periodic vertical stratification, as well as mixing via counterclockwise gyre throughout the Late Cretaceous (Fisher et al. 1994; Slingerland et al. 1996; Leckie et al. 1998; Longman et al. 1998; Steel et al. 2012; Lowery et al. 2018), may contribute to marine invertebrate taxonomic stability despite 25° of geographic distance. The oceanographic conditions throughout the WIS that lead to a breakdown or lack of development of a latitudinal diversity gradient may also have promoted functional homogenization and marine invertebrate stability in functional diversity.

### *Provincialism in Functional Assemblages:*

The WIS and GCP regions are functionally distinct and evolve towards independent provinces as habitat differences become exacerbated through time, conflicting with our second hypothesis of stability between regions. Five FEs are absent in the WIS which are present in the GCP when sampling intensities are sufficient, including photosymbiotic and boring taxa consistent with previous studies that observed few reef-associated taxa in the WIS (Gill and Cobban 1966b; Sohl 1967; Caldwell 1968; Kauffman 1984; Kauffman and Caldwell 1993). While these unoccupied FEs are consistently rare in the GCP, their absence over approximately 18 My suggests that the WIS is less functionally diverse than the GCP, most likely due to differences in their ocean environment (i.e., restricted epicontinental sea versus open-ocean facing shelf environment; Caldwell, 1968; Gill & Cobban, 1966; Kauffman, 1984; Kauffman & Caldwell, 1993; Sohl, 1967; Purcell et al., in press) (Figure 6).

Basinal differences like these have been observed in oceans today as well, resulting in functional turnover and distinct FE assemblages in different regional basins (Schumm et al. 2019). Low salinity conditions in present-day oceans have also been associated with poor development and survival of photosynthetic corals and bivalves (Soo and Todd 2014; Aguilar et al. 2019); the non-normal salinity and stratification of the WIS (Cochran et al. 2003a; Fricke et al. 2010; Petersen et al. 2016) may have thus inhibited the survival of photosymbionts and other reef-associated taxa (Kauffman 1984; Kauffman and Caldwell 1993).

Only the Late Maastrichtian networks show strong basin provincialism (Figure 14 and S29), suggesting higher spatial resolution comparisons only detect geographic FE differences when sea levels fell dramatically in the Maastrichtian. This provincialism is driven by taxonomic differences, particularly among gastropods, between the two regions, consistent with observations by Sohl (1967) among others (Figure 8). Furthermore, most FEs extirpated from the Late Maastrichtian WIS were found there in previous intervals, especially gastropod-type FEs (Sohl 1967) (Appendix B, Figure S5). Of the 11 possible gastropod-type FEs, all are present in at least two earlier substages of the WIS, but only six are present in the Late Maastrichtian WIS, where only one of which is herbivorous (Figure 8; Appendix B, Table S32 and S33). Only gastropods show clear network provincialism when molluscan clades were analyzed independently (Appendix B, Figure S37-S38). Even when considering all taxa together, WIS richness values declined (-13% αF) and SME increased (+23% SME), between the Late Campanian and Early Maastrichtian, and from the Early to Late Maastrichtian following the T9 transgression (-30% αF, +54% SME; Figure 6 and S4). These patterns are consistent with the timing of sea level fall during the Late Campanian that

exacerbated water mass differences and promoted taxonomic provincialism either through habitat change or restrictions to taxon dispersal (Figure 6; Purcell et al., 2023). Increasingly non-normal marine conditions likely stressed WIS taxa and promoted their extirpation (Cochran et al. 2003b; He et al. 2005; Fricke et al. 2010; Petersen et al. 2016). Lack of functional redundancy plausibly made these FEs susceptible to loss while more dominant FEs experienced only generic abundance decline or even increase (Figure 8 & S5). It should be noted, however, that network-based provincialism may also be influenced by sampling biases in gastropods in particular due to their lesser preservation potential and few diagnostic characters (Smith et al., 2001; Smith, 2001; Dean et al., 2019)

Smaller scale differences between the specific FE assemblages present within the study region are most likely also due to habitat differences. Several studies have found that distinct habitats have not only unique levels of functional diversity, but also different collections of specific FEs within different assemblages (de Arruda Almeida et al. 2018; Sulemana et al. 2022; Walsh et al. 2022). In all intervals except the Early Campanian, maximum FE assemblage similarity (i.e., the similarity between the specific FEs present) weakly decreases across distance (Appendix B, Figure S31 and S32). WIS data analyzed independently follows a similar pattern in all substages. That the maximum similarity of FE assemblages at both unit size 4 and 5 resolutions decreases as distances increase, even within the WIS, suggests some spatial control on ecological relationships. Similar assemblages of FEs are more likely to occur close to one another, which suggests that habitat distributions influence functional community structure (Schumm et al. 2019), even within a region with a weak latitudinal temperature change. Future studies that assess the influence of abiotic variables on these patterns will potentially help determine their influence on this dataset, but

are beyond the scope of this paper. These patterns are furthermore not observed at spatial resolutions lower than 360km units (i.e., in latitudinal bins or regionally), thus, functional diversity does not appear stable across space at all scales, partially contradicting hypothesis 1. This may in part be due to the way network models are created using a threshold method. Network connections are always present in nodes with shared FEs, so only a complete lack of similarity will cause the network to break down at high thresholds. As nodes become smaller, they contain fewer FEs, and therefore are either well connected (i.e., potentially containing the same common FEs) or have no similarity (i.e., containing rare, different FEs). This result adds support to the notion that not only does the study of functionally interesting paleo-"communities" have a higher resolution limit, but also a lower resolution limit, which suggests spatial scale sensitivity analysis should be considered in studies like this. Future studies that focus on the influence of spatial unit size and similarity should be conducted to better determine how well networks such as these assess paleo-"communities" and the specific FEs that make them up, but that is beyond the scope of this analysis.

### *Packing and Functional Redundancy:*

Spatiotemporal functional stability in these results is well explained based on functional redundancy, supporting hypothesis 3. When generic richness increased across time and space, taxa are preferentially "packed" into existing FEs before adding new FEs, promoting functional redundancy in specific FE categories (Figure 7). For example, between the Middle and Late Campanian, as generic richness increased by ~81% (147 new genera), $\alpha_F$ only increased by ~10% (three new FEs). Across this interval, the most common FE, mobile unattached epifaunal carnivores (MU-E-C in Figure 8), rose in proportional generic diversity from ~8% to ~30% while the majority of other FEs fell in proportional generic diversity

(Figure 8). High functional redundancy in specific groups diminishes functional evenness, as can be seen throughout our results. Prior studies of functional diversity patterns across space have also found that functional richness increases with area, but evenness decreases, as specific FEs become more dominant (Karadimou et al. 2016). Moreover, this is observed in our data where functional evenness responds inversely to changes in generic richness in almost every comparison across either time or space, indicating preferential packing in specific FEs, such as mobile unattached epifaunal carnivores and facultatively mobile unattached infaunal suspension feeders (MU-E-C and FU-I-SF in Figure 8, respectively); these are generally the most dominant FEs in all substages and experience some of the most significant changes in proportional generic diversity(both positive *and* negative) (Figure 8). Differential packing of genera into FEs can be viewed either from the perspective of sampling bias (i.e., poorer sampled units are more even because the most common FEs are likely to be represented by the few genera present, but not likely to have many genera in each FE) or as an ecological phenomenon, wherein the primary ecological niches are filled first, and then those with greater resource availability are packed with more genera as diversity increases. The distinction between these two interpretations is not clear, and further research would be necessary to determine if these patterns represent ecological processes.

Dominance of specific FEs is also consistent through time (Figure 8). Similar trends have been documented within other taxonomic groups, indicating that either some trait characteristics allow for greater subdivision or partitioning, or that some FEs have greater capacity for taxonomic diversity due to greater resource abundance (Halpern and Floeter 2008; Oliveira et al. 2016; Schumm et al. 2019; Pimiento et al. 2020). In this case, mobile epifaunal carnivores, facultatively mobile infaunal suspension feeders, and immobile

attached epifaunal suspension feeders are the most dominant throughout all intervals (MU-E-C, F-I-SF, and IA-E- SF in Figure 8, respectively). These groups may have contributed to greater resource partitioning and/or utilized resources that were more abundant in the WIS/GCP at this time. Among marine groups like bivalves, functional redundancy is most likely caused by greater resource availability rather than subdivision of trait space (Stanley 2008; Schumm et al. 2019). In the case of MU-E-Cs, which are consistently dominant throughout the study interval, taxa would be less limited by spatial constraints since they are mobile and unattached, making prey abundance a more likely limiting factor.

Functional redundancy appears to have buffered most FEs from complete loss even when generic diversity declined. Loss of generic richness by 8.8% between the Late Campanian and Early Maastrichtian is associated with the loss of only one FE: immobile attached epifaunal photosymbionts (IA-E-P in Figure 8). This FE had very low proportional generic richness in the Late Campanian (0.3% of genera) and is spatially restricted to the GCP (Appendix B, Table S4). These results are consistent with patterns observed in mollusks across the Miocene to Pleistocene by Pimiento et al. (2020), wherein functional diversity was sustained despite taxonomic loss by the preferential extinction of functionally redundant hypothesized competitors. Functional evenness increased from the Late Campanian to the Early Maastrichtian (SME proportional increase of ~11%) as the relative generic richness of the most dominant FE (MU-E-C) declined from ~30% to ~26%. No other FE experienced such a large decrease. Thus, whereas functional redundancy in this dominant FE likely buffered it against regional extinction, less dominant FEs were generally more likely to be extirpated. Along these lines, unaffected, less dominant FEs were more likely buffered by reduced resource competition among fewer taxa (Pimiento et al. 2020).

## Conclusions

Analysis of Campanian and Maastrichtian Western Interior Seaway (WIS) and Gulf Coastal Plain (GCP) of North American invertebrates indicates stability in functional richness ($\alpha_F$) and instability in functional evenness as additional genera are preferentially packed into a few dominant functional entities (FEs). Additionally, functional redundancy is found to buffer against FE loss across both time and space. Sampling biases these results at finer, and potentially very broad, spatial scales; although comparisons of FE assemblages across space are still viable given that they can capture the influence of habitat variability. The WIS and GCP region lack a present-day-style LDG, potentially caused by a flat latitudinal temperature gradient across these latitudes. Instead, the WIS and GCP regions display distinct FE assemblages and functional provinciality by the Late Maastrichtian which mirrors taxonomic provinciality with a lag time of one substage (Purcell et al. 2023). Provinciality is likely caused by oceanographic differences between the basins, particularly non-normal marine conditions in the WIS which were probably exacerbated by falling sea levels in the Maastrichtian. Non-normal marine conditions stress photosymbiotic taxa, including both corals and bivalves, boring taxa, and FEs occupied by herbivorous gastropods. This is observed herein by a decrease in generic richness or complete loss of these FEs in the WIS. The lag between provinces formed in the Early Maastrichtian for taxa but the Late Maastrichtian for FEs may suggest either a data-based difference between the two analyses, or an ecological pattern wherein functional diversity remained stable for a period of time after taxonomic turnover produced taxonomic provincialism. Functional diversity trends also differ by taxonomic class, discouraging future use of single clades for interpretation of spaciotemporal trends in assemblage (paleo-"community") structure. That said, overall this

investigation provides evidence for FE paleo-"community" stasis during background intervals not characterized by a global climate upheaval (e.g., mass extinction events). Results also provide additional information regarding the stability of functional diversity and redundancy through changing environmental conditions and at different spatial scales, improving our understanding about long-term and broad-scale functional diversity dynamics for both current and ancient systems.

**Acknowledgements**

CP acknowledges funding support from NSF #2021744; CEM acknowledges support from NSF #1924807, and #1601878. CP would also like to thank the many researchers who have worked to identify and describe the functional characteristics of marine invertebrates, without whom this work would not be possible.

# CHAPTER 4: TESTING FUNDAMENTAL NICHE CHARACTERISTICS USING HIGH-RESOLUTION LATE CREATACEOUS WESTERN INTERIOR MARINE INVERTEBRATES


## Abstract

Ecological niche models (ENMs) are a common tool used to analyze various aspects of ecology, evolution, biotic distribution, and community structure by correlating environmental variables with species occurrence distributions based on the concept of a realized niche. Similarly, paleoENM analysis is applied to address these same questions across long intervals of Earth's history and species' durations. This enables the exploration of fundamental assumptions about the niche itself (e.g., whether it is stable over species' durations and/or conserved across evolutionary lineages). However, very little analysis has been performed testing the fidelity of these methods under the unique biases in paleontological data. In particular, the influence of temporal aggregation and assumed ecological similitude between individual species and their genera have been poorly explored. Using a high-resolution, geographically widespread, deep time dataset, this paper analyzes the impacts of temporal resolution on paleoENM-based niche prediction, and performs basic tests of phylogenetic niche conservation between sister species of marine bivalves to draw conclusions about appropriate application of paleoENM analysis. The results of these tests indicate that great attention should be paid to environmental data sources, interpolation processes, and ecological interpretation when using fossil data, and that substage-level aggregation is unlikely to produce meaningful results within intervals characterized by highly variable environments. Furthermore, we find that individual species-level niche predictions do not match well that of the generic-level estimation. These inherent biases in fossil data likely give rise to researchers utilizing generalist taxa rather than specialists, resulting in

apparent ecological homogenization. However, our results also indicate that paleoENM is useful for assessing general aspects of ecological shift or differentiation and for analyzing niche abiotic habitat characteristics of functional entities, particularly when primary ecological characteristics and environmental factors are appropriately scaled.

## Introduction

The ecological niche is a fundamental species' characteristic that is frequently used to address macroecological and evolutionary questions (Brown 1995; Peterson et al. 2011). Although an ecological "niche" can be defined in many ways, here we use the Hutchinson niche: an n-dimensional hypervolume composed of environmental factors that allow a population to survive and reproduce in a given area (Hutchinson 1957; Peterson 2001, 2011; Soberón and Peterson 2005). Hutchinson divided niches into two forms: (1) fundamental niche (FN) – encompassing all abiotic factors necessary for survival and reproduction, and (2) realized niche (RN) – a subset of the FN wherein the FN is restricted by biotic interactions and dispersal capacity (Soberón and Peterson 2005; Peterson et al. 2011). Many modern studies have attempted to quantify the FN/RN to elucidate ecological and evolutionary concepts, including species' life histories, biotic interactions, geographic range patterns, invasion potential, biotic response to environmental shifts, and extinction vulnerability (i.e., Adhikari et al., 2015; Benito Garzón et al., 2011; Liu et al., 2022; Lockwood et al., 2013; Peterson, 2003; Planas et al., 2014; Purcell & Stigall, 2021; Valencia-Rodríguez et al., 2021; Varela et al., 2010; etc.). Given the wide variety of species, environments, and hypotheses being tested, each dataset is unique, making the identification of universal or generalizable "rules" of FN/RN interactions with evolutionary or ecological processes difficult to define or quantify comparatively.

The application of these techniques to the deep time fossil record adds additional dimensions, first and foremost the element of time, to using niche estimation to test macroevolutionary and macroecological hypotheses unique to paleontological data. Modern datasets usually involve higher spatial resolution, denser species occurrence data, can incorporate genome-based phylogenetic information, and use well-established climatic variables that can produce detailed estimation of niche dynamics (Peterson et al. 2011). In contrast, ecological niche analyses using deep time datasets have the advantage of using fossil and sedimentological data to compare long-term patterns of FN/RN not achievable in modern datasets (Maguire and Stigall 2009; Dudei and Stigall 2010; Malizia and Stigall 2011; Brame and Stigall 2014; Saupe et al. 2014, 2015, 2019; Stigall 2014, Myers et al. 2015b; Purcell and Stigall 2021). Given that species are known to survive on million-year (Myrs) timescales, the extension of these methods to deep time is imperative to truly understand how niche dynamics relate to evolutionary and ecological trajectories of species and clades.

However, paleo-niche analyses must also consider the loss of spatiotemporal detail inherent in the geologic record (reviewed in Myers et al., 2015). Primarily, the process of sedimentation homogenizes both temporal and, to a lesser degree, spatial information. At best, geologic units may record $10^4$ years of environmental information per centimeter within a single defined sedimentary layer (Schindel, 1980). Fine-scale analysis of ecology and environmental characteristics at less than this resolution is therefore rarely viable. In practice, most niche estimation in deep time relies on temporal aggregation at the geological substage to stage level (i.e., generally 100s-1000s kyrs). Moreover, temporal aggregation increases as the geographic scope of analyses gets larger – for example, when using regional or global

data, which is particularly important for "big data" paleobiological analyses so common today (e.g., Maguire and Stigall 2009; Saupe et al. 2019; Purcell and Stigall 2021). Therefore, while fossil analyses allow for long-term tests of niche dynamics and their relationship to macroevolutionary and macroecological process that are not possible using modern species, by necessity they assume either localized environmental stability across long time intervals or that long-term averages of environmental conditions are consistent with a species' preferred habitat.

Paleontological data is also biased both by the fossilization process itself which preserves hard-bodied taxa significantly more often than soft-bodied taxa (Kidwell and Flessa 1996), by the incomplete exposure of strata, and the erosion of exposed material (Antell et al. 2024). Spatial uncertainty, present in all datasets, requires that species' occurrences be spatially thinned to a single occurrence per grid cell (already likely larger than in modern datasets). Given that fossil species' occurrences, rarely extensive or dense to begin with, are often spatially clumped, spatial thinning typically leads to substantial reduction in occurrence observations (Shcheglovitova and Anderson 2013; Aiello-Lammens et al. 2015). For example, Purcell and Stigall (2021) used a spatial error of 30kms, based on the spatial error margin for the fossil taxa analyzed and linked to 30km geographic grid cell size used. This results in very few fossil species with sufficient distributions to perform robust paleo-niche analysis using popular methods, such as the Maxent algorithm (Phillips et al. 2006), often regarded as greater than five occurrences (Hernandez et al. 2006). To obtain sufficient occurrence numbers after spatial thinning, most paleo-niche analyses are restricted to genus-level comparisons, particularly in deep time (e.g., Brame & Stigall, 2014; Dudei & Stigall, 2010b; Hopkins et al., 2014; Maguire & Stigall, 2009; Malizia & Stigall, 2011b;

Myers et al., 2013; Nurnberg & Aberhan, 2013; Purcell & Stigall, 2021). How well a genus represents species-level ecological traits, however, is still debated (Hendricks et al. 2014).

Unfortunately, very few datasets allow for the high-resolution, densely fossiliferous, and well document comparisons necessary to test these issues directly. How well temporally-aggregated geological and paleontological information represents real-world biological phenomenon has yet to be well established. Even assuming that temporal resolution is not a factor in accurate niche estimation (e.g., if FNs are stable over species' durations), niche characteristics may not be conserved between related species and at the genus-level, potentially resulting in less robust paleo-niche analyses. Given the advantages of using fossil data to assess overall community structure (Fara and Benton 2000; Finnegan and Droser 2008), the relationship between environmental change and ecological response (Purcell and Stigall 2021), and various other macroecological questions (Myers et al. 2015), addressing these issues is vital to the production of accurate scientific study.

Using a high-resolution sedimentary record with regionally correlated Western Interior Seaway (WIS) marine units and a well-vetted paleontological database representing approximately 17 million years of marine life, we tested fundamental aspects of niche estimation in deep time. By comparing niche overlap in a simplified environmental space, these data were used to address three main questions: 1) how temporal resolution influences niche estimation, 2) how well fossil data of phylogenetically related taxa (sister species) present phylogenetic niche conservation, and 3) how well the genus-level niche estimation conserves species-level niche characteristics. The results of these analysis inform the utility of paleo-niche estimation in deep time to test macroevolutionary and macroecological hypotheses. We provide best practices for more accurate paleo-niche estimation and

ultimately support the application of ecological niche modeling in the fossil record (paleoENM).

## Methods:

Our study used marine invertebrate fossil occurrence data from the North American WIS during the Campanian and Maastrichtian stages of the Late Cretaceous. Occurrence data was compiled from online records downloaded from the Paleobiology Database (08/25/21 download) and iDigBio (08/30/21 download), and records from museum collections at the Black Hills Institute, the USGS-Washington Cobban Collection, and a thesis database (Mackenzie 2007). All data were vetted to remove incomplete, erroneous, or poorly spatiotemporally resolved occurrences and to update taxonomic nomenclature (see Purcell et al. 2023 for additional details and references used in dataset vetting). Stratigraphic ages were vetted to the biozone level based on stratigraphic and biozone information in the database itself or from literature sources (Cobban et al. 2006; Merewether and McKinney 2015), resulting in twelve temporal bins (Figure 15, Table 9). To remove spatial ambiguity, taxonomic occurrences with spatial resolution less than 30 km of uncertainty were removed. To include only WIS data, records east of -80° longitude and south of 31.6° latitude were removed (see Purcell et al., 2023).

The final WIS dataset of 21,156 fossil occurrences includes a total of 279 genera of bivalves, cephalopods, gastropods, echinoderms, corals, brachiopods, bryozoans, and crustaceans (Figure 2). Tests confirming signal over noise in this dataset were previously published in Purcell et al. (2023). In the cleaned database, potential taxa for use in niche estimation were identified by greater than six fossil occurrences within at least one biozone interval. More than six fossil occurrences identified at the species-level was selected as an

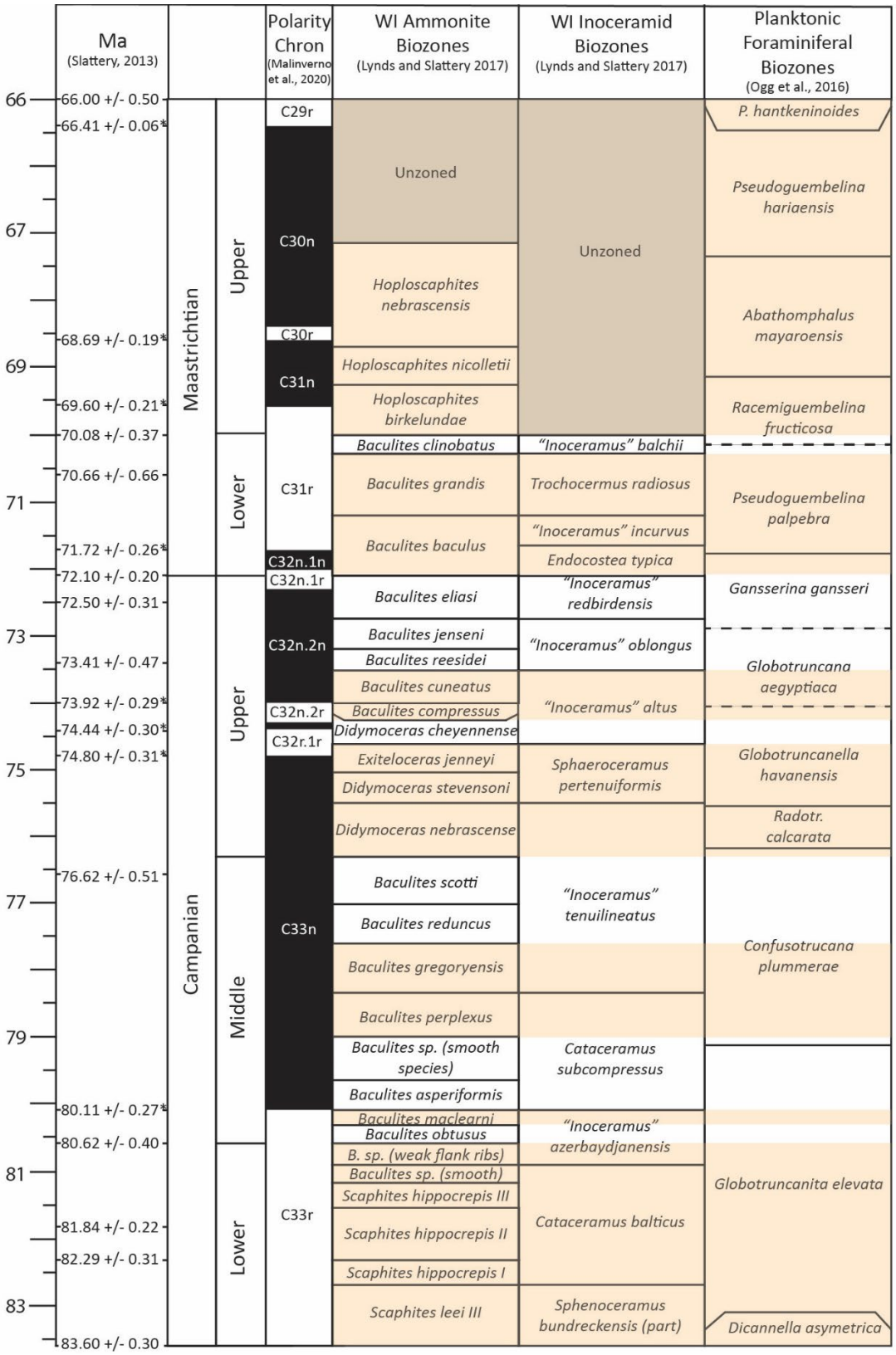



**Figure 15.** Biostratigraphic chart showing collated biozone intervals; shading delineates each interval used in analyses (modified from Purcell et al. 2023).

arbitrarily threshold because it is slightly greater than the minimum value recommended for creating ecological niche models (ENMs) using the popular Maxent algorithm (Hernandez et al. 2006), while still providing a relatively relaxed threshold for sparce fossil data. Selected taxa were spatially thinned to the 30 km geographic grid cell size determined by allowed spatial uncertainty in occurrences. After spatial thinning, only the *Inoceramus* genus had sufficient species-level occurrences for phylogenetic niche conservation tests; in this case, seven *Inoceramus* species had six or more occurrences in at least one time interval. These were also evaluated with three "comparison taxa" species from the *Ctena, Lucina*, and *Chlamys* genera that had sufficient occurrences in most of the same biozone intervals (Table 9). All ten bivalve species were used for analysis of temporal sensitivity.

**Table 9.** Summary of the number of species occurrences in each biozone interval. Note that some biozone intervals were not used for analysis, due to a lack of adequate species occurrence information.

| | *S. leei-B. sp. (weak flank ribs)* | *B. obtusus* | *B. maclearni-B. sp. (smooth)* | *B. perplexus-B. gregoryensis* | *B. reduncus-B. scotti* | *D. nebrascense-E. jenneyi* | *D. cheyennense* | *B. compressus-B. cuneatus* | *B. reesidei-B. eliasi* | *B. baculus-B. grandis* | *B. clinolobatus* | *H. birkelundae-H. nebrascensis* |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **INGROUP** | | | | | | | | | | | | |
| *Inoceramus barabini* | | | | | | 7 | | | 9 | | | |
| *Inoceramus convexus* | | | | | 8 | 7 | | | | | | |
| *Inoceramus sagensis* | | | | | | | | 15 | 11 | | | |
| *Inoceramus sublaevis* | | | | | 11 | 7 | | | | | | |
| *Inoceramus oblongus* | | | | | | | | | 9 | | | |
| *Inoceramus saskatchewanensis* | | | | | 9 | | | | | | | |
| *Inoceramus azerbaydjanensis* | | | 7 | | | | | | | | | |
| **COMPARISON TAXA** | | | | | | | | | | | | |
| *Chlamys nebrascensis* | | | | | | 7 | | | 11 | 10 | | |
| *Ctena imbricatula* | | | | 10 | 7 | 8 | | | 11 | | | |
| *Lucina subundata* | | | | | | 8 | | 9 | | | | |

In order to estimate environmental niches, 11 spatially continuous environmental layers were reconstructed based on sedimentary proxies, distance from shore, and distance from hydrothermal seep deposits (following Myers et al 2015). Stratigraphic data used to characterize environmental proxy variables were collected across WIS localities from literature sources, including theses and dissertations, published journal articles, and field guides (see SI for references). The sedimentary environmental proxy variables included: percent grain sizes in siliciclastic rocks (ranging from mud to coarse sand), percent

limestone, stratigraphic bedding style, bedding thickness, and degree of bioturbation (Table S28). Environmental layer values were averaged at individual localities based on three levels of temporal resolution: biozone, substage—Middle and Late Campanian, and stage—Maastrichtian only. The Point Statistics tool and the Inverse Distance Weighting method in ArcGIS Pro was used to interpolate values at locations without sedimentary data and create a spatially continuous, smoothed set of raster layers with 30 km resolution. Raster layers for the distance-from-shore variable were created using paleo-shorelines published in previous works (Gill and Cobban 1973; Cobban et al. 1994; Roberts and Kirschbaum 1995; Slattery et al. 2013) and updated based on fossil localities and stratigraphic information from the current database. The distance-from-hydrothermal-seeps variable was based on the locations of individual or clusters of seeps published in various literature sources (see SI for details on sedimentary data collection, aggregation, and interpolation). Environmental layers were then intersected with outcrop masks to restrict continuous variables only to areas of potential fossil occurrence, thus reducing model overfitting. These were created using relevant Late Cretaceous outcrop polygons buffered an additional 30 km to allow for spatial error in the dataset (see SI for references). Outcrop masks were clipped to strandline boundaries for the appropriate interval and then used to clip raster data for each variable (Figure 16). This resulted in environmental layer characterization for the twelve biozone-level intervals, five substages (early, middle, and late Campanian, and the early and late Maastrichtian), and the Maastrichtian stage. Based on the taxa selected for analysis, only seven biozone intervals, the early and late Campanian, the early Maastrichtian, and the Maastrichtian were utilized in the following analyses.

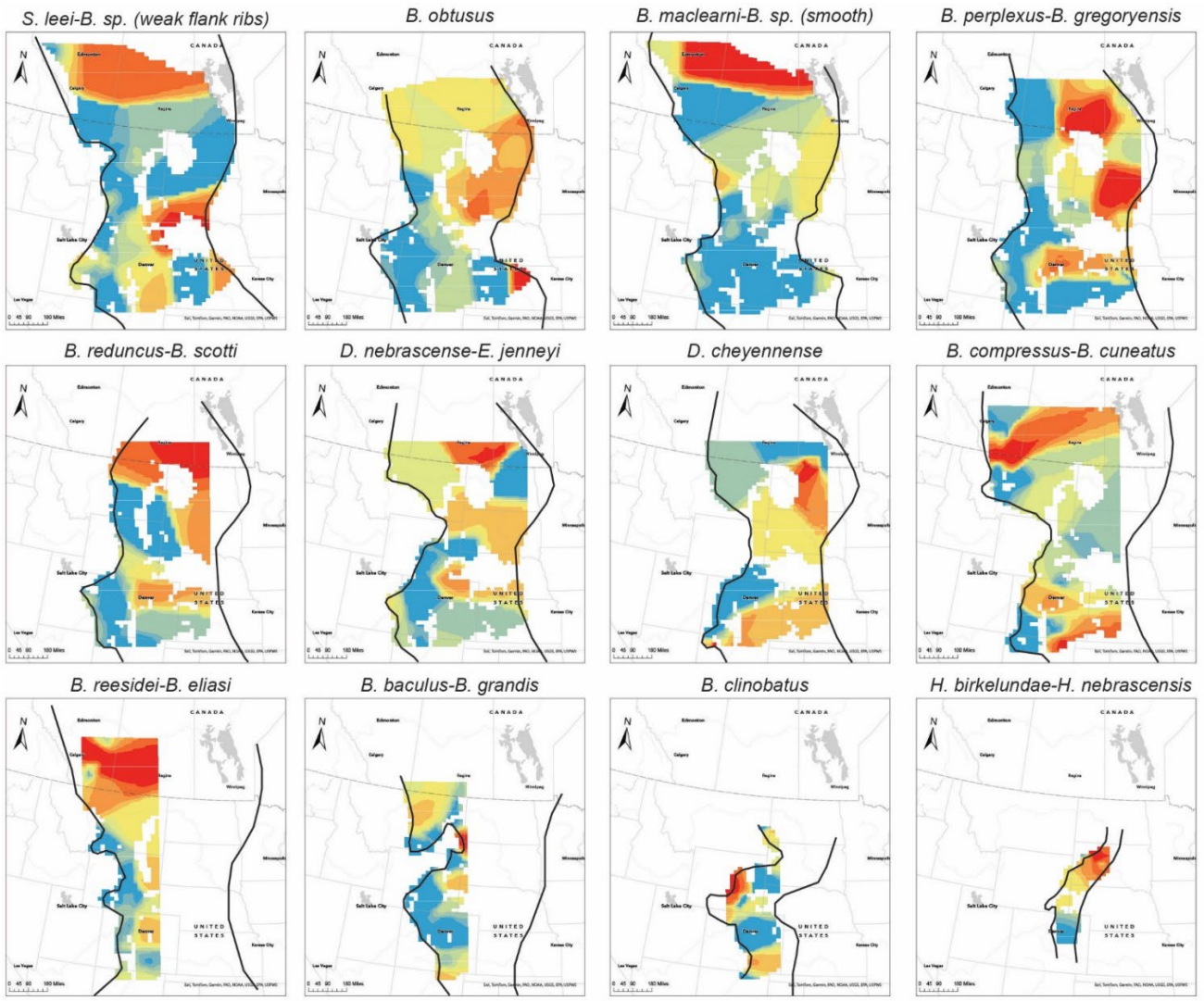



**Figure 16.** Example interpolated maps within the twelve biozone intervals, representing percent mud. Cool colors indicate low percentage and warmer colors indicate high percentage of mud. Shorelines are denoted with black. Note that limitations in sedimentary data collected resulted in unmapped portions of the basin in some intervals.

Niche overlap was assessed using comparisons of observed occurrences within a simplified two-dimensional environmental principal components space (PCS) using the ecospat package in R (Broennimann et al. 2012; Di Cola et al. 2017). Niche overlap is a metric describing the degree of similarity between two niches in environmental space (e-space). When two temporal bins were compared, only shared, non-correlated environmental proxy variables were used (Appendix C, Figure S2 and S3, Table S7). Niche overlap is quantified using two metrics of similarity, a version of the Hellinger distance, known as the I statistic, and Schoener's (1968) D statistic (Warran et al., 2008). These metrics range from 0 to 1 (indicating zero to complete overlap, respectively) and were calculated by comparing smoothed kernel density functions representing taxonomic occupancy for taxon in each PCS grid cell (Broennimann et al. 2012; Di Cola et al. 2017). Both metrics produce similar results, though Schoener's D has a greater ecological implication than the I statistic (see Warren et al., 2008 for discussion). In this analysis both metrics were calculated for comparison, but given the similarity of their results, only the D statistic is discussed in detail.

Three other metrics of niche change were also calculated: niche stability, expansion, and unfilling (Di Cola et al. 2017). Niche stability is the amount of niche space within the PCS which is maintained or conserved between the two taxa being compared (or between the same taxon in two intervals). Niche unfilling and expansion are used to determine how a niche has changed either by taxa moving away from previously occupied e-space or expanding into new e-space, respectively. Both metrics are interpreted here as instability, given that these analyses do not assess alteration of the niche across time or space. Calculations of niche overlap, stability, unfilling, and expansion were performed using both the niche equivalency and niche similarity tests in ecospat to determine if values were greater

or less than would be expected at random (Di Cola et al. 2017). The niche equivalency test is used to determine if two distributions are more or less identical than the overlap of a random set of simulated distributions created using pooled data. The niche similarity test determines if the distributions are more or less similar than the overlap of a randomized subset of the data within analogous e-space (Di Cola et al. 2017). Niche equivalency is a very strict test of niche overlap, whereas niche similarity is a more lenient test and, in general, niche equivalency is rarely observed in nature (Warren et al. 2010; Aguirre-Gutiérrez et al. 2015).

Data sensitivity to temporal resolution was analyzed by comparing niche overlap and stability between the highest resolution temporal bins (biozones) with lower resolution bins (substage and stage) for all ten bivalve species, resulting in a total of twenty-one comparisons. It was hypothesized that niche overlap and stability would fail to be statistically high between different temporal bin resolutions. Phylogenetic niche conservation (PNC) was analyzed by comparing niche overlap and stability between the species of *Inoceramus* bivalves (*I. barabini*, *I. convexus*, *I. sagensis*, *I. sublaevis, I. saskatchewanensis, I. oblongus, and I. azerbaydjanensis*). These analyses were contrasted with comparisons between *Inoceramus* species and three non-*Inoceramus* bivalve species (*Ctena imbricatula*, *Lucina subundata*, and *Chlamys nebrascensis*); in total, twenty-eight PNC comparisons were tested. It was hypothesized that niche overlap and stability would be statistically higher, and higher on average, between *Inoceramus* species than between *Inoceramus* species and non-*Inoceramus* taxa. Finally, niche overlap and stability between *Inoceramus* species and the genus-level *Inoceramus* niche estimation were analyzed to assess how well the genus represents the environmental niche space of its component species, resulting in nine comparisons. Partitioning of the genus-level niche by individual species was visualized in R

using convex hulls in a two-dimensional PCS of shared, non-correlated environmental variables within a single biozone.

## Results

### *Temporal bin Size Comparison*

Table 10 provides niche overlap results for taxon comparisons at variable temporal resolution. Of the twenty-one temporal comparisons, three (~ 14.3%) had statistically more equivalent niche overlap, and only 7 (~33% had statistically high overlap similarity (see bolded D-values and stability scores in Table 10, S9, and S11). Niche stability overall in these comparisons averaged 72% (Table 11 and S8). Only five of the 21 comparisons (~ 24%) had statistically higher niche stability values (Table 10). Unfilling and expansion were statistically lower than expected at random four and five times, respectively, in niche similarity tests (Appendix C, Table S11). A weak positive correlation between the interval duration ratio (biozone to substage and stage) and niche overlap furthermore indicates that as temporal aggregation increases niche overlap values decrease, as predicted (Figure 17). Statistically low niche values were calculated but given that their biological significance cannot be well constrained, they are not included in these results (see SI for details).

Non-analogous environmental conditions were common between temporal bins of different resolution (Appendix C, Figure S4). Furthermore, niche distributions were disjointed within PCS at the substage level for two of the eleven taxa-intervals assessed (Appendix C, Figure S4). Only one of the biozone-level niches had a disjointed distribution. These results suggest that non-analogous conditions are introduced by aggregating variables at lower temporal resolutions and the estimation of niche space occupancy for taxa within these environments is artifactually modified through this process. Furthermore, results of

overlap comparisons are somewhat unique to the species and intervals compared, suggesting that lower temporal resolution in these analyses do not produce consistent results across time interval sizes.

**Table 10.** Results (p-value, alpha=0.5) of greater niche similarity tests for temporal resolution comparisons. Bolded values indicate statistical significance at the p-value < 0.05 level.

| Species | Intervals Compared | D value | Stability |
|---|---|---|---|
| *C. nebrascensis* | *B. baculus-B. grandis* to lower Maastrichtian | **0.001*** | **0.005*** |
| *C. nebrascensis* | *B. baculus-B. grandis* to Maastrichtian | 0.105 | 0.118 |
| *I. sagensis* | *B. compressus-B. cuneatus* to upper Campanian | 0.133 | 0.166 |
| *I. azerbaydjanensis* | *B. maclearni-B. sp. (smooth)* to middle Campanian | 0.165 | 0.161 |
| *C. imbricatula* | *B. perplexus-B. gregoryensis* to middle Campanian | **0.002*** | 0.191 |
| *I. convexus* | *B. reduncus-B. scotti to* middle Campanian | 0.243 | 0.171 |
| *I. sublaevis* | *B. reduncus-B. scotti* to middle Campanian | 0.105 | **0.003*** |
| *I. saskatchewanensis* | *B. reduncus-B. scotti* to middle Campanian | **0.028*** | **0.024*** |
| *C. imbricatula* | *B. reduncus-B. scotti* to upper Campanian | 0.102 | 0.239 |
| *I. barabini* | *B. reesidei-B. eliasi* to upper Campanian | 0.034 | 0.080 |
| *I. sagensis* | *B. reesidei-B. eliasi* to upper Campanian | 0.113 | 0.205 |
| *I. oblongus* | *B. reesidei-B. eliasi* to upper Campanian | **0.014*** | 0.191 |
| *C. imbricatula* | *B. reesidei-B. eliasi* to upper Campanian | 0.085 | 0.168 |
| *L. subundata* | *B. reesidei-B. eliasi* to upper Campanian | 0.337 | 0.139 |
| *C. nebrascensis* | *B. reesidei-B. eliasi* to upper Campanian | **0.040*** | **0.021*** |
| *I. convexus* | *D. nebrascense-E. jenneyi* to upper Campanian | **0.001*** | **0.022*** |
| *I. barabini* | *D. nebrascense-E. jenneyi* to upper Campanian | 0.059 | 0.133 |
| *I. sublaevis* | *D. nebrascense-E. jenneyi* to upper Campanian | **0.003*** | 0.126 |
| *C. imbricatula* | *D. nebrascense-E. jenneyi* to upper Campanian | 0.093 | 0.153 |
| *L. subundata* | *D. nebrascense-E. jenneyi* to upper Campanian | 0.076 | 0.133 |
| *C. nebrascensis* | *D. nebrascense-E. jenneyi* to upper Campanian | **0.003*** | **0.016*** |

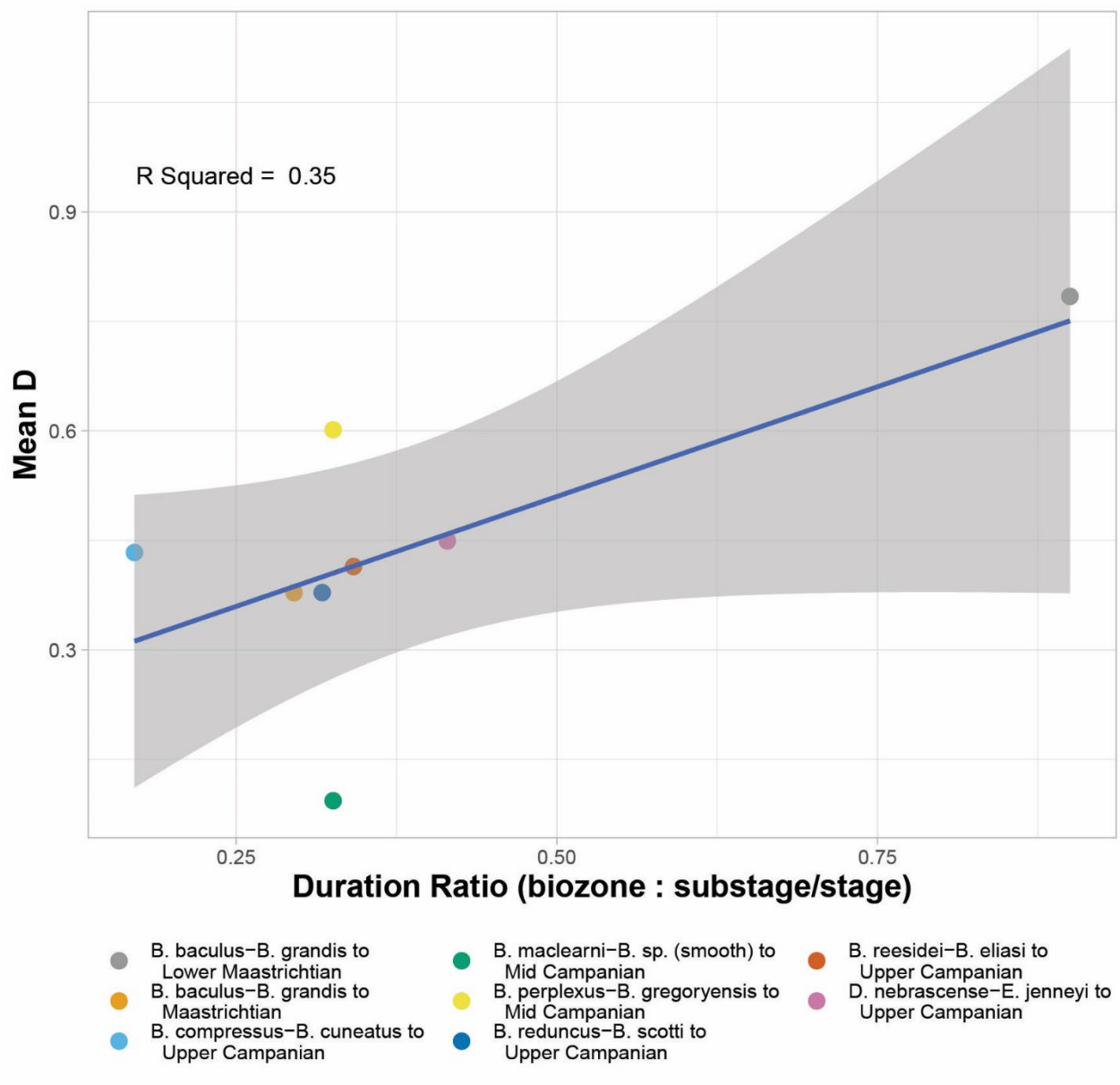



**Figure 17.** Correlation plot of duration ratio to mean niche overlap (D value).

**Table 11.** Summary of niche overlap and stability results. Sbst. D = duration of the substage/stage; b. D. = duration of biozone interval; D Ratio = ratio of the biozone duration; N = number of comparisons.

| | | | | | | D-value | | Stability | |
|---|---|---|---|---|---|---|---|---|---|
| | **Intervals Compared** | **sbst. D** | **b. D.** | **D Ratio** | **N** | **Mean** | **Median** | **Mean** | **Median** |
| **Temporal Bin Comparison** | | | | | | | | | |
| Maastrichtian | B. baculus-B. grandis to Maastrichtian | 6.1 | 1.8 | 0.3 | 1 | 0.38 | 0.38 | 0.60 | 0.60 |
| Lower Maastrichtian | B. baculus-B. grandis to lower Maastrichtian | 2 | 1.8 | 0.9 | 1 | 0.78 | 0.78 | 0.82 | 0.82 |
| Upper Campanian | B. reesidei-B. eliasi to upper Campanian | 4.1 | 1.4 | 0.3 | 6 | 0.41 | 0.45 | 0.73 | 0.72 |
| | B. compressus-B. cuneatus to upper Campanian | 4.1 | 0.7 | 0.2 | 1 | 0.43 | 0.43 | 0.71 | 0.71 |
| | D. nebrascense-E. jenneyi to upper Campanian | 4.1 | 1.7 | 0.4 | 6 | 0.45 | 0.42 | 0.73 | 0.77 |
| | B. reduncus-B. scotti to upper Campanian | 4.1 | 1.3 | 0.3 | 4 | 0.38 | 0.37 | 0.82 | 0.95 |
| Middle Campanian | B. perplexus-B. gregoryensis to middle Campanian | 4.3 | 1.4 | 0.3 | 1 | 0.60 | 0.60 | 0.74 | 0.74 |
| | B. maclearni-B. sp. (smooth) to middle Campanian | 4.3 | 1.4 | 0.3 | 1 | 0.09 | 0.09 | 0.31 | 0.31 |
| **TOTAL** | | | | | **21** | **0.43** | **0.43** | **0.72** | **0.74** |
| **Phylogenetic Comparison** | | | | | | | | | |
| Ingroup | B. reesidei-B. eliasi | - | 1.4 | - | 3 | 0.67 | 0.68 | 0.71 | 0.66 |
| | D. nebrascense-E. jenneyi | - | 1.7 | - | 3 | 0.33 | 0.10 | 0.52 | 0.41 |
| | B. reduncus-B. scotti | - | 1.3 | - | 3 | 0.58 | 0.52 | 0.63 | 0.77 |
| | **TOTAL** | **-** | | | **9** | **0.52** | **0.54** | **0.62** | **0.66** |
| Comparison taxa | B. reesidei-B. eliasi | - | 1.4 | - | 6 | 0.65 | 0.63 | 0.92 | 0.93 |
| | B. compressus-B. cuneatus | - | 0.7 | - | 1 | 0.29 | 0.29 | 0.64 | 0.64 |
| | D. nebrascense-E. jenneyi | - | 1.7 | - | 9 | 0.48 | 0.57 | 0.76 | 0.83 |
| | B. reduncus-B. scotti | - | 1.3 | - | 3 | 0.36 | 0.38 | 0.50 | 0.53 |
| | **TOTAL** | | | | **19** | **0.50** | **0.57** | **0.76** | **0.86** |
| **Niche Partitioning Comparison** | | | | | | | | | |
| Inoceramus Species to Genus | B. reesidei-B. eliasi | - | 1.4 | - | 3 | 0.45 | 0.46 | 0.63 | 0.67 |
| | D. nebrascense-E. jenneyi | - | 1.7 | - | 3 | 0.39 | 0.55 | 0.53 | 0.49 |
| | B. reduncus-B. scotti | - | 1.3 | - | 3 | 0.71 | 0.72 | 0.82 | 0.89 |
| **TOTAL** | | | | | **9** | **0.52** | **0.48** | **0.66** | **0.67** |

## *Phylogenetic Niche Conservation Comparisons*

Only *Inoceramus oblongus* vs. *I. sagensis* comparison in the *B. reesidei-B. eliasi* interval had statistically greater niche equivalence than would be expected at random (Appendix C, Table S14). Four of the nine *Inoceramus* species (~ 44%) showed greater niche similarity than would be expected at random (Table 12 and Appendix C, Table S16). Only one species comparison (*I. barabini* vs. *I. sagensis* in the *B. reesidei-B. eliasi* biozone interval) had greater niche stability than would be expected at random using the niche similarity test. Overlap plots support these results (Figure 18 and S5).

Comparisons of the *Inoceramus* species with the three non-*Inoceramus* species indicate that statistically high overlap and stability is only slightly more common between related species than unrelated ones based on the niche similarity test (Table 12). When using the niche equivalency test, seven of the nineteen non-*Inoceramus* to *Inoceramus* species comparisons (~ 26%) had greater niche equivalence than would be expected at random and three of the nineteen comparisons (~16%) had greater stability and lower niche expansion than would be expected at random (Appendix C, Table S19). Two non-*Inoceramus* comparisons showed statistically greater unfilling based on the niche equivalency test (Appendix C, Table S20). Seven of the nineteen comparisons (~ 37%) had statistically high D-values, high stability, and low expansion, and four had statistically low unfilling when using niche similarity tests (Appendix C, Table S21). Therefore, the proportion of species with statistically similar niches is slightly higher between related taxa (Appendix C, Table S16), but statistical equivalence was more common between unrelated taxa. Mean niche stability for the non-*Inoceramus* to *Inoceramus* species comparisons (0.76) was higher than within *Inoceramus* species comparisons (0.62; Appendix C, Table 11). Mean expansion was

lower between within *Inoceramus* comparisons relative to non-*Inoceramus* to *Inoceramus* comparisons (0.24 versus 0.38, respectively), and mean unfilling was approximately the same (0.32 and 0.34 respectively; Tables S18 and S13). Plots of niche comparisons for non-*Inoceramus* taxa support these results (Figures 19 and S6). Surprisingly, these analyses suggest that stability and overlap was higher on average between evolutionarily more distant taxa vs. between sister taxa. Statistical assessments furthermore do not indicate that niche equivalency was higher between related taxa than unrelated taxa, and proportions of statistically high overlap and stability based on niche similarity were not distinct.

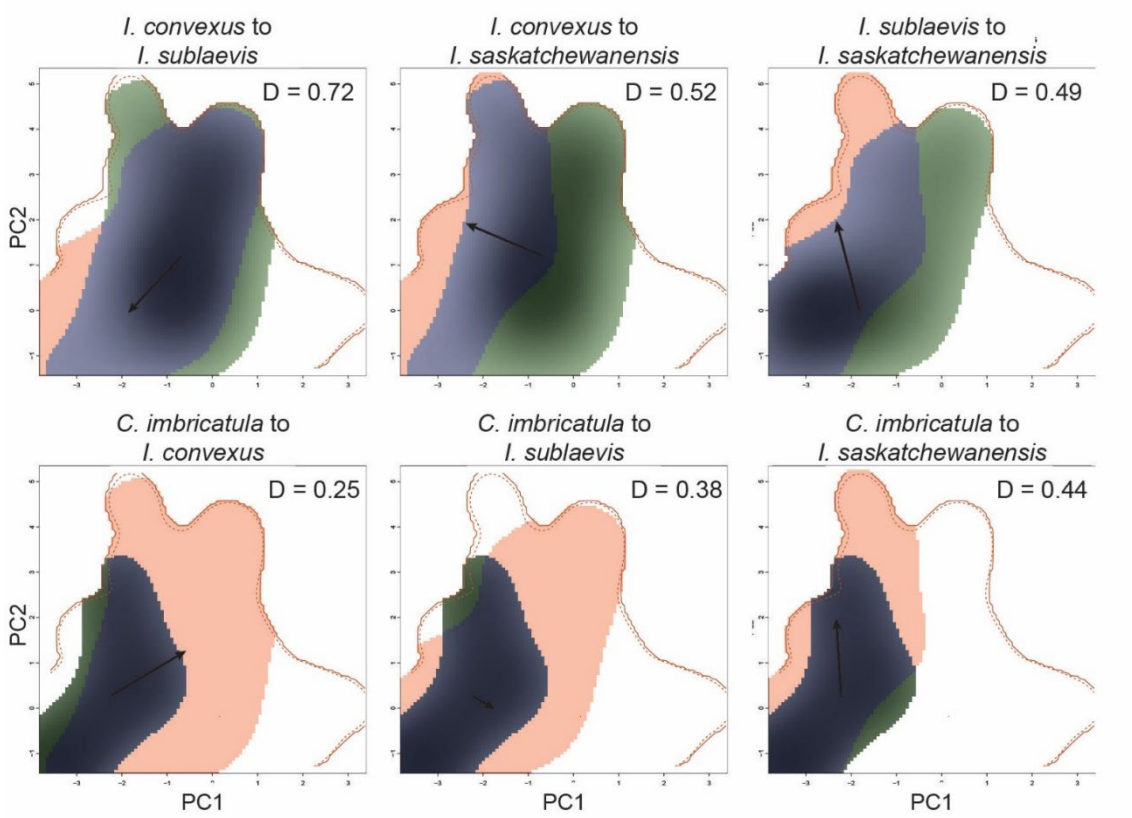



**Figure 18.** Example overlap plots of Inoceramus species against each other and against *Ctena imbricatula* during the *B. reduncus-B. scotti* biozone interval. Blue represents regions of niche stability, red regions of niche expansion (i.e., where the second listed species has expanded beyond the niche of the first listed species), and green represents regions of niche unfilling (i.e., where the second listed species niche has moved away from space occupied by the first listed species). Red lines represent the extent of environmental space realized within the interval and the black arrow indicates the shift of the niche centroid between the two species. See Figures S5 and S6 for all species overlap plots.

**Table 12.** Results (p-value, alpha=0.5) of greater niche similarity tests for phylogenetic comparisons. Bold text and asterisks indicate significant results.

| Phylogenetic Conservation Comparisons | | | |
|---|---|---|---|
| Species Compared | Interval | D value | Stability (higher) |
| Ingroup Comparisons | | | |
| I. convexus vs. I. saskatchewanensis | B. reduncus-B. scotti | 0.112 | 0.381 |
| I. sublaevis vs. I. saskatchewanensis | B. reduncus-B. scotti | 0.221 | 0.353 |
| I. convexus vs. I. sublaevis | B. reduncus-B. scotti | 0.431 | 0.392 |
| I. convexus vs. I. barabini | D. nebrascense-E. jenneyi | 0.429 | 0.209 |
| I. convexus vs. I. sublaevis | D. nebrascense-E. jenneyi | 0.431 | 0.392 |
| I. barabini vs. I. sublaevis | D. nebrascense-E. jenneyi | **0.001*** | 0.251 |
| I. barabini vs. I. sagensis | B. reesidei-B. eliasi | **0.015*** | **0.026*** |
| I. barabini vs. I. oblongus | B. reesidei-B. eliasi | **0.038*** | 0.072 |
| I. oblongus vs. I. sagensis | B. reesidei-B. eliasi | **0.048*** | 0.058 |
| Comparison taxa Comparisons | | | |
| C. imbricatula vs. I. convexus | B. reduncus-B. scotti | 0.411 | 0.295 |
| C. imbricatula vs. I. sublaevis | B. reduncus-B. scotti | 0.306 | 0.147 |
| C. imbricatula vs. I. saskatchewanensis | B. reduncus-B. scotti | 0.112 | 0.074 |
| C. nebrascensis vs. I. barabini | D. nebrascense-E. jenneyi | **0.001*** | **0.001*** |
| C. imbricatula vs. I. convexus | D. nebrascense-E. jenneyi | 0.398 | 0.305 |
| C. imbricatula vs. I. sublaevis | D. nebrascense-E. jenneyi | 0.356 | 0.626 |
| C. nebrascensis vs. I. barabini | D. nebrascense-E. jenneyi | **0.001*** | **0.001*** |
| C. nebrascensis vs. I. convexus | D. nebrascense-E. jenneyi | 0.449 | **0.015*** |
| C. nebrascensis vs. I. sublaevis | D. nebrascense-E. jenneyi | **0.015*** | **0.045*** |
| L. subundata vs. I. barabini | D. nebrascense-E. jenneyi | 0.101 | 0.138 |
| L. subundata vs. I. convexus | D. nebrascense-E. jenneyi | 0.501 | 0.356 |
| L. subundata vs. I. sublaevis | D. nebrascense-E. jenneyi | **0.028*** | 0.448 |
| L. subundata vs. I. sagensis | B. compressus-B. cuneatus | 0.265 | 0.158 |
| C. nebrascensis vs. I. barabini | B. reesidei-B. eliasi | **0.006*** | 0.098 |
| C. nebrascensis vs. I. sagensis | B. reesidei-B. eliasi | 0.067 | **0.045*** |
| C. nebrascensis vs. I. oblongus | B. reesidei-B. eliasi | 0.101 | 0.083 |
| C. imbricatula vs. I. barabini | B. reesidei-B. eliasi | **0.017*** | 0.172 |
| C. imbricatula vs. I. sagensis | B. reesidei-B. eliasi | **0.044*** | **0.035*** |
| C. imbricatula vs. I. oblongus | B. reesidei-B. eliasi | 0.076 | **0.018*** |

### *Species to Genus Comparisons (Niche Partitioning)*

Only *I. barabini* in the *B. reesidei-B. eliasi* interval had a statistically greater high niche overlap when compared with the genus-level niche based on niche similarity tests, as well as greater than random niche stability and lower than random niche expansion (Tables 13 and S26). In the *D. nebrascense-E. jenneyi* interval, *I. barabini* also had lower niche unfilling than would be expected at random, but no other comparison of overlap or stability was statistically high (Appendix C, Table S26). In the *D. nebrascense-E. jenneyi* interval, *I. sublaevis* had higher than expected niche unfilling based on the niche equivalency test (Appendix C, Table S25). No other values were statistically significant based on the niche equivalency test.

Mean niche stability for all species to genus comparisons was approximately equal to the mean stability observed among *Inoceramus* species themselves, 52% (Table 11). Unfilling was notably very low for species to genus comparisons, ranging from 0% to 24% and averaging only 5% (Appendix C, Table S23). Niche expansion was more notable, ranging from 9% to 68% and averaging 34% (Appendix C, Table S23). Niche expansion is expected to be higher and niche unfilling extremely low for species to genus comparisons since the genus niche (always analyzed as the "secondary" niche) is expected to be larger. Our results are consistent with these expectations and suggest that niche stability and overlap were not high overall between species and the genus. In cases where the species niche did not overlap with the genus niche (unfilling), it is plausible that spatial thinning removed relevant occurrence points of specific species from the genus, resulting in a broader species' ecological extent from the genus-level data. This is observed for *I. sublaevis* in the *D.*

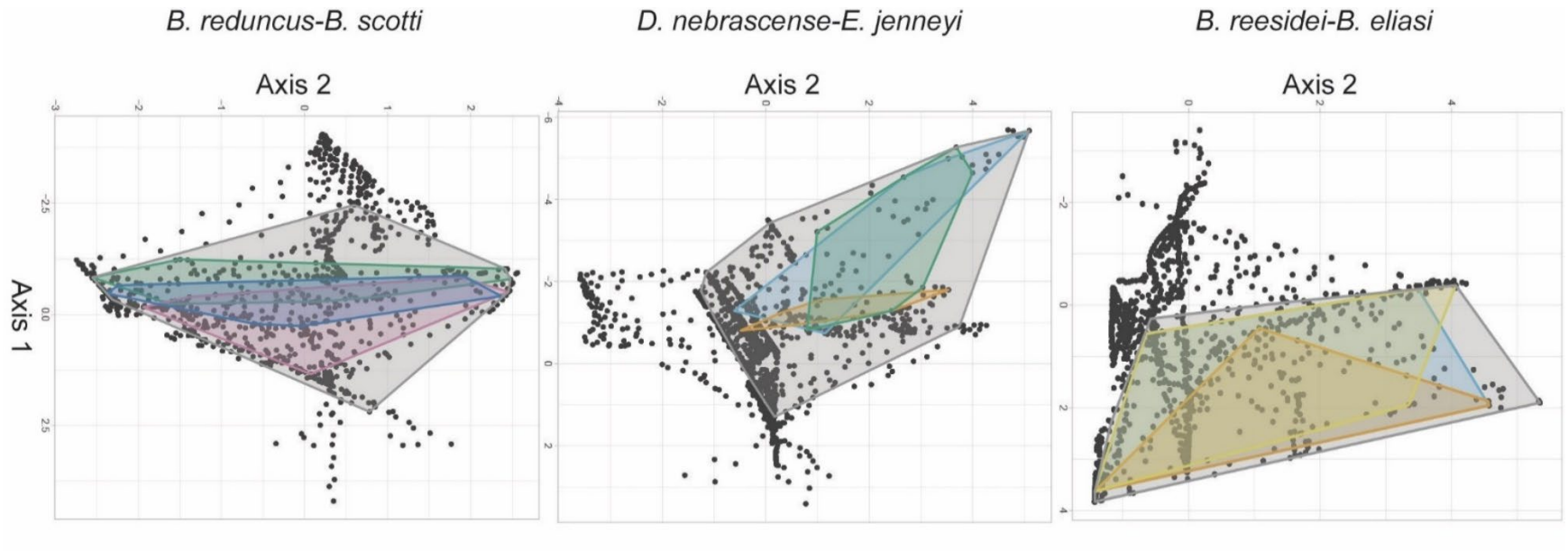



**Figure 19.** Convex hull comparisons of the Inoceramus genus and Inoceramus species niches in PCS for three biozone intervals. Polygons represent occupied environmental spaces for each taxa within PC environmental space. Dark grey points represent environmental data observed in the interval. Colors indicate *Inoceramus* species (green = *I. barabini,* light blue = *I. convexus*, pink = *I. sagensis*, dark blue = *I. oblongus*, orange = *I. sublaevis*, yellow = *I. saskatchewanensis*); grey = *Inoceramus* genus estimation.

*nebrascense-E. jenneyi* comparison with *Inoceramus*, which had statistically high unfilling based on niche similarity test results, caused by a gap in the *Inoceramus* niche distribution, presumably from spatial thinning (Appendix C, Figure S7). Principal component comparisons of multiple *Inoceramus* species to the genus niche using convex hulls visualizes the structure of species niches within the generic niche (Figure 19). These plots show species' niche partitioning to some degree within the genus-level niche, but also that overlap is common between species.

**Table 13.** Results (p-value, alpha=0.5) of greater niche similarity tests for phylogenetic comparisons. Bold text and asterisks indicate significant results.

| Species | Interval | D value | Stability (higher) |
|---|---|---|---|
| *I. convexus* | *B. reduncus-B. scotti* | 0.114 | 0.225 |
| *I. sublaevis* | *B. reduncus-B. scotti* | 0.165 | 0.144 |
| *I. saskatchewanensis* | *B. reduncus-B. scotti* | 0.060 | 0.145 |
| *I. convexus* | *D. nebrascense-E. jenneyi* | 0.292 | 0.202 |
| *I. sublaevis* | *D. nebrascense-E. jenneyi* | 0.137 | 0.347 |
| *I. barabini* | *D. nebrascense-E. jenneyi* | 0.152 | 0.170 |
| *I. barabini* | *B. reesidei-B. eliasi* | **0.033*** | **0.041*** |
| *I. sagensis* | *B. reesidei-B. eliasi* | 0.206 | 0.089 |
| *I. oblongus* | *B. reesidei-B. eliasi* | 0.212 | 0.076 |

## Discussion

### *The Effects of Temporal Aggregation in Paleo-niche Analyses*

The temporal resolution used for deep-time analysis is typically constrained by logistical factors such as geologic age constraints, inadequate taxonomic abundances, and issues with regional correlation. The combination of these factors frequently forces researchers to aggregate sedimentary data at the lower resolutions (substage or higher) when analyzing paleo-niche parameters (e.g., Maguire and Stigall 2009; Hopkins et al. 2014; Saupe et al. 2019; Purcell and Stigall 2021). Higher-resolution studies can be achieved but generally

involve investigating less geographically extensive regions, such as a single, well-sampled sedimentary basin with well correlated high-resolution stratigraphic data (e.g., Brame & Stigall, 2014; Dudei & Stigall, 2010; Malizia & Stigall, 2011). Generally, these studies are forced to assume that data homogenization will increase analytical noise, but not modify the overall signal of niche patterns, if they exist (e.g., Purcell & Stigall, 2021). Here we specifically tested for the influence of temporal aggregation and found that niche characteristics are not well conserved at lower temporal resolutions, although temporal aggregation does not typically introduce unique niche e-space occupation.

Across the Late Cretaceous, temporal biozone intervals ranging in duration from ~ 0.7 Myrs (*B. compressus-B. cuneatus*) to ~ 1.8 Myrs (*D. nebrascense-E. jenneyi)* were compared with longer intervals lasting ~ 2 Myrs (early Maastrichtian) to ~ 4.3 Myrs (late Campanian). We were also able to compare the *B. baculus-B. grandis* biozone interval with the Maastrichtian stage overall, which represents ~ 6.1 Myrs of deposition. Across these comparisons, differences in temporal duration were found to correlate with decreased niche overlap (Figure 17). Lower temporal resolution of the original data therefore did not well-characterize the niche of species at comparatively higher resolutions. Within these temporal comparisons, species niches had high overlap (mean D-value = 0.43) and stability (mean = 72%) but values were often statistically insignificant either for the niche equivalency or similarity tests (Table 9). Only ~ 33% of the taxa analyzed had statistically high D-values of overlap and only ~ 24% had statistically high stability.

Furthermore, map interpolations at different temporal resolutions show notable differences between biozone-level bins within most environmental proxy variables (Appendix C, Figure S2). For example, percent mud is high (greater than 50%) only in the

northeastern portion of the basin and lower (less than 30%) in the rest of the basin during the *D. nebrascense-E. jenneyi* interval. During the succeeding *D. cheyennense* interval, however, percent mud is relatively high (greater than 30%) in the majority of the southern and easter portions of the basin (Figure 16). These patterns are in turn distinct from the distribution of mud within the other two biozones that make up the Late Campanian. When aggregated at the Late Campanian substage level, the regions with highest percent mud are in the northern portion of the basin, within two distinct "patches" (Appendix C, Figure S2). Even in higher-resolution studies using fossil data that are limited to shorter duration aggregations, such as 0.5 Myrs temporal bins, environmental variables are aggregated across intervals potentially much longer than the duration of the species occupation (i.e., Dudei and Stigall, 2010). Basin geometry is also highly variable across the study interval, with marine areas shifting rapidly with changing sea levels (both eustatic and tectonic), particularly across the Late Campanian and Maastrichtian (Figure 16) (Cobban et al., 1994; Gill & Cobban, 1973; Roberts & Kirschbaum, 1995; Slattery et al., 2013). This almost certainly leads to a distortion of environmental averages within regions that shift from marine to non-marine within a single interval. Changes to shorelines are likely occurring even within biozone intervals themselves, though these alternations are too high-resolution to detect in this dataset (and perhaps any). Thus, aggregating data at a higher geochronological level necessarily homogenizes and distorts otherwise localized or regional characteristics. This results not only in a loss of spatiotemporal detail, but also potentially erroneous spatial correlations between shorter-lived taxa and long-term environmental averages.

While changes in shoreline are not ubiquitous to all regions or intervals of Earth history, periods of environmental change are common targets for deep-time paleo-niche

analysis. Consequently, researchers must pay close attention to the influence that temporal aggregation has on the interpretation of their results, and not assume homogenization will represent niche characteristics accurately. In all paleoecological analyses, the highest temporal resolution available should be used whenever possible to avoid over-generalizing variables, and changes to broader paleoceanography and data distributions should be carefully considered before attempting or interpreting these kinds of analyses. Paleo-niche analysis can provide useful information provided the resolution of the data matches the question being asked, however. For example, analyses that are conducted to assess how species shift their niche space through time in response to changes in temperature can constrain their area of analyses to consistently marine environments that lack dramatically shifting shoreline.

*Implications for Testing Phylogenetic Niche Conservation in Deep Time*

Phylogenetic niche conservation is a debated biotic pattern relevant to both ecological and evolutionary trends (Crisp and Cook 2012; Pyron et al. 2015). Related taxa are generally considered to have more ecologically similar environmental requirements than unrelated taxa, given the fact that they share more recent ancestral traits (Hadly et al. 2009; Nürnberg and Aberhan 2013). Indeed, PNC has been observed in both allopatric species (Peterson et al. 1999) and in sympatric species (Lovette and Hochachka 2006), indicating that both physical separation of populations and genetic divergence can result in conservation of fundamental ecological traits. If PNC is widespread over macroevolutionary timescales, then it becomes a primary justification supporting the application of genus-level niche traits in paleo-niche modeling and other deep time applications (Dudei and Stigall 2010; Malizia and Stigall 2011; Nürnberg and Aberhan 2013; Brame and Stigall 2014; Hendricks et al. 2014; Purcell

and Stigall 2021). Our secondary hypothesis predicted niche conservatism between sister species for this reason. However, niche comparisons between unrelated taxa produced higher average stability than between related taxa (76% and 62%, respectively) and very similar overlap values (50% and 52%, respectively). The proportion of comparisons with statistically greater niche stability was lower for within-genus species comparisons than species comparisons across different genera (~ 11% and 37%, respectively), and statistically high overlap was only slightly more common for within-genus versus between-genus species comparisons (~ 44% and 37%, respectively). These results reject our hypothesis of greater niche conservation between related versus among less closely related taxa and suggest that either PNC is not present within the taxa considered, or substantial biases prevent PNC from being observed. A lack of PNC may be, and was here, observed as within-genus species niche partitioning, which is generally regarded as important to promote speciation (Cavender-Bares et al. 2004; Graham et al. 2004); it is also likely that evidence for PNC is taxon- and/or spatiotemporal scale-specific. Differentiating niche characteristics may furthermore simply not be observable at a given level of analysis or method of niche estimation (Lovette and Hochachka 2006), caused by environmental variable biases and/or the fundamental characteristics in the taxa selected for analyses.

Specifically, low spatial and temporal resolution, including potentially biozone-level resolutions, and the broad-scale distribution of the taxa analyzed likely culminates in an over-generalization of environmental characteristics and a bias towards preservation of generalist taxa. Deep-time environmental variables are especially prone to over-generalization in these kinds of analyses, since they are by necessity time-averaged, autocorrelated proxies (Appendix C, Figure S3). For example, oxygen levels are critical to habitat restrictions of

marine biota (Deutsch et al. 2015), but oxygen levels can only be approximated in sedimentary units by a combination of variables such as sedimentary structures (indicating energy level and water depth) and bioturbation, which is broadly associated with oxygen levels sufficient to allow for activity below the sediment-water interface (Leszczyński 1991). However, neither sedimentary structure nor degree of bioturbation are characterized in sufficient detail in most datasets to capture much more than "oxic-to-some-degree" vs. "anoxic" at each specific location, often due to insufficient preservation and/or the idiosyncrasies of documentation. Generalizations like these culminate in a simplification of real-world patterns such that higher resolution niche differentiation may be difficult to observe, resulting in high niche similarity. Furthermore, other aspects of species' realized niches, such as biotic interactions, cannot be well-characterized in most cases, even using more recent datasets (Peterson et al. 2011).

These results furthermore indicate that the current method of approximating environmental characteristics using continuous sedimentary variable proxies may be less useful than manually classifying relevant habitat regions on a more categorical level. For example, anoxic and oxygen-rich marine habitats can be interpreted directly by a researcher based on a collection of sedimentological variables and mapped accordingly. While these variable maps may also include generalizations, they have the potential to more accurately represent cumulative environmental variables that are directly relevant to taxonomic distributions in a way that cannot be easily captured using sedimentary data alone. Niche models created using ground truthed categorical habitats such as these may therefore produce more definitive results when comparing niche characteristics between taxa or across time, and should be tested in future analyses.

In addition to the issues with environmental generalization, taxonomic preservation and sampling bias in the fossil record further complicate the resolution issue, since only a very small subset of a population (even across a long temporal interval) is fossilized, collected, and documented (Kidwell and Flessa 1996). This means that it is challenging to collect species occurrences across their entire geographic and temporal range. Niche studies on modern animals and plants have shown that not using the full biogeographic range of species' distributions can obscure accurate niche estimation, particularly when using ENM algorithms (i.e., Owens et al. 2013). Furthermore, when conducting deep-time niche analysis, these occurrences are further reduced by spatial thinning, which again promotes analysis of common, widespread, and long-lived species rather than taxa with patchy or localized distributions (whether those be real or artifactual). It is commonly assumed that wide-ranging taxa have large niches and are therefore more ecologically generalized (Kammer et al. 1997; Slatyer et al. 2013). Even under ideal circumstances with high-resolution, modern data, generalist taxa have been found to have similar ecological characteristics to one another (Denelle et al. 2020). Given that subtle environmental differences are unlikely to be preserved when using deep-time environmental proxy data, ecological distinctions between generalist taxa occupying similar overall habitats are also unlikely to be observed. The high similarity between both related and unrelated species observed here supports this interpretation. Given that this analysis includes some of the highest resolution and best-vetted data available for paleo-niche analysis, these results suggest that PNC is unlikely to be easily distinguished in fossil data unless higher resolution datasets with a greater biogeographic diversity of taxa and their environments can be achieved. Instead, fossil data is more applicable to analyzing broad-scale changes in niche parameters within generalist species.

*Within-species partitioning of generic niches*

In paleontological studies, the genus-level, or even above, is often used for ecological analyses since because (1) genus-level data is so much more abundant and (2) genus-level patterns are considered a decent proxy of species-level biotic processes (Hadly et al. 2009; Malizia and Stigall 2011; Nürnberg and Aberhan 2013; Brame and Stigall 2014; Tong et al. 2021). Others argue that the species is the fundamental biological unit of analysis (Vrba 1980, 1984, 1989), and that higher taxonomic levels do not inherently capture ecologically relevant biotic responses (Hendricks et al. 2014). For example, while a genus' niche might show expansion across time due to a few species on the extremities of the genus niche moving into new abiotic space, individual species that do not occupy marginal regions of the genus' niche may instead be experiencing substantial reduction of their niche space occupation in a way that cannot be observed at the genus-level (Hendricks et al. 2014; Purcell and Stigall 2021). This can be visualized here with convex hull reconstructions (Figure 19). In the *B. reduncus-B. scotti* interval, for example, the extremities of the first principal component axis of the genus-level niche are not occupied by the three species analyzed directly, and any change to these species across this axis could not be observed at the genus-level (Figure 19). Furthermore, our results of niche comparisons between *Inoceramus* species and the *Inoceramus* genus-level niche estimation suggest that the genus-level niche does not represent individual species well statistically (Table 11). Average niche overlap for species-level analysis compared to the genus-level is only 0.52, and only one species (*I. barabini*) had statistically high overlap and stability based on the niche similarity test in only one biozone interval of two (Table 13). This contradicts several studies using fossil data that have found similar patterns of niche overlap and conservation between both species and genera (e.g., Hadly et al. 2009; Malizia and Stigall 2011; Nürnberg and Aberhan

2013; Brame and Stigall 2014). How time influences a genus-to-species niche relationship is beyond the scope of this analysis but given the poor overlap in this study between individual species and the genus-level analysis, it is very possible that genus-level paleo-niche patterns reflect ecological phenomenon at the species-level.

Conversely, if the genus is considered a functional entity, rather than an evolutionary entity, its utility is vastly increased. Firstly, there are distinct advantages to utilizing higher-level taxonomic information for addressing biological and ecological traits in general. Obviously, using the genus greatly expands the utility of sparce datasets because it allows researchers to incorporate fossil data that has only be identified to the genus-level. Additionally, studies have found that higher taxonomic levels may present distinct biotic responses relevant to life history traits (Hadly et al. 2009; Smith et al. 2019; Tong et al. 2021). More importantly for using paleontological data for niche analyses, genera classifications are typically based on morphological similarity, which is directly linked in most cases with the functional traits relevant to an organism's lifestyle (i.e., Weller 1949). While this often means that genera cannot be considered evolutionarily entities, since they may represent convergent morphologies rather than phylogenetically related taxa, this principle enables researchers to classify genera into functional ecological entities that occupy a specific ecological role in their community strcture and assess their temporal and spatial distribution relative to abiotic factors on a macroecological level (Foster and Twitchett 2014; Dunhill et al. 2018; Edie et al. 2018).

Therefore, if we consider a genus niche to represent a *functional entity* rather than an evolutionary entity, the issues related to interpreting the genus relative to the species niche are largely resolved. Analyses using the genus niche can instead enhance our understanding

of how functional entities are distributed relative to fundamental abiotic characteristics and how they respond across space and time. The actual species-level changes occurring in these analyses may not be clear, depending on the genus and number of species included within it, but any genus-level alteration may be indicative of taxonomic or ecological instability (Purcell and Stigall 2021). For example, contraction of a genus niche requires either species extinction or contraction of species' niches (Figure 20). Depending on the degree of overlap between species within the niche (i.e., PNC), these alterations must furthermore be simultaneous across the same interval of comparison. Ignoring the specific species-level changes however, this pattern would indicate an overall ecological responses by the functional entity represented by the genus, and can lead to further, more detailed analyses related to abiotic pressures and ecological processes. Only genus niche stability, in fact, may be entirely nonspecific about the species niche patterns (Figure 20C), while still indicating that either niche characteristics of the functional entity are being maintained through stability at the species level or through compensation for any alteration by individual species by the others within the genus. Therefore, though the genus niche is not representative of individual species ecological niches, it is in fact highly relevant to understanding functional diversity patterns at a broad scale.

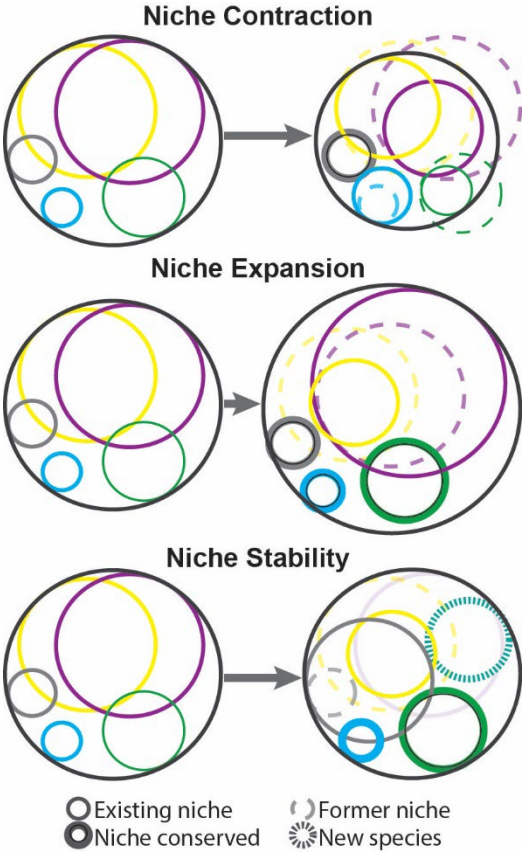



**Figure 20.** Examples of niche dynamics through time at the species and genus-level, which can be interpreted as representing the niche of an ecological entity. The genus-level niche is in black, species niches are colored.

## Conclusions

This dataset represents the highest-resolution, most geographically extensive, and thoroughly cleaned deep-time dataset used for niche comparisons to date. We find that paleo-niche analyses should be considered individualistically and wholistically to avoid over-generalizing biological phenomenon. The size of a temporal bin, particularly during intervals with substantial changes to paleoceanography and paleogeography, can strongly influence the outcome of niche estimation in deep time, and accuracy diminishes with environmental data aggregated over longer intervals. High niche similarity values resulting from such low-resolution analysis are therefore potentially misleading, given that they likely result from the over-generalization of already biased environmental proxy data and the use of geographically

widespread, generalist taxa. The case study comparing genus-level niche estimation to the species-level did not find that the generic-based niche represented individual species' niches well, despite the fact that this genus is composed of abundant, generalist species with relatively high niche overlap between them. This suggests that the structure of species' paleo-niches within genera is an exciting research area in need of more and specific analysis. Furthermore, if the genus is considered as a functional entity, genus-level niche analyses may open up an extremely useful avenue for further research into how abiotic factors across space and time influence the fundamental functional characteristics of ecological communities.

# CHAPTER 5: CONCLUSION

The Western Interior Seaway has been studied for over 200 years (Everhart, 2017), resulting in the collection of vast numbers of fossils and geologic data all collected with specific, and not necessarily complementary research goals. The Western Interior Seaway persisted from approximately the late Jurassic to the early Paleocene (approximately 100 Ma) and is one of the best preserved and studied sedimentary basins in the world (Miall et al., 2008). Not only is the WIS well studied as a whole, but the digital age has provided a previously unprecedented venue for cataloguing and disseminating paleontological and sedimentological data. The evolution of quantitative methods applicable to fossil data has also greatly advanced, and computer software has made exploring "big data" questions related to Earth history very accessible such that researchers can investigate large biostratigraphic regions like the WIS with ease relative to the methods available to previous researchers. Geologic data, recording aspects of the interacting biotic and abiotic aspects of the Earth system, can now more than ever before be used to address questions related to how the Earth functions in the past and evaluate underlying principles that may explain how it will progress in the future.

This expanding scientific horizon requires that we carefully assess the match between data uncertainty and the questions we are asking, but also the methods we are using to answer them. Limitations in paleontological data, while somewhat mitigated by our ever-growing datasets and increased understanding about biases, still exist and must be taken into account. Previously established paradigms regarding biotic patterns and processes, while not inherently wrong, require reassessment using updated information and methods to support their accuracy. For example, the analyses documented here indicate that biotic subprovinces

were not an established characteristic of the WIS, as previously established by Kauffman (1980). In short, the fundamental relationships between biotic and abiotic data that paleontologists and other researchers explore should be carefully considered not just within the confines of *a priori* factors, but also within the limits of what paleontological and geological data can achieve using current methods. One goal of this dissertation is to explicitly address some of these issues using a high-resolution, well establish dataset from an extensively studied, geographically extensive region; thus testing potential biases in as comprehensive a dataset as exists for deep time marine ecosystems.

The first research chapter (Chapter 2) uses relatively novel quantitative methods to reassess a long-established assumption regarding the paleobiogeography of the WIS, and the relationship between taxonomic distributions and abiotic factors. This chapter finds that the current dataset is incompatible with previous conclusions regarding WIS provinciality (Kauffman, 1980), supporting further research into the biogeographic structure of continental-sized regions in the fossil record. The second research chapter (Chapter 3) takes these methods further, addressing a popular topic in paleontology today: that functional diversity patterns result from ecological processes unique from taxonomic diversity, and tests the influence of spatial constraints on paleobiogeographic analytical results. This study compares functional diversity patterns with major abiotic shifts present at the same time at different spatial scales and contrasts these patterns with modern patterns. The last research chapter (Chapter 4) explores the validity of ecological niche modeling using deep-time datasets by comparing the influence that temporal resolution has on niche analyses and assessing the conservation of niche characteristics between related specific taxa and their genus. It furthermore provides an extensive and detailed stratigraphic dataset that

compliments the taxonomic data used in previous chapters and lays the groundwork for using stratigraphic data to create sedimentological maps of the WIS at the biostratigraphic level. This dataset will contribute significantly to continued studies linking paleobiogeographical and paleoenvironmental patterns in the WIS.

This research, while presenting novel data related to the biotic and abiotic system present in the WIS, are first and foremost linked to testing how methods shape research results (and consequent interpretations). They are meant to highlight the importance of considering data from multiple angles and scales, while also taking biases into account whenever feasible. The Late Cretaceous WIS represents one of the best studied, most vetted, and extensively documented datasets in the world, allowing researchers to not only explore highly relevant questions related to abiotic and biotic interactions, but also to establish principles for paleontological research moving forward. It is my hope that, regardless of whether the conclusions of this research remain eternally accurate, that these data and the methods used to explore them can create a foundation for future research, improving our understanding of the Earth system and its exciting history as a whole.

# REFERENCES


Aberhan, M., and W. Kiessling. 2015: Persistent ecological shifts in marine molluscan assemblages across the end-Cretaceous mass extinction. Proceedings of the National Academy of Sciences of the United States of America 112:7207–7212.

Adhikari, D., R. Tiwary, and S. K. Barik. 2015: Modelling hotspots for invasive alien plants in India. PLoS ONE 10.

Aguilar, C., J. B. Raina, S. Fôret, D. C. Hayward, B. Lapeyre, D. G. Bourne, and D. J. Miller. 2019: Transcriptomic analysis reveals protein homeostasis breakdown in the coral Acropora millepora during hypo-saline stress. BMC Genomics 20.

Aguirre-Gutiérrez, J., H. M. Serna-Chavez, A. R. Villalobos-Arambula, J. A. Pérez de la Rosa, and N. Raes. 2015: Similar but not equivalent: Ecological niche comparison across closely-related Mexican white pines. Diversity and Distributions 21:245–257.

Aiello-Lammens, M. E., R. A. Boria, A. Radosavljevic, B. Vilela, and R. P. Anderson. 2015: spThin: An R package for spatial thinning of species occurrence records for use in ecological niche models. Ecography 38:541–545.

Alroy, J. 2010: Geographical, environmental and intrinsic biotic controls on Phanerozoic marine diversification. Palaeontology 53:1211–1235.

Antell, G. T., R. B. J. Benson, and E. E. Saupe. 2024: Spatial standardization of taxon occurrence data—a call to action. Paleobiology:1–17.

de Arruda Almeida, B., A. J. Green, E. Sebastián-González, and L. dos Anjos. 2018: Comparing species richness, functional diversity and functional composition of waterbird communities along environmental gradients in the neotropics. PLoS ONE 13.

Bambach, R. K., A. M. Bush, and D. H. Erwin. 2007: Autecology and the filling of ecospace: Key metazoan radiations. Palaeontology 50:1–22.

Barron, E. J. 1989: Studies of Cretaceous Climate. Pp.*in* A. Berger, R. E. Dickson, and J. W. Kidson, eds. Understanding Climate Change. Vol. 52. .

Baselga, A., and C. D. L. Orme. 2012: Betapart: An R package for the study of beta diversity. Methods in Ecology and Evolution 3:808–812.

Benito Garzón, M., R. Alía, T. M. Robson, and M. A. Zavala. 2011: Intra-specific variability and plasticity influence potential tree species distributions under climate change. Global Ecology and Biogeography 20:766–778.

Braile, L. W., W. J. Hinze, G. R. Keller, E. G. Lidiak, and J. L. Sexton. 1986: Tectonic development of the New Madrid rift complex, Mississippi embayment, North America. Tectonophysics 131:1–21.

Brame, H.-M. R., and A. L. Stigall. 2014: Controls on niche stability in geologic time: congruent responses to biotic and abiotic environmental changes among Cincinnatian (Late Ordovician) marine invertebrates. Paleobiology 40:70–90.

Brodie, J. F., and P. D. Mannion. 2022: The hierarchy of factors predicting the latitudinal diversity gradient. Trends in Ecology and Evolution 38:15–23.

Broennimann, O., M. C. Fitzpatrick, P. B. Pearman, B. Petitpierre, L. Pellissier, N. G. Yoccoz, W. Thuiller, M. J. Fortin, C. Randin, N. E. Zimmermann, C. H. Graham, and A. Guisan. 2012: Measuring ecological niche overlap from occurrence and spatial environmental data. Global Ecology and Biogeography 21:481–497.

Brown, J. H. 1995: Macroecology. University of Chicago Press, Chicago and London, p.

Burgener, L., E. Hyland, E. Griffith, H. Mitášová, L. E. Zanno, and T. A. Gates. 2021: An extreme climate gradient-induced ecological regionalization of the Upper Cretaceous Western Interior Basin of North America. GSA Bulletin 133:2125–2136.

Bush, A. M., and R. K. Bambach. 2011: Paleoecologic megatrends in marine metazoa. Annual Review of Earth and Planetary Sciences 39:241–269.

Caldwell, W. G. E. 1968: The Late Cretaceous Bearpaw Formation in the South Saskatchewan River valley. Geology Division Report No. 8, Saskatchewan Research Council, Canada.

Caldwell, W. G. E. 1974: Western Interior Cretaceous. Geosc1ence Canada 2:11–20.

Cavender-Bares, J., D. D. Ackerly, D. A. Baum, and F. A. Bazzaz. 2004: Phylogenetic overdispersion in Floridian oak communities. American Naturalist 163:823–843.

Cobban, W. A., E. A. Merewether, T. D. Fouch, and J. D. Obradovich. 1994: Some Cretaceous shorelines in the Western Interior of the United States. Mesozoic Systems of the Rocky Mountain Region, USA:393–414.

Cobban, W. A., I. Walaszczyk, J. D. Obradovich, K. C. Mckinney, and M. D. Myers. 2006: A USGS Zonal Table for the Upper Cretaceous Middle Cenomanian−Maastrichtian of the Western Interior of the United States Based on Ammonites, Inoceramids, and Radiometric Ages. USGS Open-File Report 2006-1250:1–46.

Cochran, J. K., N. H. Landman, K. K. Turekian, A. Michard, and D. P. Schrag. 2003a: Paleoceanography of the Late Cretaceous (Maastrichtian) Western Interior Seaway of North America: evidence from Sr and O isotopes. Palaeogeography, Palaeoclimatology, Palaeoecology 191:45–64.

Cochran, J. K., N. H. Landman, K. K. Turekian, A. Michard, D. P. Schrag, and J. K. Cochran. 2003b: Paleoceanography of the Late Cretaceous (Maastrichtian) Western Interior Seaway of North America: evidence from Sr and O isotopes. Palaeogeography, Palaeoclimatology, Palaeoecology 191:45–64.

Di Cola, V., Olivier Broennimann, Blaise Petitpierre, Frank T. Breiner, Manuela D'Amen, Christophe Randin, Robin Engler, Julien Pottier, Dorothea Pio, Anne Dubuis, Loic Pellissier, Rubén G. Mateo, Wim Hordijk, Nicolas Salamin, and Antoine Guisan. 2017: ecospat: an R package to

support spatial analyses and modeling of species niches and distributions. Ecography 40:774–787.

Crisp, M. D., and L. G. Cook. 2012: Phylogenetic niche conservatism: What are the underlying evolutionary and ecological causes? New Phytologist 196:681–694.

Csárdi, G., and T. Nepusz. 2006: The igraph software package for complex network research. InterJournal Complex Systems:1–1695.

Dean, C. D., P. A. Allison, G. J. Hampson, and J. Hill. 2019: Aragonite bias exhibits systematic spatial variation in the Late Cretaceous Western Interior Seaway, North America. Paleobiology 45:571–597.

DeCelles, P. G. 2004: Late Jurassic to Eocene evolution of the Cordilleran Thrust Belt and Foreland Basin System, Western United States of America. American Journal of Science 304:105–168.

Denelle, P., C. Violle, and F. Munoz. 2020: Generalist plants are more competitive and more functionally similar to each other than specialist plants: insights from network analyses. Journal of Biogeography 47:1922–1933.

Dennis, K. J., J. K. Cochran, N. H. Landman, and D. P. Schrag. 2013: The climate of the Late Cretaceous: New insights from the application of the carbonate clumped isotope thermometer to Western Interior Seaway macrofossil. Earth and Planetary Science Letters 362:51–65.

Deutsch, C., A. Ferrel, B. Seibel, H. O. Pörtner, and R. B. Huey. 2015: Climate change tightens a metabolic constraint on marine habitats. Science 348:1132–1135.

Dudei, N. L., and A. L. Stigall. 2010: Using ecological niche modeling to assess biogeographic and niche response of brachiopod species to the Richmondian Invasion (Late Ordovician) in the Cincinnati Arch. Palaeogeography, Palaeoclimatology, Palaeoecology 296:28–43.

Dunhill, A. M., W. J. Foster, J. Sciberras, and R. J. Twitchett. 2018: Impact of the Late Triassic mass extinction on functional diversity and composition of marine ecosystems. Palaeontology 61:133–148.

Dwyer, C. H. 2019: Constraining the oxygen values of the Late Cretaceous Western Interior Seaway using marine bivalves.Thesis, University of New Mexico, Albuquerque, 1–92pp.

Edie, S. M., D. Jablonski, and J. W. Valentine. 2018: Contrasting responses of functional diversity to major losses in taxonomic diversity. Proceedings of the National Academy of Sciences of the United States of America 115:732–737.

Elderbak, K., and R. M. Leckie. 2016: Paleocirculation and foraminiferal assemblages of the Cenomanian-Turonian Bridge Creek Limestone bedding couplets: Productivity vs. dilution during OAE2. Cretaceous Research 60:52–77.

Fara, E., and M. J. Benton. 2000: The Fossil Record of Cretaceous Tetrapods. Vol. 15.p.

Finnegan, S., and M. L. Droser. 2008: Reworking diversity: Effects of storm deposition on evenness and sampled richness, Ordovician of the basin and range, Utah and Nevada, USA. Palaios 23:87–96.

Fisher, C. G., W. W. Hay, and D. L. Eicher. 1994: Oceanic front in the Greenhorn Sea (Late Middle through Late Cenomanian). Paleoceanography 9:879–892.

Floyd, M., M. Mizuyama, M. Obuchi, B. Sommer, M. G. Miller, I. Kawamura, H. Kise, J. D. Reimer, and M. Beger. 2020: Functional diversity of reef molluscs along a tropical-to-temperate gradient. Coral Reefs 39:1361–1376.

Forsyth, L. Z., and B. Gilbert. 2021: Parallel responses of species diversity and functional diversity to changes in patch size are driven by distinct processes. Journal of Ecology 109:793–805.

Foster, W. J., and R. J. Twitchett. 2014: Functional diversity of marine ecosystems after the Late Permian mass extinction event. Nature Geoscience 7:233–238.

Foster, W. J., C. L. Garvie, A. M. Weiss, A. D. Muscente, M. Aberhan, J. W. Counts, and R. C. Martindale. 2020: Resilience of marine invertebrate communities during the early Cenozoic hyperthermals. Scientific Reports 10.

Fricke, H. C., B. Z. Foreman, and J. O. Sewall. 2010: Integrated climate model-oxygen isotope evidence for a North American monsoon during the Late Cretaceous. Earth and Planetary Science Letters 289:11–21.

Gale, A. S., J. Mutterlose, S. J. Batenburg, F. M. Gradstein, F. P. Agterberg, J. G. Ogg, and M. R. Petrizzo. 2020: The Cretaceous Period. Pp.1023–1086 *in* Geologic Time Scale 2020. Elsevier.

Gill, J. R., and W. A. Cobban. 1966: The Red Section of the Upper Cretaceous Pierre Shale in Wyoming. Geological Survey Professional Paper 393-A:1–76.

Gill, J. R., and W. A. Cobban. 1973: Stratigraphy and Geologic History of the Montana Group and Equivalent Rocks, Montana, Wyoming, and North and South Dakota. Geological Survey Professional Paper 776:1–37.

Graham, C. H., S. R. Ron, J. C. Santos, C. J. Schneider, and C. Moritz. 2004: Integrating Phylogenetics and Environmental Niche Models to Explore Speciation Mechanisms in Dendrobatid Frogs. Vol. 58.p.

Graham, R. L., and P. Hell. 1985: On the History of the Minimum Spanning Tree Problem. Annals of the History of Computing 7:43–57.

Hadly, E. A., P. A. Spaeth, and C. Li. 2009: Niche conservatism above the species level. PNAS 106:19707–19714.

Halpern, B. S., and S. R. Floeter. 2008: Functional diversity responses to changing species richness in reef fish communities. Marine Ecology Progress Series 364:147–156.

Haq, B. U. 2014: Cretaceous eustasy revisited. Global and Planetary Change 113:44–58.

He, S., T. K. Kyser, and W. G. E. Caldwell. 2005: Paleoenvironment of the Western Interior Seaway inferred from δ18O and δ13C values of molluscs from the Cretaceous Bearpaw marine cyclothem. Palaeogeography, Palaeoclimatology, Palaeoecology 217:67–85.

Hendricks, J. R., E. E. Saupe, C. E. Myers, E. J. Hermsen, and W. D. Allmon. 2014: The Generification of the Fossil Record. Paleobiology 40:511–528.

Hernandez, P. A., C. H. Graham, L. L. Master, and D. L. Albert. 2006: The effect of sample size and species characteristics on performance of different species distribution modeling methods. Ecography 29:773–785.

Hijmans, R. J. 2021: geosphere: Spherical Trigonometry. R package version 1.5-14, https://CRAN.R-project.org/package=geosphere.

Hopkins, M. J., C. Simpson, and W. Kiessling. 2014: Differential niche dynamics among major marine invertebrate clades. Ecology Letters 17:314–323.

Huang, S., K. Roy, and D. Jablonski. 2014: Do past climate states influence diversity dynamics and the present-day latitudinal diversity gradient? Global Ecology and Biogeography 23:530–540.

Huber, B. T., D. A. Hodell, and C. P. Hamilton. 1995: Middle-Late Cretaceous climate of the southern high latitudes: Stable isotopic evidence for minimal equator-to-pole thermal gradients. GSA Bulletin 107:1164–1191.

Hutchinson, G. E. 1957: Concluding Remarks. Cold Spring Harbor Symposia on Quantitative Biology 22:415–427.

Jeletzky, J. A. 1971: Marine Cretaceous Biotic Provinces and Paleogeography of Western and Arctic Canada: Illustrated by a Detailed Study of Ammonites. Geological Survey Canada Paper 70–22:1–92.

Kammer, T. W., T. K. Baumiller, and W. I. Ausich. 1997: Species longevity as a function of niche breadth: Evidence from fossil crinoids. Geology 25:219–222.

Karadimou, E. K., A. S. Kallimanis, I. Tsiripidis, and P. Dimopoulos. 2016: Functional diversity exhibits a diverse relationship with area, even a decreasing one. Scientific Reports 6.

Kauffman, E. G. 1973: Cretaceous Bivalvia. Pp.353–383 *in* Hallam, ed. Atlas of Paleobiogeography. Elsevier Publishing Company, Amsterdam.

———. 1984: Paleobiogeography and evolutionary response dynamic in the Cretaceous Western Interior Seaway of North America. Jurassic-Cretaceous Biochronology and Paleogeography of North America, Geological Association of Canada Special Paper 27:273–306.

Kauffman, E. G., and W. G. E. Caldwell. 1993: The Western Interior Basin in space and time. Pp.1–30 *in* W. G. E. Caldwell and E. G. Kauffman, eds. Evolution of the Western Interior Basin. Vol. 39. Geological Association of Canada, Special Paper.

Kidwell, S. M., and K. W. Flessa. 1996: The quality of the fossil record: Populations, species, and communities. Annual Review of Earth Planetary Science 24:433–464.

Kiel, S. 2016: A biogeographic network reveals evolutionary links between deep-sea hydrothermal vent and methane seep faunas. Proceedings of the Royal Society B: Biological Sciences 283:1–9.

———. 2017: Using network analysis to trace the evolution of biogeography through geologic time: A case study. Geology 45:711–714.

Kivelä, M., S. Arnaud-Haond, and J. Saramäki. 2015: EDENetworks: A user-friendly software to build and analyse networks in biogeography, ecology and population genetics. Molecular Ecology Resources 15:117–122.

Kocsis, Á. T., and N. B. Raja. 2020: chronosphere: Earth system history variables. .

Kocsis, Á. T., C. J. Reddin, and W. Kiessling. 2018: The biogeographical imprint of mass extinctions. Proceedings of the Royal Society B: Biological Sciences 285:1–6.

Kocsis, Á. T., C. J. Reddin, C. R. Scotese, P. J. Valdes, and W. Kiessling. 2021: Increase in marine provinciality over the last 250 million years governed more by climate change than plate tectonics. Proceedings of the Royal Society B: Biological Sciences 288:1–7.

Leckie, R. M., R. F. Yuretich, O. L. O. West, D. Finkelstein, and M. Schmidt. 1998: Paleoceanography of the Southwestern Western Interior Sea During the Time of the Cenomanian-Turonian Boundary (Late Cretaceous). Pp.*in* W. E. Dean and M. A. Arthur, eds. Stratigraphy and Paleoenvironments of the Cretaceous Western Interior Seaway, USA. .

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

Leszczyński, S. 1991: Oxygen-related controls on predepositional ichnofacies in turbidites, Guipúzcoan Flysch (Albian-Lower Eocene), Northern Spain. PALAIOS 6:271–280.

Linnert, C., S. A. Robinson, J. A. Lees, P. R. Bown, I. Perez-Rodriguez, M. R. Petrizzo, F. Falzoni, K. Littler, J. A. Artz, and E. E. Russell. 2014: Evidence for global cooling in the Late Cretaceous. Nature Communications 5:1–7.

Liu, C., C. Wolter, F. Courchamp, N. Roura-Pascual, and J. M. Jeschke. 2022: Biological invasions reveal how niche change affects the transferability of species distribution models. Ecology 103.

Lockwood, J. L., M. F. Hoopes, and M. P. Marchetti. 2013: Invasive Ecology. John Wiley & Sons, Ltd., Publication, p.

Longman, M. W., B. A. Luneau, and S. M. Landon. 1998: Nature and Distribution of Niobrara Lithologies in the Cretaceous Western Interior Seaway of the Rocky Mountain Region. The Mountain Geologist 35:137–170.

Lovette, I. J., and W. M. Hochachka. 2006: Simultaneous effects of phylogenetic Niche conservatism and competition on avian community structure. Ecology 87:14–28.

Lowery, C. M., R. M. Leckie, R. Bryant, K. Elderbak, A. Parker, D. E. Polyak, M. Schmidt, O. Snoeyenbos-West, and E. Sterzinar. 2018: The Late Cretaceous Western Interior Seaway as a model for oxygenation change in epicontinental restricted basins. Earth-Science Reviews 177:545–564.

Lynds, R. M., and J. S. Slattery. 2017: Correlation of the Upper Cretaceous Strataa of Wyoming. Wyoming State Geological Survey Open Files Report 2017-3.

Mackenzie, R. A. I. 2007: Exploring Late Cretaceous Western Interior ammonoid geographic range and its relationship to diversity dynamic using geographic information systems (GIS).Masters Thesis, Bowling Green Sate University, 1–335pp.

Magneville, C., N. Loiseau, C. Albouy, N. Casajus, T. Claverie, A. Escalas, F. Leprieur, E. Maire, D. Mouillot, and S. Villéger. 2022: mFD: an R package to compute and illustrate the multiple facets of functional diversity. Ecography 2022.

Maguire, K. C., and A. L. Stigall. 2009: Using Ecological Niche Modeling for Quantitative Biogeographic Analysis: A Case Study of Miocene and Pliocene Equinae in the Great Plains. Vol. 35.p.

Magurran, A. E. 2003: Measuring Biological Diversity. Blackwell Science Ltd., Malden, MA, p.

Malinverno, A., K. W. Quigley, A. Staro, and J. Dyment. 2020: A Late Cretaceous-Eocene geomagnetic polarity timescale (MQSD20) that steadies spreading rates on multiple mid-ocean ridge flanks. Journal of Geophysical Research: Solid Earth 125:1–19.

Malizia, R. W., and A. L. Stigall. 2011: Niche stability in Late Ordovician articulated brachiopod species before, during, and after the Richmondian Invasion. Palaeogeography, Palaeoclimatology, Palaeoecology 311:154–170.

Mannion, P. D., P. Upchurch, R. B. J. Benson, and A. Goswami. 2014: The latitudinal biodiversity gradient through deep time. Trends in Ecology and Evolution 29:42–50.

Mannion, P. D., R. B. J. Benson, P. Upchurch, R. J. Butler, M. T. Carrano, and P. M. Barrett. 2012: A temperate palaeodiversity peak in Mesozoic dinosaurs and evidence for Late Cretaceous geographical partitioning. Global Ecology and Biogeography 21:898–908.

Mareš, M. 2008: The saga of minimum spanning trees. Computer Science Review 2:165–221.

Menegotto, A., and T. F. Rangel. 2018: Mapping knowledge gaps in marine diversity reveals a latitudinal gradient of missing species richness. Nature Communications 9.

Merewether, E. A., and K. C. McKinney. 2015: Composite biostratigraphic outcrop sections for Cretaceous Formations along a South-trending transect from western Montana, western Wyoming, eastern Utah, northeastern Arizona, and northwestern New Mexico, U.S.A. U.S. Geological Survey Open-File Report 2015-1087:1–10.

Miall, A. D., O. Catuneanu, B. K. Vakarelov, and R. Post. 2008: The Western interior basin. Pp.329–362 *in* A. D. Miall, ed. Sedimentary Basins of the World. Vol. 5. .

Minor, K. P., A. Wroblewski, R. J. Steel, C. Olariu, and J. P. Crabaugh. 2022: Facies partitioning of fluvial, wave, and tidal influences across the shoreline-to-shelf architecture in the Western Interior Campanian Seaway, USA. Geological Society of London, Special Publications 523:487–523.

Moalic, Y., D. Desbruyeres, C. M. Duarte, A. F. Rozenfeld, Bachraty Charleyne, and S. Arnaud-Haond. 2012: Biogeography revisited with network theory: Retracing the history of hydrothermal vent communities. Systematic Biology 61:127–137.

Monger, J. W. H. 1993: Cretaceous tectonics of the North American Cordillera. Pp.31–47 *in* W. G. E. Caldwell and E. G. Kauffman, eds. Evolution of the Western Interior Basin: Geological Association of Canada Special Paper. Vol. 39. .

Muscente, A. D., A. Prabhu, H. Zhong, A. Eleish, M. B. Meyer, P. Fox, R. M. Hazen, and A. H. Knoll. 2018: Quantifying ecological impacts of mass extinctions with network analysis of fossil communities. Proceedings of the National Academy of Sciences of the United States of America 115:5217–5222.

Myers, C. E., R. A. MacKenzie, and B. S. Lieberman. 2013: Greenhouse biogeography: the relationship of geographic range to invasion and extinction in the Cretaceous Western Interior Seaway. Paleobiology 39:135–148.

Myers, C. E., A. L. Stigall, and B. S. Lieberman. 2015: PaleoENM: Applying ecological niche modeling to the fossil record. Paleobiology 41:226–244.

Nawrot, R., D. Scarponi, M. Azzarone, T. A. Dexter, K. M. Kusnerik, J. M. Wittmer, A. Amorosi, and M. Kowalewski. 2018: Stratigraphic signatures of mass extinctions: Ecological and sedimentary determinants. Proceedings of the Royal Society B: Biological Sciences 285.

Newman, M. E. J. 2012: Communities, modules and large-scale structure in networks. Nature Physics 8:25–31.

Nicholson, D. B., P. A. Holroyd, P. Valdes, and P. M. Barrett. 2016: Latitudinal diversity gradients in mesozoic non-marine turtles. Royal Society Open Science 3.

Nickols, K. J., J. Wilson White, J. L. Largier, and B. Gaylord. 2015: Marine population connectivity: Reconciling large-scale dispersal and high self-retention. American Naturalist 185:196–211.

Nürnberg, S., and M. Aberhan. 2013: Habitat breadth and geographic range predict diversity dynamics in marine Mesozoic bivalves. Paleobiology 39:360–372.

Ogg, J. G. (James G., Gabi. Ogg, and F. M. Gradstein. 2016: A concise geologic time scale 2016. Elsevier, p.

Oliveira, B. F., A. Machac, G. C. Costa, T. M. Brooks, A. D. Davidson, C. Rondinini, and C. H. Graham. 2016: Species and functional diversity accumulate differently in mammals. Global Ecology and Biogeography 25:1119–1130.

Opsahl, T. 2009: Structure and Evolution of Weighted Networks.PhD Dissertation, Queen Mary College University of London, London, 1–151pp.

Owens, H. L., L. P. Campbell, L. L. Dornak, E. E. Saupe, N. Barve, J. Soberón, K. Ingenloff, A. Lira-Noriega, C. M. Hensz, C. E. Myers, and A. T. Peterson. 2013: Constraints on interpretation of ecological niche models by limited environmental ranges on calibration areas. Ecological Modelling 263:10–18.

Petersen, S. v., C. R. Tabor, K. C. Lohmann, C. J. Poulsen, K. W. Meyer, S. J. Carpenter, J. M. Erickson, K. K. S. Matsunaga, S. Y. Smith, and N. D. Sheldon. 2016: Temperature and salinity of the late cretaceous western interior seaway. Geology 44:903–906.

Peterson, A. T. 2001: Predicting Species' Geographic Distributions Based on Ecological Niche Modeling. Vol. 103.p.

———. 2003: Predicting the geography of species' invasions via ecological niche modeling. The Quarterly Review of Biology 78:419–433.

Peterson, A. T. 2011: Ecological niche conservatism: A time-structured review of evidence. Journal of Biogeography 38:817–827.

Peterson, A. T., J. Soberón, and V. Sánchez-Cordero. 1999: Conservatism of ecological niches in evolutionary time. Science 285:1265–1267.

Peterson, A. T., J. Soberón, R. G. Pearson, R. P. Anderson, E. Martínez-Meyer, M. . Nakamura, and M. B. Araújo. 2011: Ecological niches and geographic distributions. Princeton University Press, Princeton, NJ, p.

Phillips, S. B., V. P. Aneja, D. Kang, and S. P. Arya. 2006: Modelling and analysis of the atmospheric nitrogen deposition in North Carolina. International Journal of Global Environmental Issues 6:231–252.

Pimiento, C., C. D. Bacon, D. Silvestro, A. Hendy, C. Jaramillo, A. Zizka, X. Meyer, and A. Antonelli. 2020: Selective extinction against redundant species buffers functional diversity. Proceedings of the Royal Society B: Biological Sciences 287.

Planas, E., E. E. Saupe, M. S. Lima-Ribeiro, A. T. Peterson, and C. Ribera. 2014: Ecological niche and phylogeography elucidate complex biogeographic patterns in Loxosceles rufescens (Araneae, Sicariidae) in the Mediterranean Basin. BMC Evolutionary Biology 14.

Poulsen, C. J. 2004: A balmy Arctic. Nature 432:814–815.

Purcell, C., L. Scuderi, and C. Myers. 2023: Faunal provinciality in the Late Cretaceous Western Interior Seaway using network modeling. Geology 51:839–844.

Purcell, C. K. Q., and A. L. Stigall. 2021: Ecological niche evolution, speciation, and feedback loops: Investigating factors promoting niche evolution in Ordovician brachiopods of eastern Laurentia. Palaeogeography, Palaeoclimatology, Palaeoecology 578.

Pyron, R. A., G. C. Costa, M. A. Patten, and F. T. Burbrink. 2015: Phylogenetic niche conservatism and the evolutionary basis of ecological speciation. Biological Reviews 90:1248–1262.

Roberts, L. N. R., and M. A. Kirschbaum. 1995: Paleogeography of the Late Cretaceous of the Western Interior of middle North America- Coal distribution and sediment accumulation. U.S. Geological Survey Professional Paper 1561:1–115.

Rojas, A., P. Patarroyo, L. Mao, P. Bengtson, and M. Kowalewski. 2017: Global biogeography of Albian ammonoids: A network-based approach. Geology 45:659–662.

Roopnarine, P. D., and K. D. Angielczyk. 2015: Community stability and selective extinction during the Permian-Triassic mass extinction. Science 350:90–93.

Saramäki, J., M. Kivelä, J. P. Onnela, K. Kaski, and J. Kertesz. 2007: Generalizations of the clustering coefficient to weighted complex networks. Physical Review E 75:027105.

Saupe, E. E., A. Farnsworth, D. J. Lunt, N. Sagoo, K. V Pham, and D. J. Field. 2019: Climatic shifts drove major contractions in avian latitudinal distributions throughout the Cenozoic. PNAS 116:12895–12900.

Saupe, E. E., J. R. Hendricks, R. W. Portell, H. J. Dowsett, A. Haywood, S. J. Hunter, and B. S. Lieberman. 2014: Macroevolutionary consequences of profound climate change on niche evolution in marine molluscs over the past three million years. Proceedings of the Royal Society B: Biological Sciences 281.

Saupe, E. E., H. Qiao, J. R. Hendricks, R. W. Portell, S. J. Hunter, J. Soberón, and B. S. Lieberman. 2015: Niche breadth and geographic range size as determinants of species survival on geological time scales. Global Ecology and Biogeography 24:1159–1169.

Schoener, W. 1968: The Anolis Lizards of Bimini: Resource Partitioning in a Complex Fauna. Ecology 49:704–726.

Schumm, M., S. M. Edie, K. S. Collins, V. Gómez-Bahamón, K. Supriya, A. E. White, T. D. Price, and D. Jablonski. 2019: Common latitudinal gradients in functional richness and functional evenness across marine and terrestrial systems. Proceedings of the Royal Society B: Biological Sciences 286.

Scott, G. R., and W. A. Cobban. 1986: Geologic and biostratigraphic map of the Pierre Shale in the Colorado Springs–Pueblo area, Colorado. U.S. Geological Survey Miscellaneous Investigation Series I–1627.

Sessa, J. A., T. J. Bralower, M. E. Patzkowsky, J. C. Handley, and L. C. Ivany. 2012: Environmental and biological controls on the diversity and ecology of Late Cretaceous through early Paleogene marine ecosystems in the U.S. Gulf Coastal Plain. Paleobiology 38:218–239.

Shcheglovitova, M., and R. P. Anderson. 2013: Estimating optimal complexity for ecological niche models: A jackknife approach for species with small sample sizes. Ecological Modelling 269:9–17.

Signor, P. W., and J. H. Lipps. 1982: Sampling bias, gradual extinction patterns and catastrophes in the fossil record. Geological Society of America Special Papers 190:291–296.

Slattery, J. S., W. A. Cobban, K. C. Mckinney, P. J. Harries, and A. L. Sandness. 2013: Early Cretaceous to Paleocene paleogeography of the Western Interior Seaway: The interaction of eustasy and tectonism. Wyoming Geological Association Guidebook:22–60.

Slatyer, R. A., M. Hirst, and J. P. Sexton. 2013: Niche breadth predicts geographical range size: A general ecological pattern. Ecology Letters 16:1104–1114.

Slingerland, R., L. R. Kump, M. A. Arthur, P. J. Fawcett, B. B. Sageman, and E. J. Barron. 1996: Estuarine circulation in the Turonian Western Interior seaway of North America. GSA Bulletin 108:941–952.

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, A. B., A. S. Gale, and N. E. A. Monks. 2001: Sea-Level Change and Rock-Record Bias in the Cretaceous: A Problem for Extinction and Biodiversity Studies. Paleobiology 27:241–253.

Smith, A. B., W. Godsoe, F. Rodríguez-Sánchez, H. H. Wang, and D. Warren. 2019: Niche Estimation Above and Below the Species Level. Trends in Ecology and Evolution 34:260–273.

Soberón, J., and A. T. Peterson. 2005: Interpretation of models of fundamental niches and species' distributional areas. Biodiversity Informatics 2:1–10.

Sohl, N. F. 1967: Upper Cretaceous Gastropods From the Pierre Shale at Red Bird, Wyoming. Geological Survey Professional Paper 393-B:1–46.

Sohl, N. F. 1971: North American Cretaceous Biotic Provinces Delineated by Gastropods. North America Paleontological Convention, Proceedings 2:1610–1638.

Sohl, N. F. 1987: Cretaceous Gastropods: Contrasts between Tethys and the Temperate Provinces. Journal of Paleontology 61:1085–1111.

Somerfield, P. J. 2008: Identification of the Bray-Curtis similarity index: Comment on Yoshioka (2008). Marine Ecology Progress Series 372:303–306.

Soo, P., and P. A. Todd. 2014: The behaviour of giant clams (Bivalvia: Cardiidae: Tridacninae). Marine Biology 161:2699–2717.

Stanley, S. M. 2008: Predation Defeats Competition on the Seafloor. Vol. 34. Winter, p.

Steel, R. J., P. Plink-Bjorklund, and J. Aschoff. 2012: Tidal deposits of the Campanian Western Interior Seaway, Wyoming, Utah and Colorado, USA. Pp.437–471 *in* R. A. Jr. Davis and R. W. Dalrymple, eds. Principles of Tidal Sedimentology. Springer, Netherlands.

Stigall, A. L. 2014: When and how do species achieve niche stability over long time scales? Ecography 37:1123–1132.

Sulemana, A., K. A. Monney, and J. P. Deikumah. 2022: Variations in Avian Species and Functional Diversity in Different Habitat Types in a Vulnerable Savannah Ecosystem in Ghana. International Journal of Ecology 2022.

Super, J. R., K. Chin, M. Pagani, H. Li, C. Tabor, D. M. Harwood, and P. M. Hull. 2018: Late Cretaceous climate in the Canadian Arctic: Multi-proxy constraints from Devon Island. Palaeogeography, Palaeoclimatology, Palaeoecology 504:1–22.

Tong, Y. J., H. D. Yang, J. J. Shaw, X. K. Yang, and M. Bai. 2021: The relationship between genus/species richness and morphological diversity among subfamilies of jewel beetles. Insects 12:1–15.

Valencia-Rodríguez, D., L. Jiménez-Segura, C. A. Rogéliz, and J. L. Parra. 2021: Ecological niche modeling as an effective tool to predict the distribution of freshwater organisms: The case of the Sabaleta Brycon henni (Eigenmann, 1913). PLoS ONE 16.

Varela, S., J. M. Lobo, J. Rodríguez, and P. Batra. 2010: Were the Late Pleistocene climatic changes responsible for the disappearance of the European spotted hyena populations? Hindcasting a species geographic distribution across time. Quaternary Science Reviews 29:2027–2035.

Vilhena, D. A., and A. B. Smith. 2013: Spatial bias in the marine fossil record. PloS one 8.

Vrba, E. S. 1980: Evolution, species and fossils: how does life evolve? South African Journal of Science 76:61–84.

———. 1984: Individuals, hierarchies and processes: towards a more complete evolutionary theory. Paleobiology 10:123–164.

———. 1989: Levels of selection and sorting, with special reference to the species level. Vol. 6. Oxford Surveys of Evolutionary Biology, p.

Walsh, G., A. A. Pease, D. J. Woodford, M. L. J. Stiassny, J. Y. Gaugris, and J. South. 2022: Functional diversity of afrotropical fish communities across river gradients in the Republic of Congo, west central Africa. Frontiers in Environmental Science 10.

Wang, Y., O. V. Akeju, and T. Zhao. 2017: Interpolation of spatially varying but sparsely measured geo-data: A comparative study. Engineering Geology 231:200–217.

Warren, D. L., R. E. Glor, and M. Turelli. 2008: Environmental niche equivalency versus conservatism: Quantitative approaches to niche evolution. Evolution 62:2868–2883.

———. 2010: ENMTools: A toolbox for comparative studies of environmental niche models. Ecography 33:607–611.

Webster, R., and M. A. Oliver. 2007: Geostatistics for Environmental Scientists. John Wiley & Sons, Hoboken, New York., p.

Weller, J. M. 1949: Paleontologic Classification. Journal of Paleontology 23:680–690.

Wickham, H. 2016: Ggplot2 Elegant Graphics for Data Analysis. Springer, Houston, Texas, p.

Williams, G. D., and C. R. Stelek. 1975: Speculations on the Cretaceous Palaeogeography of North America. Pp.1–20 *in* The Cretaceous System in the Western Interior of North America: Geological Association of Canada Special Paper 13. .

Zhao, T., Y. Hu, and Y. Wang. 2018: Statistical interpretation of spatially varying 2D geo-data from sparse measurements using Bayesian compressive sampling. Engineering Geology 246:162–175.

Ziegler, A. M., and D. B. Rowley. 1998: The Vanishing record of eperic seas, with emphasis on the late creatacous “Hudson Seaway.” Pp.147–166 *in* T. J. Crowley and K. C. Burke, eds. Tectonic Boundary Conditions for Climate Reconstructions. Oxford University Press.

2021: Environmental Systems Research Institute (ESRI). ArcGIS Pro 3.0, Redlands, CA, p.

# APPENDICES

# APPENDIX A:
# Supplementary Materials for "Faunal provinciality in the Late Cretaceous Western Interior Seaway using a network modeling approach" by Purcell et al.:

Appendix A-1: Detailed Methods and Results

*Database Vetting:*

Original downloads from the PBDB and iDigBio were configured to include all taxa from the Late Cretaceous and then vetted to remove incomplete, erroneous, or poorly spatially and temporally resolved data. All genera denoted with "?", "c.f.", "aff." and taxa not identified to the genus-level were removed to avoid occurrences with insufficient taxonomic resolution. All taxonomic names were vetted using peer-reviewed literature to identify taxonomic updates, synonymies, and to correct any spelling errors (Appendix B provides all references used to vet taxonomic designations). Stratigraphic ages were updated and vetted to geologic substage temporal resolution based on stratigraphic and biozone information included in the original database file, primary literature sources cited for each occurrence, and the WIS stratigraphic database of Myers et al. (2015). Occurrence records with spatial resolution > 30km of uncertainty were removed to avoid spatially ambiguous data, and records south of 23° latitude and east of -80° longitude were removed to avoid records from outside the WIS and GCP regions. The resulting geographic range of the dataset extends from 23° to 54° N and 83° to 118° W (Figure S1).

The final vetted database contained 33,168 fossil occurrences (Table S1; Figure S1) representing 574 unique genera and 1113 unique species of marine macroinvertebrates including bivalves, ammonites, gastropods, echinoderms, corals, brachiopods, polychaetes, and crustaceans. These groups represent a wide variety of lifestyle modes both within clades and between distinct phylogenetic groups. They are furthermore important components in both modern and ancient shallow marine ecosystems (Bush and Bambach 2011) and therefore allow comparison to modern systems.

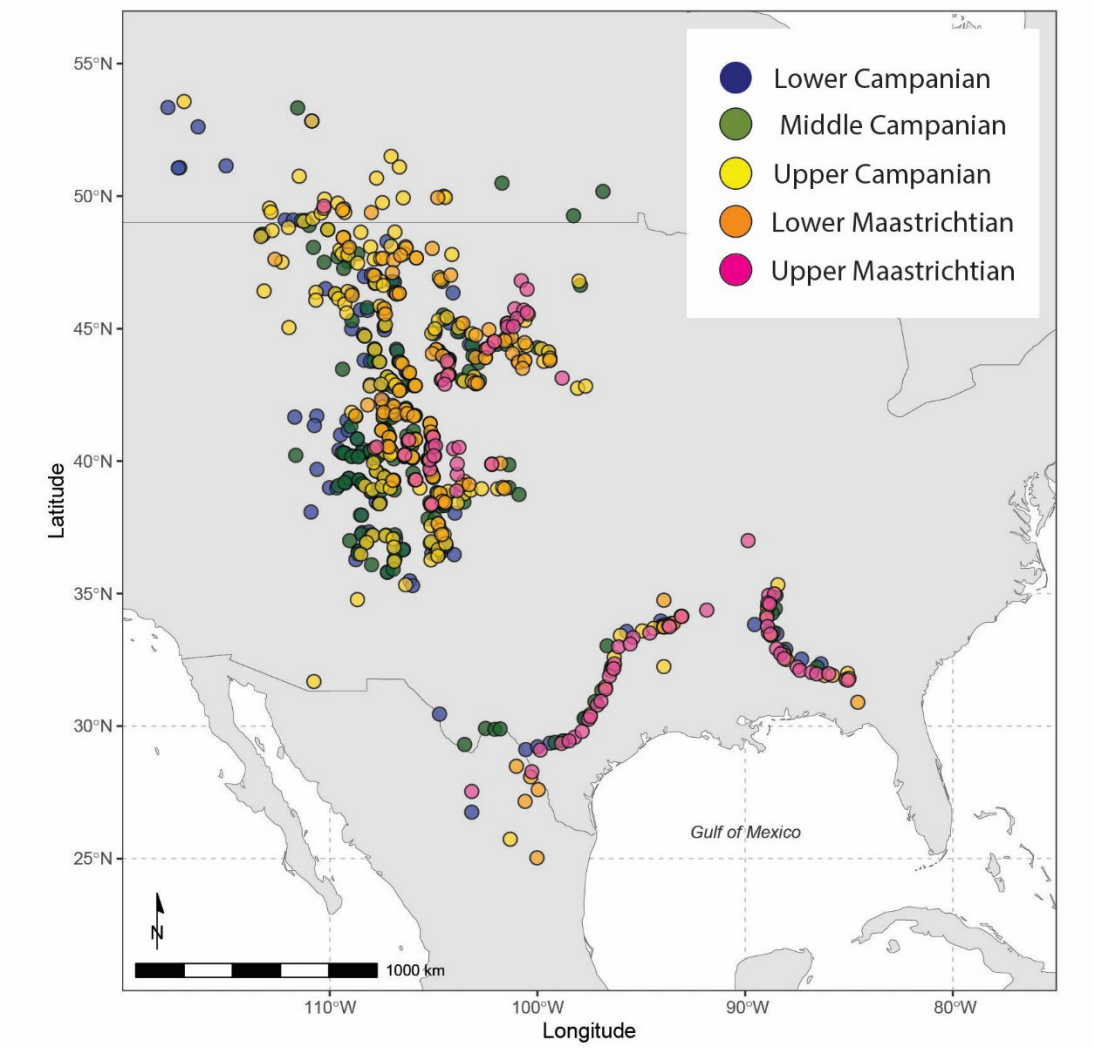



**Figure S1**. Map of fossil occurrence data information. Point colors indicate substage.

*Database Evaluation:*

Prior to network modeling, the vetted database was analyzed to assess for taphonomic and sampling bias (Vilhena and Smith 2013). This included comparisons of taxonomic richness with sampling bias, or the number of fossil occurrences per geographic unit. Since taxonomic information is categorical and autocorrelation of this data type is complex and frequently includes duplicated locations, alternative methods for analyzing general sampling bias spatial autocorrelation were utilized. This was achieved by creating a 60km grid in ArcGIS Pro using the Grid Index Tool (ESRI, 2021) that was overlain on the data. This grid size was chosen to account for uncertainty in fossil location information, while still maintaining localized community ranges that would be consistent across geographic space. An alternative 30km grid size at our occurrence resolution was tested and produced primarily singleton grid cells (i.e., a grid cell with only 1 fossil occurrence) and was therefore uninformative for our analysis of fossil communities. The genus- and species-level data were summarized by number of fossil occurrences per 60km grid cell as well as the number of unique genera and species per grid cell, though only genus-level data were analyzed using network modeling (see Appendix C for R code).

The data, once aggregated, was found to be highly right-skewed (Figure S2 and S3) and to have a strong correlation between the number of fossil occurrences in a grid cell (i.e., sampling) and the taxonomic richness (Figure S4). Thus, the data is non-normal and there is a strong positive relationship between taxonomic richness and the number of occurrences sampled in a given 60-km pixel, or node, herein regarded as sampling effort. This correlation is used in subsequent analyses, including MST assessments, to determine the influence that sampling bias has on network outputs by allowing for indirect comparisons between the number of fossils present in a node as indicated by generic richness, and network values or patterns.

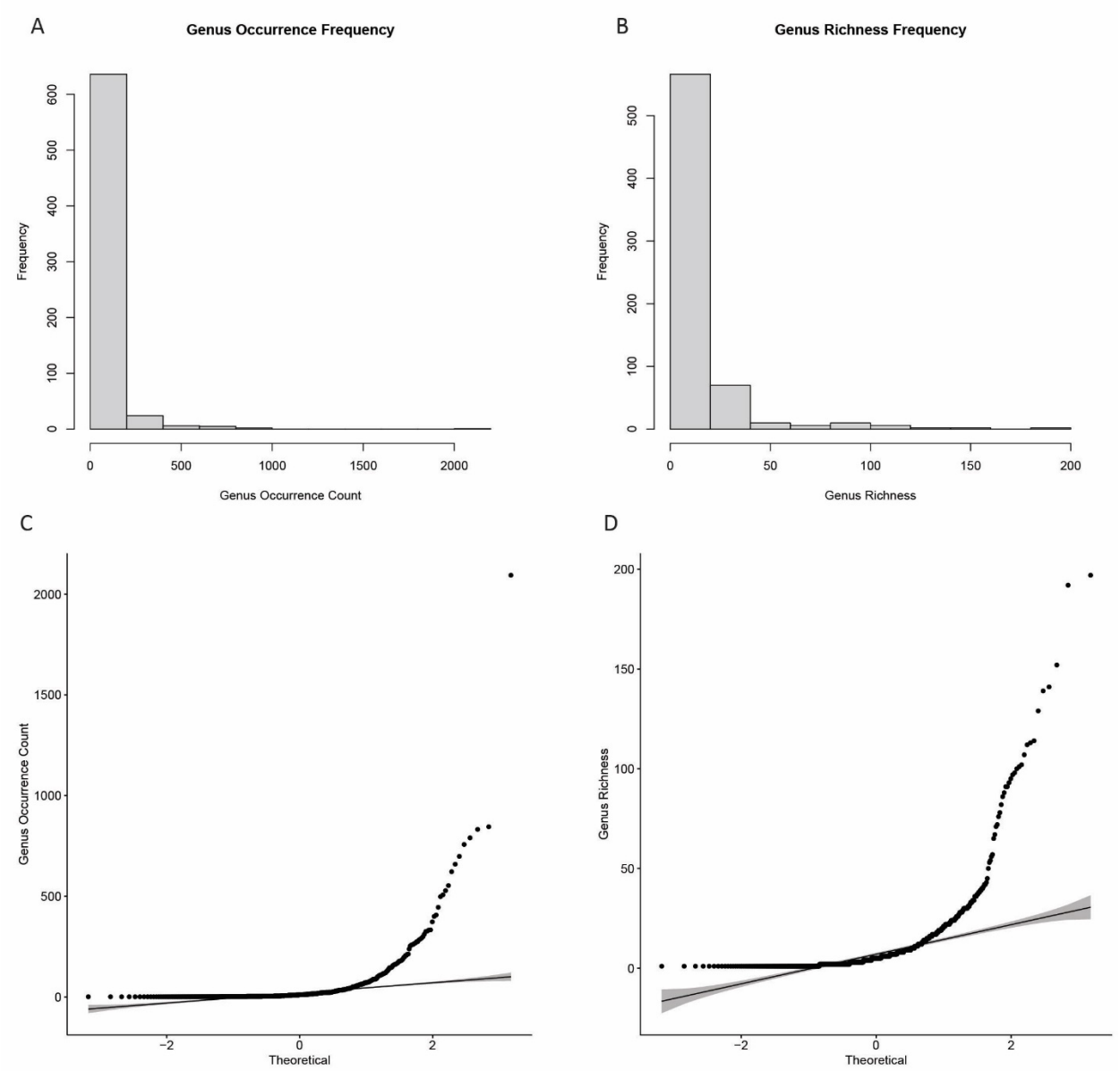



**Figure S2.** Histograms and Q-Q plots of genus-level occurrence counts and richness for the data aggregated in 60-km grid cells. The data is highly right-skewed and deviates greatly from the trend line, indicating that it is non-normal.

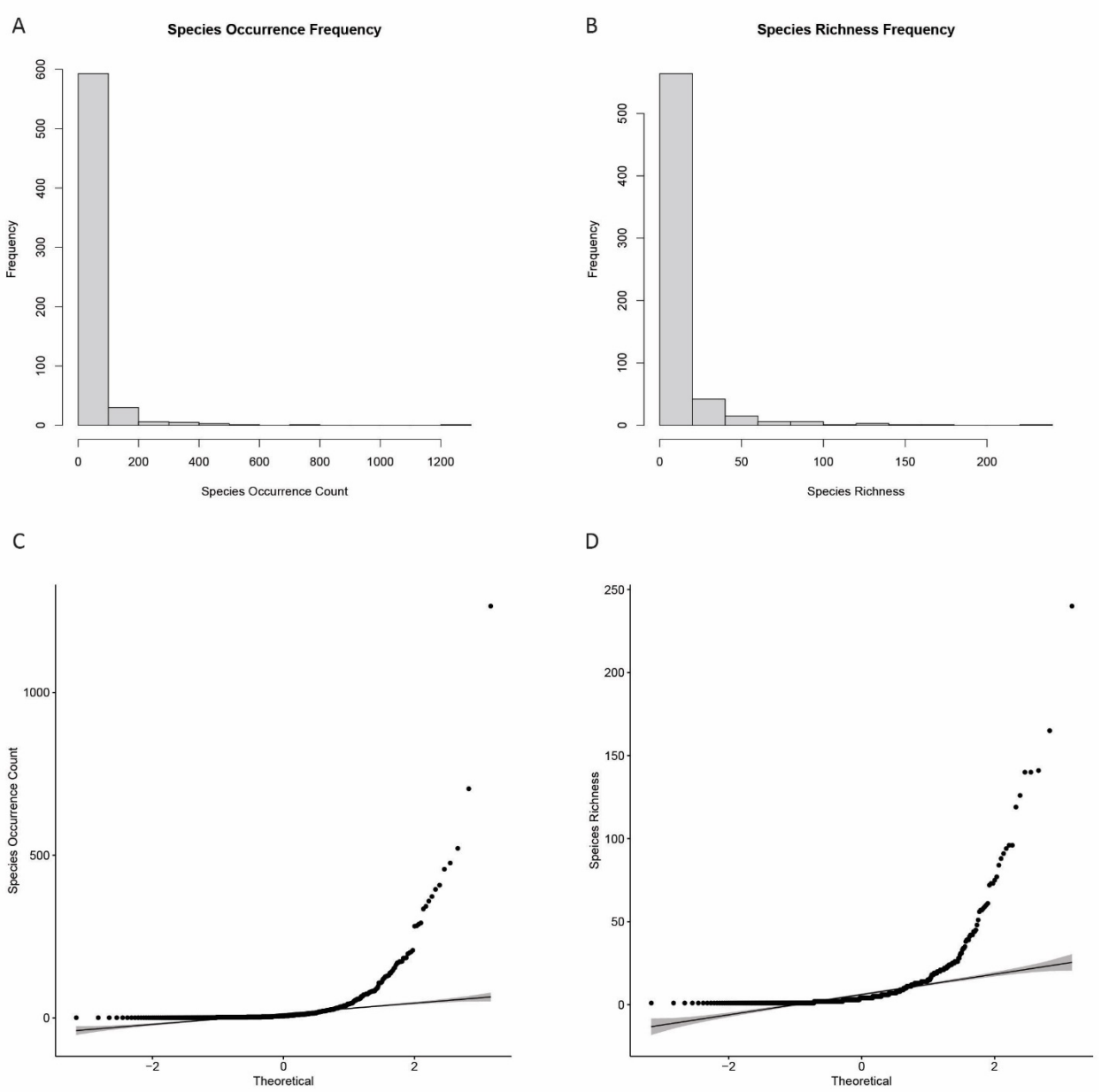



**Figure S3.** Histograms and Q-Q plots of species-level occurrence counts and richness for the data aggregated in 60-km grid cells. The data is highly right-skewed and deviates greatly from the trend line, indicating that it is non-normal.

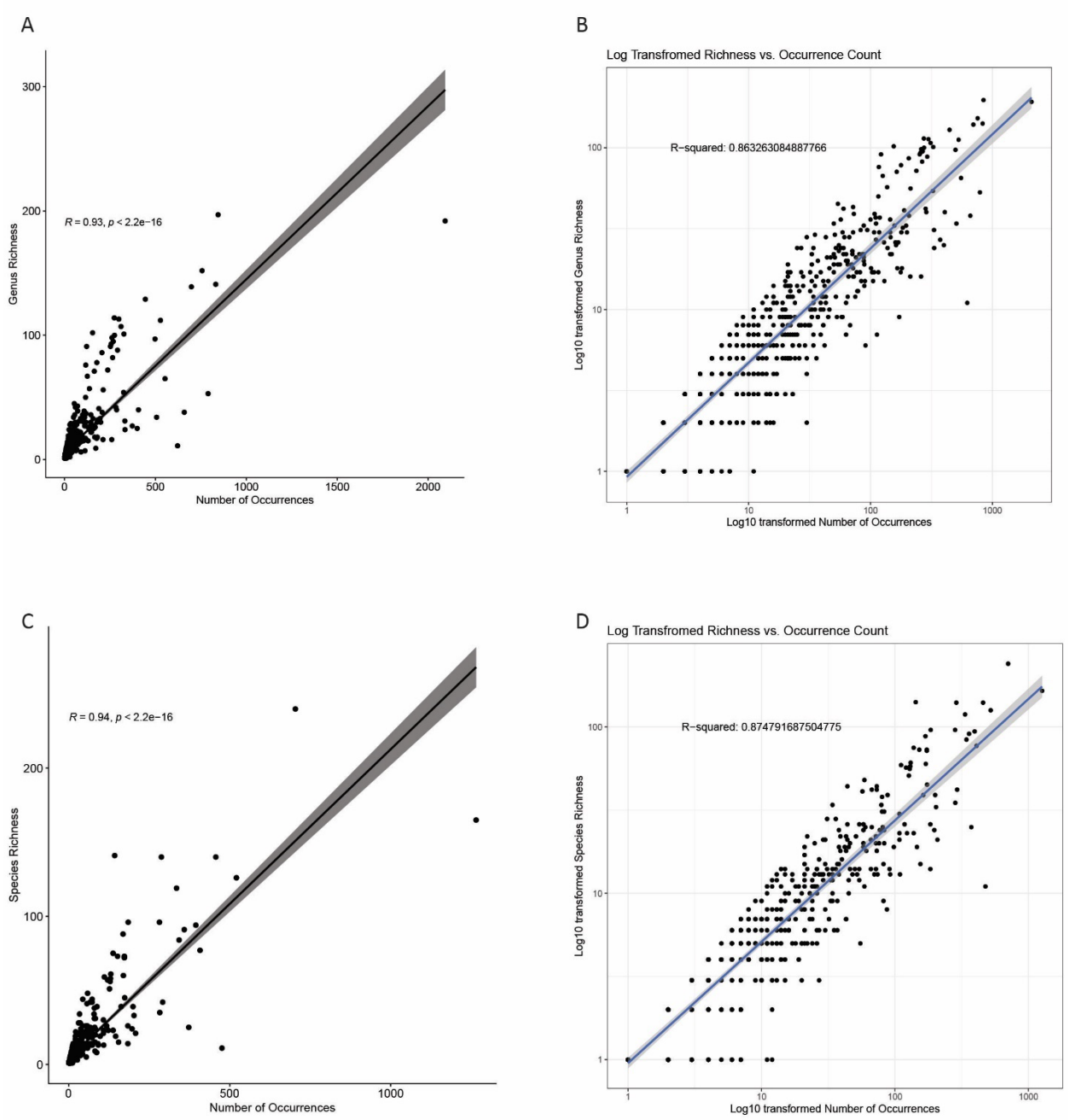



**Figure S4.** Scatter plots and log-transformed plots of taxonomic richness and the number of occurrences for the data aggregated in 60-km grid cells. The data is strongly correlated prior to $\log_{10}$ normalization. Once normalized, the correlation is smaller though still high, and the trend line fits with much smaller residual error.

The nature of the geographic distribution and autocorrelation embedded in the dataset was analyzed with ArcGIS Pro (Spatial Autocorrelation (Global Moran's I) (Spatial Statistics) tool) and

the Cluster and Outlier Analysis (Anselin Local Moran's I (Spatial Statistics) tool; ESRI, 2021) to determine the degree of local and global spatial autocorrelation of genus and species-level richness and total number of fossils collected at each location. Global and local spatial autocorrelation of these data was then analyzed using the Moran's I statistic and local Anselin Moran's I statistic, respectively, in ArcGIS Pro (ESRI, 2021). Overall, results of spatial autocorrelation analyses (both global and local) indicate high spatial clustering within the WIS and GCP with some high-low and low-high outliers based on region (Figures S4-S8). Visual comparisons between Anselin Moran's I with network results, however, did not indicate that any obvious parallels between autocorrelation patterns of clustering/outliers with either link weights or betweenness centrality. **These results indicate that spatial bias is not a primary control on general network topology.**

The Global Moran's I analysis indicates that all substages are statistically clustered (p-values ranging from 0 to 0.078), with only the distribution of species-level occurrence counts in the Lower Campanian falling just outside the 95% confidence interval. All z-scores for the tests are positive, indicating clustering rather than dispersal. These results suggest that the data overall is highly clustered globally when summarized based on the 60km grid, both in terms of taxonomic richness and sampling effort. The results of the Anselin Local Moran's I analysis indicates similar clustering at local scales (Figure S4-S8). For each substage, many summarized 60km grid points have statistically high clustering of either high values with other high values, or low values with other low values for both unique taxa and total taxa occurrence counts.

**Table S2:** Results of Global Moran's I analysis. Distance threshold values were calculated automatically to ensure that each occurrence had at least one neighbor and have been rounded up to the nearest km.

| | **Number Unique Genera/Grid Cell** | | **Number Genus Occ/Grid Cell** | | | **Number Unique Species/Grid Cell** | | **Number Species Occ/Grid Cell** | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Substage** | Moran's I p | Moran's I z | Moran's I p | Moran's I z | Distance Threshold (nearest km) | Moran's I p | Moran's I z | Moran's I p | Moran's I z | Distance Threshold (nearest km) |
| Low CAM | 0.003 | 2.997 | 0.006 | 2.724 | 432 | 0.006 | 2.757 | 0.078 | 1.762 | 416 |
| Mid CAM | 0.000 | 14.046 | 0.000 | 8.393 | 492 | 0.000 | 10.389 | 0.000 | 7.006 | 498 |
| Up CAM | 0.000 | 5.577 | 0.000 | 6.468 | 559 | 0.000 | 3.662 | 0.000 | 4.428 | 645 |
| Low MAA | 0.000 | 8.280 | 0.000 | 5.163 | 371 | 0.000 | 7.604 | 0.000 | 5.019 | 371 |
| Up MAA | 0.000 | 10.486 | 0.000 | 5.786 | 289 | 0.000 | 13.393 | 0.000 | 7.144 | 746 |

Local spatial relationships related to sampling and taxonomic richness were assessed using the Cluster and Outlier Analysis (Anselin Local Moran's I) (Spatial Statistics) tool in ArcGIS Pro. Aneslin Local Moran's I analysis determines if the values of a particular feature are more or less similar to its neighbors than would be expected at random using a 95% confidence interval. The results can indicate spatial outliers and local clusters that are not assessed using the global statistic. Overall, Aneslin Local Moran's I analysis indicates that the data from the Campanian substages are concentrated in the WIS region rather than the GCP based both on the number of unique taxa (i.e., taxonomic richness) and the total fossil occurrence counts (i.e., sampling count). These analyses also strongly suggest that the distribution of genera/species richness and sampling counts are not random but show high clustering. Analyses of the Campanian substages in general show that the dominance of high-high clusters for richness shifts from the more northern grid cells of the WIS in the Lower Campanian, to the more central portion of the WIS in the Middle and Upper Campanian. The Maastrichtian substages, on the other hand, indicate that higher richness and sampling are concentrated in the GCP and that the WIS is dominated by low-low clustering for all analysis.

It also appears that clustering/outlier patterns for sampling counts parallel that of taxonomic richness, except in the Middle Campanian substage where the generic richness and the total genus

sampling count results appear inverse. These results indicate that, despite the fact that the southern portion of the WIS (Colorado and southern Wyoming) has clusters of low sampling, this area is dominated by high-high clusters of generic richness. Similarly, the northern portion of the WIS (Montana and Canada) is dominated by low-low clusters of generic richness, despite the fact that it is dominated by high sampling. This suggests that the few occurrence records in the southern region of the WIS that do exist have a disproportionately high genus richness, and the reverse for the northern region. Interestingly, this pattern does not hold when the data are analyzed based on species for the Middle Campanian. Details for each substage are listed below.

**Visual comparisons of networks relative to Aneslin Local Moran's I analysis indicate that spatial distribution of unique genera counts within grid cells did produce noticeable bias in the substage networks (Figure S5 – S9).** In the Lower Campanian, the northern portion of the WIS, which is dominated by high-high clusters of unique genera counts had overall weaker faunal associations and stronger links in the southern portion where low-low clusters dominate. In the Middle Campanian, this pattern is reversed: the southern portion of the basin, dominated by high-high clusters, maintains relatively stronger links than the northern portion, which is characterized by low-low clusters. The Upper Campanian network has the strongest links between regions with low-low clustering in the northwest and east-central regions, but relatively strong faunal associations were found across the majority of the WIS, connecting the southwestern, northwestern, and east-central portions (see Figure S10, S12, S14, S16, and S18). The Lower Maastrichtian network maintains its strongest links between the southern and east-central portions of the WIS, which are characterized by low-low clustering, but relatively strong links are maintained across the whole basin and are weakest in the north where few clusters or outliers were observed. The Lower Maastrichtian network also presents strong links within the GCP between grid cells along the eastern margin, where high-high clusters dominate. Finally, the Upper Maastrichtian network is characterized by strong links both in the WIS, where low-low clusters dominate, and in the GCP, particularly along the eastern margin

where high-high clusters are the most common. The faunal associations are somewhat stronger in the WIS during this substage than they are in the GCP.

*Lower Campanian:* Data from the Lower Campanian summarized by the 60-km grid cells indicate that the WIS has clusters of both high-high and low-low counts of unique genera (Figure S5). Clusters of high-high unique genera are mostly found in central Wyoming and Montana, while low-low clusters are mostly distributed in northern New Mexico and southern Utah, as well as Alberta, Canada. Outliers of low unique genera counts surrounded by high (low-high outliers) are also fairly common and distributed across Montana and Wyoming. Only three grid cells were found to be statistical outliers of high unique genera counts surrounded by low (high-low outliers), and these are located in southern Utah and northern New Mexico. The majority of grid cells in Utah, Colorado, and the eastern regions of Montana and Wyoming were not statistically significant clusters or outliers regarding unique genera counts. The GCP region had no significant clusters or outliers for the Lower Campanian data. Patterns found using the summarized number of unique species were very similar to those of the unique genera, except that the number of high-high clusters was far fewer in central Wyoming and Montana and that a greater number of high-low outliers were found in the western Colorado. One grid cell was determined to be a high-low outlier for the number of unique species in the GCP of Alabama. The data summarized by the number of total genera indicate that there are both high-high clusters and low-high outliers distributed throughout the central WIS, primarily in Wyoming and northern Utah. Low-low clusters of genus occurrence counts are located primarily in Alberta and northern New Mexico. Only two grid cells were determined to be high-low outliers, located in Alabama and northern New Mexico. Patterns for the summarized number of species occurrence counts were very similar to those of genus occurrence counts, except that the number of high-high clusters and low-high outliers in central WIS were far fewer and that four total grid cells were identified as high-low outliers in northern New Mexico and Colorado.

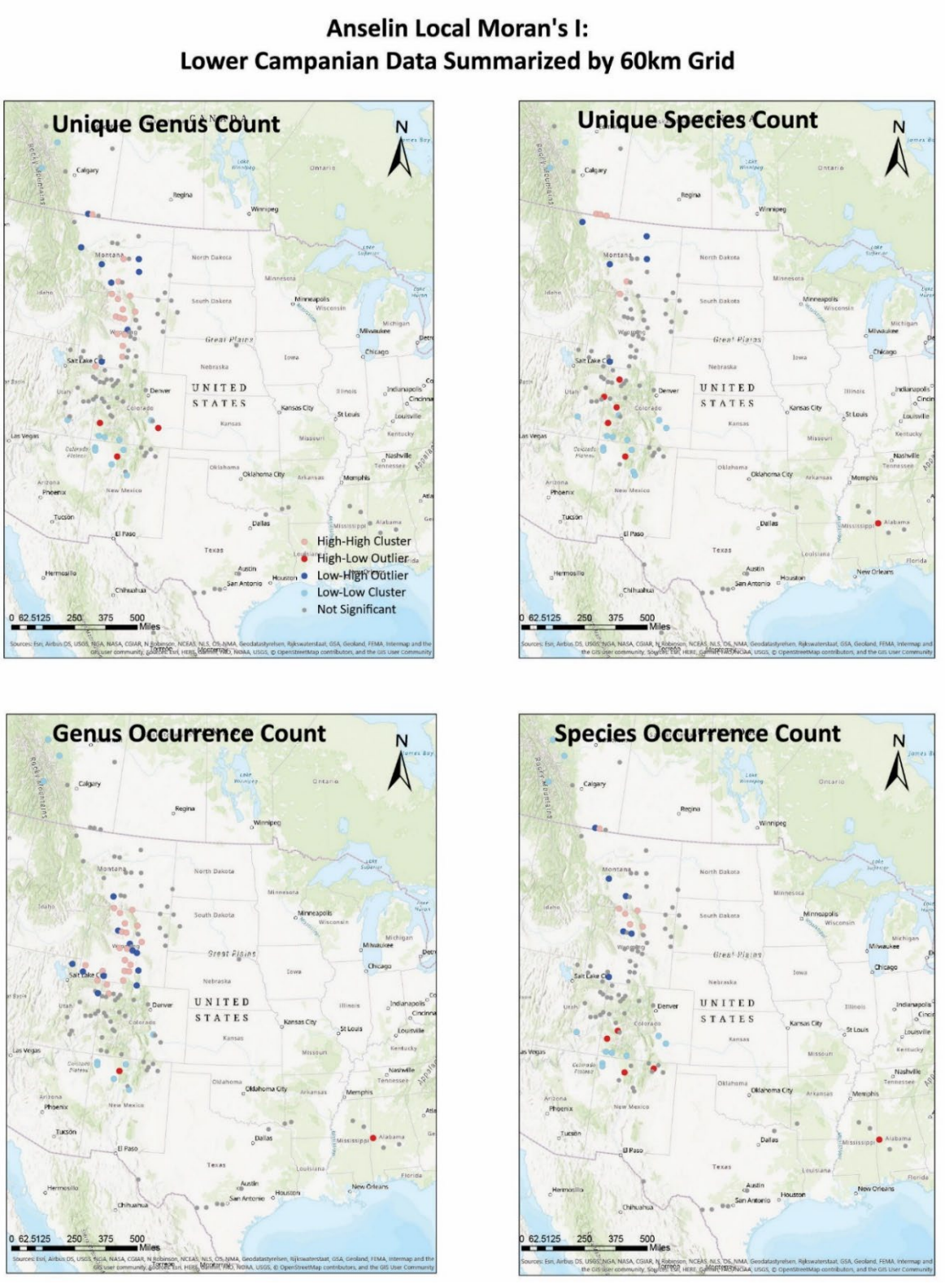



**Figure S5.** Map results of Anselin Local Moran's I analysis showing the distribution of clustering and outliers for number of unique genera summarized using the 60km grid for the Lower Campanian.

*Middle Campanian:* Data from the Middle Campanian summarized by the 60-km grid cells indicate that the northern WIS in Montana is dominated by low-low clusters with few high-low outliers and that the central WIS in Colorado, Utah, and Wyoming is dominated by both high-high clusters and low-high outliers for both the number of unique genera and species (Figure S6). Grid cells located in New Mexico, eastern Wyoming, North Dakota, and most of the GCP were predominantly non-significant, excluding a few low-low clusters. When the data was analyzed based on the number of species occurrences, a very similar pattern emerges, except that the GCP has a much higher number of low-low clusters. When analyzed based on the number of genus occurrences, however, the distribution of clusters and outliers is almost completely opposite that of the number of species occurrences. For data summarized by genus occurrences, the Montana region is dominated by high-high clusters with a single cell that represents a low-high outlier, the central region of the WIS is dominated by both low-low clusters and high-low outliers, and grid cells in the GCP region were almost all found to be low-low clusters.

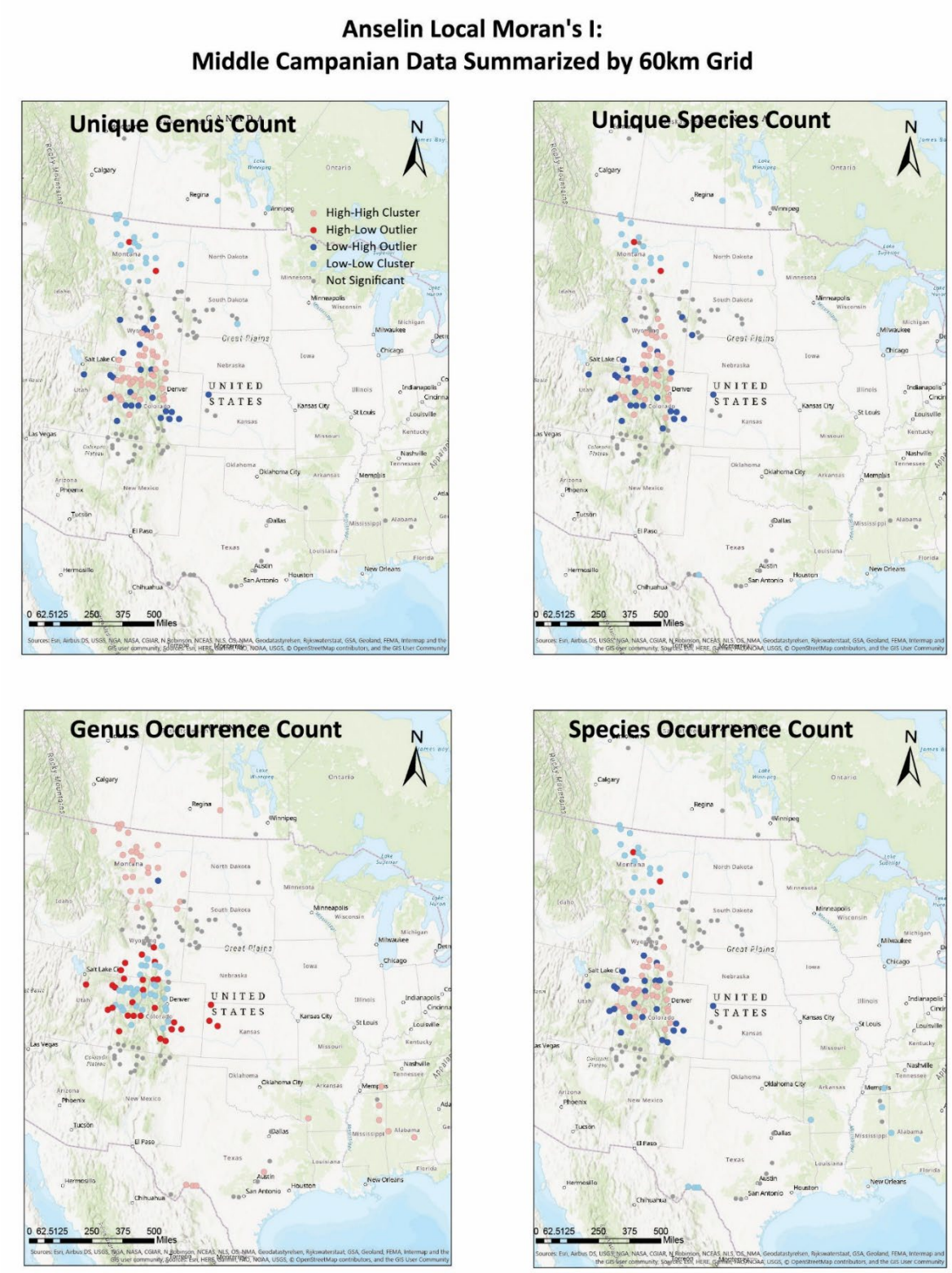



**Figure S6.** Map results of Anselin Local Moran's I analysis showing the distribution of clustering and outliers for number of unique genera summarized using the 60km grid for the Middle Campanian.

*Upper Campanian:* Data from the Upper Campanian summarized by the 60-km grid cells indicate that the northern WIS in Montana, Alberta, and Saskatchewan is dominated by low-low clusters with few high-low outliers located in Montana and that the central WIS in Colorado and southern Wyoming is dominated by both high-high clusters and low-high outliers for both the number of unique genera and the number of unique species (Figure S7). Grid cells located in New Mexico, eastern Wyoming, North Dakota, Nebraska, South Dakota, and most of the GCP were predominantly non-significant. However, the data analyzed based on unique genera indicate a concentration of low-low clusters in South Dakota, northern Nebraska, and North Dakota that is not found when analyzed based on unique species counts. There are also several grid cells identified as low-high outliers across the GCP for unique species, and one high-low outlier for unique genera counts in eastern Texas. The distribution of significant clusters and outliers identified when the data was analyzed based on both the total number of genus and species count is extremely similar to those found for the number of unique genera and species in the Upper Campanian.

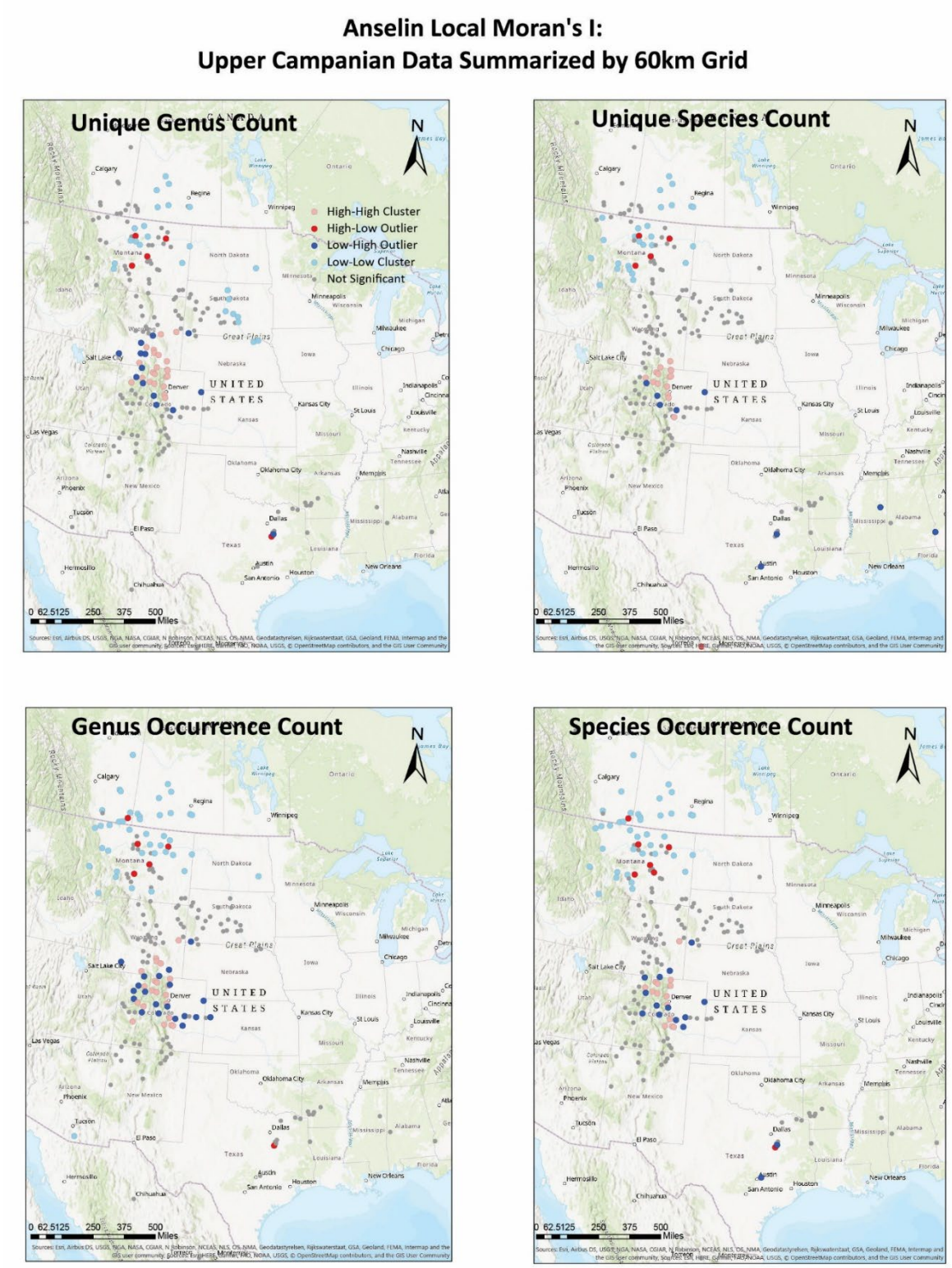



**Figure S7.** Map results of Anselin Local Moran's I analysis showing the distribution of clustering and outliers for number of unique genera summarized using the 60km grid for the Upper Campanian.

*Lower Maastrichtian:* Data from the Lower Maastrichtian summarized by the 60-km grid cells indicate that the South Dakota and southern Colorado regions are dominated by low-low clusters of unique genera and unique species, but that the majority of the WIS region has no significant clustering or outliers (Figure S8). The GCP region, however, is dominated by low-low clustering in northern Mexico and Arkansas for the number of unique genera and mix of high-high clusters with low-high outliers in the east. This pattern holds for the GCP region when analyzed based on the number of unique species, but with only a single low-low cluster grid cell in northern Mexico. Two grid cells were identified as high-low outliers for both the unique genus and species analysis, one in eastern Montana and another in eastern Texas. The distribution of clusters and outliers for analysis performed based on the total species and genus occurrence counts were very similar to that of the number of unique genera overall, except that Colorado does not have significant low-low cluster grid cells while Montana and southern Canada regions do.

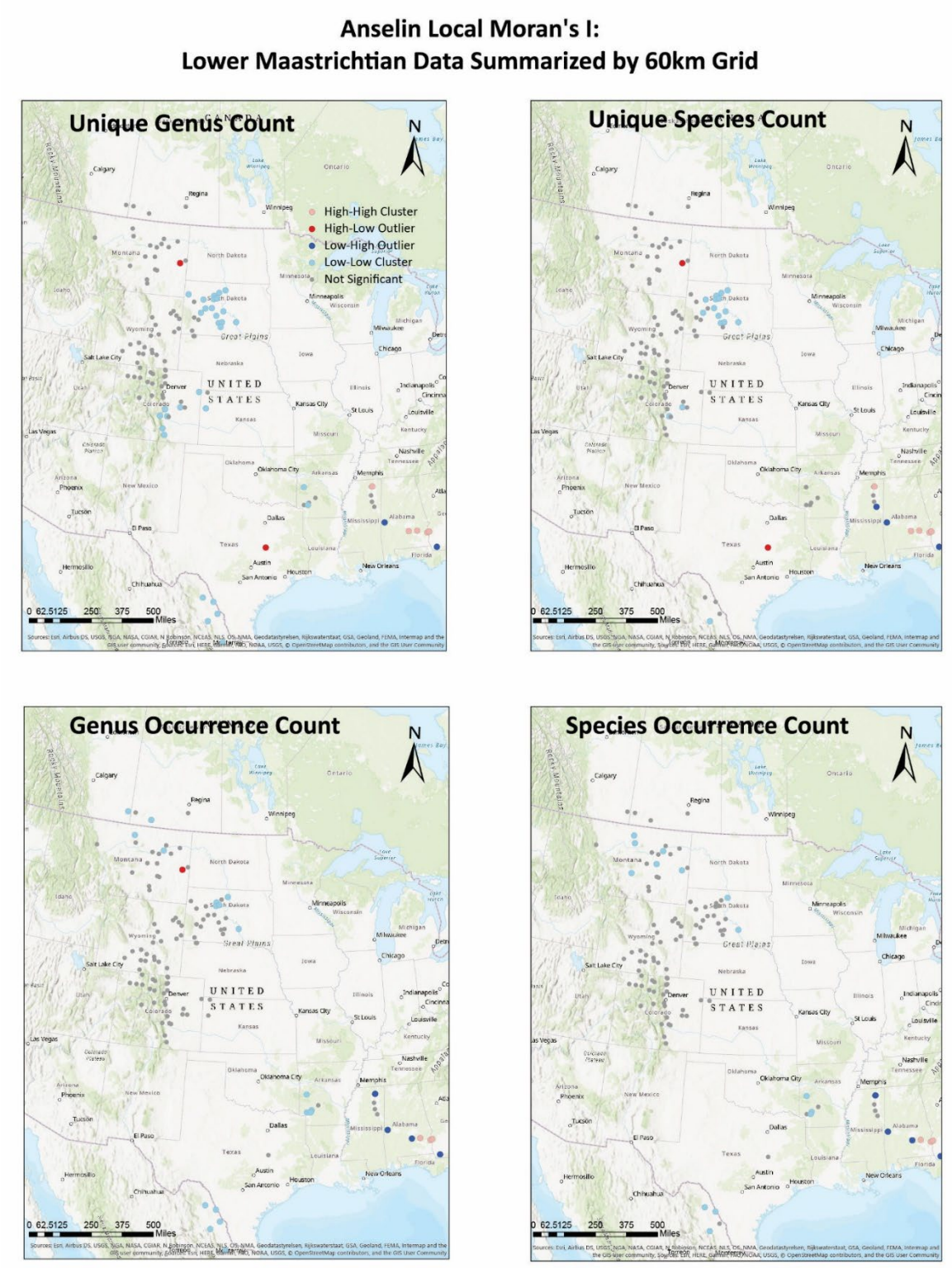



**Figure S8.** Map results of Anselin Local Moran's I analysis showing the distribution of clustering and outliers for number of unique genera summarized using the 60km grid for the Lower Maastrichtian.

*Upper Maastrichtian:* Data from the Upper Maastrichtian summarized by the 60-km grid cells indicate that the WIS is dominated by low-low clusters of unique genera and unique species, and that the western portion of the GCP, including Texas and northern Mexico have no significant outliers or clusters (Figure S9). The eastern GCP region, however, is dominated by low-low clustering, particularly in Alabama and Georgia, and low-high outliers for both the number of unique genera and species. This general pattern holds when analyzed based on both genus and species occurrence counts, excepting that the WIS region in North and South Dakota are predominantly non-significant.

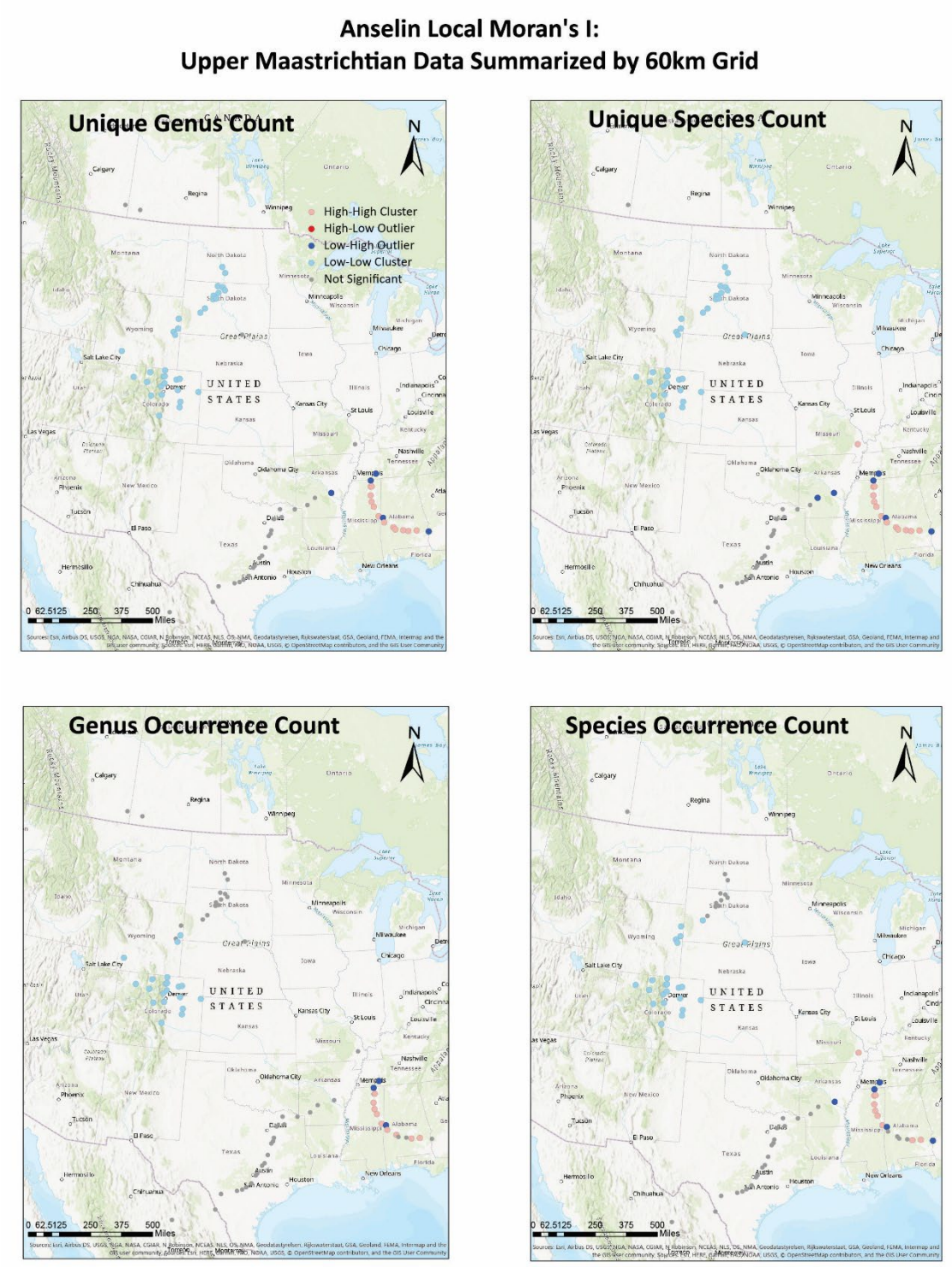



**Figure S9.** Map results of Anselin Local Moran's I analysis showing the distribution of clustering and outliers for number of unique genera summarized using the 60km grid for the Upper Maastrichtian.

*Network Analysis:*

Faunal occurrence records of genera were aggregated into "paleogeographic analysis units" or nodes based on the 60km-grid that was created in ArcGIS Pro using the Grid Index tool with cell sizes set to 60x60km (see above section on data evaluation). Aggregated data in the form of a presence-absence matrix were input into the EDENetworks version 2018 software package (Kivelä et al. 2015). Nodes with fewer than three genera were removed from the analysis to allow for better comparisons of assemblages and to avoid singletons (i.e., Kiel, 2017). Weighted distances between nodes, or links, were calculated based on the Bray–Curtis dissimilarity index, which calculates how similar two nodes are based on the number of shared taxa,

$$D(A,B) = \frac{2\sum_{i=1}^{k}\min(A_i, B_i)}{\sum_{i=1}^{k}(A_i + B_i)} \qquad (EQ\ 1)$$

where A and B are two nodes, calculated over a vector with length K where every $i^{th}$ element represents the presence or absence of a specific taxa. This calculation is computationally identical to the Sorensen or Dice coefficient (Somerfield 2008). The index value ranges from 0, indicating that the two nodes share identical taxa, to 1, indicating that they share none. R Code used to configure the database into tables that could be input into EDENetworks can be found in Appendix C.

All data locations were converted to paleo coordinates using the R package chronosphere (Kocsis and Raja 2020) to reproject the data back to their original positions and relative distances. Faunal provinces were delimited using these distance weight measures in a thresholding network approach (Moalic et al. 2012; Kivelä et al. 2015; Kiel 2016) where the threshold represents the maximum weight (level of dissimilarity) of the links used to construct the network; weights that fall above this value (i.e., those with greater faunal dissimilarity) were removed from the network (Kivelä et al. 2015). Thus, at lower thresholds the network is more fractured (more and smaller network components) and at higher thresholds the network is more connected (fewer and larger components).

The point at which all components are connected without forming a unitary component is referred to as the percolation point (Kivelä et al. 2015). Components which become disconnected at thresholds at and below the percolation point can be interpreted as representing distinct "community" groups (Newman 2012) and therefore can be used to draw conclusions about the distribution of faunal provinces.

Network models were created using EDENetworks for all substages individually and for the complete database, which combines all substages together. Networks were produced between the 0.99 threshold and 0.40 threshold to analyze the consistency of topological patterns at intervals of 0.05, and an additional network was produced for each percolation point (Table S3). General patterns in network connection across all substages for the complete database were also visually assessed by using a coarser spatial aggregation of the data, by a 360km grid aggregation (see below). This aggregation produced lower resolution comparison with fewer nodes but allowed for easier visualization of network connections across both space and time simultaneously.

Networks and their figures were produced for each substage between the 0.99 threshold and 0.40 threshold to analyze the consistency of topological patterns at intervals of 0.05, and an additional network was produced for each percolation point (Table S3). The complete database, containing the data from all substages, was only analyzed at thresholds between 0.76 and 0.40 due to computational time constraints. Topological patterns were generally consistent at all thresholds for all networks, and no distinct minor components or subprovinces become more obvious within the primary WIS and GCP provinces at lower thresholds (Figure S10-S20).

Results of the network analysis indicate a single faunal province, representing the WIS throughout the Campanian, and two primary faunal provinces in the Maastrichtian (one circumscribing the WIS and the other delineating the GCP) (Figure S16-S19); the two provinces were also observed in the complete database network (Figure S20). The lack of a GCP province in the Campanian (Figure S10-S15) is probably driven by a lack of GCP data; only in the Maastrichtian

when GCP data is more abundant does the GCP province become clear. Topological patterns were generally consistent at all thresholds for all networks, and no distinct minor components become more obvious within the primary WIS and GCP provinces at lower thresholds.

**Table S3.** Summary table of network analysis results for all substages and complete database including summary values for individual threshold analysis.

| | Low CAM | Mid CAM | Up CAM | Low MAA | Up MAA | Complete DB |
|---|---|---|---|---|---|---|
| Size | 75 | 109 | 138 | 68 | 50 | 440 |
| Distance Measure | Jaccard | Jaccard | Jaccard | Jaccard | Jaccard | Jaccard |
| Avg Distance | 0.79 | 0.8 | 0.83 | 0.84 | 0.87 | 0.86 |
| Min Distance | 0 | 0.2 | 0 | 0 | 0 | 0 |
| Max Distance | 1 | 1 | 1 | 1 | 1 | 1 |
| **PERCOLATION POINT** | **0.61** | **0.57** | **0.5** | **0.8** | **0.86** | **0.76** |
| Edges | 367 | 175 | 174 | 674 | 447 | 16965 |
| Avg Degree (<k>) | 9.79 | 3.21 | 2.52 | 19.82 | 17.88 | 77.11 |
| Max Degree (kmax) | 33 | 11 | 17 | 41 | 29 | 119 |
| Avg Clustering (<c>) | 0.41 | 0.25 | 0.18 | 0.68 | 0.8 | 0.58 |
| Avg Distance (<d>) | 0.51 | 0.48 | 0.44 | 0.69 | 0.69 | 0.67 |
| Connectivity (components) | 18 | 25 | 74 | 3 | 9 | 22 |
| Connectivity (largest) | 57 | 84 | 61 | 66 | 44 | 415 |
| | **MINIMUM SPANNING TREE** | | | | | |
| Edges | 74 | 108 | 137 | 67 | 49 | 439 |
| Avg Degree (<k>) | 1.97 | 1.98 | 1.99 | 1.97 | 1.96 | 2 |
| Max Degree (kmax) | 6 | 7 | 6 | 5 | 4 | 8 |
| Avg Clustering (<c>) | 0 | 0 | 0 | 0 | 0 | 0 |
| Avg Distance (<d>) | 0.49 | 0.5 | 0.52 | 0.56 | 0.58 | 0.48 |
| Connectivity (components) | connected | connected | connected | connected | connected | connected |
| Connectivity (largest) | | | | | | |
| | **THRESHOLD 0.99** | | | | | |
| Edges | 2322 | 5741 | 8761 | 2066 | 987 | - |
| Avg Degree (<k>) | 61.92 | 105.34 | 126.97 | 60.76 | 39.48 | - |
| Max Degree (kmax) | 71 | 108 | 135 | 65 | 48 | - |
| Avg Clustering (<c>) | 0.93 | 0.98 | 0.97 | 0.94 | 0.89 | - |
| Avg Distance (<d>) | 0.75 | 0.8 | 0.81 | 0.82 | 0.84 | - |
| Connectivity (components) | 3 | connected | 2 | 2 | connected | - |
| Connectivity (largest) | 73 | | 137 | 67 | | - |
| | **THRESHOLD 0.95** | | | | | |
| Edges | 2299 | 5492 | 8235 | 1832 | 606 | - |
| Avg Degree (<k>) | 61.31 | 100.77 | 119.35 | 53.88 | 24.24 | - |
| Max Degree (kmax) | 70 | 108 | 133 | 65 | 38 | - |
| Avg Clustering (<c>) | 0.93 | 0.96 | 0.94 | 0.91 | 0.87 | - |
| Avg Distance (<d>) | 0.75 | 0.79 | 0.8 | 0.8 | 0.75 | - |
| Connectivity (components) | 3 | connected | 2 | 3 | connected | - |
| Connectivity (largest) | 73 | | 137 | 66 | | - |
| | **THRESHOLD 0.90** | | | | | |
| Edges | 2042 | 4731 | 6901 | 1505 | 497 | - |

| | | | | | | |
|---|---|---|---|---|---|---|
| Avg Degree (<k>) | 54.45 | 86.81 | 100.01 | 44.26 | 19.88 | - |
| Max Degree (kmax) | 67 | 102 | 126 | 58 | 29 | - |
| Avg Clustering (<c>) | 0.88 | 0.89 | 0.85 | 0.83 | 0.85 | - |
| Avg Distance (<d>) | 0.72 | 0.77 | 0.78 | 0.78 | 0.71 | - |
| Connectivity (components) | 4 | connected | 2 | 3 | 3 | - |
| Connectivity (largest) | 72 | | 137 | 66 | 48 | - |
| | **THRESHOLD 0.85** | | | | | |
| Edges | 1684 | 3651 | 4958 | 1070 | 431 | - |
| Avg Degree (<k>) | 44.91 | 66.99 | 71.86 | 31.47 | 17.24 | - |
| Max Degree (kmax) | 63 | 97 | 109 | 53 | 28 | - |
| Avg Clustering (<c>) | 0.79 | 0.79 | 0.74 | 0.75 | 0.8 | - |
| Avg Distance (<d>) | 0.69 | 0.73 | 0.74 | 0.74 | 0.69 | - |
| Connectivity (components) | 4 | 3 | 4 | 3 | 7 | - |
| Connectivity (largest) | 72 | 106 | 135 | 66 | 31 | - |
| | **THRESHOLD 0.80** | | | | | |
| Edges | 1400 | 2711 | 3586 | 674 | 357 | 26242 |
| Avg Degree (<k>) | 37.33 | 49.74 | 51.97 | 19.82 | 14.28 | 119.28 |
| Max Degree (kmax) | 60 | 84 | 91 | 41 | 26 | 268 |
| Avg Clustering (<c>) | 0.75 | 0.7 | 0.68 | 0.68 | 0.73 | 0.65 |
| Avg Distance (<d>) | 0.66 | 0.7 | 0.71 | 0.69 | 0.66 | 0.71 |
| Connectivity (components) | 5 | 4 | 4 | 3 | 9 | 15 |
| Connectivity (largest) | 69 | 104 | 135 | 66 | 30 | 423 |
| | **THRESHOLD 0.75** | | | | | |
| Edges | 1086 | 18.51 | 2333 | 506 | 300 | 16253 |
| Avg Degree (<k>) | 28.96 | 33.96 | 33.81 | 14.88 | 12 | 73.88 |
| Max Degree (kmax) | 53 | 65 | 71 | 35 | 25 | 192 |
| Avg Clustering (<c>) | 0.65 | 0.62 | 0.58 | 0.63 | 0.68 | 0.58 |
| Avg Distance (<d>) | 0.63 | 0.67 | 0.67 | 0.66 | 0.75 | 0.67 |
| Connectivity (components) | 8 | 5 | 11 | 4 | 11 | 24 |
| Connectivity (largest) | 66 | 103 | 128 | 59 | 28 | 377 |
| | **THRESHOLD 0.70** | | | | | |
| Edges | 789 | 1111 | 1347 | 309 | 215 | 9009 |
| Avg Degree (<k>) | 21.04 | 20.39 | 19.52 | 9.09 | 8.6 | 40.95 |
| Max Degree (kmax) | 45 | 42 | 53 | 22 | 23 | 135 |
| Avg Clustering (<c>) | 0.6 | 0.51 | 0.46 | 0.54 | 0.59 | 0.5 |
| Avg Distance (<d>) | 0.59 | 0.62 | 0.62 | 0.61 | 0.6 | 0.62 |
| Connectivity (components) | 12 | 6 | 14 | 8 | 12 | 36 |
| Connectivity (largest) | 62 | 102 | 125 | 56 | 27 | 336 |
| | **THRESHOLD 0.65** | | | | | |
| Edges | 523 | 579 | 710 | 189 | 145 | 4737 |
| Avg Degree (<k>) | 13.95 | 10.62 | 10.29 | 5.56 | 5.8 | 21.53 |
| Max Degree (kmax) | 40 | 24 | 33 | 17 | 18 | 85 |
| Avg Clustering (<c>) | 0.48 | 0.43 | 0.38 | 0.41 | 0.51 | 0.43 |

| | | | | | | |
|---|---|---|---|---|---|---|
| Avg Distance (<d>) | 0.54 | 0.57 | 0.56 | 0.57 | 0.56 | 0.56 |
| Connectivity (components) | 15 | 11 | 19 | 15 | 16 | 54 |
| Connectivity (largest) | 60 | 98 | 119 | 50 | 24 | 353 |
| | **THRESHOLD 0.60** | | | | | |
| Edges | 304 | 274 | 352 | 79 | 97 | 2310 |
| Avg Degree (<k>) | 8.11 | 5.03 | 5.1 | 2.32 | 3.88 | 10.5 |
| Max Degree (kmax) | 32 | 16 | 21 | 10 | 12 | 60 |
| Avg Clustering (<c>) | 0.36 | 0.29 | 0.28 | 0.26 | 0.41 | 0.34 |
| Avg Distance (<d>) | 0.494 | 0.52 | 0.5 | 0.5 | 0.52 | 0.5 |
| Connectivity (components) | 24 | 15 | 43 | 25 | 18 | 88 |
| Connectivity (largest) | 48 | 94 | 92 | 40 | 22 | 317 |
| | **THRESHOLD 0.55** | | | | | |
| Edges | 203 | 142 | 195 | 51 | 55 | 1372 |
| Avg Degree (<k>) | 5.41 | 2.61 | 2.83 | 1.5 | 2.2 | 6.24 |
| Max Degree (kmax) | 19 | 11 | 17 | 9 | 8 | 48 |
| Avg Clustering (<c>) | 0.33 | 0.2 | 0.2 | 0.16 | 0.3 | 0.26 |
| Avg Distance (<d>) | 0.45 | 0.46 | 0.45 | 0.46 | 0.48 | 0.46 |
| Connectivity (components) | 27 | 36 | 64 | 37 | 25 | 142 |
| Connectivity (largest) | 47 | 54 | 73 | 16 | 17 | 261 |
| | **THRESHOLD 0.50** | | | | | |
| Edges | 181 | 112 | 174 | 42 | 33 | 1199 |
| Avg Degree (<k>) | 4.83 | 2.06 | 2.52 | 1.24 | 1.32 | 5.45 |
| Max Degree (kmax) | 18 | 11 | 17 | 9 | 6 | 48 |
| Avg Clustering (<c>) | 0.31 | 0.17 | 0.18 | 0.1 | 0.27 | 0.25 |
| Avg Distance (<d>) | 0.43 | 0.45 | 0.44 | 0.45 | 0.45 | 0.45 |
| Connectivity (components) | 29 | 47 | 74 | 43 | 34 | 195 |
| Connectivity (largest) | 46 | 51 | 61 | 15 | 10 | 207 |
| | **THRESHOLD 0.45** | | | | | |
| Edges | 80 | 42 | 77 | 12 | 5 | 433 |
| Avg Degree (<k>) | 2.13 | 0.77 | 1.12 | 0.35 | 0.2 | 1.97 |
| Max Degree (kmax) | 12 | 6 | 10 | 3 | 2 | 21 |
| Avg Clustering (<c>) | 0.23 | 0.07 | 0.09 | 0.03 | 0 | 0.13 |
| Avg Distance (<d>) | 0.36 | 0.36 | 0.36 | 0.34 | 0.26 | 0.36 |
| Connectivity (components) | 45 | 74 | 95 | 57 | 45 | 281 |
| Connectivity (largest) | 30 | 24 | 35 | 5 | 3 | 128 |
| | **THRESHOLD 0.40** | | | | | |
| Edges | 58 | 31 | 47 | 10 | 5 | 303 |
| Avg Degree (<k>) | 1.55 | 0.57 | 0.68 | 0.29 | 0.2 | 1.38 |
| Max Degree (kmax) | 10 | 6 | 9 | 2 | 2 | 18 |
| Avg Clustering (<c>) | 0.22 | 0.05 | 0.08 | 0.04 | 0 | 0.11 |
| Avg Distance (<d>) | 0.33 | 0.34 | 0.32 | 0.32 | 0.26 | 0.33 |
| Connectivity (components) | 47 | 83 | 108 | 59 | 45 | 311 |
| Connectivity (largest) | 28 | 18 | 31 | 3 | 3 | 108 |

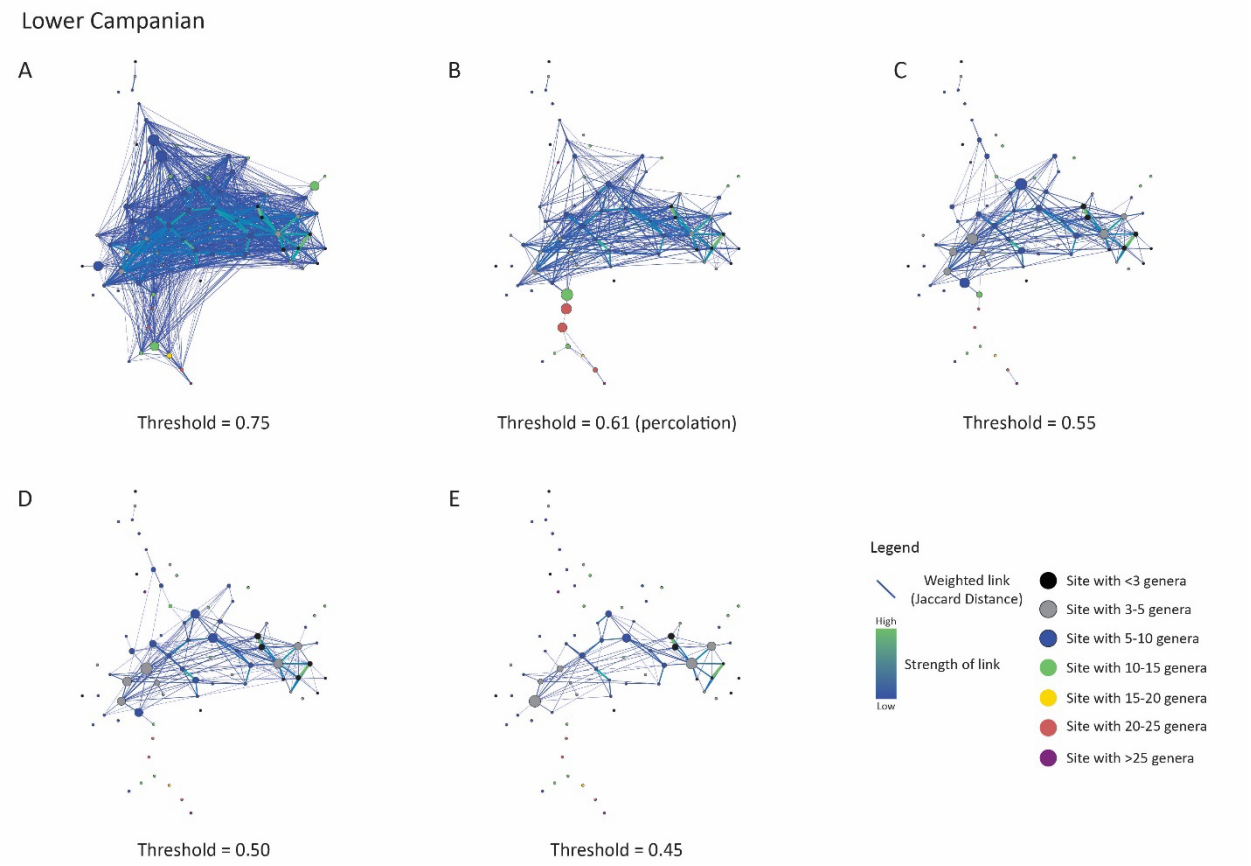



**Figure S10.** Set of network figures for the Lower Campanian. Node sizes indicate betweenness centrality, color denotes the number of unique genus occurrences present at the site, and link weights between nodes denoted by line color and thickness. No obvious structure except in WIS, where the network is densely interconnected without substructure. Only one major component in WI. All other interconnections not included in this major component have less than 2 nodes each. No configuration seemed to make any subprovinces more visible, so the Minimum Spanning Tree configuration were used to plot the network figures. GCP region is not well connected at or below the percolation threshold. High betweenness centrality values in WIS. No obvious connection between low generic richness and weak linkages.

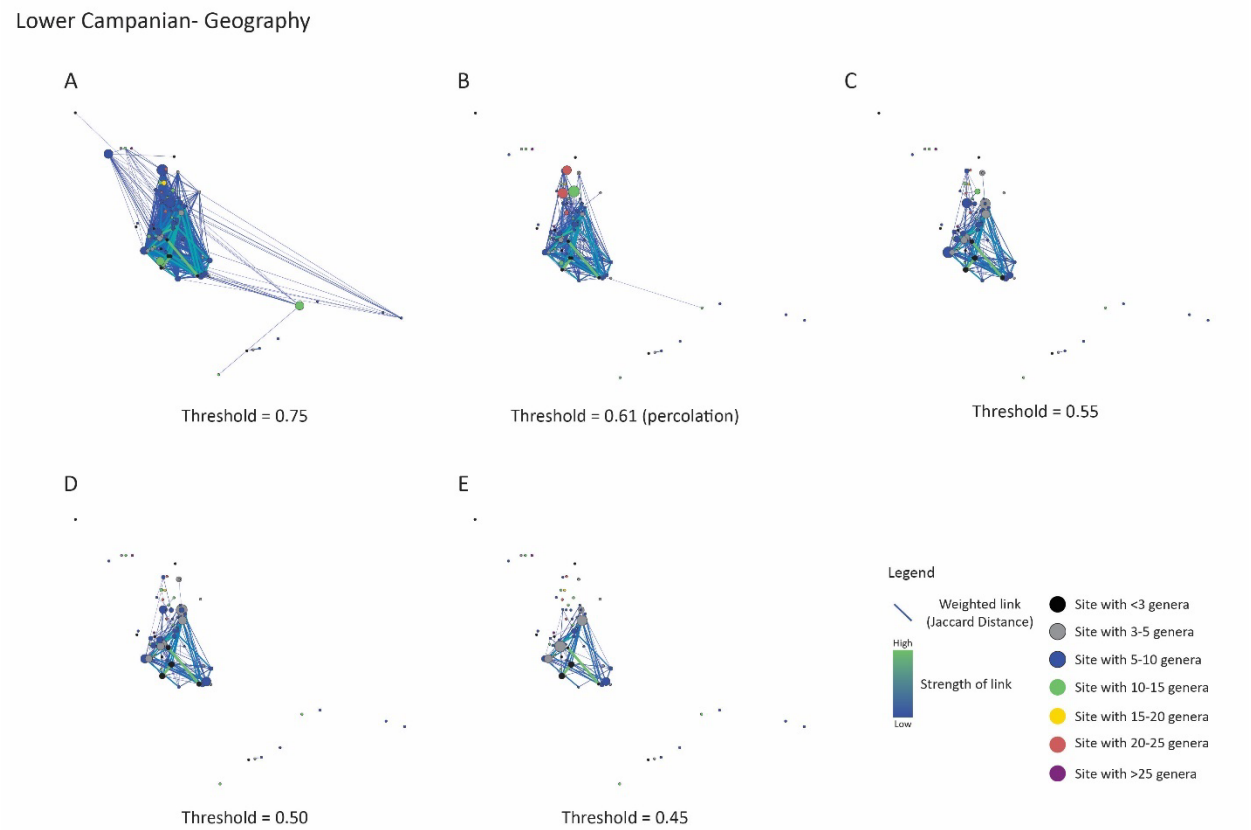



**Figure S11.** Set of network figures for the Lower Campanian with nodes placed based on geographic coordinates. Node sizes indicate betweenness centrality, color denotes the number of unique genus occurrences present at the site, and link weights between nodes denoted by line color and thickness. No obvious structure except in WIS, where the network is densely interconnected without substructure. Only one major component in WI with weak connections between WI and GCP nodes. GCP region is not well connected at or below the percolation threshold. Northern sites were also poorly connected and become less so with lower thresholds. High betweenness centrality values in WIS and strong links between southeast, central-east, and southwest nodes of the WI.

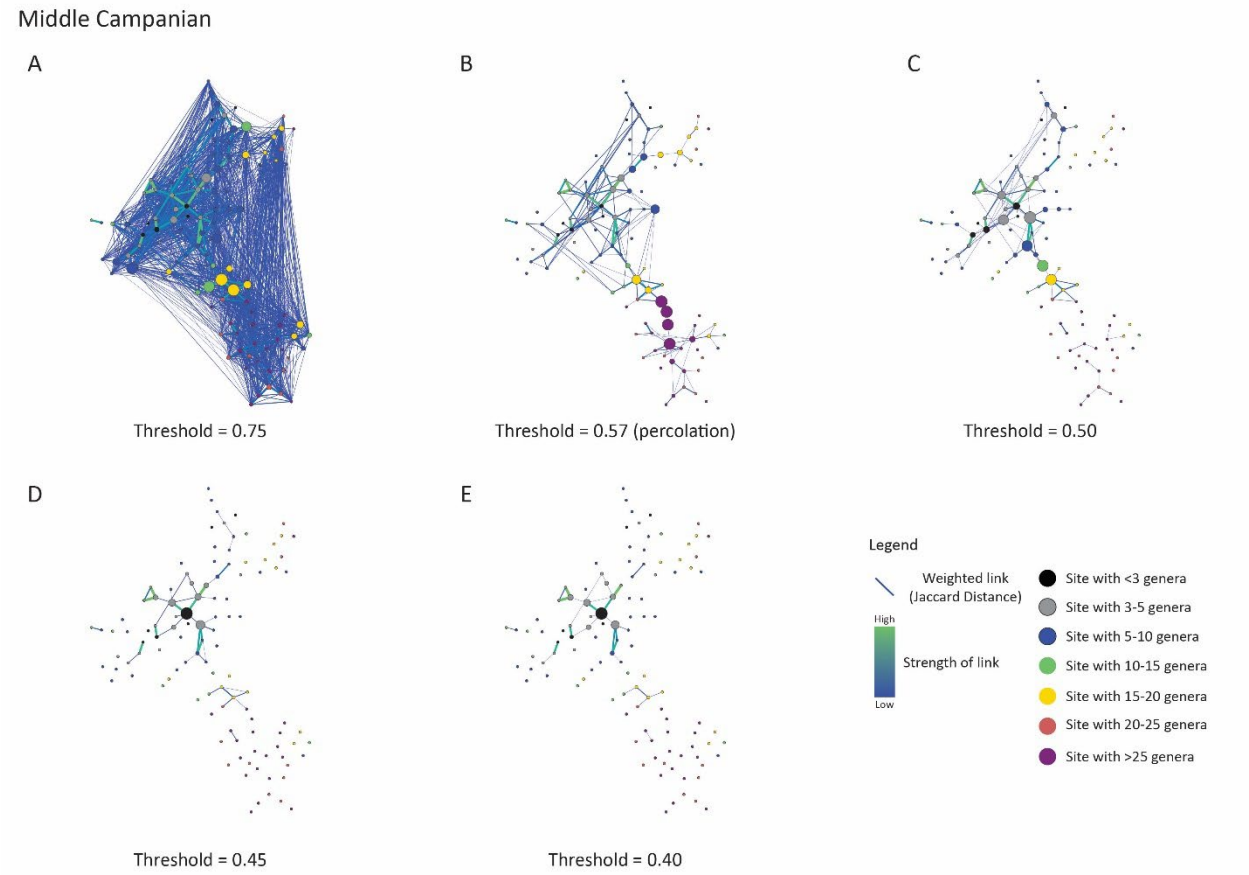



**Figure S12.** Set of network figures for the Middle Campanian. Node sizes indicate betweenness centrality, color denotes the number of unique genus occurrences present at the site, and link weights between nodes denoted by line color and thickness. No obvious structure except in WIS, where nodes interconnected without substructure. Only one major component in WIS and all other components have less than 2 nodes each. No configuration seemed to make any subprovinces more visible, so used Minimum Spanning Tree configuration to plot figures. High betweenness centrality values in the WIS. At lower thresholds, linkages break down somewhat in WIS but still strong connections across whole region. No obvious connection between low generic richness and weak linkages. In fact, strongest links appear to be between those sites with less than 5 genera.

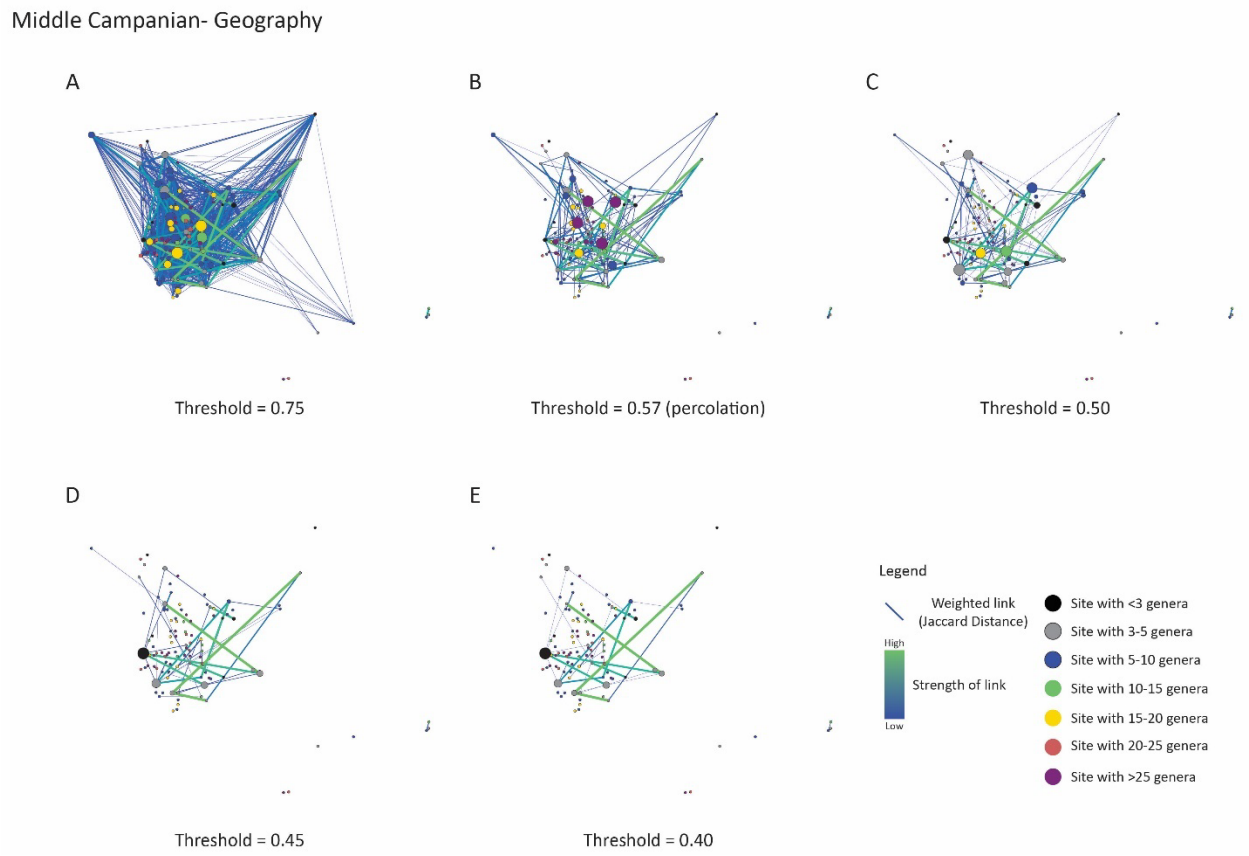



**Figure S13.** Set of network figures for the Middle Campanian with nodes placed based on geographic coordinates. Node sizes indicate betweenness centrality, color denotes the number of unique genus occurrences present at the site, and link weights between nodes denoted by line color and thickness. No obvious structure except in WIS, where nodes interconnected without substructure. Only one major component in WIS and all other components have less than 2 nodes each. GCP region is not well connected at all. At lower thresholds, linkages break down somewhat in WIS but still strong connections across whole region, particularly in the south, south-western, and eastern nodes.

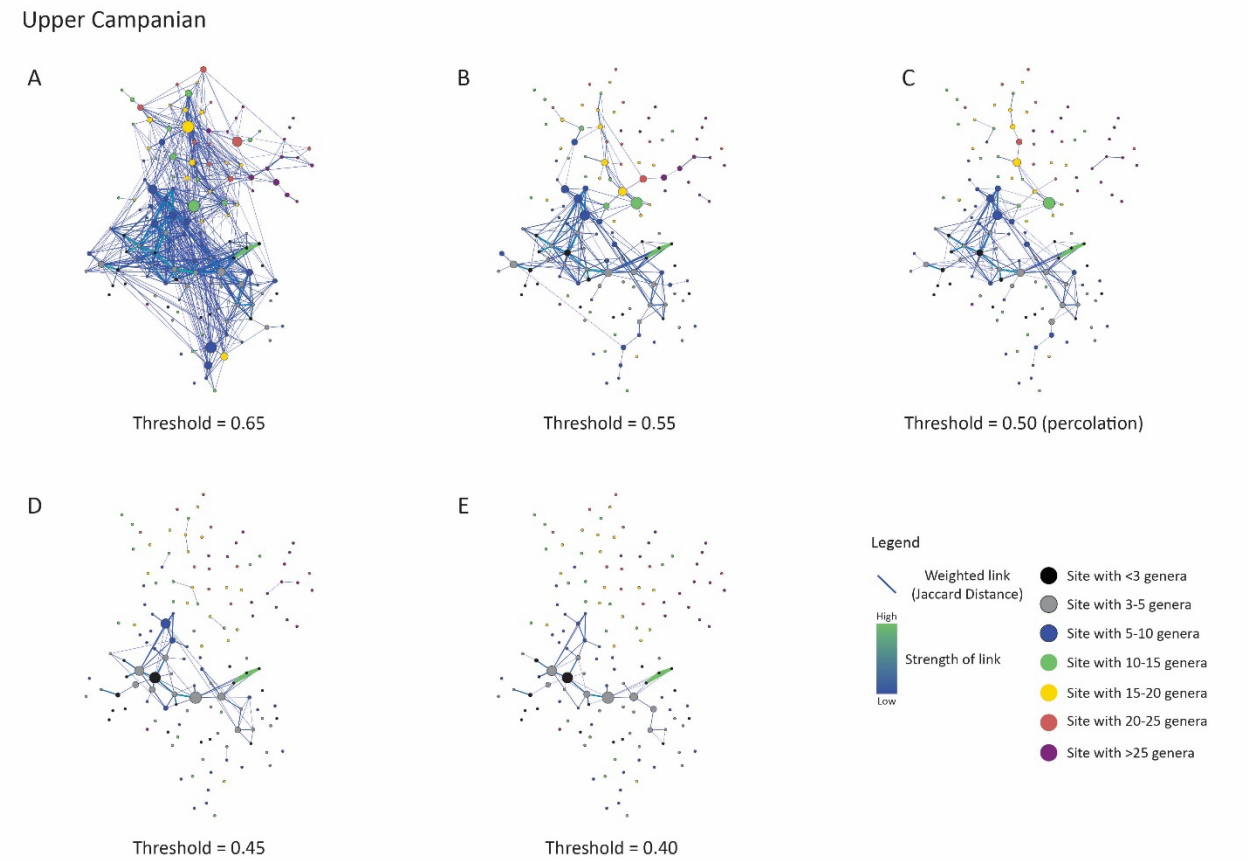



**Figure S14.** Set of network figures for the Upper Campanian. Node sizes indicate betweenness centrality, color denotes the number of unique genus occurrences present at the site, and link weights between nodes denoted by line color and thickness. No obvious structure except in WI, where nodes are interconnected without substructure. Only one major component in WIS and any others have less than 3 nodes each. No configuration seemed to make any subprovinces more visible, so used Minimum Spanning Tree configuration to plot figures. At lower thresholds, linkages break down somewhat in WIS but still strong connections across whole region. No obvious connection between low generic richness and weak linkages. In fact, strongest links appear to be between those sites with less than 10 genera.

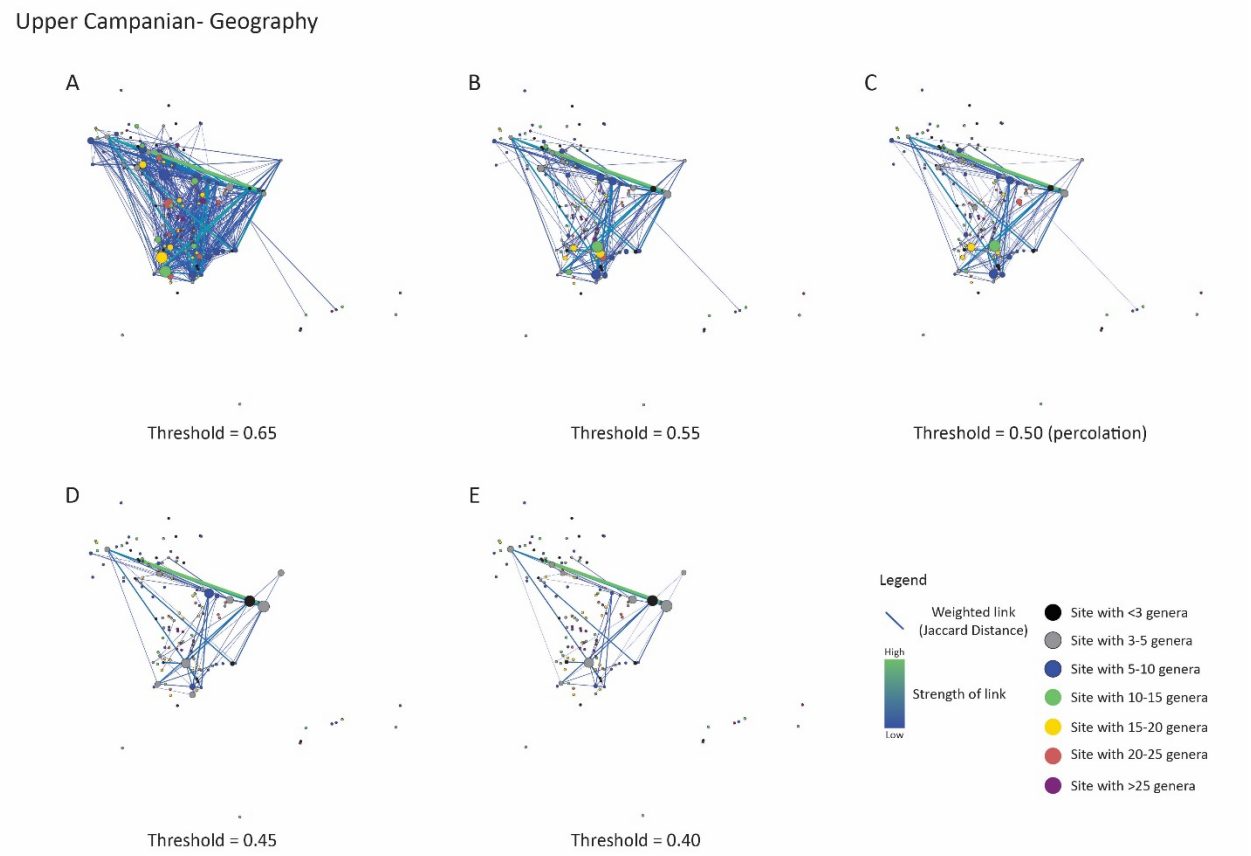



**Figure S15.** Set of network figures for the Upper Campanian with nodes placed based on geographic coordinates. Node sizes indicate betweenness centrality, color denotes the number of unique genus occurrences present at the site, and link weights between nodes denoted by line color and thickness. No obvious structure except in WI, where nodes are interconnected without substructure. GCP region is not well connected at all (only one weak connection between the WIS and GCP nodes). At lower thresholds, linkages break down somewhat in WIS but still strong connections across whole region, particularly between and northwest to eastern nodes.

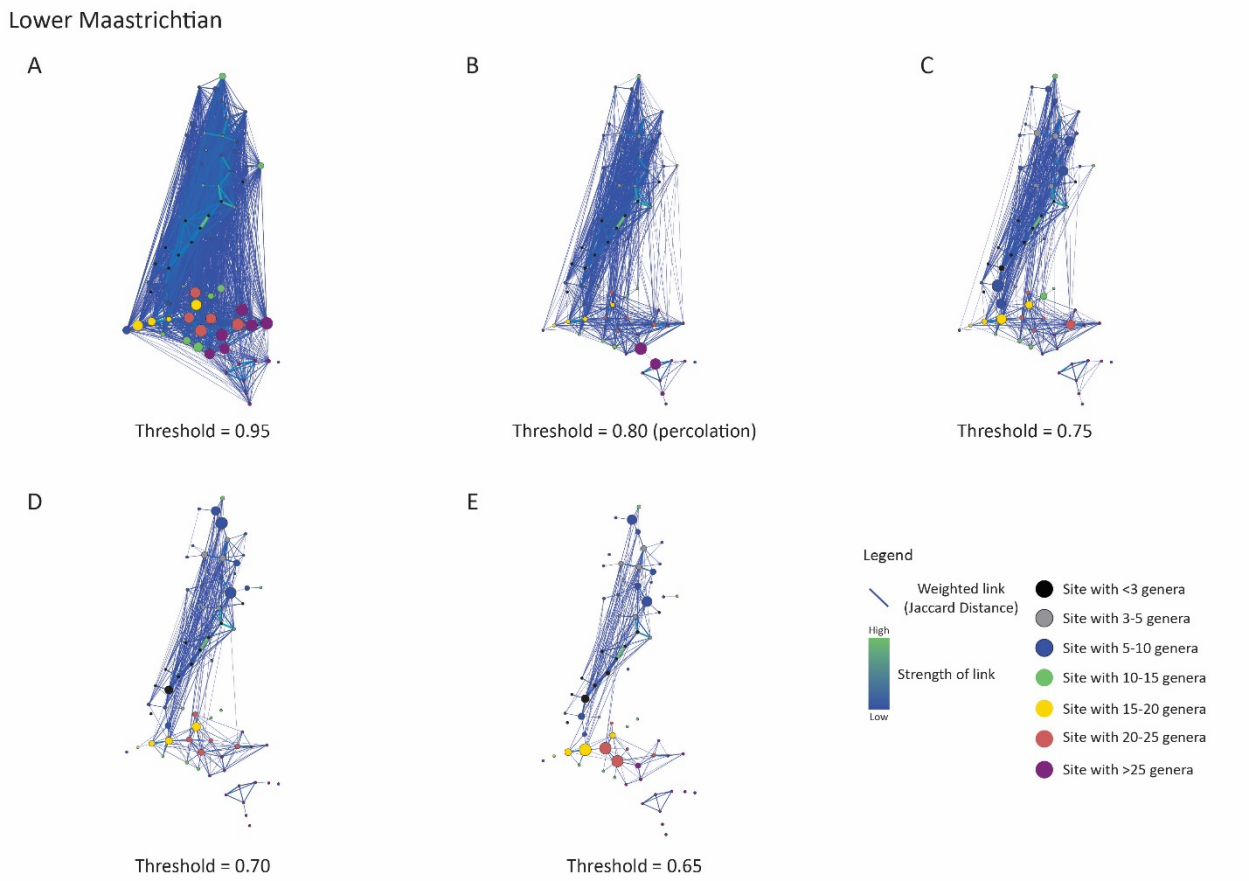



**Figure S16.** Set of network figures for the Lower Maastrichtian. Node sizes indicate betweenness centrality, color denotes the number of unique genus occurrences present at the site, and link weights between nodes denoted by line color and thickness. Two major components in the WIS and GCP regions. One weak connection between GCP site (isolated otherwise) and WIS. No configuration seemed to make any subprovinces more visible, so used Minimum Spanning Tree configuration to plot figures. GCP region on the east side maintains its connections throughout thresholds. High betweenness centrality values in WIS, mostly at lower thresholds, not at percolation point. At lower thresholds, linkages break down somewhat in WIS but still strong connections across whole region. No obvious connection between low generic richness and weak linkages. In fact, strongest links appear to be between those sites with less than 10 genera.

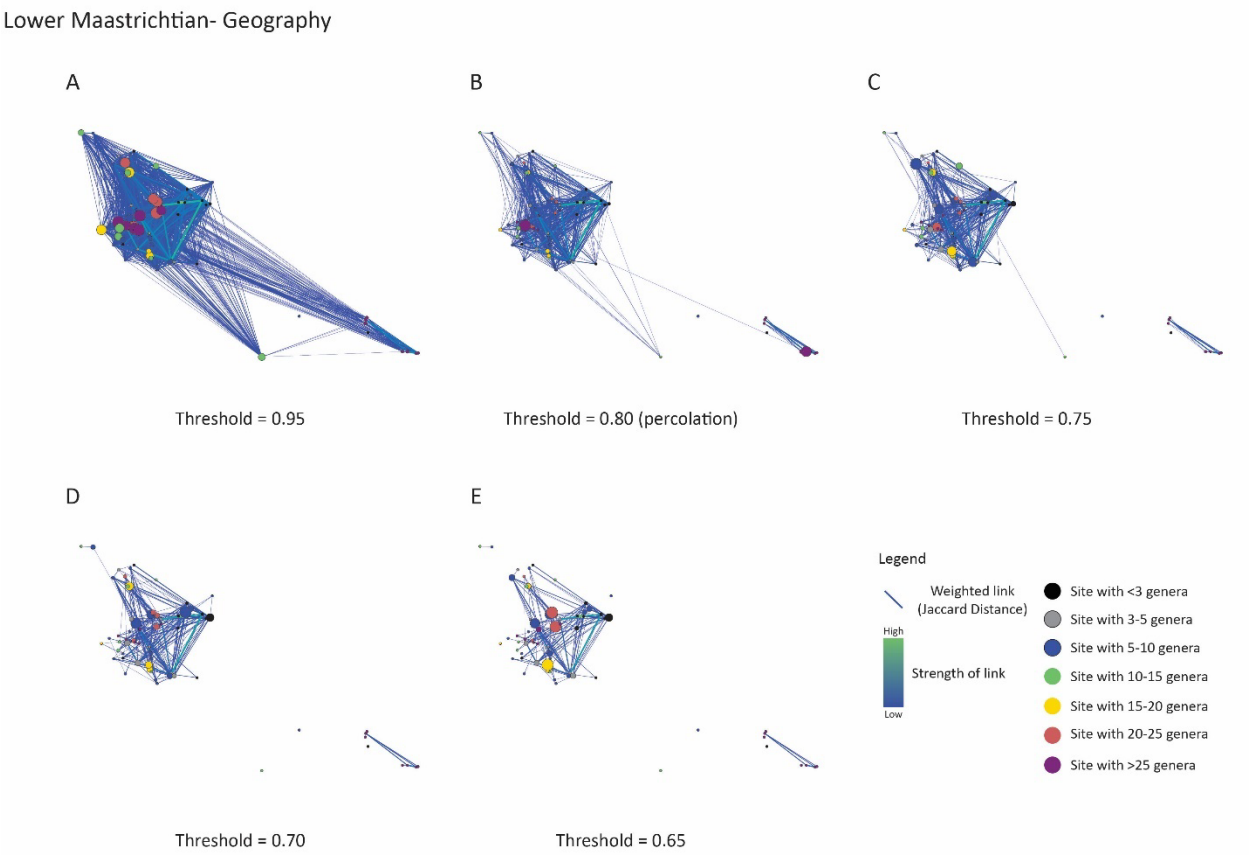



**Figure S17.** Set of network figures for the Lower Maastrichtian with nodes placed based on geographic coordinates. Node sizes indicate betweenness centrality, color denotes the number of unique genus occurrences present at the site, and link weights between nodes denoted by line color and thickness. One weak connection between GCP site and WIS below the percolation point. GCP region on the east side maintains its connections throughout thresholds. High betweenness centrality values in WIS. At lower thresholds, linkages break down somewhat in WIS but still strong connections across whole region, though they are strongest between the southern, central-east, and central nodes.

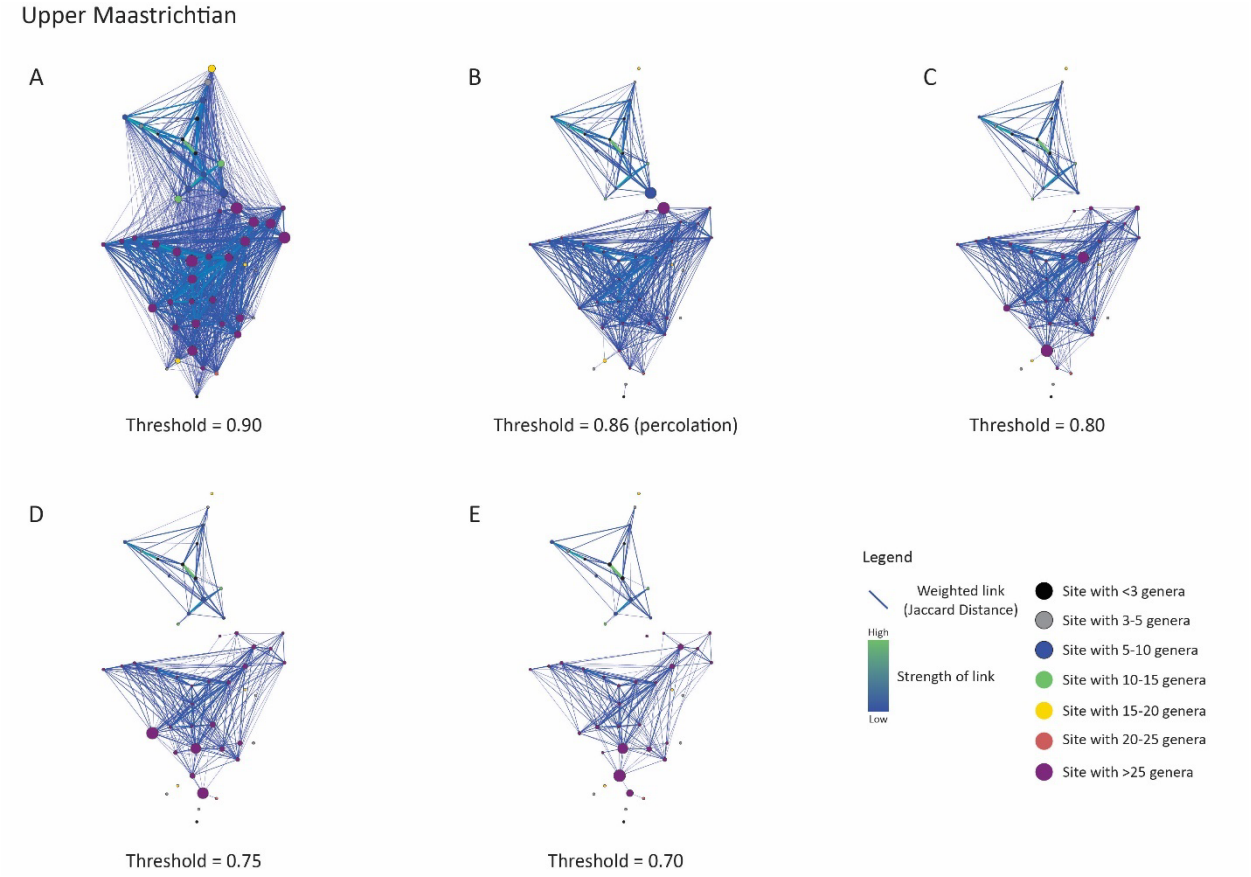



**Figure S18.** Set of network figures for the Upper Maastrichtian. Node sizes indicate betweenness centrality, color denotes the number of unique genus occurrences present at the site, and link weights between nodes denoted by line color and thickness. Two major components, the WIS and GCP regions, connected by one weak connection between the WIS and GCP that is gone by 0.80 threshold. No configuration seemed to make any subprovinces more visible in WI, so used Minimum Spanning Tree configuration to plot figures. GCP region has strong connections across region. High betweenness centrality values in GCP, mostly at lower thresholds, not at percolation point. At very low thresholds (0.5), linkages break down somewhat in GCP and WIS but still strong connections across both regions, particularly in eastern side of GCP. No obvious connection between low generic richness and weak linkages. In fact, some of strongest links appear to be between those sites with less than 5 genera in WIS.

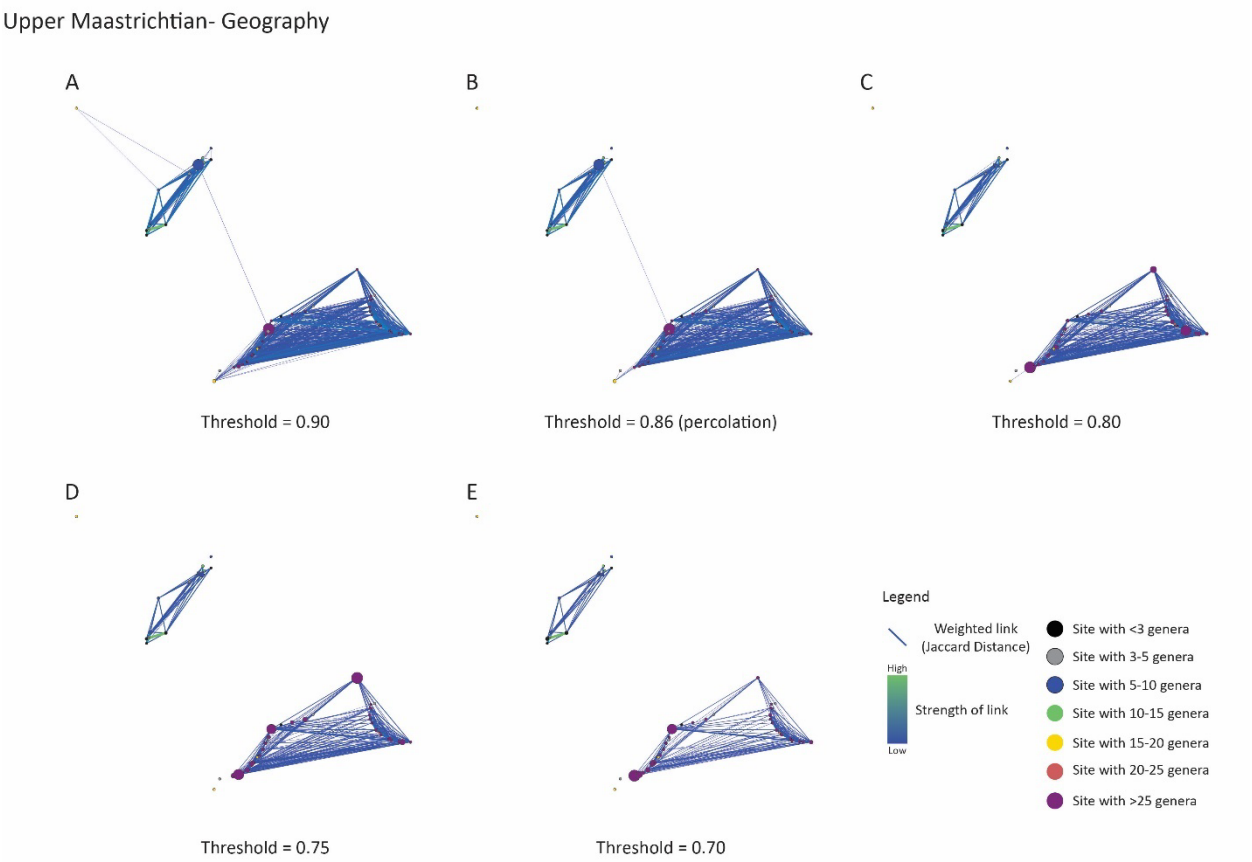



**Figure S19.** Set of network figures for the Upper Maastrichtian with nodes placed based on geographic coordinates. Node sizes indicate betweenness centrality, color denotes the number of unique genus occurrences present at the site, and link weights between nodes denoted by line color and thickness. Two major components, the WIS and GCP regions, connected by one weak connection between the WIS and GCP that is gone by 0.80 threshold. GCP region has strong connections across region. High betweenness centrality values in GCP, mostly at lower thresholds, not at percolation point. At very low thresholds (0.50), linkages break down somewhat in GCP and WIS but still strong connections within both regions, particularly in eastern side of GCP.

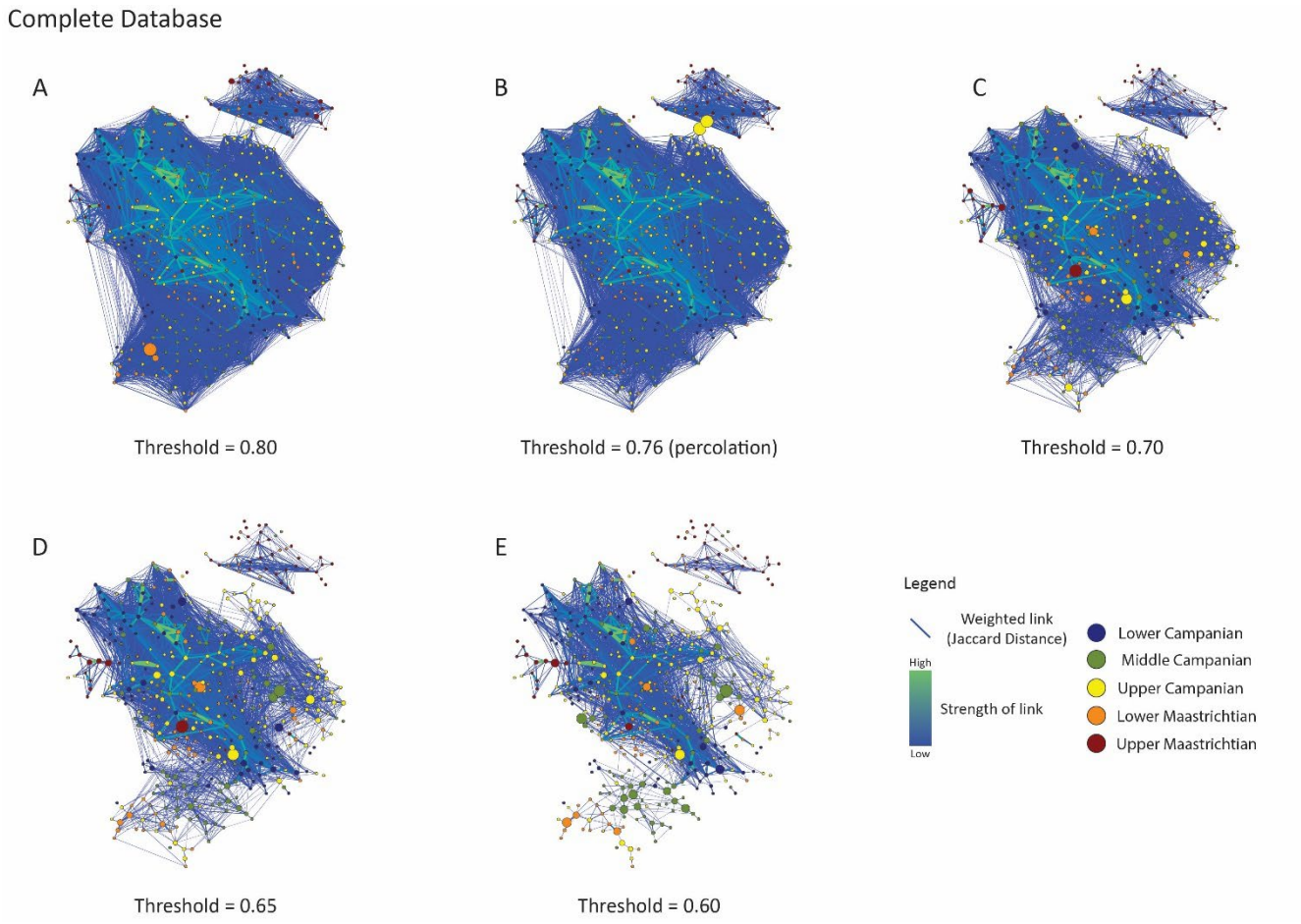



**Figure S20.** Set of network figures for the entire database. Node sizes indicate betweenness centrality, color denotes node substage, and link weights between nodes denoted by line color and thickness. Only 2 major components, not divided otherwise, one is the WIS region and the other is GCP region (mostly MAA). The WIS and GCP regions connected mainly by Upper Campanian sites. Same age sites tend to be more closely connected to each other than different age sites on Minimum Spanning Tree. Topology maintained throughout thresholds.

*Network Randomization Comparisons*

Clustering coefficient values, which describe the ratio of the number existing connections between a node and its neighbors to the maximum possible such connections, were calculated and then averaged over the network as a whole (Saramäki et al. 2007). Clustering coefficients for networks created at the percolation point were compared with a null model of randomized networks that contained the same number of nodes and links to determine if the network was more or less clustered than would be expected from a random distribution (Kiel, 2016; Table S4). Average clustering coefficient values that are distinct from random indicate substructure in the network, suggesting that the network represents aspects of the real-world system, such as community groups or faunal provinces, that differ in assemblage along distinct network components (Newman 2012).

Each clustering coefficient was compared with a null model of randomized networks that contain the same number of nodes and links to determine if the network is more or less clustered than would be expected at random (Kiel 2016). Networks were randomized in R using the igraph and tnet packages (Csárdi & Nepusz, 2006; Opsahl, 2009; see Appendix C for R codes). Randomization for each substage network were run 1000 times, rewiring the same numbers of nodes and links while maintaining link weight distributions. Average clustering coefficient distributions for the randomized networks were then compared with the clustering coefficient of the original network (Table S4). **Results indicate that all network clustering coefficients are greater or less than three standard deviations from the mean, and therefore highly distinct from random.**

**Table S4.** Results of randomization analysis to compare clustering coefficients. 1000 randomized networks created for each network using the same numbers of nodes and links (with maintained link weights) rewired.

| | Networks <CC> | Mean Randomized <CC> | 1st S.D. Rand. <CC> | 2nd S.D. Rand. <CC> | 3rd S.D. Rand. <CC> |
|---|---|---|---|---|---|
| **Lower Campanian** | 0.531 | 0.395 | 0.404 | 0.414 | 0.423 |
| **Middle Campanian** | 0.307 | 0.065 | 0.079 | 0.093 | 0.107 |
| **Upper Campanian** | 0.440 | 0.171 | 0.189 | 0.206 | 0.223 |
| **Lower Maastrichtian** | 0.660 | 0.498 | 0.503 | 0.507 | 0.512 |
| **Upper Maastrichtian** | 0.923 | 0.595 | 0.601 | 0.606 | 0.612 |
| **Complete Database** | 0.569 | 0.368 | 0.369 | 0.370 | 0.371 |

*Complete 360km$^2$ Grid Cell Aggregation Network*

We also spatially aggregated the data to a 360km grid cell size, both to check for sensitivity and to allow for easier spatiotemporal visualization of network connections. This analysis was performed using the same methods as the 60km$^2$-aggregation using instead a 360km$^2$ grid index layer created in ArcGIS Pro using the Grid Index Tool (ESRI, 2021). The 360km aggregation produced lower resolution networks that compared fewer nodes. Furthermore, networks were only created at the percolation point for the composite database as a whole rather than for individual substages and at different thresholds to save time on computation (and following the consistency of network topology patterns found in using the 60km resolution nodes, above).

Results indicate that the patterns observed at the higher resolution 60km-grid-aggregation were maintained at this courser aggregation. The WIS and GCP components were still distinct from one another in the Maastrichtian. One advantage of coarser aggregation is that it allowed for easier visual comparisons across substages when the nodes were mapped geographically (Figure S21). The WIS component is well connected across all substages excluding the Late Maastrichtian, and connections are common between non-consecutive substages from the Middle Campanian through the Early Maastrichtian, though node links are most common between consecutive substages.

These patterns are consistent with patterns results from the 60km-grid aggregation, and furthermore suggest that the WIS was a distinct faunal province that persisted for a significant period of time. During the Late Maastrichtian, the WIS component was still well interconnected and distinct from the GCP component, but it does not share strong connections with the earlier WIS components (i.e., the three Campanian substages or the Lower Maastrichtian). This suggests turnover of fauna in the region while maintaining internal spatial homogeneity within the Late Maastrichtian itself. These patterns support interpretations discussed in the main text, that sea levels fell towards the end of the interval, potentially altering climate and circulation patterns, and altering faunal associations in the Late Maastrichtian.

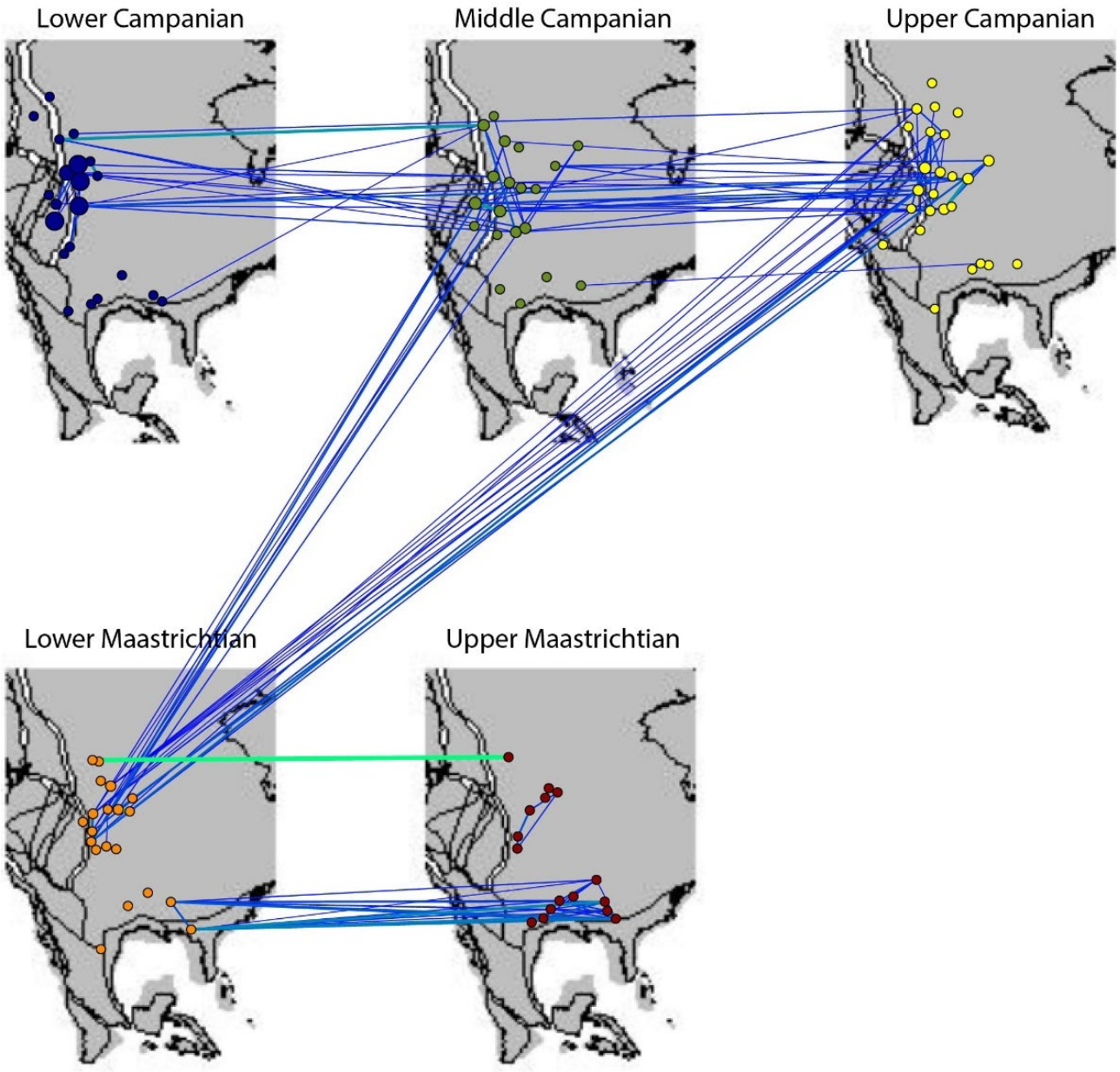



**Figure S21.** Network maps of the composite network created using a 360km2 grid aggregation. Note the well-connected WIS component throughout all substages excluding the Upper Maastrichtian.

*Detailed Subcomponent Analysis*

To determine if much smaller but spatially consistent subcomponents (potential communities at lower spatial scale than "subprovinces") existed in any of the substage networks, identified subcomponents containing more than four interconnected nodes were analyzed at thresholds below the percolation point. A limit of at least four nodes was chosen arbitrarily for this analysis; few subcomponents were observed in any network with fewer than 4 connected nodes. These subcomponents were mapped using the ggplot2 package (Wickham 2016) in R to determine if they represent consistent geographic subprovinces of highly connected nodes. R code for this analysis can be found in Appendix C. Only the Middle Campanian and the Late Maastrichtian were found to contain subcomponents with greater than four interconnected nodes, at threshold levels below the percolation points. Figure S22 and S23 display the major and subcomponents from these networks plotted on a U.S. map and coded by color.

Manual subcomponent analysis did not find any subcomponents in the WIS that were geographically or temporally stable. In the Middle Campanian, the secondary component observed is not consistent through all thresholds, instead only representing the same nodes between the 0.45 and 0.40 thresholds. The third component observed at the 0.45 threshold is furthermore not observed at any other threshold analyzed here. **These results indicate that these components do not represent consistent network features, but are instead lesser faunal associations that shift with changing thresholds.** It is possible, however, that the secondary component observed between the 0.45 and 0.40 thresholds in the WIS represents a significant faunal association, perhaps related to previously described biogeographic features like Kauffman (1984)'s "endemic center" or Central Interior Province. The geographic position of this subcomponent does correspond roughly with both units based on Kauffman's (1984) map (Fig. 7, p. 286). Further work may be able to elucidate if the faunal assemblages in this region are indeed related to a distinct endemic center or a weak subprovince in the WIS but cannot be conclusively determined here. In the Lower Maastrichtian, the secondary

component observed in the 0.75 and 0.60 thresholds indicates that the GCP region, along its eastern margin, is characterized by a consistent subcomponent observed in the overall network topology. However, this component breaks down at the 0.50 threshold. Thus, no other subcomponents are observed in the network that could correspond with subprovinces.

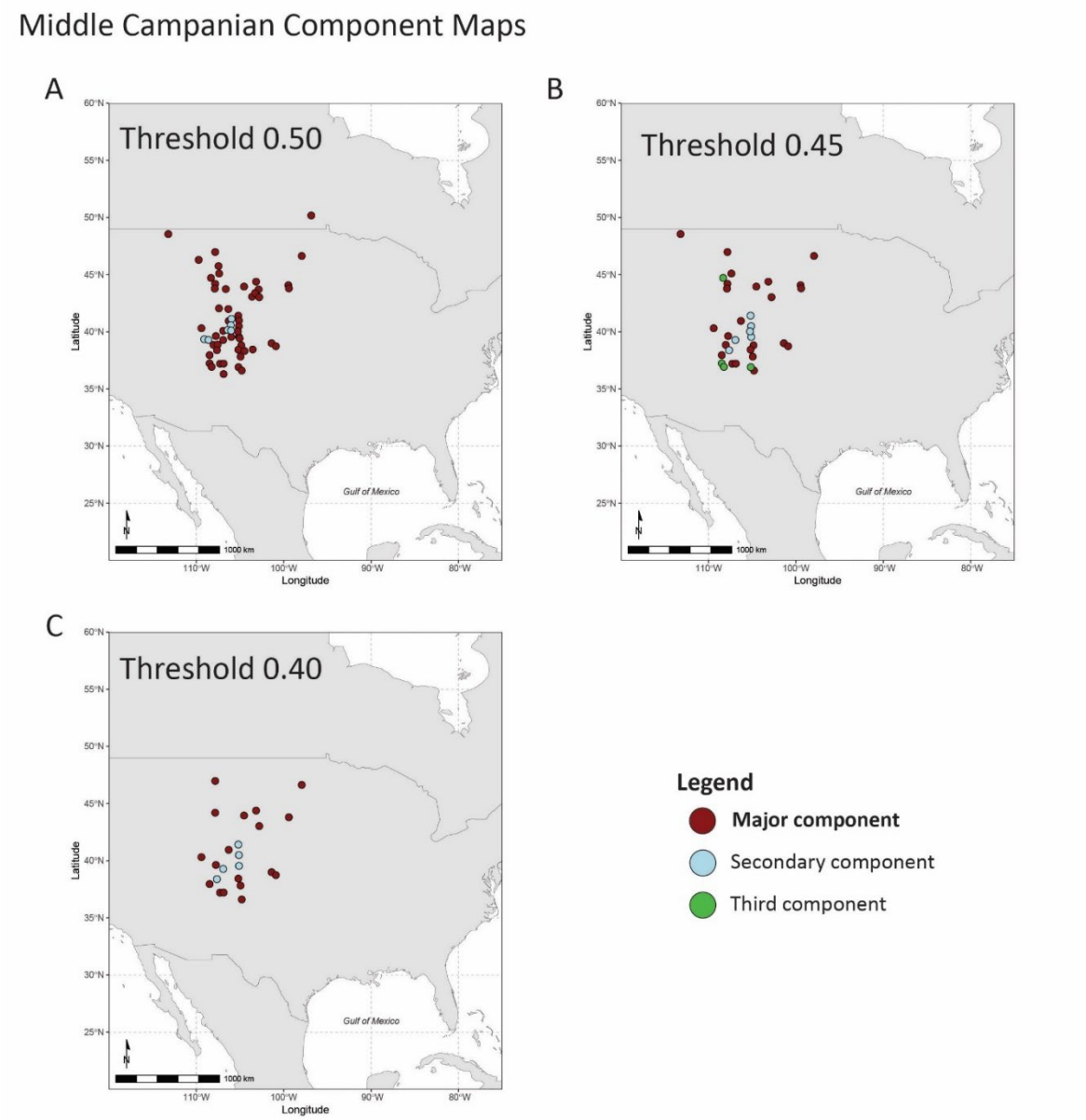



**Figure S22.** Maps of major and minor components (containing >3 interconnected nodes) from network analysis of the Middle Campanian for thresholds 0.40, 0.45, and 0.50. All other node components containing less than 4 nodes were not mapped. The major component for all thresholds covers the majority of the WIS region while the secondary component occupies the middle WI at all threshold levels. However, the secondary component observed at the 0.50 does not consist of the same nodes as those of the 0.45 and 0.40 thresholds. This indicates that a consistent subcomponent only emerges at the 0.45 threshold while the secondary component observed in the 0.50 threshold breaks down into smaller pieces. The third subcomponent is only observed at the 0.45 threshold, indicating that it is a portion of the major component that becomes distinct at the 0.45 threshold level before breaking down into smaller pieces at lower thresholds.

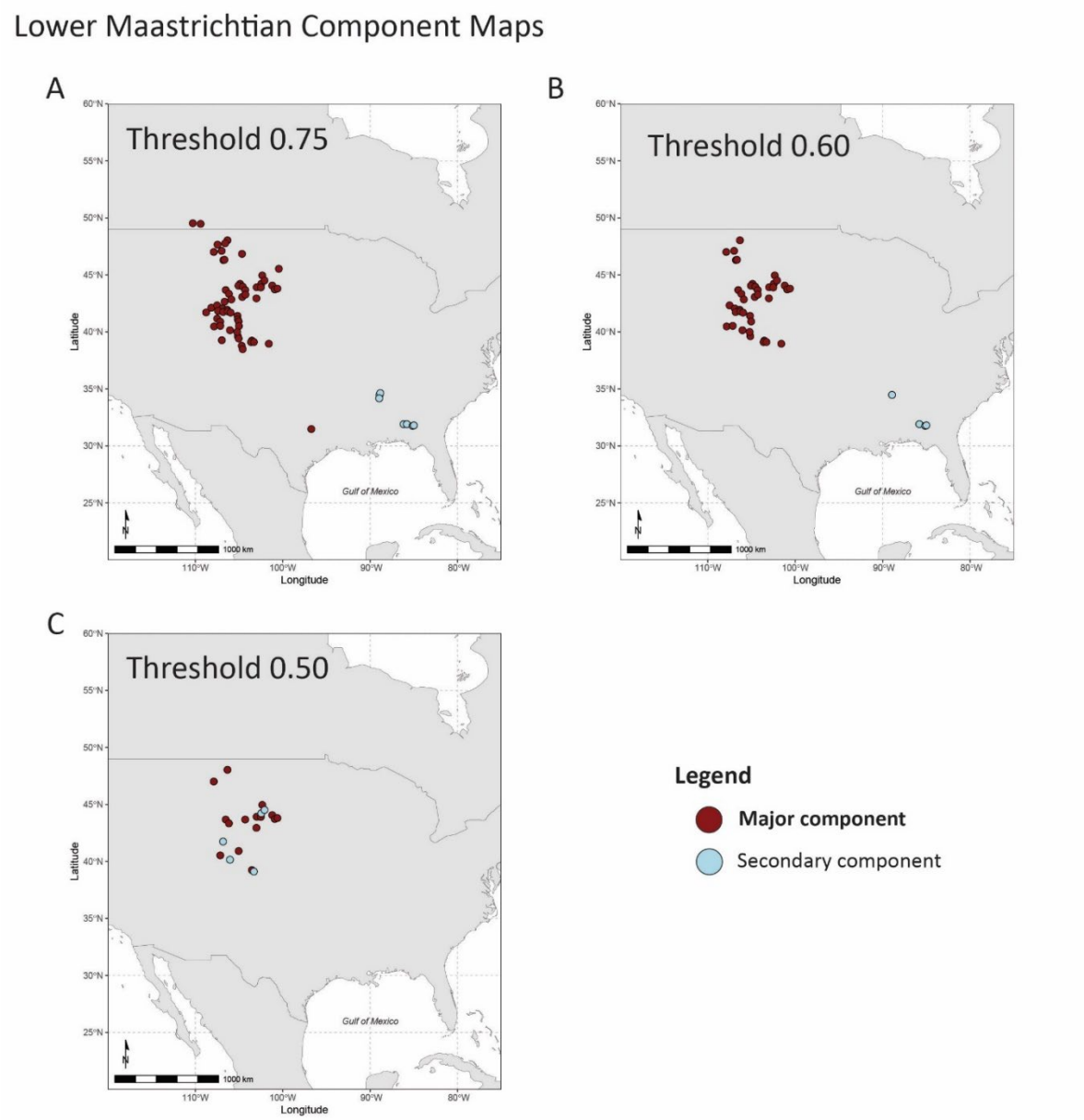



**Figure S23.** Maps of major and minor components (containing >3 nodes) from network analysis of the Lower Maastrichtian for thresholds 0.50, 0.60, and 0.75. All other node components containing less than 4 nodes were not mapped. The major component for all thresholds covers the majority of the WIS region. The secondary component is located in the eastern GCP from the 0.75 to 0.60 thresholds, but shifts to the WIS region at the 0.50 threshold level as the GCP region subcomponent breaks down into ones with less than 4 interconnected nodes. This indicates that the secondary component is fairly consistent and represents the eastern GCP region for most threshold levels, but no other consistent subcomponents exist within the network.

*Average Betweenness Centrality:*

Within each network, the betweenness centrality (BC) metric was measured for each node, describing the importance of that node in connecting other nodes through a shortest path (Kivelä et al. 2015). Betweenness centrality of a node is calculated as

$$BC(i) = \sum_{s \neq v \neq t} \frac{\sigma_{st}(i)}{\sigma_{st}} \qquad (EQ2)$$

Where $\sigma_{st}$ is the number of shortest paths connecting nodes s and t and $\sigma_{st}$(i) is the number of those paths which pass through node *i*. This metric indicates the degree to which a node/region/province acts as potential geographic or phylogenetic connection between otherwise dissimilar regions/provinces. Nodes were binned into 5-degree paleolatitudinal bins to compare average BC ($BC_{ave}$) across latitudinal space.

It should be noted that, given that the GCP does not contain as many latitudinal bins as the WIS (only existing in the study area between approximately 25-37°N), latitudinal binning comparisons across the GCP are inappropriate at this scale of analysis. Further, there is some latitudinal overlap between the southernmost WIS and the northernmost GCP, around 35°N, due to the Mississippi Embayment which may influence these results. Since we found that generic richness was correlated strongly with sampling effort (Figure S4) we assessed its impact on network communication by comparing $BC_{ave}$ values calculated using generic richness bins (Figure S25). Generic richness values did not have a strong relationship with $BC_{ave}$ values, indicating that generic richness and, by that proxy, sampling effort did not strongly control $BC_{ave.}$ Had sampling effort and generic richness correlated strongly with $BC_{ave}$, it would be indicative of sampling bias.

Betweenness centrality varied between substages and at different threshold levels for each substage (Figure S24). The values of $BC_{ave}$ were consistently higher at the percolation threshold for all substages than at all thresholds below the percolation points, indicating that the number of shortest paths running through nodes was particularly high in the network at the percolation point. Higher

$BC_{ave}$ values at thresholds above the percolation point are present but not common, and most substages show a decrease in $BC_{ave}$ at thresholds above the percolation point (Figure S24). **These results suggest that nodes within each substage network are particularly important for forming connections between other nodes at the percolation point.** However, it should be noted that betweenness centrality is greatly dependent on network density, or the number of links between nodes, and therefore the level of influence a set of nodes has on network connectivity can only be directly compared within the same threshold level for a substage. This enables comparisons of BC within substage thresholds based on different latitude bins and binned numbers of unique fossil occurrences (Figure S25).

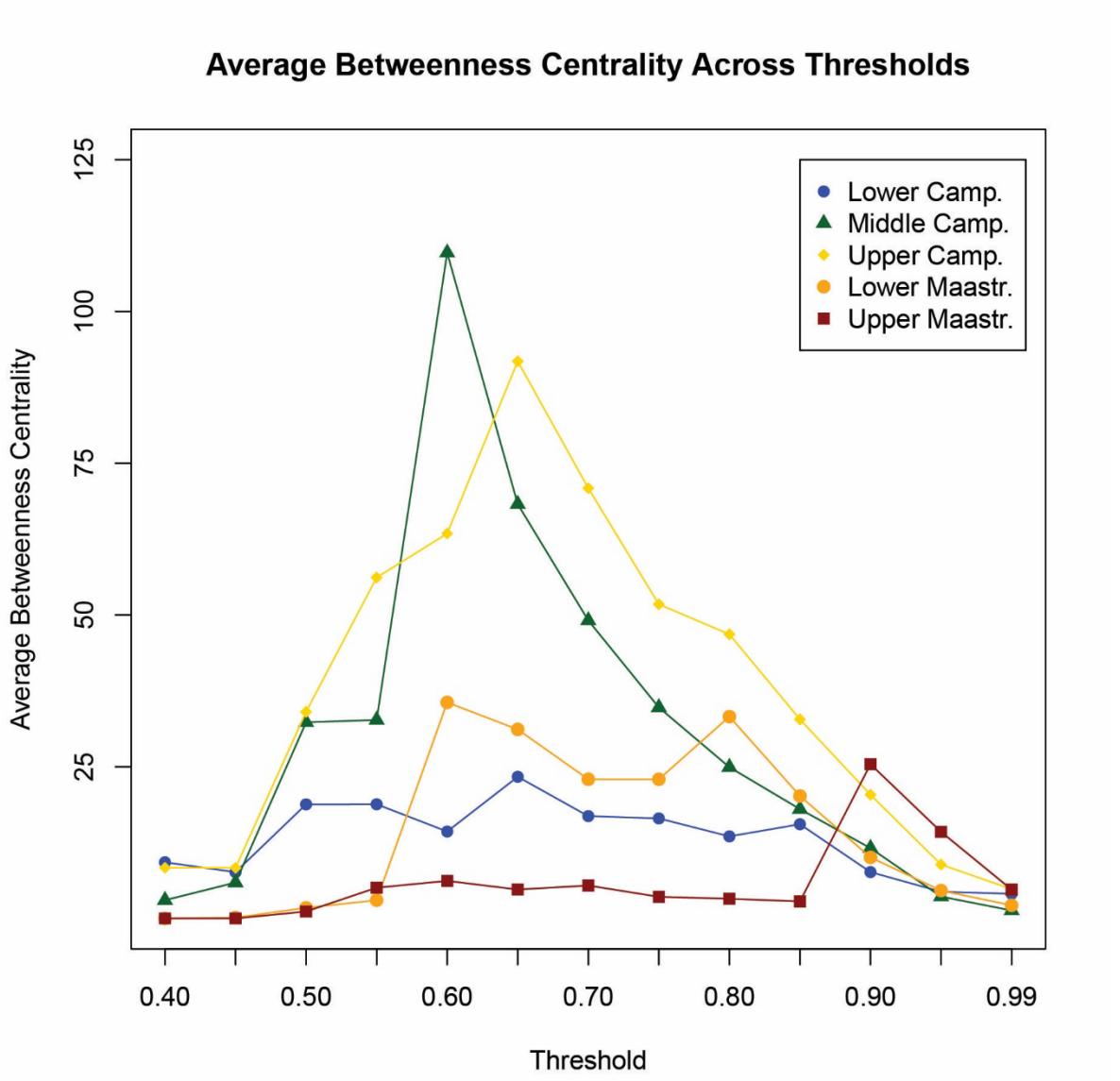


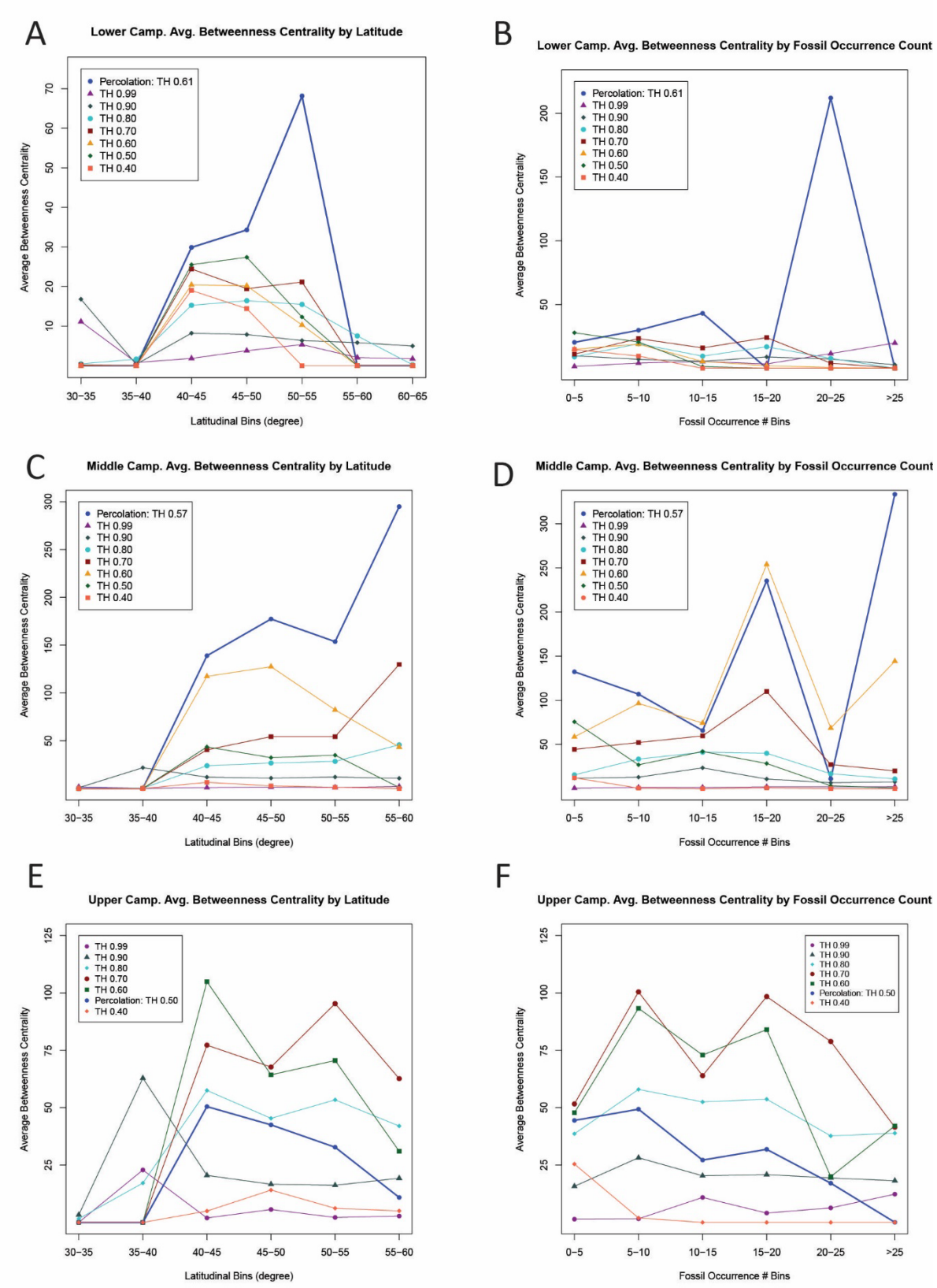

**Figure S24.** Average betweenness centrality values for the individual substage networks at different threshold levels at and below their percolation points. The highest value of BC was recorded at the percolation point for the Middle Campanian (0.57), at 156.6. This value suggests that the network was highly structured. All other substages have values less than 40, and all substages show a general decrease in BC values at thresholds below their percolation points excluding the Lower Maastrichtian. The Lower Maastrichtian has a relatively stable BC value between its percolation point (0.80) and 0.60, with some variation, indicating that structure remains relatively high in these networks up until the 0.55 threshold level. Line color and symbol indicates substage, dashed lines are interpolated across several threshold levels.

Figure S25 shows the results of $BC_{ave}$ for each substage at different thresholds based on latitudinal bins and sampling count bins. Average BC values in the Campanian substages are highest in the mid- to northern latitudes (40-60°N; Figure 5), consistent with higher sampling in the WIS vs. GCP during these intervals. However, at the percolation point for each substage, the latitudinal bin with the highest $BC_{ave}$ shifts south through time. In the Lower and Middle Campanian, the highest $BC_{ave}$ is found in the 50-55°N and 55-60°N bins, respectively, but in the Upper Campanian the highest $BC_{ave}$ value was observed in the 40-45°N bin. This pattern is consistent but dampened at other thresholds, except in the Upper Campanian where the pattern is amplified at thresholds just above the percolation point (Figure S24). The highest average BC value for the Lower Maastrichtian substage is in the 35-40°N latitude bin; this shifts in the Upper Maastrichtian to highest average BC values in the 45–50°N bin with a second highest value in the 35–40°N bin. The fact that the second-highest $BC_{ave}$ value in the Lower Maastrichtian exists in the 35-40°N bin is probably a product of increased sampling effort for that latitude, and suggests network patterns are somewhat influenced by sampling distributions in geographic space. Overall, patterns observed at the percolation point for each substage were paralleled but dampened at other thresholds, except for the Upper Campanian and both Maastrichtian substages where these trends were amplified but not substantially distinct. **These patterns indicate that network topology is consistent at different thresholds.**

Average BC binned by generic richness (i.e., the number of unique genera per 60-km grid cell) (vs. generic richness) do not indicate that sampling bias impacts the ability for nodes to act as communicating links between other areas. If greater sampling bias and therefore higher generic richness greatly influenced network communication, we would expect there to be a positive relationship between $BC_{ave}$ and the number of unique genera. However, nodes with lower richness appear to have higher general $BC_{ave}$ than nodes with greater richness for most substages and their thresholds, though this is not a universal trend and specifics depend on threshold level as well as substage. At thresholds below the percolation point, peaks in $BC_{ave}$ for richness bins are diminished,

excluding the Lower Maastrichtian substage. **Average BC of generic richness bins therefore suggests no strong correlation exists between the number of unique taxa and a node's ability to communicate information within the network (Figure S25).**

Overall, these results indicate that lower numbers of unique taxa do not diminish a node's ability for be positioned along the shortest connecting paths between other nodes, communicating faunal information succinctly, and may in fact result in increasing the likelihood for these positions. Linkage patterns within individual substage Minimum Spanning Tree (MST) furthermore indicate that, while associations the number of unique genera in a node are non-random, they are not dominated by links between nodes of the same bin of unique genera counts (see more on MST below). This suggests that the number of unique genera within a grid cell has some influence on network topology, but that it is not the primary deciding factor. Nodes with fewer unique genera are probably characterized by highly common taxa, making them important for connecting nodes with more distinct and rarer genera. This interpretation corresponds well with the results of MST assessments, since nodes with a lower number of unique genera appear to form stronger links than those with greater number of unique genera. This may indicate that the nodes with fewer unique taxa that are common form strong links to one another and form links with nodes containing a more unique genera, while nodes with more and rarer taxa differ more, forming weaker links. **Thus, sampling bias was not found to strongly impact network connections.**


A
Lower Camp. Avg. Betweenness Centrality by Latitude
Percolation: TH 0.61
TH 0.99
TH 0.90
TH 0.80
TH 0.70
TH 0.60
TH 0.50
TH 0.40
Average Betweenness Centrality
10
20
30
40
50
60
70
30–35
35–40
40–45
45–50
50–55
55–60
60–65
Latitudinal Bins (degree)
B
Lower Camp. Avg. Betweenness Centrality by Fossil Occurrence Count
Percolation: TH 0.61
TH 0.99
TH 0.90
TH 0.80
TH 0.70
TH 0.60
TH 0.50
TH 0.40
Average Betweenness Centrality
50
100
150
200
0–5
5–10
10–15
15–20
20–25
>25
Fossil Occurrence # Bins
C
Middle Camp. Avg. Betweenness Centrality by Latitude
Percolation: TH 0.57
TH 0.99
TH 0.90
TH 0.80
TH 0.70
TH 0.60
TH 0.50
TH 0.40
Average Betweenness Centrality
50
100
150
200
250
300
30–35
35–40
40–45
45–50
50–55
55–60
Latitudinal Bins (degree)
D
Middle Camp. Avg. Betweenness Centrality by Fossil Occurrence Count
Percolation: TH 0.57
TH 0.99
TH 0.90
TH 0.80
TH 0.70
TH 0.60
TH 0.50
TH 0.40
Average Betweenness Centrality
50
100
150
200
250
300
0–5
5–10
10–15
15–20
20–25
>25
Fossil Occurrence # Bins
E
Upper Camp. Avg. Betweenness Centrality by Latitude
TH 0.99
TH 0.90
TH 0.80
TH 0.70
TH 0.60
Percolation: TH 0.50
TH 0.40
Average Betweenness Centrality
25
50
75
100
125
30–35
35–40
40–45
45–50
50–55
55–60
Latitudinal Bins (degree)
F
Upper Camp. Avg. Betweenness Centrality by Fossil Occurrence Count
TH 0.99
TH 0.90
TH 0.80
TH 0.70
TH 0.60
Percolation: TH 0.50
TH 0.40
Average Betweenness Centrality
25
50
75
100
125
0–5
5–10
10–15
15–20
20–25
>25
Fossil Occurrence # Bins

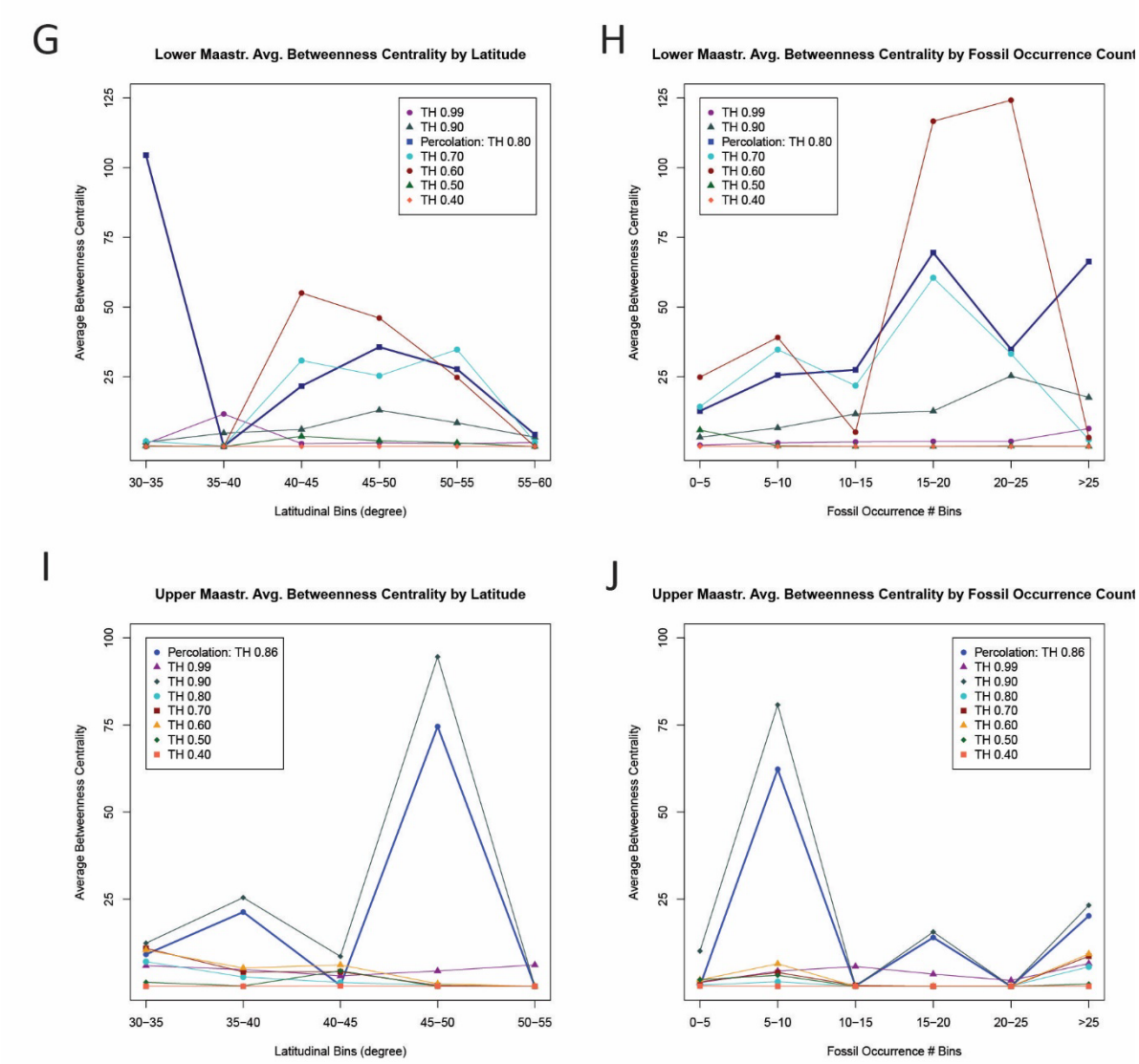



**Figure S25.** Average BC values for the different substages binned by latitude and by the number of fossil occurrences aggregated into each grid cell or node.

*Average Link Weights through Time and by Latitude:*

A comparison of average/median network link weights with latitude revealed significant variability within each substage and no clear trend across substages (Figure S26, A and B; Table S5), suggesting that faunal similarity does not vary consistently with latitude through time or space. In other words, network similarity neither increases or decreases across latitude in any interval. Generally, the lowest average similarity was observed in the 35-40°N bin. Highest similarity shifts north from the 40-45°N bins in the Lower and Middle Campanian to 45-50°N in the Upper Campanian; similarly, the Lower Maastrichtian showed the highest average similarity at 30-35°N, shifting to 45-50°N in the Upper Maastrichtian. Faunal similarity is overall greater between nodes within a latitudinal bin and those nodes without and higher between nodes in the WIS bins (40-50°N) to those without than the GCP (Figure S26, C and D).

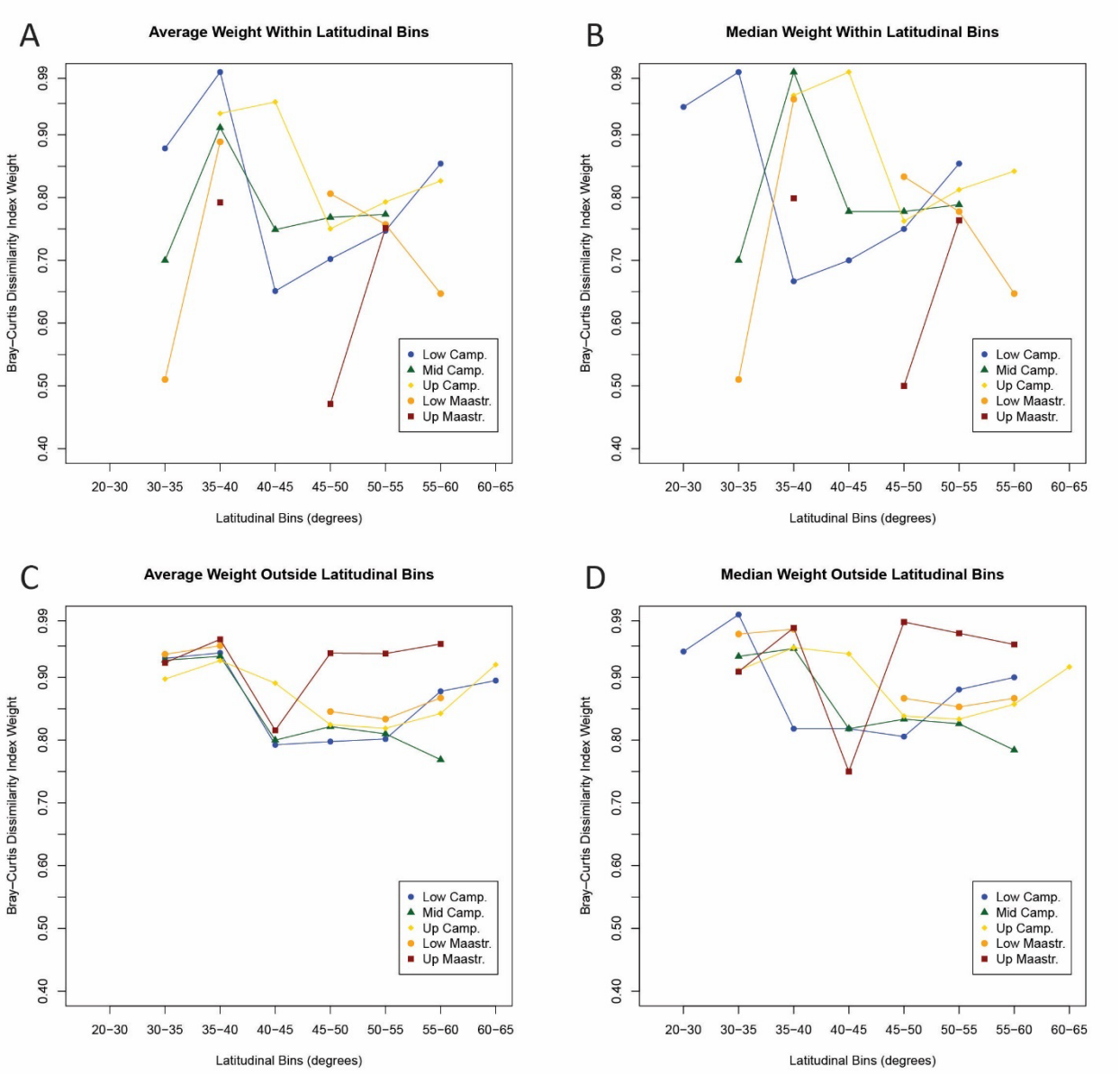



**Figure S26.** Plots of average and median weights (Bray–Curtis dissimilarity index) within bins of latitude.

*Distance Comparisons:*

The relationship between network links and geographic distance was analyzed to determine how distance impacts faunal similarity across the study region. Under the null hypothesis it is expected that, as distance increases, link dissimilarity values or weights will increase as well given an assumption that faunal similarity is primarily spatially controlled. Distance values were calculated for the completely connected networks (Threshold = 0.99) for each substage in order to assess all faunal links based on paleocoordinates using the geodesic distance calculation from the geosphere package in R (Hijmans, 2021; code in Appendix C). Link weights between different nodes were binned by 50- to 500-kms and used to calculate average and median weight values. Networks were also analyzed by binning link values into 5-degree paleolatitudinal bins as mentioned previously to compare link weight change across latitude.

Comparisons between geographic distance and faunal similarity determined that the similarity of nodes is greater in general when the two nodes are closer together, but weakens significantly at ~1000 km distance. At very great distances (i.e., >2000 km), the correlation similarity loses power in most substages (Figure S27). Results show little increase in weight with distance until the 1000-1500 km bin at which weights increase sharply. Average network link weight did not consistently continue to increase beyond the 2000 – 2500 km bin range. Overlap in confidence intervals indicate that most substages do not have significantly different similarity values. However, mean similarity notably decreases from the Lower to the Middle Campanian (Figure S27 A). Median values showed similar trends, with generally higher average values in each bin.

To test for differences in similarity across distance within the different basins, the WIS and GCP components identified by the networks were also analyzed individually, following the same procedure as use for overall basin link weight averages (see section Average Dissimilarity by Network-Identified Components below). Patterns observed in the collective database were largely maintained in the WIS and GCP analyzed individually, but distances greater than ~2000 km were not

frequently observed (Figure S29). In the WIS, only the Lower Maastrichtian is distinctly less similar than other substages across most distance bins (Figure S29 A). In the GCP, confidence interval values in almost all distance bins overlap, suggesting that no significant difference in similarity is observed (Figure S29 C).

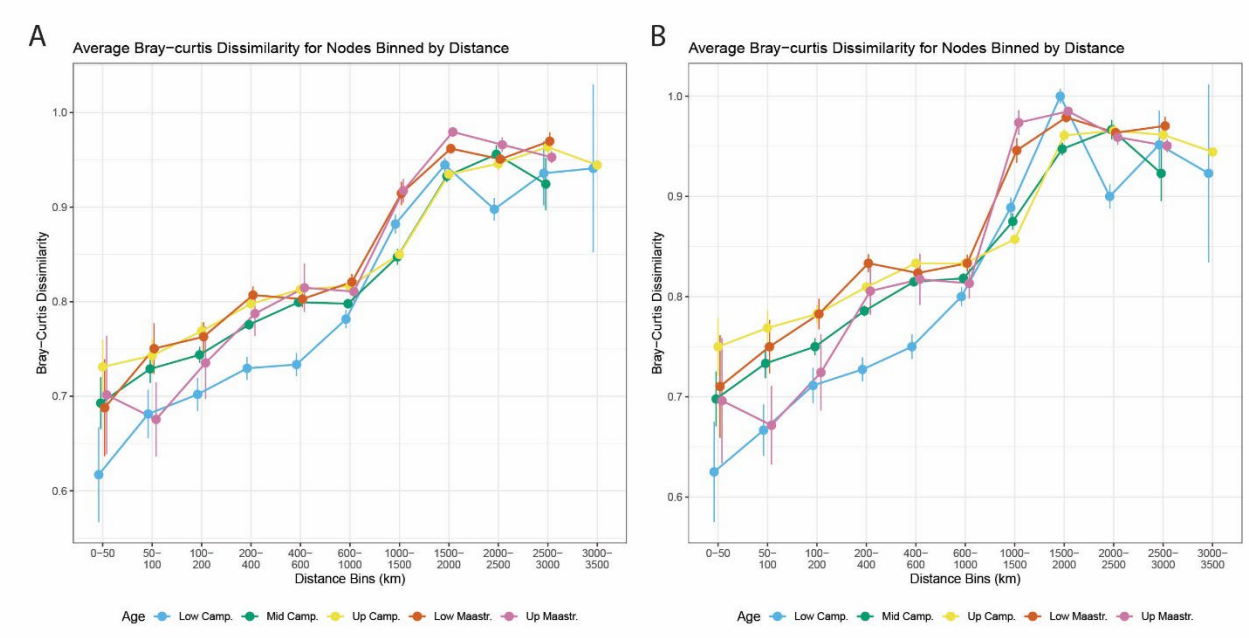



**Figure S27.** Plots of average (A) and median (B) weights (Bray–Curtis dissimilarity index) within bins of geographic distance (geodesic distance using the WGS84 ellipsoid). 95% Confidence Intervals indicated by vertical bars.

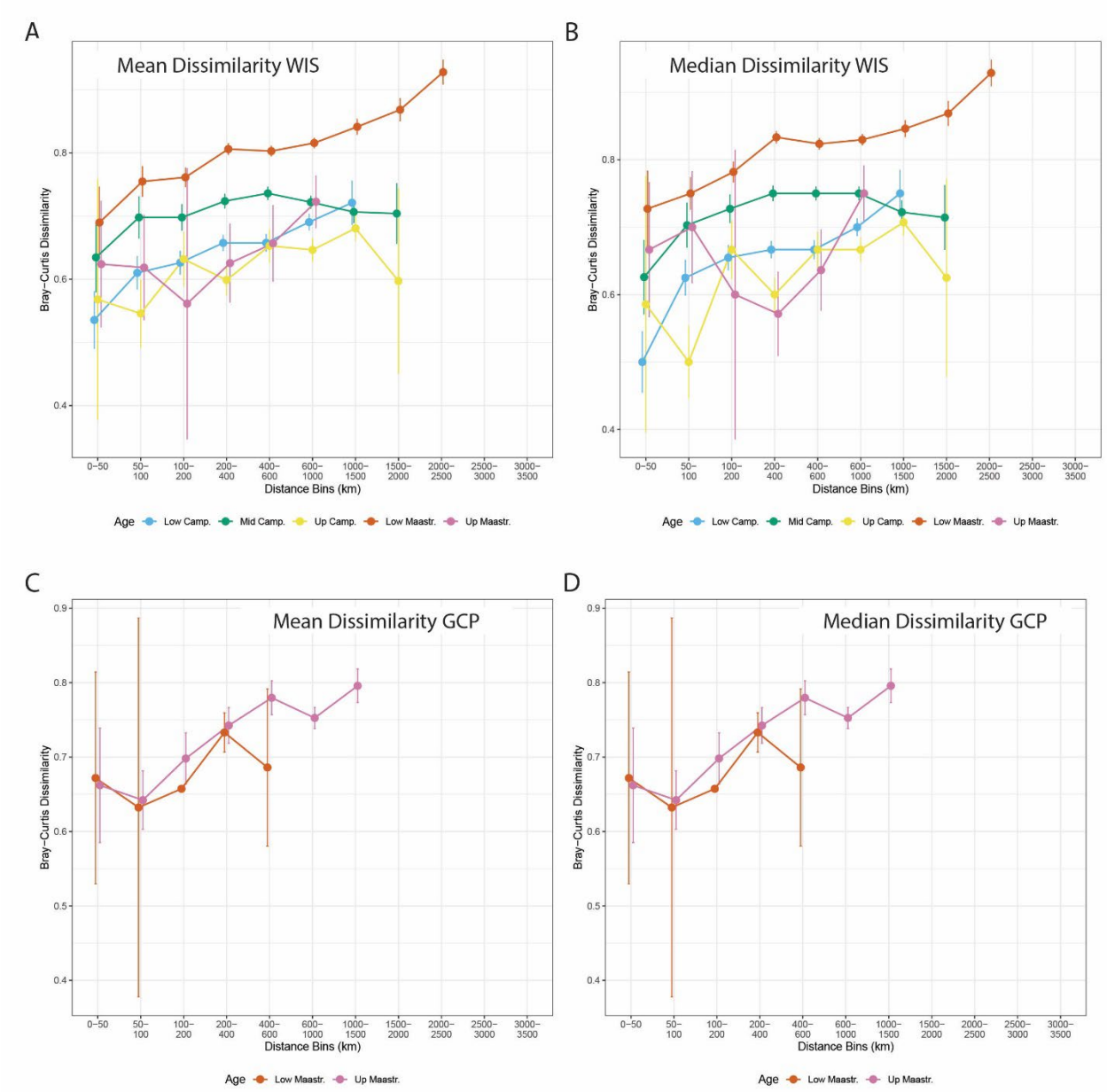



**Figure S28.** Plots of average (A) and median (B) WIS weights (Bray–Curtis dissimilarity index) and average (C) and median (D) GCP weights within bins of geographic distance (geodesic distance using the WGS84 ellipsoid). 95% Confidence Intervals indicated by vertical bars.

**Table S5**. Table of mean/median weights (Bray–Curtis dissimilarity index) binned by latitude for each substage. NA values indicate that the bin lacked occupied grid cells.

| Degree Bins | Low CAM | Mid CAM | Up CAM | Low MAA | Up MAA |
|---|---|---|---|---|---|
| **30-35** | 0.88/0.94 | 0.70/0.70 | NA | 0.51/0.51 | NA |
| **35-40** | 1.0/1.0 | 0.91/1.0 | 0.93/0.96 | 0.89/0.96 | 0.79/0.80 |
| **40-45** | 0.65/0.67 | 0.75/0.78 | 0.95/1.0 | NA | NA |
| **45-50** | 0.70/0.70 | 0.77/0.78 | 0.75/0.76 | 0.81/0.83 | 0.47/0.50 |
| **50-55** | 0.75/0.75 | 0.77/0.79 | 0.79/0.81 | 0.76/0.78 | 0.75/0.76 |
| **55-60** | 0.85/0.85 | NA | 0.83/0.84 | 0.65/0.65 | NA |
| **60-65** | NA | NA | NA | NA | NA |
| **Total** | 0.78/0.80 | 0.79/0.82 | 0.82/0.84 | 0.83/0.86 | 0.85/0.95 |

*Average Dissimilarity by Network-Identified Components:*

The GCP region is defined as the submerged continent below the TA (~ northern TX) and extends east to overlap in latitude the southern WIS in the Mississippi Embayment, which might render latitude-based comparisons of the region highly uninformative. It is furthermore possible that the WIS and GCP provinces experienced different faunal association patterns across time which cannot be observed when the two are assessed in the same network. To test if network patterns differ between the WIS and GCP basins, we examined each provincial component individually. We separated the WIS from the GCP based on network components identified just below the percolation point for each substage (when components become disassociated from one another), which are geographically associated with the WIS and GCP in R. Faunal dissimilarity values represented by the links between nodes present in these two components were then used to calculate average faunal dissimilarity for both the WIS and GCP at each substage. WIS and GCP component values were averaged to include all links (i.e., a threshold of 100) but based on network-identified components at a level just below the percolation point for each, when the two components are disconnected completely. The GCP component was only present and therefore only calculated for in the Maastrichtian substages..

Results of average link weights for the WIS and GCP as distinct component not restricted to latitude show that WIS average faunal similarity was stable during the Campanian then decreased in the Maastrichtian while similarity increased within the GCP over the Maastrichtian (Table S6).

**Table S6.** Results of dissimilarity comparisons between network identified components representing the WIS and GCP in each substage. The threshold level used to identified nodes within each component are listed, as well as average dissimilarity values for links between identified nodes in the fully connected network for each substage.

| Substage | Threshold of ID'ed components | WIS Avg Dissimilarity | GCP Avg Dissimilarity |
|---|---|---|---|
| Lower Camp. | 0.60 | 0.726 | NA |
| Middle Camp. | 0.55 | 0.771 | NA |
| Upper Camp. | 0.45 | 0.779 | NA |
| Lower Maastr. | 0.75 | 0.810 | 0.919 |
| Upper Maastr. | 0.85 | 0.882 | 0.816 |

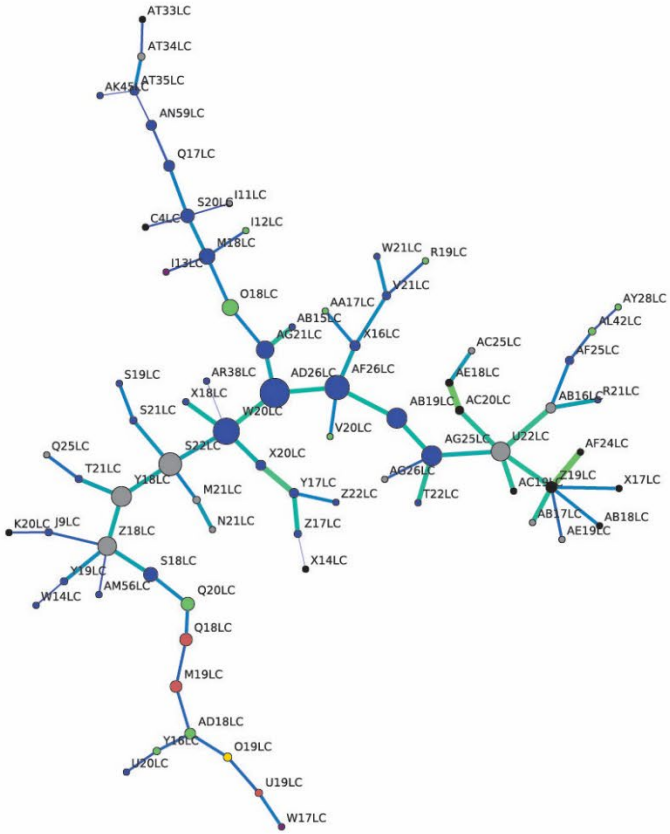
*Minimum Spanning Trees:*

Beyond the previous analysis binning $BC_{ave}$ by generic richness demonstrating that sampling bias did not control network component results (i.e., identified WIS and GCP faunal provinces), we tested for sampling bias by comparing the influence of generic richness on network topology using minimum spanning trees (MST). MST are subnetworks that connect nodes using the minimum link weights possible while preventing loops (forcing the network to produce only branching trees where branches cannot reconnect via looping), therefore representing fundamental aspects of network topology (Graham and Hell 1985; Mareš 2008). The topology of MSTs test sampling bias by calculating the degree of clustering between nodes containing similar generic richness. If sampling bias is the primary factor controlling topology, MSTs should show extremely high clustering between nodes of the same richness bin, a lower relative proportion of links between adjacent bins, and virtually no links between non-adjacent bins. Essentially, there should be a high dominance of links between nodes of the same richness level, were richness (i.e., sampling bias) the primary factor controlling how similar nodes are to one another.

Results of MST link proportion comparisons for substages (Figure S29, A-E) indicate that a high and relatively even proportion of links exist between nodes with the same or similar generic richness, and very low proportions of non-adjacent bin links (Tables S9-S13). This result suggests that while sampling bias influenced network topology, it was not the *primary* factor delineating network connections. No MST shows a majority of clustering between nodes of the same richness bin with a lower proportion between adjacent bins and virtually none between non-adjacent bins, excluding perhaps the Upper Maastrichtian. **Therefore, sampling bias is unlikely to have strongly influenced network topology in all substages, except perhaps the Upper Maastrichtian**. The MST for the complete database indicates that most substages are dominated by within-interval connections, rather than randomly or geographically sharing links with other intervals (Figure S29, F). This is expected based on an assumption of temporal autocorrelation and supports age-defined

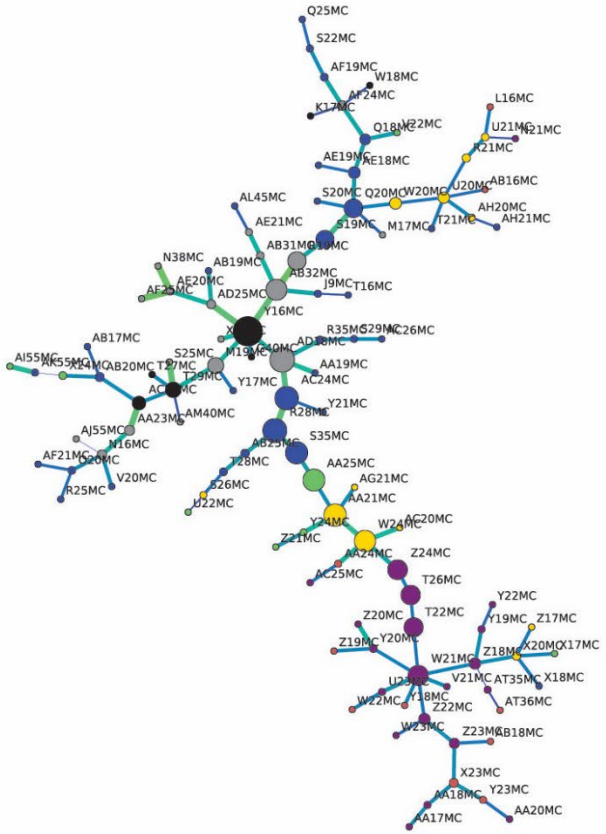
communities as a primary control on topology. Substages are also linked with adjacent substages, rather than to non-adjacent ones, consistent with the expectation that communities will become less similar through time (Table S14).

A



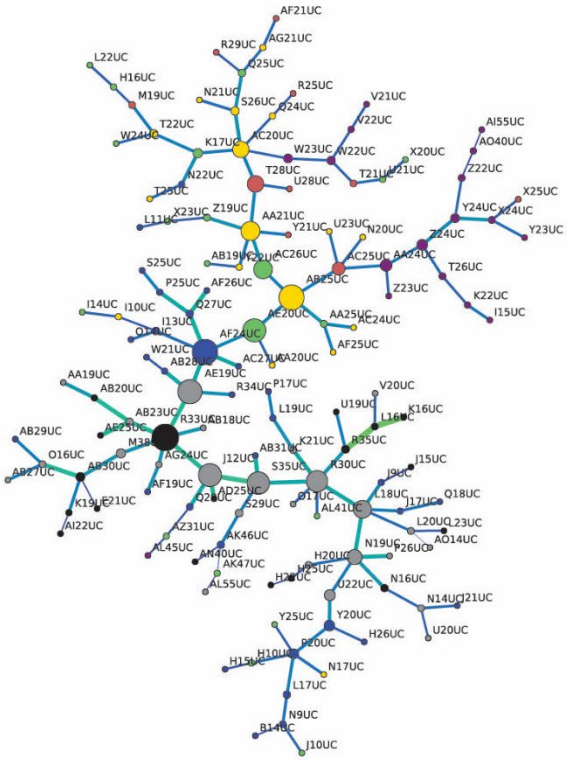

Lower Campanian

B



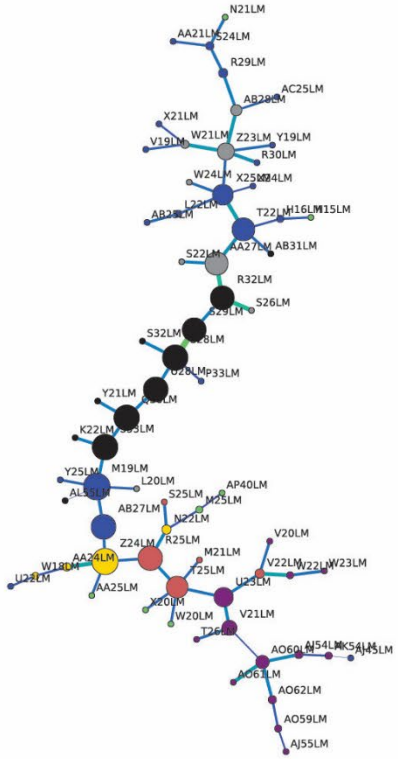

Middle Campanian

C



Upper Campanian

D



Lower Maastrichtian

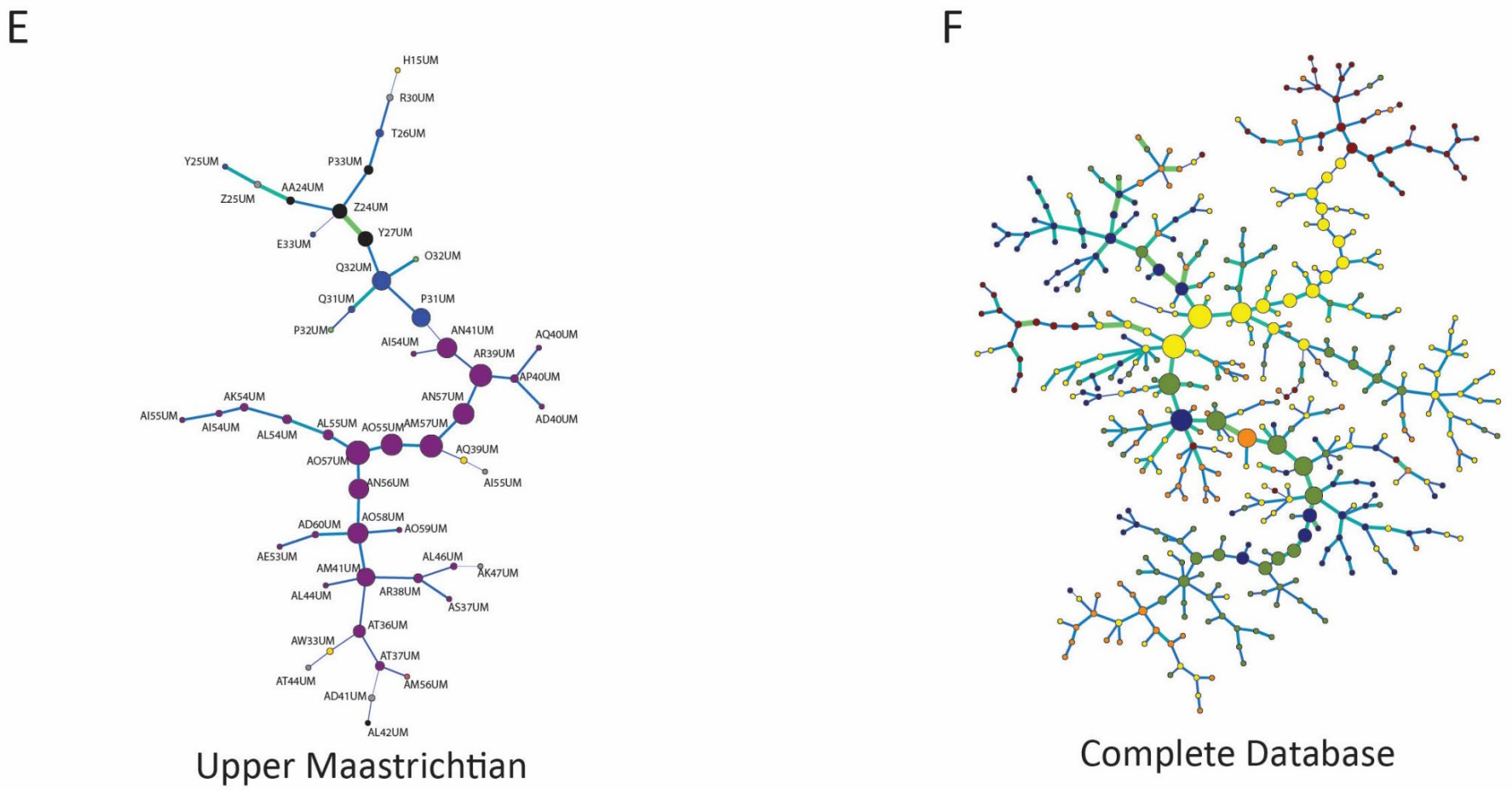



**Figure S29.** Minimum spanning trees for each substage and the collective network. For substages (A-E), node color indicates the number of unique fossil genera present in each node, node size indicates betweenness centrality, and link size and color denotes the weight between two nodes (Bray–Curtis dissimilarity index). See Figures S10-S20 for legends regarding nodes and edges. Nodes for the complete database (F) are colored based on the substage they belong to. Nodes are labeled with their grid cell ID.

**Table S7.** List of the number of links between unique fossil occurrence count bins in the minimum spanning tree of the Lower Campanian. There are substantially more links between nodes of the same bin count and similar bin counts (adjacent bins) than between non-adjacent bin counts.

| Lower Campanian – Minimum Spanning Tree Links | | |
|---|---|---|
| ***Type of Link*** | ***Count*** | ***Percentage*** |
| Same Bin: | 36 | 47.37% |
| Adjacent Bin | 35 | 46.05% |
| Non-Adjacent Bin | 5 | 6.6% |

**Table S8**. List of the number of links between unique fossil occurrence count bins in the minimum spanning tree of the Middle Campanian. There are substantially more links between nodes of the same bin count and similar bin counts (adjacent bins) than between non-adjacent bin counts.

| Middle Campanian – Minimum Spanning Tree Links | | |
|---|---|---|
| ***Type of Link*** | ***Count*** | ***Percentage*** |
| Same Bin: | 54 | 49.54% |
| Adjacent Bin | 45 | 41.28% |
| Non-Adjacent Bin | 10 | 13.16% |

**Table S9.** List of the number of links between unique fossil occurrence count bins in the minimum spanning tree of the Upper Campanian. There are substantially more links between nodes of the same bin count and similar bin counts (adjacent bins) than between non-adjacent bin counts.

| Upper Campanian – Minimum Spanning Tree Links | | |
|---|---|---|
| ***Type of Link*** | ***Count*** | ***Percentage*** |
| Same Bin: | 59 | 43.07% |
| Adjacent Bin | 67 | 48.90% |
| Non-Adjacent Bin | 11 | 8.03% |

**Table S10.** List of the number of links between unique fossil occurrence count bins in the minimum spanning tree of the Lower Maastrichtian. There are substantially more links between nodes of the same bin count and similar bin counts (adjacent bins) than between non-adjacent bin counts.

| Lower Maastrichtian – Minimum Spanning Tree Links | | |
|---|---|---|
| ***Type of Link*** | ***Count*** | ***Percentage*** |
| Same Bin: | 33 | 49.25% |
| Adjacent Bin | 25 | 37.31% |
| Non-Adjacent Bin | 9 | 13.43% |

**Table S11.** List of the number of links between unique fossil occurrence count bins in the minimum spanning tree of the Upper Maastrichtian. There are substantially more links between nodes of the same bin count and similar bin counts (adjacent bins) than between non-adjacent bin counts.

| **Upper Maastrichtian – Minimum Spanning Tree Links** | | |
|---|---|---|
| ***Type of Link*** | ***Count*** | ***Percentage*** |
| Same Bin: | 31 | 63.27% |
| Adjacent Bin | 7 | 14.29% |
| Non-Adjacent Bin | 11 | 22.45% |

**Table S12.** List of the number of links between substages in the minimum spanning tree of the complete database. There are substantially more links between nodes of the same substage, less than half as many between adjacent substages, and less than a third between non-adjacent substages.

| **Complete Database – Minimum Spanning Tree Links** | | |
|---|---|---|
| ***Type of Link*** | ***Count*** | ***Percentage*** |
| Same Substage: | 270 | 61.36% |
| Adjacent Substages: | 109 | 24.77% |
| Non-Adjacent Substages: | 61 | 13.86% |

*Extended Biostratigraphic Chart (Extension of Figure 1):*

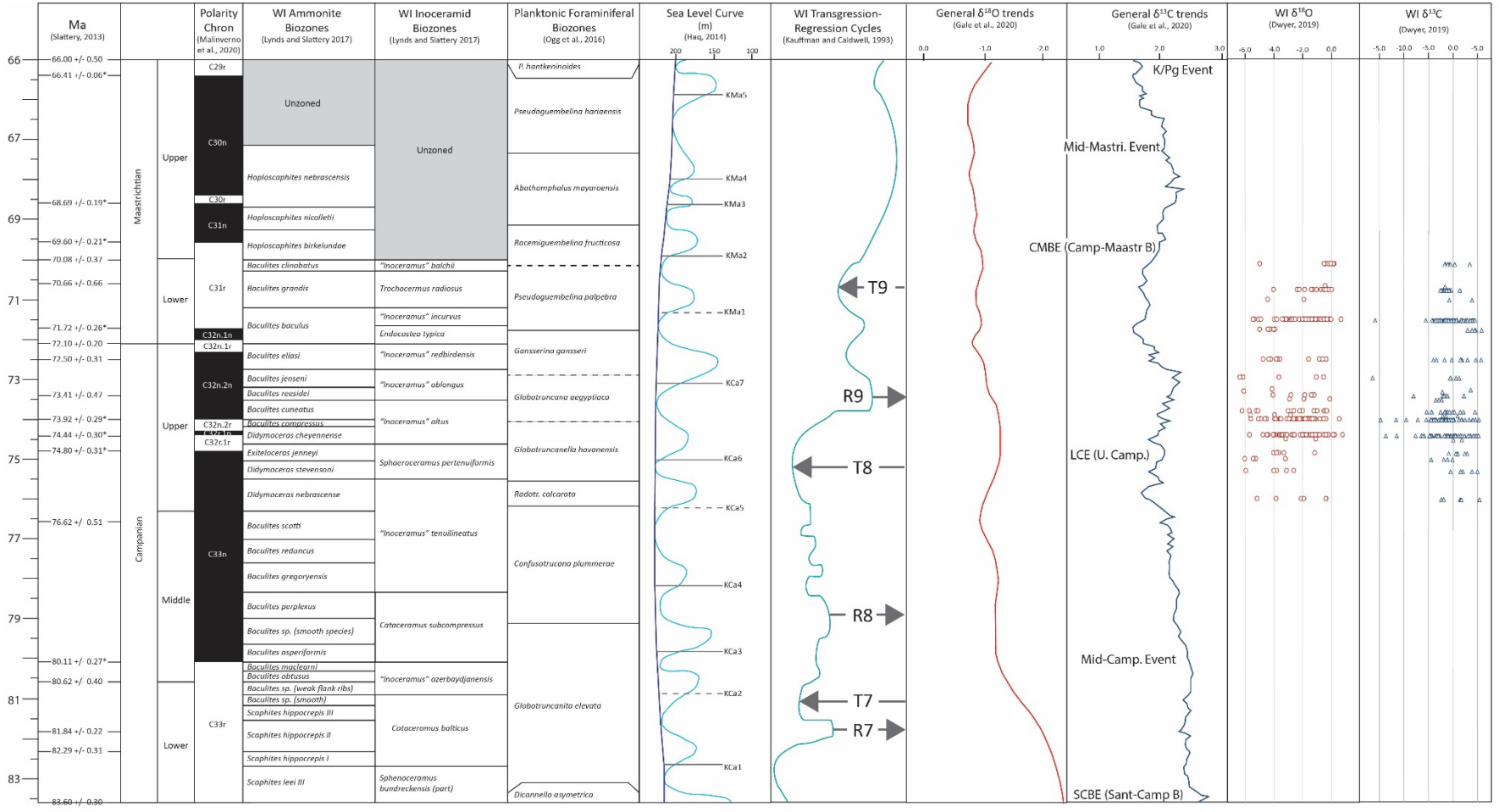



Figure S30. Campanian to Maastrichtian biostratigraphic chart showing polarity chon intervals (Malinverno et al. 2020), biozones (Ogg et al. 2016; Lynds and Slattery 2017), sea level fluctuations (Haq 2014), global stable carbon and oxygen isotope trends (Gale et al. 2020), and stable oxygen and carbon isotope data collated for the Western Interior region (Dwyer, 2019). Ages marked with an asterisk have been added or updated based on Malinverno et al. (2020) from Slattery et al. (2013). WI isotope values modified from unpublished data by Dwyer (2019) which span the R9 regressive interval of the Bearpaw Cyclothem (D. cheyennenses – B. baculus ammonite zones).

SI References:

Bush, A.M., and Bambach, R.K., 2011, Paleoecologic megatrends in marine metazoa: Annual Review of Earth and Planetary Sciences, v. 39, p. 241–269, doi:10.1146/annurev-earth-040809-152556.

Csárdi, G., and Nepusz, T., 2006, The igraph software package for complex network research: InterJournal, v. Complex Systems, p. 1–1695, https://www.researchgate.net/publication/221995787.

Dwyer, C.H., 2019, Constraining the oxygen values of the Late Cretaceous Western Interior Seaway using marine bivalves [Thesis]: University of New Mexico, 1–92 p.

Environmental Systems Research Institute (ESRI), 2021, Redlands, CA, ArcGIS Pro 3.0.

Gale, A.S., Mutterlose, J., Batenburg, S.J., Gradstein, F.M., Agterberg, F.P., Ogg, J.G., and Petrizzo, M.R., 2020, The Cretaceous Period, in Geologic Time Scale 2020, Elsevier, p. 1023–1086, doi:10.1016/B978-0-12-824360-2.00027-9.

Graham, R.L., and Hell, P., 1985, On the History of the Minimum Spanning Tree Problem: Annals of the History of Computing, v. 7, p. 43–57, doi:10.1109/MAHC.1985.10011.

Haq, B.U., 2014, Cretaceous eustasy revisited: Global and Planetary Change, v. 113, p. 44–58, doi:10.1016/j.gloplacha.2013.12.007.

Hijmans, R.J., 2021, geosphere: Spherical Trigonometry: R package version 1.5-14, https://CRAN.R-project.org/package=geosphere, doi:10.1007/s00190012.

Kauffman, E.G., 1984, Paleobiogeography and evolutionary response dynamic in the Cretaceous Western Interior Seaway of North America: Jurassic-Cretaceous Biochronology and Paleogeography of North America, Geological Association of Canada Special Paper, v. 27, p. 273–306.

Kiel, S., 2016, A biogeographic network reveals evolutionary links between deep-sea hydrothermal vent and methane seep faunas: Proceedings of the Royal Society B: Biological Sciences, v. 283, p. 1–9, doi:10.1098/rspb.2016.2337.

Kiel, S., 2017, Using network analysis to trace the evolution of biogeography through geologic time: A case study: Geology, v. 45, p. 711–714, doi:10.1130/G38877.1.

Kivelä, M., Arnaud-Haond, S., and Saramäki, J., 2015, EDENetworks: A user-friendly software to build and analyse networks in biogeography, ecology and population genetics: Molecular Ecology Resources, v. 15, p. 117–122, doi:10.1111/1755-0998.12290.

Kocsis, Á.T., and Raja, N.B., 2020, chronosphere: Earth system history variables:

Lynds, R.M., and Slattery, J.S., 2017, Correlation of the Upper Cretaceous Strataa of Wyoming: Wyoming State Geological Survey Open Files Report 2017-3,.

Malinverno, A., Quigley, K.W., Staro, A., and Dyment, J., 2020, A Late Cretaceous-Eocene geomagnetic polarity timescale (MQSD20) that steadies spreading rates on multiple mid-ocean ridge flanks: Journal of Geophysical Research: Solid Earth, v. 125, p. 1–19.

Mareš, M., 2008, The saga of minimum spanning trees: Computer Science Review, v. 2, p. 165–221, doi:10.1016/j.cosrev.2008.10.002.

Moalic, Y., Desbruyeres, D., Duarte, C.M., Rozenfeld, A.F., Bachraty Charleyne, and Arnaud-Haond, S., 2012, Biogeography revisited with network theory: Retracing the history of hydrothermal vent communities: Systematic Biology, v. 61, p. 127–137.

Myers, C.E., Stigall, A.L., and Lieberman, B.S., 2015, PaleoENM: Applying ecological niche modeling to the fossil record: Paleobiology, v. 41, p. 226–244, doi:10.1017/pab.2014.19.

Newman, M.E.J., 2012, Communities, modules and large-scale structure in networks: Nature Physics, v. 8, p. 25–31, doi:10.1038/nphys2162.

Ogg, J.G. (James G., Ogg, Gabi., and Gradstein, F.M., 2016, A concise geologic time scale 2016: Elsevier.

Opsahl, T., 2009, Structure and Evolution of Weighted Networks [PhD Dissertation]: Queen Mary College University of London, 1–151 p.

Saramäki, J., Kivelä, M., Onnela, J.P., Kaski, K., and Kertesz, J., 2007, Generalizations of the clustering coefficient to weighted complex networks: Physical Review E, v. 75, p. 027105.

Slattery, J.S., Cobban, W.A., Mckinney, K.C., Harries, P.J., and Sandness, A.L., 2013, Early Cretaceous to Paleocene paleogeography of the Western Interior Seaway: The interaction of eustasy and tectonism: Wyoming Geological Association Guidebook, p. 22–60.

Somerfield, P.J., 2008, Identification of the Bray-Curtis similarity index: Comment on Yoshioka (2008): Marine Ecology Progress Series, v. 372, p. 303–306, doi:10.3354/meps07841.

Vilhena, D.A., and Smith, A.B., 2013, Spatial bias in the marine fossil record.: PloS one, v. 8, doi:10.1371/journal.pone.0074470.

Wickham, H., 2016, ggplot2 Elegant Graphics for Data Analysis: Houston, Texas, Springer.

Appendix A-2. Taxonomic References List (See Table S13)

The following list of references were used to vet and update taxonomic nomenclature for the fossil data included in Table S1 (excel file). Table S15 lists the original names provided by the primary databases (i.e., iDigBio, PBDB, etc.), the updated name, if one is necessary, the reference used to check the name, and any additional notes. Taxa were removed if they were not marine invertebrates, identified down to at least the genus level, or if they included "c.f.", "aff.", or "?" following the genus name. Most references were found using basic google searches, but others were researched in specific available volumes (i.e., Akers and Akers, 2002a).

**Taxonomic References**

Aberhan, M., 2004, Early Jurassic Bivalvia of northern Chile. Part II. Subclass Anomalodesmata: Beringeria, v. 34, p. 117–154.

Aberhan, M., 2010, Jurassic bivalves: Taxonomic opinions:

Aberhan, M., and Kiessling, W., 2015, Persistent ecological shifts in marine molluscan assemblages across the end-Cretaceous mass extinction: Proceedings of the National Academy of Sciences of the United States of America, v. 112, p. 7207–7212, doi:10.1073/pnas.1422248112.

Akers, R.E., and Akers, T.J., 1997, Texas Cretaceous gastropods: Houston, Texas, Houston Gem and Mineral Society, 1–340 p.

Akers, R.E., and Akers T J, 2002a, Texas Cretaceous Bivavles 2, Descriptions and illustrations of all named Texas genera and species: Houston Gem and Mineral Society, Paleontology Section, Texas Paleontology Series, v. Publication Number 7, 1–516 p.

Akers, R.E., and Akers T J, 2002b, Texas Cretaceous Bivavles 2, Descriptions and illustrations of all named Texas genera and species: Houston Gem and Mineral Society, Paleontology Section, Texas Paleontology Series, v. Publication Number 7, 1–516 p.

Alencáster, Gloria, O., and Lourdes, 2006, Maastrichtian Inoceramid bivalves from Central Chiapas, Southeastern Mexico: Journal of Paleontology, v. 80, p. 946–957, doi:10.1666/0022.

Alexander, A.N., 1929, Ostracoda of the Cretaceous of North Texas: University of Texas Bulletin, v. 2907, p. 1–137.

Alstine, J.B. van, 1974, Paleontology of brackish-water faunas in two tongues of the Cannonball Formation (Paleocene, Danian), Slope and Golden Valley counties, southwestern North Dakota [Master' s Thesis]: University of North Dakota.

Anderson, F.M., 1945, Knoxville Series in the California Mesozoic: Bulletin of the Geological Society of America, v. 56, p. 909–1014.

Anderson, F.M., 1958a, Upper Cretaceous of the Pacific Coast: The Geological Society of America, Memoir, v. 71, p. 1–378.

Anderson, F.M., 1958b, Upper Cretaceous of the Pacific Coast: The Geological Society of America, Memoir, v. 71, p. 1–378.

Anderson, R.C., Chabaud, A.G., and Willmott, S., 1974a, CIH Keys to the nematode parasites of vertebrates: Commonwealth Agricultural Bureaux.

Anderson, R.C., Chabaud, A.G., and Willmott, S., 1974b, CIH Keys to the nematode parasites of vertebrates: Commonwealth Agricultural Bureaux.

Andrade, C., and Brey, T., 2014, Trophic ecology of limpets among rocky intertidal in Bahia Laredo, Strait of Magellan (Chile): Anales del Instituto de la Patagonia, v. 42, p. 65–70, doi:10.4067/s0718-686x2014000200006.

Angas, G.F., 1879, Descriptions of ten species of marine shells from the province of South Australia: Proceedings of the Zoological Society of London, p. 861–864.

Antia, J., and Fielding, C.R., 2011, Sequence stratigraphy of a condensed low-accommodation succession: Lower upper cretaceous dakota sandstone, henry mountains, southeastern Utah: AAPG Bulletin, v. 95, p. 413–447, doi:10.1306/06301009182.

Arkadiev, V. v, Atabekian, A.A., Baraboshkin, E.Y., and T. N. Bogdanova, 2000, Stratigraphy and ammonites of Cretaceous deposits of south-west Crimea: Palaeontographica Abteilung A, v. 255, p. 85–125.

Arnold, R., 1908, Descriptions of new Cretaceous and Tertiary fossils from the Santa Cruz Mountains, California: Proceedings of the United States National Museum, v. 34, p. 345–389.

Aurele La Rocque, 1960, Molluscan Faunas of the Flagstaff Formation of Central Utah: The Geological Society of America Memoir, v. 78, p. 1–99.

Ayoub-Hannaa, W.S., Bengtson, P., Fürsich, F.T., and Andrade, E.J., 2015, Cenomanian–Coniacian (Upper Cretaceous) bivalves of the Sergipe Basin, Brazil: Order Pholadomyida: Revista Brasileira de Paleontologia, v. 18, p. 31–70, doi:10.4072/rbp.2015.1.03.

Bandel, K., and Dockery, D.T., 2016, Mollusca of the Coon Creek Formation in Tennessee and Mississippi with a Systematic Discussion of the Gastropoda: Bulletin of the Alabama Natural History Museum, v. 33, p. 34–96.

Baron-Szabo, R.C., 2006, Corals of the K/T-boundary: Scleractinian corals of the suborders Astrocoeniina, Faviina, Rhipidogyrina and Amphiastraeina: Journal of Systematic Palaeontology, v. 4, p. 1–108, doi:10.1017/S1477201905001689.

Baron-Szabo, R.C., 2008, Corals of the K/T-boundary : Scleractinian corals of the suborders Dendrophylliina, Carophylliina, Fungiina, Microsolenina, and Stylinina: Zootaxa, p. 1–244.

Barroso-Barcenilla, F., 2007, Revision and new data of the ammonite family Acanthoceratidae de Grossouvre, 1894, from the Lower Turonian of the Iberian Trough, Spain: Palaeontographica Abteilung A, v. 280, p. 123–163.

Bassler, R.S. (Ray S., Berry, E.W., Berry, E.W., Clark, W.B., Gardner, J.A., Goldman, M.I., Pilsbry, H.A., and Stephenson, L.W., 1916, Maryland Geological Survey Upper Cretaceous Text and Plates: Baltimore and The Johns Hopkins Press, v. 2, doi:https://doi.org/10.5962/bhl.title.22326.

Beikirch, D.W., and Feldmann, R.M., 1980, Decapod Crustaceans from the Pflugerville Member, Austin Formation (Late Cretaceous: Campanian) of Texas:, https://www.jstor.org/stable/1304064?seq=1&cid=pdf-.

Bennington, J.B., 2003, Transcending Patchiness in the Comparative Analysis of Paleocommunities: A Test Case from the Upper Cretaceous of New Jersey: PALAIOS, v. 18, p. 22–33, https://www.jstor.org/stable/3515706.

Berry, C.T., 1941, Cretaceous ophiurans from Texas: Journal of Paleontology, v. 15, p. 61–67.

Berry, C.T., 1939, Some fossil Amphineura from the Atlantic coastal plain of North America: Proceedings of the Academy of Natural Sciences of Philadelphia, v. 91, p. 207–217.

Berry, E.W., 1925, The flora of the Ripley formation: U.S. Geological Survey Professional Paper 136, p. 1–94.

Beu, A.G., and Maxwell, P.A., 1987, A revision of the fossil and living gastropods related to Plesiotriton Fischer, 1884 (Family Cancellariidae, Subfamily Plesiotritoninae n. subfam.). With an Appendix: Genera of Buccinidae Pisaniinae related to Colubraria Schumacher, 1817: New Zealand Geological Survey Palaeontological Bulletin, v. 54, p. 1–140.

Bickel, D., 1973, Nonmarine molluscs and two new marine species of Pisidium from the Tongue River Formation (Paleocene), North Dakota: The Nautilus, v. 87, p. 13–21.

Bicknell, R.D.C., Błażejowski, B., Wings, O., Hitij, T., and Botton, M.L., 2021, Critical re-evaluation of Limulidae uncovers limited Limulus diversity: Papers in Palaeontology, v. 7, p. 1525–1556, doi:10.1002/spp2.1352.

Bishop, G.A., 1985, A New Crab, Prehepatus harrisi (Crustacea Decapoda) from the Coon Creek and Prairie Bluff Formations, Union County, Mississippi: Journal of Paleontology, v. 59, p. 1028–1032.

Bishop, G.A., 1982, Homolopsis mendryki: A New Fossil Crab (Crustacea, Decapoda) from the Late Cretaceous Dakoticancer Assemblage, Pierre Shale (Maastrichtian) of South Dakota: Journal of Paleontology, v. 56, p. 221–225, https://about.jstor.org/terms.

Bishop, G.A., 1988, New Fossil Crabs, Plagiophthalmus izetti, Latheticocarcinus shapiroi, And Sagittiformosus carabus (Crustacea, Decapoda), From The Western Interior Cretaceous, USA: Proceedings of the Biological Society of Washington, v. 101, p. 375–381.

Bishop, G.A., 1986, Two new crabs Parapaguristies tuberculatus and Palaeoxantho libertiensis from the Prairie Bluff Formation (Middle Maastrichtian), Union County Mississippi, USA: Proceedings of the Biological Society of Washington, v. 99, p. 604–611.

Bishop, G.A., 1983, Two new species of crabs, Notopocorystes (Eucorystes) Eichhorni and Zygastroc griesi (Decapoda: Brachyura) from the Bearpaw Shale (Campanian) of north-central Montana: Journal of Paleontology, v. 57, p. 900–910.

Bishop, G.A., 1991, Xanthosia occidentalis Bishop, 1985 and Xanthosia spinosa, new speceies, two late Cretaceous crabs from the Pierre Shale of the western interior: Journal of Crustacean Biology, v. 11, p. 305–314.

Bishop, G.A., and Williams, A.B., 2000, Fossil crabs from the Tepee Buttes, submarine seeps of the Late Cretaceous Pierre Shale, South Dakota and Colorado, U.S.A.: Journal of Crustacean Biology, v. 2, p. 286–300, https://academic.oup.com/jcb/article/20/5/286/2419522.

Blake, D.B., and Sprinkle, J., 1996, Astrocratis acutispina, new genus and species, a new asteroid from the Late Cretaceous of Texas: Eclogae Geologicae Helvetiae, v. 89, p. 1311–1320.

Bock, P., 2022, World List of Bryozoa. Hippoporina Neviani, 1895. Accessed through: World Register of Marine Species at https://www.marinespecies.org/aphia.php?p=taxdetails&id=110825 on 2022-11-18:

Bose, E., 1928, Cretaceous ammonites from Texas and northern Mexico.: University of Texas Bulletin, v. 2748, p. 143–312.

Bose, E., 1906, La fauna de moluscos del Senoniano de Cardenas,  San Luis Potosi: Mexico Institute Geologico Boletm, v. 24, p. 1–94.

Bouchet, P., Rocroi, J.-P., Bieler, R., and Carter, J.G., 2010, Nomenclator of Bivalve Families with a Classification of Bivalve Families: Malacologia, v. 52, p. 1–184.

Bouchet, Ph. (Philippe), Rocroi, J.-Pierre., and Frýda, J., 2005, Classification and nomenclator of gastropod families: ConchBooks, 397 p.

Bowles, E., 1939, Eocene and Paleocene Turritellidae of the Atlantic and Gulf Coastal Plain of North America: Journal of Paleontology, v. 13, p. 267–336, https://about.jstor.org/terms.

Boyle, C.B., 1893, Catalogue and Bibliography of North American Mesozoic Invertebrata.:

Braunberger, W.F., and Hall, R.L., 2001, Ammonoid faunas from the Cardium Formation (Turonian-Coniacian, Upper Cretaceous) and contiguous units, Alberta, Canada: I. Scaphitidae: Canadian Journal of Earth Sciences, v. 38, p. 333–346, doi:10.1139/cjes-38-3-333.

Brown, L.G., and Neville, B.D., 2015, Catalog of the recent taxa of the families Epitoniidae and Nystiellidae (Mollusca: Gastropoda) with a bibliography of the descriptive and systematic literature : Zootaxa, v. 3907.

Bryan, J.R., and Jones, D.S., 1989, Fabric of the cretaceous-tertiary marine macrofaunal transition at Braggs, Alabama: Palaeogeography, Palaeoclimatology, Palaeoecology, v. 69, p. 279–301.

Canu, F., and Bassler, R.S., 1929, Bryozoa of the Philippine region (100 Bulletin of the United States National Museum, Ed.): Washington, U.S. G.P.O., v. 9, 1–685 p.

Canu, F., and Bassler, R.S., 1926, Class Bryozoa, *in* The fauna of the Ripley Formation on Coon Creek, Tennessee. U.S. Geological survey Professional Paper, v. 137, p. 32–39.

Canu, F., and Bassler, R.S., 1920, North American early Tertiary Bryozoa: United States National Museum Bulletin, v. 106, p. 1–879.

Chan, B.K.K., Dreyer, N., Gale, A.S., Glenner, H., Ewers-Saucedo, C., Pérez-Losada, M., Kolbasov, G.A., Crandall, K.A., and Høeg, J.T., 2021, The evolutionary diversity of barnacles, with an updated classification of fossil and living forms: Zoological Journal of the Linnean Society, v. 193, p. 789–846, http://zoobank.org/.

Christensen, W.K., 1993, Actinocamax cobbani n. sp. from the Coniacian of Montana and Wyoming and the Occurrence of Late Cretaceous Belemnites in North America and Greenland: Journal of Paleontology, v. 67, p. 434–446, https://about.jstor.org/terms.

Cicimurri, D.J., Bell, G.L.Jr., and Stoffer, P.W., 1999, Vertebrate Paleontology of the Pierre Shale and Fox Hills Formations (Late Campanian-Late Maastrichtian) of Badlands National Park, South Dakota, *in* Santucci, V.L. and McClelland, L. eds., National Park Service Paleontological Research, v. 4, p. 1–7.

Clark, W.B., 1916, The Upper Cretaceous deposits of Maryland: Upper Cretaceous Volume,: Maryland Geological Survey,.

Clark, W.B., and Martin, G.C., 1901, MoUuscoidea: Brachiopods: Systematic Paleontology, Eocene, Maryland Geological Survey, Eocene, p. 203–204.

Clark, W.B., and Twitchell, M.W., 1915, The Mesozoic and Cenozoic Echinodermata of the United States:

Clarke, F.W., and Washington, H.S., 1924, The composition of the Earth’s crust: Professional Paper, v. 127, p. 1–127.

Coan, E. v, and Valentich-Scott, P., 2012, Bivalve seashells of tropical West America: Marine bivalve mollusks from Baja California to northern Peru, v. 1–2, p. 1–1259.

Cobban, W.A., 1972, New and little-known ammonites from the Upper Cretaceous (Cenomanian and Turonian) of the western interior of the United States: United States Geological Survey Professional Paper, v. 699, p. 1–24.

Cobban, W.A., 1977a, A new curved Baculite from the Upper Cretaceous of Wyoming: Journal Research U.S. Geol. Survey, v. 5, p. 457–462.

Cobban, W., 2016, A survey of the Cretaceous ammonite Placenticeras Meek, 1876, in the United States Western Interior, with notes on the earliest species from Texas: Acta Geologica Polonica, v. 66, p. 587–608, doi:10.1515/agp-2016-0031.

Cobban, W.A., 1974, Ammonites from the Navesink Formation at Atlantic Highlands, New Jersey: Geological Survey Professional Paper, v. 845.

Cobban, W.A., 1962a, Baculites from the Lower Part of the Pierre Shale and Equivalent Rocks in the Western Interior:, https://about.jstor.org/terms.

Cobban, W.A., 1952, Cenomanian Ammonite Fauna from the Mosby Sandstone of Central Montana: Geological Survey Professional Paper, v. 243-D, p. 1–55.

Cobban, W.A., 1977b, Characteristic Marine molluscan Fossils from the Dakota Sandstone and Intertongued Mancos Shale, West-Central New Mexico: Geological Survey Professional Paper 1009.:

Cobban, W.A., 1984a, Mid-Cretaceous ammonite zones, Western Interior, United States.: Bulletin - Geological Society of Denmark, v. 33, p. 71–89, doi:10.37570/bgsd-1984-33-06.

Cobban, W.A., 1983, Molluscan fossil record from the northeastern part of the Upper Cretaceous seaway, Western Interior: United States Geological Survey Professional Paper, v. 1253, p. 1–25.

Cobban, W.A., 1984b, Molluscan Record from a Mid-Cretaceous Borehole in Weston County, Wyoming: Geological Survey Professional Paper, v. 1271, p. 1–24.

Cobban, W.A., 1962b, New Baculites from the Bearpaw Shale and Equivalent Rocks of the Western Interior: Journal of Paleontology, v. 36, p. 126–135, https://www.jstor.org/stable/1301240.

Cobban, W.A., 1951, New Species of Baculites from the Upper Cretaceous of Montana and South Dakota: Journal of Paleontology, v. 25, p. 817–821, https://about.jstor.org/terms.

Cobban, W.A., 1990, Sciponoceras gracile (Shumard) Gommon Upper Cretaceous guide fossil in New Mexico: New Mexico Geology, p. 90–91.

Cobban, W.A., 1961, The Ammonite Family Binneyitidae Reeside in the Western Interior of the United States Author.:

Cobban, W.A., 1969, The Late Cretaceous Ammonites Scaphites leei Reeside and Scaphites hippocrepis (DeKay) in the Western Interior of the United States: Geological Survey Professional Paper, v. 619, p. 1–29.

Cobban, W.A., 1964, The Late Cretaceous Cephalopod Haresiceras Reeside and Its Possible Origin: Geological Survey Professional Paper 454-I: Shorter Contributions to General Geology, p. 1–33.

Cobban, 1987, The Upper Cretaceous ammonite Rhaeboceras Meek in the western interior of the United States: United States Geological Survey Professional Paper, p. 1–15.

Cobban, W.A., Dyman, T.A., Pollock, G.L., Takahashi, K.I., Davis, L.E., and Riggin, D.B., 2000, Inventory of dominantly marine and brackish-water fossils from Late Cretaceous rocks in and near Grand Staircase-Escalante National Monument, Utah, *in* Sprinkel, D.A., Chidsey, T.C.Jr., and Anderson, P.B. eds., Geology of Utah's Parks and Monuments, Utah Geological Association Publication, v. 28, p. 579–589.

Cobban, W.A., Dyman, T.S., and Porter, K.W., 2005, Paleontology and stratigraphy of upper Coniacian–middle Santonian ammonite zones and application to erosion surfaces and marine transgressive strata in Montana and Alberta: Cretaceous Research, v. 26, p. 429–449.

Cobban, W.A., Erdmann, C.E., Lemke, R.W., and Maughan, E.K., 1976, Type Sections and Stratigraphy of the Metnbers of the Blackleaf and Marias River Formations (Cretaceous) of the Sweetgrass Arch, Montana: Geological Survey Professional Paper 974, p. 1–66.

Cobban, W.A., and Hook, S.C., 1983a, Mid-Cretaceous (Turonian) ammonite fauna from Fence Lake area of west-central New Mexico: New Mexico Bureau of Mines and Mineral Resources Memoir, v. 40, p. 1–50.

Cobban, W.A., and Hook, S.C., 1983b, Mid-Cretaceous (Turonian) ammonite fauna from Fence Lake area of west-central New Mexico: New Mexico Bureau of Mines and Mineral Resources Memoir, v. 40, p. 1–50.

Cobban, W.A., and Hook, S.C., 1980, The Upper Cretaceous (Turonian) Ammonite Family Coilopoceratidae Hyatt in the Western Interior of the United States: Geological Survey Professional Paper, v. 1192, p. 1–28.

Cobban, W.A., Hook, S.C., and Kennedy, W.J., 1989, Upper Cretaceous rocks and ammonite faunas of southwestern New Mexico: New Mexico Bureau of Mines & Mineral Resources, 1–70 p.

Cobban, W.A., and Jeletzky, J.A., 1965, A New Scaphite from the Campanian Rocks of the Western Interior of North America: Journal of Paleontology, v. 39, p. 794–801, https://about.jstor.org/terms.

Cobban, W.A., and Kennedy, W.J., 1992a, Campanian Trachyscaphites springer ammonite fauna in North-East Texas: Paleontology, v. 35, p. 63–93.

Cobban, W.A., and Kennedy, W.J., 1991, Chapter C: Baculites thomi Reeside, 1927, an Upper Cretaceous Ammonite in the Western Interior of the United States, *in* U.S. Geological Survey Bulletin 1934: Shorter Contributions to Paleontology and Stratigraphy,.

Cobban, W.A., and Kennedy, W.J., 1995, Maastrichtian Ammonites Chiefly from the Prairie Bluff Chalk in Alabama and Mississippi: Journal of Paleontology, v. 69, p. 1–40, https://about.jstor.org/terms.

Cobban, W.A., and Kennedy, W.J., 1992b, The Last Western Interior Baculites from the Fox Hills Formation of South Dakota: Journal of Paleontology, v. 66, p. 682–684, https://www.jstor.org/stable/1305854.

Cobban, W.A., Kennedy, W.J., and Landman, N.H., 1999a, Platyscaphites, a new ammonite from the Lower Campanian (Upper Cretaceous) of the United States Western Interior: Bulletin de L'Institut Royal Des Sciences Naturelles de Belgique , v. 69-Supp, p. 47–54.

Cobban, W.A., Kennedy, W.J., and Landman, N.H., 1999b, Platyscaphites, a new ammonite from the Lower Campanian (Upper Cretaceous) of the United States Western Interior.:

Cobban, W.A., and Merewether, E.A., 1983, Stratigraphy and Paleontology of Mid-Cretaceous Rocks in Minnesota and Contiguous Areas: Geological Survey Professional Paper 1253,.

Cobban, W.A., and Scott, G.R., 1964a, Multinodose Scaphitid Cephalopods from the Lower Part of the Pierre Shale and Equivalent Rocks in the Conterminous United States: Geological Survey Professional Paper 483-E, p. E1–E13.

Cobban, W.A., and Scott, G.R., 1964b, Multinodose Scaphitid Cephalopods from the Lower Part of the Pierre Shale and Equivalent Rocks in the Conterminous United States: Geological Survey Professional Paper 483-E, p. 1–29.

Cobban, W.A., and Scott, G.R., 1972, Stratigraphy and ammonite fauna of the Graneros Shale and Greenhorn Limestone near Pueblo, Colorado: U.S. Geological Survey Professional Paper 645, p. 1–41.

Cobban, W.A., Scott, G.R., and Mckelvey, V.E., 1972, Stratigraphy and Ammonite Fauna of the Graneros Shale and Greenhorn Limestone Near Pueblo, Colorado: Geological Survey Professional Paper 645, p. 1–108.

Collins, J.S., 1986, A new Stramentum (Cirripedia) from the Lower Turonian of Nigeria: Bulletin of the British Museum of Natural History, v. 40, p. 125–131.

Collins, J.S.H., and Mellen, F.F., 1973, Cirripedes from the Upper Cretaceous of Alabama and Mississippi, Eastern Gulf Region, U.S.A.: Bulletin of the British Museum of Natural History, v. 23, p. 351–388.

Cooke, C.W., 1953, American Upper Cretaceous Echinoidea.:

Cooper, G.A., 1973, Fossil and recent Cancellothyridacea (Brachiopoda).:

Cooper, D.A., Cooper, R.W., Stevens, J.B., Stevens, M.S., Cobban, W.A., and Walaszczyk, I., 2017, The Boquillas Formation of the Big Bend National Park, Texas, USA, a reference Cenomanian through Santonian (Upper Cretaceous) carbonate succession at the southern end of the Western Interior Seaway: Acta Geologica Polonica, v. 67, p. 547–565, doi:10.1515/agp-2017-0033.

Correa, M.L., Freiwald, A., Hall-Spencer, J., and Taviani, M., 2005, Distribution and habitats of Acesta excavata (Bivalvia: Limidae) with new data on its shell ultrastructure, *in* Freiwald A. and

Roberts, J.M. eds., Cold-Water Corals and Ecosystems, Berlin Heidelberg, Springer-Verlag, p. 173–205, doi:10.1007/3-540-27673-4_9.

Coryell, H.N., and Salmon, E.S., 1934, A molluscan faunule from the Pierre Formation in Eastern Montana.:

Cummings, K.S., and Graf, D.L., 2010, Mollusca: Bivalvia, *in* Thorp, J.H. and Covich, A.P. eds., Ecology and Classification of North American Freshwater Invertebrates, Amsterdam, The Netherlands, Elsevier.

Cunha, C.M., and Salvador, R.B., 2018, Type specimens of fossil "Architectibranchia" and Cephalaspidea (Mollusca, Heterobranchia) in the Academy of Natural Sciences of Philadelphia: Zoosystematics and Evolution, v. 94, p. 505–527, doi:10.3897/ZSE.94.27401.

Cushman, J.A., 1921, American species of Operculina and Heterostegina and their faunal relations: US Geological Survey Professional Paper, v. 128-E, p. 125–137.

Cvancara, A.M., 1966, Revision of the faunal of the Cannonball Formation (Paleocene) of North and South Dakota: Contributions of the Museum of Paleontology, v. XX, p. 1–97.

Darragh, T.A., 1970, Catalogue of Australian Tertiary Mollusca (Except Chitons): Memoirs of the National Museum of Victoria, v. 31, p. 125–212.

DeVries, T.J., 2007, Cenozoic Turritellidae (Gastropoda) from Southern Peru: Journal of Paleontology, v. 81, p. 331–351, doi:10.1666/0022.

Dhondt, A. v, 1973, Systematic Revision of the Subfamily Neitheinae (Pectinidae, Bivalvia, Mollusca) of the European Cretaceous: Institut Royal Des Sciences Naturalles De Belgique Memoire, v. 176, p. 1–100.

Dhondt, A. v, Walaszczyk, I., Tchegliakova, N., and Jaillard, E., 2004, What is Inoceramus peruanus Bruggen, 1910? Acta Geological Polonica, v. 54, p. 535–359.

Dochev, D., 2015, Turonian (Upper Cretaceous) inoceramid bivalves of the genus Mytiloides from the Sredna Gora Mountains, north-western Bulgaria: Acta Geologica Polonica, v. 65, p. 101–120, doi:10.1515/agp-2015-0004.

Dockery III, D.T., 1993, The Streptoneuran gastropods, exclusive of the Stenoglossa, of the Coffee Sand (Campanian) of Northeastern Mississippi.:

d'Orbigny, A., 1850, Prodrome de Paléontologie stratigraphique universelle des animaux mollusques et rayonnés: v. 2, 1–427 p.

Dunker, W., 1845, Diagnoses Molluscorum quorundam novorum, quae ex itinere ad oras Africae occidentalis reportavit cl. G. Tams, Med. Dr.: Zeitschrift für Malakozoologie,.

Dushane, H., 1977, A New Species of Amaea (Scalina) from the Pliocene of Baja California Sur, Mexico (Mollusca: Gastropoda):, https://www.jstor.org/stable/1303766.

Dyni, J.R., and Cullins, H.L., 1965, Meeker and Loyd Sandstone Members of the Mancos Shale Moffat and Rio Blanco Counties, Colorado: Geological Survey Bulletin 1164-J, p. J1–J7.

Elder, W.P., 1991, An unusual Late Cretaceous fauna from an oyster-rich interval in the Santa Cruz Mountains of California: United States Geological Survey Bulletin 1934, p. E1–E18.

Elder, W.P., 1996, Bivalves and Gastropods from the Middle Campanian Anacacho Limestone, South Central Texas: Journal of Paleontology, v. 70, p. 247–271, https://www.jstor.org/stable/1306388.

Elder, W.P., 1989, Molluscan Extinction Patterns Across the Cenomanian-Turonian Stage Boundary in the Western Interior of the United States: Paleobiology, v. 15, p. 299–320, https://www.jstor.org/stable/2400859.

Elder, W.P., 1994, Some macrofossils from the Cretaceous (Campanian) Anacacho Limestone of Texas: US Geological Survey Open File Report 94-551, p. 1–22.

Elder, W.P., and Box, S.E., 1992, Late Cretaceous inoceramid bivalves of the Kuskokwim Basin, southwestern Alaska, and their implications for basin evolution: Paleontological Society Memoir, v. 26, p. 1264–1269.

Elias, M.K., 1933, Cephalopods of Pierre Formation of Wallace County, Kansas, and Adjacent Area: The University of Kansas Science Bulletin, v. XXI.

Emerson, B.L., Emerson, J.H., Akers, R.E., and Akers, T.J., 1994, Texas Cretaceous Ammonites and Nautiloids: Houston Gem and Mineral Society, Texas Paleontology Series, v. 5, 1–439 p.

Evans, J., and Shumard, B.F., 1854, Descriptions of new fossil species from the Cretaceous formation of Sage Creek, Nebraska, collected by the North Pacific Railroad Expedition under Gov. J.J. Stevens: Proceedings of the Academy of Natural Sciences of Philadelphia, v. 7, p. 163–164.

Evans, J., and Shumard, B.F., 1857, On some new species of fossils from Cretaceous formations of Nebraska Territory: Transactions of the Academy of Natural Science of Saint Louis, v. 1, p. 38–42.

Fauchald, K., 2007, World Register of Polychaeta: http://www.marinespecies.org/polychaeta,.

Feldmann, R.M., Schweitzer, C.E., Baltzly, L.M., Bennett, O.A., Jones, A.R., Mathias, F.F., Weaver, K.L., and Yost, S.L., 2013, New and previously known decapod crustaceans from the Late Cretaceous of New Jersey and Delaware, USA: Bulletin of the Mizunami Fossil Museum, p. 7–37.

Feldmann, R.M., Schweitzer, C.E., Hu, S., Zhang, Q., Zhou, C., Xie, T., Huang, J., and Wen, W., 2012, Macrurous Decapoda from the Luoping Biota (Middle Triassic) of China: Journal of Paleontology, v. 86, p. 425–441, doi:10.1666/11-113.1.

Feldmann, R.M., Schweitzer, C.E., and Wahl, W.R., 2008, Ekalakia (Decapoda: Brachyura): The Preservation of Eyes Links Cretaceous Crabs to Jurassic Ancestors:, https://www.jstor.org/stable/20144266.

Fiege, K., 1930, Über die Inoceramen des Oberturon mit besonderer Berücksichtigung der im Rheinland und Westfalen vorkommenden Formen: Palaeontographica, v. 73, p. 31–48.

Filkorn, H.F., Avendaño-Gil, J., Coutiño-José, M.A., and Vega-Vera, F.J., 2005, Corals from the Upper Cretaceous (Maastrichtian) Ocozocoautla Formation, Chiapas, Mexico: Revista Mexicana de Ciencias Geologicas, v. 22, p. 115–128.

Fisher, S.P., 1952, The Geology of Emmons County North Dakota: North Dakota Geological Survey Bulletin, v. 26, p. 1–36.

la Fon, N.A., 1981, Offshore Bar Deposits of Semilla Sandstone Member of Mancos Shale (Upper Cretaceous), San Juan Basin, New Mexico: The American Association of Petroleum Geologists Bulletin, v. 65, p. 706–721, http://pubs.geoscienceworld.org/aapgbull/article-pdf/65/4/706/4469148/aapg_1981_0065_0004_0706.pdf?casa_token=baV8oEqeMacAAAAA:bYEkMOJJc33ex3S1p4rVWGYxp6e-SCykFnmhdMjyVC.

Foster, W.J., Garvie, C.L., Weiss, A.M., Muscente, A.D., Aberhan, M., Counts, J.W., and Martindale, R.C., 2020, Resilience of marine invertebrate communities during the early Cenozoic hyperthermals: Scientific Reports, v. 10, doi:10.1038/s41598-020-58986-5.

Fox, J.E., 2007, Mollusks from the late Campanian upper DeGrey Formation of the Pierre Shale Group, Missouri River Valley, central South Dakota, *in* Martin, J.E. and Parris, D.C. eds., The Geology and Paleontology of the Late Cretaceous Marine Deposits of the Dakotas, Geologic Society of America, v. 427.

Francis, A., Francis, R., Bertrand, M., and Christian, D., 2016, A North American ammonite fauna from the late Middle Turonian of Vaucluse and Gard, southern France: Acta Geologica Polonica, v. 66, p. 729–736.

Franţescu, O.E., Feldmann, F.M., and Schweitzer, C.E., 2016, Cretaceous fossil Raninoida De Haan, 1839 (Crustacea, Decapoda, Brachyura) from northeast Texas: Journal of Paleontology, v. 90, p. 1118–1132.

Fursich, F.T., 1993, Palaeoecology and evolution of Mesozoic salinity-controlled benthic macroinvertebrate associations: Lethia, v. 26, p. 327–346.

Fursich, F.T., and Kirkland, J.I., 1986, Biostratinomy and paleoecology of a Cretaceous brackish lagoon: PALAIOS, v. 1, p. 543–560.

Gabb, W.M., 1859, Catalogue of the Invertebrate Fossils of the Cretaceous Formation of the United States, with References: Proceedings of the Academy of Natural Sciences of Philadelphia, v. 11, p. 1–20.

Gabb, W.M., 1876, Notes on American Cretaceous fossils with descriptions of some new species: Proceedings of the Academy of Natural Sciences of Philadelphia for 1876, p. 276–324.

Gabb, W.M., 1861, Synopsis of the Mollusca of the Cretaceous Formation, Including the Geographical and Stratigraphical Range and Synonymy: Proceedings of the American Philosophical Society, v. 8, p. 57–257.

Gardner, J.A., 1916, Mollusca, *in* Clark, W.B. ed., Upper Cretaceous, Baltimore, Maryland Geological Survey, p. 371–733.

Gardner, J.A., 1947, The molluscan fauna of the Alum Bluff Group of Florida. Part VIII. Ctenobranchia (remainder, Aspidobranchia, and Scaphopoda): United States Geological Survey Professional Paper, v. 142-H, p. 493–656.

Garvie, C.L., 2021, The macrofauna of the Tehuacana Limestone Member (Danian, Kincaid Formation) of Central Texas, with the Description of a few new taxa from the Pisgah Member: Bulletins of American Paleontology, v. 399–400, p. 1–222, doi:10.32857/bap.2021.399.01.

Gill, J.R., and Cobban, W.A., 1973, Stratigraphy and Geologic History of the Montana Group and Equivalent Rocks, Montana, Wyoming, and North and South Dakota.:

Gill, J.R., and Cobban, W.A., 1966a, The Red Bird Section of the Upper Cretaceous Pierre Shale in Wyoming: United States Government Printing Office, Washington.

Gill, J.R., and Cobban, W.A., 1966b, The Red Bird Section of the Upper Cretaceous Pierre Shale in Wyoming: Geological Survey Professional Paper 393-A.:

Gill, J.R., Merewether, E.A., and Cobban, W.A., 1970, Stratigraphy and Nomenclature of Some Upper Cretaceous and Lower Tertiary Rocks in South-Central Wyoming: Geological Survey Professional Paper 667.:

Glaubrecht, M., Brinkmann, N., and Pöppe, J., 2009, Diversity and disparity “down under”: Systematics, biogeography and reproductive modes of the “marsupial” freshwater Thiaridae (Caenogastropoda, Cerithioidea) in Australia: Zoosystematics and Evolution, v. 85, p. 199–275, doi:10.1002/zoos.200900004.

Goldfuss, G.A., 1835, Petrefacta Germaniae tam ea, quae in Museo Universitatis Regiae Borussicae Fridericiae Wilhelmiae Rhenanae servantur quam alia quaecunque in Museis Hoeninghusiano, Muensteriano aliisque extant, iconibus et descriptionibus illustrata. Zweiter Teil: Düsseldorf, Arnz & Comp, 1–312 p.

Gordon, D.P., Sutherland, J.E., Perez, B.A., Waeschenbach, A., Taylor, P.D., and di Martino, E., 2020, The bryozoan genus Conopeum (Electridae) in New Zealand, with description of a new species and discussion of the morphological and genetic characters of Conopeum seurati (Canu, 1928): Journal of Natural History, v. 54, p. 947–970.

Gordon, D.P., Taylor, P.D., and Bigey, F.P., 2009, Phylum Bryozoa: moss animals, sea mats, lace corals, *in* Gordon, D.P. ed., New Zealand inventory of biodiversity: 1. Kingdom Animalia: Radiata, Lophotrochozoa, Deuterostomia, p. 271–297.

Götz, S., and Mitchell, S., 2009, The Laluzia armini (gen et spec nov) ecosystem: understanding a deeper-water rudist lithosome from the Early Maastrichtian of Mexico: Facies, v. 55, p. 539–551.

Grier, J.C., and Grier, J.W., 1998, Paleontological Society New Findings of the Ammonite Rhaeboceras, including a New Species, from the Pierre Shale of Eastern Montana: Journal of Paleontology, v. 72, p. 473–476, https://about.jstor.org/terms.

Grier, J.W., Grier, J.C., Larson, N.L., and Petersen, J.G., 2007, Synonomy of the ammonite genus Ponteixites Warren with Rhaeboceras Meek: Rocky Mountain Geology, v. 42, p. 123–136.

Griffitts, M.O., 1949, A New Rudistid from the Niobrara of Colorado:, https://www.jstor.org/stable/1299472.

Groves, L.T., and Squires, R.L., 2018, Annotated catalog of the fossil invertebrates described by, and named for, William More Gabb (1839–1878): Zootaxa, v. 4534, p. 1–150, doi:10.11646/zootaxa.4534.1.1.

Gunther, P.R., and Hills, L. v, 1972, Megaspores and Other Palynomorphs of the Brazeau Formation (Upper Cretaceous), Nordegg Area of the Nordegg Area, Alberta: Proceedings of the Annual Meetings. American Association of Stratigraphic Palynologist, v. 3, p. 29–48.

Guschick, R.C., and Rodriguez, J., 1967, Brachiopod zonation and correlation of Sappington Formation of Western Montana: The American Association of Petroleum Geologists Bulletin, v. 51, p. 601–620, http://pubs.geoscienceworld.org/aapgbull/article-pdf/51/4/601/4399183/aapg_1967_0051_0004_0601.pdf.

Hall, J., and Meek, F.B., 1856, Descriptions of new species of fossils from the Cretaceous formations of Nebraska, with observations upon Baculites ovatus and B. compressus, and the progressive development of the septa in Baculites, Ammonites, and Scaphites: American Academy of Arts and Sciences Memoir, new series, v. 5, p. 379–411.

Hanley, J.H., and Flores, R.M., 1987, Taphonomy and Paleoecology of nonmarine Mollusca: Indicators of alluvial plain lacustrine sedimentation, upper part of the Tongue River Member, Fort Union Formation (Paleocene), Northern Powder River Basin, Wyoming and Montana: PALAIOS, v. 2, p. 479–496.

Hansson, H.G., 2001, Echinodermata, *in* Costello, M.J. and et al. eds., European register of marine species: a check-list of the marine species in Europe and a bibliography of guides to their identification, Collection Patrimoines Naturels, p. 336–351.

Harbison, A., 1945, Upper Cretaceous Mollusks of the Lower Ripley Formation near Dumas, Mississippi.:

Harries, P.J., and Kauffman, E.G., 1990, Patterns of survival and recovery following the Cenomanian-Turonian (Late Cretaceous) mass extinction in the Western Interior Basin, United States, *in* Kauffman, E.G. and Walliser, O.H. eds., Extinction Events in Earth History. Lecture Notes in Earth Sciences, Berlin, Springer, v. 30.

Harries, P.J., and Schopf, K.M., 2007, Late Cretaceous gastropod drilling intensities: Data from the Maastrichtian Fox Hills Formation, Western Interior Seaway, USA: Palaios, v. 22, p. 35–46, doi:10.2110/palo.2005.p05-016r.

Hartman, J.H., 1981, Early Tertiary nonmarine Mollusca of New Mexico: A review: Geological Society of America Bulletin, v. 92, p. 942–950.

Hartman, J.H., 2015a, New viviparid gastropods from the end Cretaceous and early Paleogene of the Williston Basin, USA and Canada: The Nautilus, v. 129, p. 1–22, https://www.biodiversitylibrary.org/item/274716.

Hartman, J.H., 2015b, New viviparid gastropods from the end Cretaceous and early Paleogene of the Williston Basin, USA and Canada: The Nautilus, v. 129, p. 1–22, https://www.biodiversitylibrary.org/item/274716.

Hartman, J.H., and Kirkland, J.I., 2002, Brackish and marine mollusks of the Hell Creek Formation of North Dakota: Evidence for a persisting Cretaceous seaway, *in* Hartman, J.H., Johnson, K.R., and Nichols, D.J. eds., The Hell Creek Formation and the Cretaceous-Tertiary boundary in the northern Great Plains: An Integrated continental record of the end of the Cretaceous, Geological Society of America, v. 361.

Hartman, J.H., and Roth, B., 1998, Late Paleocene and early Eocene nonmarine molluscan faunal change in the Bighorn Basin, northwestern Wyoming and south- central Montana, *in* Aubry, M.P., Lucas, S.G., and Berggren, W.A. eds., Late Paleocene-Early Eocene Climatic and Biotic Events in the Marine and Terrestrial Records, Columbia University Press, p. 323–379.

Hartmann, J.D.W., 1821, System der Erd- und Flußschnecken der Schweiz. Mit vergleichender Aufzählung aller auch in den benachbarten Ländern, Deutschland, Frankreich und Italien sich vorfindenden Arten, *in* Neue Alpina, Winterthur, A writing of the Swiss natural history, dedicated to Alpine and agriculture, v. 1, p. 194–268.

Hasenmueller, W.A., and Hattin, D.E., 1990, New Species of the Bivalve Anomia from Lower and Middle Turonian Parts of the Greenhorn Limestone, Central Kansas: Journal of Paleontology, v. 64, p. 104–110.

Hass, O., 1943, Some abnormally coiled ammonites from the Upper Cretaceous of Angola: American Museum Novitates, v. 1222, p. 1–17.

Hattin, D.E., 1986, Carbonate Substrates of the Late Cretaceous Sea, Central Great Plains and Southern Rocky Mountains:, https://www.jstor.org/stable/3514473?seq=1&cid=pdf-.

Hattin, D.E., and Hirt, D.S., 1991, Paleoecology of Scalpellomorph Cirripeds in the Fairport Member, Carlile Shale (Middle Turonian), of Central Kansas: PALAIOS, v. 6, p. 553–563, https://about.jstor.org/terms.

Hayden, F.V., 1876, Volume II: Bulletin of the United States Geological and Geographical Survey of the Territories, 1876: U.S. Geological Survey Bulletin, v. 222, p. 1–392.

Hayward, P.J., 2001, Bryozoa, *in* Costello, M.J., Emblow, C., and White, R. eds., European register of marine species: a check-list of the marine species in Europe and a bibliography of guides to their identification, Collection Patrimoines Naturels, v. 50, p. 325–333.

Heidelberger, D., 2001, Mitteldevonische (Givetische) Gastropoden (Mollusca) aus der Lahnmulde (suedliches Rheinisches Schiefergebirge): Geologische Abhandlungen Hessen, v. 106, p. 1–291.

Henderson, J., 1918, A mollusk hunt in Wyoming: The Nautilus, v. XXXII.

Henderson, J., 1935, Fossil non-marine Mollusca of North America: Geological Society of America Special Papers No. 3, p. 1–313.

Henderson, R.A., 1990, Late albian ammonites from the northern territory, Australia: Alcheringa, v. 14, p. 109–148, doi:10.1080/03115519008527815.

Henderson, J., 1908, New species of Cretaceous invertebrates from northern Colorado: Proceedings of the United States National Museum, v. 34, p. 259–264.

Henderson, J., 1934, Some New Mesozoic Mollusca from the Rocky Mountain Region: Journal of Paleontology, v. 8, p. 259–263.

Hendy, A.J.W., Jones, D.S., Moreno, F., Zapata, V., and Jaramillo, C., 2015, Neogene molluscs, shallow marine paleoenvironments, and chronostratigraphy of the Guajira Peninsula, Colombia: Swiss Journal of Palaeontology, v. 134, p. 45–75, doi:10.1007/s13358-015-0074-1.

Hinde, G.J., 1890, Notes on the palaeontology of Western Australia 2. Corals and Polyzoa: Geological Magazine, New Series, Decade III, v. 7, p. 194–204.

Holland, F.D., Erickson, J.M., and O'brien, D.E., 1975, Casterolimus: A new Late Cretaceous generic link in Limulid lineage.:

Hook, S.C., 2012, Evolution of the Late Cretaceous oyster genus Cameleolopha Vyalov 1936 in central New Mexico: New Mexico Geology, v. 34, p. 76–95.

Hook, S.C., 1983, New Mexico Geological Society Stratigraphy, paleontology, depositional framework, and nomenclature of marine Upper Cretaceous rocks, Socorro County, New Mexico, *in* Chaplin, C.E. and Callender, J.F. eds., New Mexico Geological Society Guidebook, 34th Field Conference, Socorro Region II, p. 165–172, http://nmgs.nmt.edu/publications/guidebooks/34.

Hook, S.C., 2011, New Mexico Geology The Late Cretaceous oyster Cameleolopha bellaplicata (Shumard 1860), guide fossil to middle Turonian strata in New Mexico: New Mexico Geology, v. 33, p. 67–95.

Hook, S.C., 2013, The Upper Cretaceous (Turonian) Juana Lopez Beds of the D-Cross Tongue of the Mancos Shale in central New Mexico and their relationship to the Juana Lopez Member of the Mancos Shale in the San Juan Basin: New Mexico Geology, v. 35, p. 59–81.

Hook, S.C., and Cobban, W.A., 1981, Lopha Sannionis- common Upper Cretaceous guide fossil in New Mexico: New Mexico Bureau of Mines and Mineral Resources, Annual Report, July 1, 1979 to June 30, 1980, p. 52–56.

Hook, S.C., and Cobban, W.A., 1977, Pycnodonte newberryi (Stanton)—common guide fossil in Upper Cretaceous of New Mexico: New Mexico Bureau of Mines and Mineral Resources, Annual Report (July 1, 1976 to June 30, 1977), p. 48–54.

Hren, M.T., Sheldon, N.D., Grimes, S.T., Collinson, M.E., Hooker, J.J., Bugler, M., and Lohmann, K.C., 2013a, Terrestrial cooling in Northern Europe during the Eocene-Oligocene transition:

Proceedings of the National Academy of Sciences of the United States of America, v. 110, p. 7562–7567, doi:10.1073/pnas.1210930110.

Hren, M.T., Sheldon, N.D., Grimes, S.T., Collinson, M.E., Hooker, J.J., Bugler, M., and Lohmann, K.C., 2013b, Terrestrial cooling in Northern Europe during the Eocene-Oligocene transition: Proceedings of the National Academy of Sciences of the United States of America, v. 110, p. 7562–7567, doi:10.1073/pnas.1210930110.

Huber, M., 2010, Compendium of bivalves. A full-color guide to 3,300 of the World's Marine Bivalves. A status on Bivalvia after 250 years of research: Hackenheim, ConchBooks.

Huber, M., Langleit, A., and Kreipl, K., 2015, Tellinidae, *in* Huber, M. ed., Compendium of bivalves 2, Harxheim, ConchBooks.

Hunt, A.P., and Lucas, S.G., 2022a, Stratigraphy, paleontology and age of the Fruitland and Kirtland Formations (Upper Cretaceous), San Juan Basin, New Mexico, *in* Lucas, S.G., Kues, B.S., Williamson, T.E., and Hunt, A.P. eds., New Mexico Geological Society43 rd Annual Fall Field Conference Guidebook, New Mexico Geological Society, p. 217–239, doi:10.56577/ffc-43.217.

Hunt, A.P., and Lucas, S.G., 2022b, Stratigraphy, paleontology and age of the Fruitland and Kirtland Formations (Upper Cretaceous), San Juan Basin, New Mexico, *in* Lucas, S.G., Kues, B.S., Williamson, T.E., and Hunt, A.P. eds., New Mexico Geological Society43 rd Annual Fall Field Conference Guidebook, New Mexico Geological Society, p. 217–239, doi:10.56577/ffc-43.217.

Hyatt, A., 1894, Phylogeny of an Acquired Characteristic: The American Naturalist, v. 27, p. 865–877.

Hyžný, M., and Klompmaker, A.A., 2015, Systematics, phylogeny, and taphonomy of ghost shrimps (Decapoda): a perspective from the fossil record: Arthropod Systematic Phylogeny, v. 73, p. 401–437.

Ifrim, C., and Stinnesbeck, W., 2007, Early Turonian ammonites from Vallecillo, north-eastern Mexico: taxonomy, biostratigraphy and palaeobiogeographical significance: Cretaceous Research, v. 28, p. 642–664, doi:10.1016/j.cretres.2006.10.004.

Ifrim, C., Stinnesbeck, W., Espinosa, B., and Ventura, J.F., 2015, Upper Campanian (Upper Cretaceous) cephalopods from the Parras Shale near Saucedas, Coahuila, Mexico: Journal of South American Earth Sciences, v. 64, p. 229–257, doi:10.1016/j.jsames.2015.07.013.

Ifrim, C., Stinnesbeck, W., and Ventura, J.F., 2013, An endemic cephalopod assemblage from the lower Campanian (Late Cretaceous) Parras Shale, western Coahuila, Mexico: Journal of Paleontology, v. 87, p. 881–901, doi:10.1666/12-123.

Ikuno, K., and Hirano, H., 2015, Nomenclatural review of Polyptychoceras and 18 related taxa (Ammonoidea: Diplomoceratidae): Swiss Journal of Palaeontology, v. 134, p. 227–232, doi:10.1007/s13358-015-0081-2.

Jäger, M., and Jäger, M., 2012, Sabellids and serpulids (Polychaeta sedentaria) from the type Maastrichtian, the Netherlands and Belgium.:

Johnson, J.H., 1930, The Benton Fauna of Eastern Colorado and Kansas and Its Recorded Geologic Range.:

Johnson, D.W., 1904, The Geology of the Cerrillos Hills New Mexico [Dissertation]: Columbia University, 1–204 p.

Johnson, P.A., and Hendy, J.W., 2005, Paleoecology of Mollusks from the Upper Cretaceous Belly River Group, *in* Currie, P.J. and Koppelhus, E.B. eds., Dinosaur Provincial Park: A Spectacular Ancient Ecosystem Revealed, Bloomington and Indianapolis, Indiana University Press, p. 139–162.

Johnston, P.A., and Hendy, A.J.W., 2005, Paleoecology of Mollusks from the Upper Cretaceous Belly River Group, *in* Currie, P.J. and Koppelhus, E.B. eds., Dinosaur Provincial Park: a spectacular ancient ecosystem revealed, Bloomington, Indiana University Press, p. 139–166.

Kafanov, A.I., and Ogasawara, K., 2005, Cenozoic molluscan (Bivavia) cenozone of Hokkaido, Northern Japan.:

Karasawa, H., Schweitzer, C.E., Feldmann, R.M., and Luque, J., 2014, Phylogeny and classification of raninoida (decapoda: Brachyura): Journal of Crustacean Biology, v. 34, p. 216–272, doi:10.1163/1937240X-00002216.

Kauffman, E.G., 1967, Cretaceous thyasira from the western interior of North America: Smithsonian Miscellaneous Collections, v. 152.

Kauffman, E.G., 1975, Dispersal and Biostratigraphic Potential of Cretaceous Benthonic Bivalvia in the Western Interior: The Cretaceous System in the Western Interior of North America: Geological Association of Canada Special Paper, p. 163–194.

Kauffman, E.G., 1978, Evolutionary Rates and Patterns Among Cretaceous Bivalvia: Philosophical Transactions of the Royal Society London B, v. 284, p. 277–304, https://www.jstor.org/stable/2418161?seq=1&cid=pdf-.

Kauffman, E.G., Sageman, B., Kirkland, J.I., Elder, W.P., Harries, P.J., and Villamil, T., 1993, Molluscan biostratigraphy of the Cretaceous Western Interior Basin, North America, *in* Caldwell, W.G.E. and Kauffman, E.G. eds., Evolution of the Western Interior Basin, Geological Association of Canada Special Paper, v. 39, p. 397–434, https://www.researchgate.net/publication/230892263.

Kennedy, W.J., 1986, Campanian and Maastrichtian ammonites from northern Aquitaine, France: London, Paleontological Association.

Kennedy, W.J., 1988, Late Cenomanian and Turonian Ammonite faunas from north-east and central Texas: Special Papers in Paleontology, p. 1–131.

Kennedy, W.J., 1984, Systematic Palaeontology and Stratigraphic Distribution of the Ammonite Faunas of the French Coniacian: Special Papers in Palaeontology, v. 31, p. 1–160.

Kennedy, W.J., 1995, Systematic palaeontology, in Ammonite faunas, biostratigraphy, and sequence stratigraphy of the Coniacian-Santonian of the Corbières (NE Pyrénées): Bulletin des Centres de Recherches Elf Exploration-Production Elf-Aquitaine, v. 19, p. 377–499.

Kennedy, W.J., 1999, Systematic paleontology, in Maastrichtian ammonites from Balochistan, Pakistan: Journal of Paleontology, v. 73, p. 641–662.

Kennedy, W.J., and Bilotte, M., 1995, A new ammonite fauna from the sub-Pyrenean Campanian (Upper Cretaceous): Geobios, v. 28, p. 359–370.

Kennedy, W.J., and Cobban, W.A., 1993, Campanian Ammonites from the Annona Chalk near Yancy, Arkansas: Journal of Paleontology, v. 67, p. 83–97, https://www.jstor.org/stable/1305965.

Kennedy, W.J., and Cobban, W.A., 1999, Campanian (Late Cretaceous) ammonites from the Bergstrom Formation in Central Texas: Acta Geologica Polonica, v. 49, p. 67–80.

Kennedy, W.J., and Cobban, W.A., 1991, Coniacian ammonite faunas from the United States Western Interior: Special Papers in Paleontology, v. 45, p. 1–96.

Kennedy, W.J., and Cobban, W.A., 1997, Paleontological Society Upper Campanian (Upper Cretaceous) Ammonites from the Marshalltown Formation-Mount Laurel Boundary Beds in Delaware: Journal of Paleontology, v. 71, p. 62–73, https://www.jstor.org/stable/1306542.

Kennedy, W.J., and Cobban, W.A., 1994, Upper Campanian Ammonites from the Mount Laurel Sand at Biggs Farm, Delaware:, https://www.jstor.org/stable/1306231?seq=1&cid=pdf-.

Kennedy, W.J., Cobban, W.A., and Hook, S.C., 1988, Middle Cenomanian (Late Cretaceous) molluscan fauna from the base of the Boquillas Formation, Cerro de Muleros, Dona Ana County, New Mexico: New Mexico Bureau of Mines & Mineral Resources Bulletin, v. 114, p. 35–44.

Kennedy, W.J., Cobban, W.A., and Landman, N.H., 2001, A revision of the Turonian members of the ammonite subfamily Collignoniceratinae from the United States Western Interior and Gulf Coast: Bulletin American Museum of Natural History, p. 148.

Kennedy, W.J., Cobban, W.A., and Landman, N.H., 1997, New Ammonoid Records from the Merchantville Formation (Upper Cretaceous) of Maryland and New Jersey: American Museum Novitates, p. 1–17.

Kennedy, W.J., Cobban, W.A., and Landman, N.H., 1999, The Heteromorph Ammonite Didymoceras cochleatum (Meek and Hayden, 1858), from the Pierre Shale of South Dakota and Wyoming: American Museum Novitates, v. 3268, p. 1–8.

Kennedy, W.J., Cobban, W.A., and Scott, G.R., 2000a, Heteromorph ammonites from the middle Campanian Baculites scotti Zone in the U.S. Western Interior: Acta Geologica Polonica, v. 50, p. 223–241.

Kennedy, W.J., Cobban, W.A., and Scott, G.R., 2000b, Heteromorph ammonites from the Upper Campanian (Upper Cretaceous) Baculites cuneatus and Baculites reesidei zones of the Pierre Shale in Colorado, USA: Acta Geologica Polonica, v. 50, p. 1–20.

Kennedy, W.J., Crame, J.A., Bengtson, P., and Thomson, M.R.A., 2007, Coniacian ammonites from James Ross Island, Antarctica: Cretaceous Research, v. 28, p. 509–531, doi:10.1016/j.cretres.2006.08.006.

Kennedy, W.J., Kauffman, E.G., and Klinger, H.C., 1973, Upper Cretaceous Invertebrate Faunas from Durban, South Africa: Geological Society of South Africa Transactions, v. 76, p. 95–111.

Kennedy, W.J., and Klinger, H.C., 2013, Scaphitid ammonites from the Upper Cretaceous of KwaZulu-Natal and Eastern Cape Province, South Africa: Acta Geologica Polonica, v. 63, p. 527–543.

Kennedy, W.J., Landman, N.H., Cobban, W.A., and Scott, G.R., 2000c, Late Campanian (Cretaceous) heteromorph ammonites from the Western Interior of the United States: Bulletin of the AMNH, p. 1–88.

Kennedy, W.J., and Simmons, M.D., 1991, Mid-Cretaceous ammonites and associated microfossils from the Central Oman Mountains: Newsletter on Stratigraphy, v. 25, p. 127–154.

Kennedy, W.J., and Summesberger, H., 2000, Additional ammonites from the Upper Campanian (Upper Cretaceous) of the Gschliefgraben (Ultrahelvetic; Austria): Annalen des Naturhistorischen Museums in Wien, Serie B, v. 102, p. 85–107, https://www.jstor.org/stable/41702010.

Keyes, C.R., 1894, Missouri Geological Survey: Paleontology of Missouri (Part II): Jefferson City, Tribune Printing Company, v. V, 1–266 p.

Khakimov, F.K., 1976, Taxonomic names: Kampanskie i Maastrikhtskie ammonity sredney Azii, p. 1–146.

Kier, P.M., and Lawson, M.H., 1978, Index of Living and Fossil Echinoids 1924-1970: Smithsonian Contributions to Paleobiology, v. 34, p. 1–182.

Kirkland, J.I., 1996, Bulletin of the New Mexico Museum of Natural History and Science: Paleontology of the Greenhorn Cyclothem (Cretaceous: Late Cenomanian to Middle Turonian) at Black Mesa, Northeastern Arizona (S. G. Lucas, Ed.): Albuquerque, NM, Bulletin of the New Mexico of Natural History and Science, 1–131 p.

Klinger, H.C., and Kennedy, W.J., 2001, Stratigraphic and geographic distribution, phylogenetic trends and general comments on the Ammonite Family Baculitidae Gill, 1871 (with an annotated list of species referred to the family): Annals of the South African Museum, v. 107, p. 1–290.

Knight, J.B., 1932, The Gastropods of the St. Louis, Missouri, Pennsylvanian Outlier: IV. The Pseudomelaniidae: Journal of Paleontology, v. 6, p. 189–202, https://www.jstor.org/stable/1298174.

Koch, C.F., 1980, Bivalve Species Duration, Areal Extent and Population Size in a Cretaceous Sea:, https://about.jstor.org/terms.

Kolesnikov, Ch.M., 1980, System, stratigraphic distribution and zoogeography of Mesozoic limnological bivalvian Mollusca in the USSR, *in* Martinson, G.G. ed., Limnobios drevnikh ozernykh basseynov Yevrazii, Leningrad, Nauka, p. 9–65.

Kosnik, M.A., 2005, Changes in Late Cretaceous-Early Tertiary Benthic Marine Assemblages: Analyses from the North American Coastal Plain Shallow Shelf: Paleobiology, v. 31, p. 459–479, https://www.jstor.org/stable/4096947.

Kosnik, M.A., 2002, personal opinion:

Kříž, J., and Steinová, M., 2009, Uppermost ordovician bivalves from the Prague Basin (Hirnantian, Perunica, Bohemia): Bulletin of Geosciences, v. 84, p. 409–436, doi:10.3140/bull.geosci.1141.

Kuehne, D.F., 1999, Upper Cretaceous macroinvertebrate faunas of the northern Atlantic Coastal Plain: The Mosasaur, v. 6, p. 29–80.

Kummel, B., 1954, Jurassic nautiloids from western North America: Journal of Paleontology, v. 28, p. 320–324.

Kuzmicheva, E.I., 1987, Upper Cretaceous and Paleogene corals from the USSR: 1–187 p.

Landau, B., Marques da Silva, C., and Gilli, C., 2009, The Early Pliocene Gastropoda (Mollusca) of Estepona, Southern Spain. Part 8: Nassariidae: Palaeontos, v. 17, p. 1–101.

Landis, E.R., Dane, C.H., and Cobban, W.A., 1973, Stratigraphic terminology of the Dakota Sandstone and Mancos Shale, West-central New Mexico: Geological Survey Bulletin 1327-J, p. J1–J44.

Landman, N.H., Johnson, R.O., Garb, M.P., Edwards, L.E., and Kyte, F.T., 2007, Cephalopods from the Cretaceous/Tertiary boundary interval on the Atlantic Coastal Plain, with a description of the highest ammonite zone in North America. Part II. Manasquan River Basin, Monmouth County, New Jersey: Bulletin of the American Museum of Natural History , p. 1–122.

Landman, N.H., Kennedy, W.J., Cobban, W.A., and Larson, N.L., 2010, Scaphites of the "Nodosus Group" from the Upper Cretaceous (Campanian) of the Western Interior of North America: Bulletin of the American Museum of Natural History, p. 1–242.

Landman, N.H., Kennedy, W.J., Cobban, W.A., Larson, N.L., and Jorgensen, S.D., 2013, A new species of Hoploscaphites (Ammonoidea: Ancyloceratina) from cold methane seeps in the Upper Cretaceous of the U.S. Western Interior: American Museum Noveitates, v. 3781, p. 1–39.

Landman, N.H., Kennedy, W.J., Grier, J., Larson, N.L., Grier, J.W., Linn, T., Tackett, L., and Jicha, B.R., 2020, Large Scaphitid ammonites (Hoploscaphites) from the Upper Cretaceous (Upper Campanian-Lower Maastrichtian) of North America: Endless variation on a single theme: Bulletin of the American Museum of Natural History, p. 1–131.

Landman, N.H., Kennedy, W.J., and Larson, N.L., 2015, A new species of scaphitid ammonite from the lower Maastrichtian of the Western Interior of North America, with close affinities to Hoploscaphites constrictus Sowerby, 1817: American Museum Novitates, v. 3833, p. 1–40.

Landman, N.H., Rye, D.M., and Shelton, K.L., 1983, Early ontogeny of Eutrephoceras compared to Recent Nautilus and Mesozoic ammonites: evidence from shell morphology and light stable isotopes: Paleobiology, v. 9, p. 269–279.

Landman, N.H., Slattery, J.S., and Harries, P.J., 2016, Encrustation of inarticulate brachiopods on scaphitid ammonites and inoceramid bivalves from the Upper Cretaceous U. S. Western Interior: Acta Geologica Polonica, v. 66, p. 645–662, https://scholarcommons.usf.edu/geo_facpub/1083.

Landman, N.H., and Waage, K.M., 1993, Scaphitid ammonites of the Upper Cretaceous (Maastrichtian) Fox Hills Formation in South Dakota and Wyoming: Bulletin of the American Museum of Natural History, v. 215, p. 257.

Larina, E., Garb, M., Landman, N., Dastas, N., Thibault, N., Edwards, L., Phillips, G., Rovelli, R., Myers, C., and Naujokaityte, J., 2016, Upper Maastrichtian ammonite biostratigraphy of the Gulf Coastal Plain (Mississippi Embayment, southern USA): Cretaceous Research, v. 60, p. 128–151, doi:10.1016/j.cretres.2015.11.010.

Larson, N.L., 2010, Fossil coleoids from the Late Cretaceous (Campanian & Maastrichtian) of the Western Interior: Ferrantia, v. 59, p. 78–113, https://www.researchgate.net/publication/309456036.

Larson, N.L., 2012, The Late Campanian (Upper Cretaceous) cephalopod fauna of the Coon Creek Formation at the type locality: The Journal of Paleontological Sciences, p. 1–40.

Larson, N.L., Jorgensen, S.D., Farrar, R.A., and Larson, P.L., 1997, Ammonites and the Other Cephalopods of the Pierre Seaway: Geoscience Press, Inc., 1–96 p.

Levey, R.A., 1985, Depositional Model for Understanding Geometry of Cretaceous Coals: Major Coal Seams, Rock Springs Formation, Green River Basin, Wyoming: The American Association of Petroleum Geologist Bulletin, v. 69, p. 1359–1380, http://pubs.geoscienceworld.org/aapgbull/article-pdf/69/9/1359/4689649/aapg_1985_0069_0009_1359.pdf?casa_token=qETWJSOqlHsAAAAA:ruUxoedyacupJFfcGhXC7GGBqcpJUchT8de1Ijz95qTJ.

Linnert, C., Robinson, S.A., Lees, J.A., Bown, P.R., Perez-Rodriguez, I., Petrizzo, M.R., Falzoni, F., Littler, K., Artz, J.A., and Russell, E.E., 2014, Evidence for global cooling in the Late Cretaceous: Nature Communications, v. 5, p. 1–7.

Lockwood, R., 2003, Abundance not linked to survival across the end-Cretaceous mass extinction: Patterns in North American bivalves: PNAS, v. 100, p. 2478–2482, www.pnas.org.

Ludbrook, N.H., 1967, Tertiary Molluscan types from Table Cape in the Tasmanian Museum, Hobart: Papers and Proceedings of the Royal Society of Tasmania, v. 101, p. 65–69.

Luque, J., Schweitzer, C.E., Santana, W., Portell, R.W., Vega, F.J., and Klompmaker, A.A., 2017, Checklist of fossil decapod crustaceans from tropical America. Part I: Anomura and Brachyura: Nauplius, v. 25, doi:10.1590/2358-2936e2017025.

Malchus, N., 1990, Revision der Kreide-Austern (Bivalvia: Pteriomorphia) Ägyptens (Biostratigraphie, Systematik): Berliner Geowissenschaftliche Abhandlungen, v. A 125.

Manso, C.L.C., and Andrade, E.J., 2008, Equinoides do Turoniano (Cretaceo Superior) de Sergipe, Brasil: São Paulo, UNESP, Geociencias, v. 27, p. 319–327.

Marcou, J.B., 1885, A list of the Mesozoic and Cenozoic types in the Collections of the U.S. National Museum.:

Marquet, R., 2002, The Neogene Amphineura and Bivalvia (Protobranchia and Pteriomorphia) from Kallo and Doel (Oost Vlaanderen, Belgium): Palaentos, v. 2, p. 1–100.

Martinelli, J.C., Lopes, H.M., Hauser, L., Jimenez-Hidalgo, I., King, T.L., Padilla-Gamiño, J.L., Rawson, P., Spencer, L.H., Williams, J.D., and Wood, C.L., 2020, Confirmation of the shell-boring oyster parasite Polydora websteri (Polychaeta: Spionidae) in Washington State, USA: Scientific Reports, v. 10, doi:10.1038/s41598-020-60805-w.

Marwick, J., 1953, Divisions and faunas of the Hokonui System (Triassic and Jurassic): New Zealand Geological Survey Paleontological Bulletin, v. 21.

Matúš, H., and Adiël, A.K., 2015, Systematics, phylogeny, and taphonomy of ghost shrimps (Decapoda): a perspective from the fossil record: Arthropod Systematic Phylogeny, v. 73, p. 401–437.

McArthur, J.M., 1994, Recent trends in strontium isotope stratigraphy: Terra Nova, v. 6, p. 331–358.

McArthur, J.M., Kennedy, W.J., Chen, M., Thirlwall, M.F., and Gale, A.S., 1994, Strontium isotope stratigraphy for Late Cretaceous time: Direct numerical calibration of the Sr isotope curve based on the US Western Interior: Palaeogeography, Palaeoclimatology, Palaeoecology, v. 108, p. 95–119.

McLearn, F.H., 1926, Contributions to Canadian Palaeontology: New species from the Coloradoan of Lower Smokey and Lower Peace Rivers, Alberta: Canada Geological Survey, v. 42, p. 117–126.

McLearn, F.H., 1929, Mesozoic Palaeontology of Blairmore Region: Cretaceous invertebrates: National Museum of Canada Bulletin, v. 58, p. 73–79.

Meek, F.B., 1876, A report on the invertebrate Cretaceous and Tertiary fossils of the upper Missouri country: United States Geological Survey of the Territories Report, v. 9, p. 1–629.

Meek, F.B., 1864, Checklist of Invertebrate Fossils of North America. Cretaceous and Jurassic: Smithsonian Miscellaneous Collections , v. 177.

Meek, F.B., 1860, Descriptions of new fossil remains collected in Nebraska and Utah, by the exploring expeditions under the command of Capt. J.H. Simpson of U.S. Topographical Engineers: Proceedings of the Academy of Natural Sciences of Philadelphia, v. 12, p. 308–315.

Meek, F.B., 1871, Preliminary paleontological report, consisting of lists of fossils, with descriptions of some new types, etc.: U.S. Geological Survey of Wyoming (Hayden), Preliminary Report, v. 4, p. 287–318.

Meek, F.B., 1873, Preliminary paleontological report, consisting of lists and descriptions of fossils, with remarks on the ages of the rocks in which they were found, etc.: U.S. Geological Survey of the Territories (Hayden), Sixth Annual Report, p. 429–518.

Meek, F.B., Hall, J., Witfield, R.P., and Ridgeway, R., 1877, Part I. Palaeontology: Washington, Govt. print. off.

Meek, F.B., and Hayden, F. v, 1857, Description of new Species and GeneraofFossils, collected by Dr. F. V. Hayden in Nebraska Territory, under the direction of Lieut. G. K. Warren, U.S. Topographical Engineer; with some remarks on the Tertiary and Cretaceous formations of the north-west, and the parallelism of the latter with those of other portions of the United States and Territories: Proceedings of the Academy of Natural Sciences of Philadelphia, v. 9, p. 117–148.

Meek, F.B., and Hayden, F. v, 1861, Descriptions of new Lower Silurian, (Primordial), Jurassic, Cretaceous, and Tertiary fossils, collected in Nebraska, by the exploring expedition under the command of Capt. Wm. F. Reynolds, U.S. Top. Engineers with some remarks on the rocks from which they were obtained: Academy of Natural Sciences of Philadelphia Proceedings, v. 13, p. 415–447.

Meek, F.B., and Hayden, F. v, 1860, Descriptions of new organic remains from the Tertiary, Cretaceous and Jurassic rocks of Nebraska: Academy of Natural Sciences of Philadelphia Proceedings, v. 12, p. 175–185.

Meek, F.B., and Hayden, F. v, 1856, Descriptions of new species of Gasteropoda and Cephalopoda from the Cretaceous formations of Nebraska Territory: Proceedings of the Academy of Natural Sciences of Philadelphia, v. 8, p. 70–72.

Merewether, E.A., and McKinney, K.C., 2015, Composite biostratigraphic outcrop sections for Cretaceous Formations along a South-trending transect from western Montana, western Wyoming, eastern Utah, northeastern Arizona, and northwestern New Mexico, U.S.A.: U.S. Geological Survey Open-File Report 2015-1087, p. 1–10, doi:10.3133/ofr20151087.

Merrill, G.P., 1905, Catalogue of the Type and Figured Specimens of Fossils, Minerals, Rocks, and Ores in the Department of Geology, United States Museum: Bulletin of the United States Museum, p. 1–704, doi:https://doi.org/10.5479/si.03629236.53.i.

Miller, H.W., 1968, Invertebrate Fauna and Environment of Deposition of the Niobrara Formation (Cretaceous) of Kansas: Fort Hays Studies Series, v. 62, https://scholars.fhsu.edu/fort_hays_studies_series.

Mitchell, S.F., 2013, A revision of selected Lower Cretaceous American caprinoid rudists: implications for phylogeny and biostratigraphy.:

Mitchell, S.F., 2010a, Morphology, taxonomy and lifestyle of the Maastrichtian rudist bivalve Thyrastylon: Turkish Journal of Earth Sciences, v. 19, p. 635–642.

Mitchell, S.F., 2010b, Revision of three large species of Barrettia from Jamaica Taxonomy of Rudist Bivalves (Bivalvia: Hippuritida) View project Geology and geological evolution of Jamaica View project Revision of three large species of Barrettia from Jamaica:, https://www.researchgate.net/publication/233924620.

Mitchell, S.F., 2005, The oldest barnacle from the Caribbean is a rudist bivalve (Maastrichtian, Jamaica): Cretaceous Research, v. 26, p. 895–897, doi:10.1016/j.cretres.2005.06.002.

Moore, R.C., 1967, Unique stalked crinoids from Upper Cretaceous of Mississippi: The University of Kansas Paleontological Contributions, p. 1–35.

Morton, S.G., 1834, Synopisis of the organic remains of the Cretaceous group of the United States: Illustrations by nineteen plates to which is added an appendix, containing a tabular view of the Tertiary fossils hitherto discovered in North America: Philadelphia, Key and Biddle, 1–88 p.

Moussavou, B.M., Yakouya-Moubamba, U.G., Eyeghe, A.C.A., and Mamidi, M.A., 2017, Nouvelles données sur les macro-invertébrés du Turonien de la région de Libreville (bassin côtier nord-gabonais): Annales de Paléontologie, v. 103, p. 283–292.

Munster, G. von, 1841, Beschreibung und Abbildung der in den Kalkmergelschichten von St. Cassian gefundenen Versteinerungen. In: H.L. Wissmann and G. von Munster (H. L. Wissmann & G. von Munster, Eds.): Buchner, Bayreuth, v. 4, 1–152 p.

Newton, H., and Jenney, W.P., 1880, Report on the Geology and Resources of the Black Hills of Dakota with Atlas.:

Nickles, J.M., and Bassler, R.S., 1900, A Synopsis of American fossil Bryozoa including Bibliography and synonymy: Bulletin of the United States Geological Survey No. 178, p. 1–663.

Niebuhr, B., Taherpour Khalil Abad, M., Wilmsen, M., Razmi, J.N., Aryaei, A.A., and Ashouri, A., 2016a, First record of late Campanian ammonites from the Abderaz Formation of the Koppeh Dagh, northeastern Iran: Cretaceous Research, v. 58, p. 202–222, doi:10.1016/j.cretres.2015.10.006.

Niebuhr, B., Taherpour Khalil Abad, M., Wilmsen, M., Razmi, J.N., Aryaei, A.A., and Ashouri, A., 2016b, First record of late Campanian ammonites from the Abderaz Formation of the Koppeh Dagh, northeastern Iran: Cretaceous Research, v. 58, p. 202–222, doi:10.1016/j.cretres.2015.10.006.

Noda, M., and Matsumoto, T., 1998, Palaeontology and stratigraphy of the inoceramid species from the mid-Turonian through upper Middle Coniacian in Japan: Acta Geological Polonica, v. 48, p. 435–482.

Nyborg, T., Ossó, À., and Vega, F.J., 2014, A new species of icriocarcinid crab (Crustacea, Portunoidea) from the uppermost Cretaceous of California, USA: palaeobiogeographic implications: Scripta Geology, v. 147, p. 83–93.

Ohta, Y., 1970, A review of some Cretaceous corbiculids in North America: Transactions and Proceedings of the Palaeontological Society of Japan, new series, v. 79, p. 291–315.

Oleinik, A.E., and Marincovich Jr., L., 2003, Biotic response to the Eocene-Oligocene Transition: Gastropod assemblages in the High-Latitude North Pacific, *in* Prothero, D.R., Ivany, L.C., and Nesbitt, E.A. eds., From Greenhouse to Icehouse: The Marine Eocene-Oligocene transition, Columbia University Press, p. 36–56.

Olsson, A.A., 1922, The Miocene of Northern Costa Rica: Bulletins of American Paleontology, v. 9, p. 1–309.

Owen, D.D., 1852, Report of a geological survey of Wisconsin, Iowa, and Minnesota; and incidentally of a portion of Nebraska Territory made under instructions of the United States Treasury Department: Philadelphia, Lippincott, Grambo, v. 2 volumes, 1–638 p.

Özdikmen, H., 2013, Substitute names for three preoccupied generic names in Gastropoda: Munis Entomology & Zoology, v. 8, p. 252–256.

Pan, H.Z., 1977, Mesozoic and Cenozoic fossil Gastropoda from Yunnan: Mesozoic Fossils from Yunnan, v. 2, p. 83–152.

Parodiz, J.J., 1969, The Tertiary non-marine Mollusca of South America: Annals of Carnegie Museum, v. 40, p. 1–242.

Perrilliat, M.D.C., Vega, F.J., Espinosa, B., and Naranjo-Garcia, E., 2008, Late Cretaceous and Paleogene freshwater gastropods from Northeastern Mexico: Journal of Paleontology, v. 82, p. 255–266.

Pojeta, J., and Sohl, N.F., 1987, Ascaulocardium armatum (Morton, 1833), New Genus (Late Cretaceous): The Ultimate Variation on the Bivalve Paradigm: Memoir (The Paleontological Society, https://about.jstor.org/terms.

Pons, J.M., García Barrera, P., Oviedo, A., and Vicens, E., 2021, Mexican Upper Cretaceous rudists (Hippuritida, Bivalvia): Taxonomic, stratigraphic, and geologic data: Journal of South American Earth Sciences, v. 109, doi:10.1016/j.jsames.2021.103237.

Pons, J.M., Vicens, E., and Garcia-Barrera, 2017, Campanian and Maastrichtian plagioptychid rudists (Hippuritida, Bivalvia) of the Chiapas Central Depression, southern Mexico: Journal of Paleontology, v. 91, p. 230–244.

Pons, J.M., Vicens, E., Oviedo, A., Aguilar, J., García-Barrera, P., and Alencáster, G., 2013, The rudist fauna of the Cárdenas Formation, Maastrichtian, San Luis Potosí State, Mexico: Journal of Paleontology, v. 87, p. 726–754, doi:10.1666/12-116.

Pons, J.M., Vicens, E., Pichardo, Y., Aguilar, J., Oviedo, A., Alencáster, G., and Garcia-Barrera, P., 2010, A New Early Campanian Rudist Fauna from San Luis Potosi in Mexico and Its Taxonomic and Stratigraphic Significance: Journal of Paleontology, v. 84, p. 974–995, doi:10.2307/40802062.

Pons, J.M., Vicens, E., and Schmidt-Effing, R., 2016, Campanian rudists (Hippuritida, Bivalvia) from Costa Rica (Central America): Journal of Paleontology, v. 90, p. 211–238, doi:10.1017/jpa.2016.27.

Powell, J.D., 1963, Turonian (Cretaceous) Ammonites from Northeastern Chihuahua, Mexico.:

Prozorova, L., 1988, A new species of the genus Aplexa (Pulmonata, Physidae) from Japan: Zoologicheskiĭ Zhurnal, v. 77, p. 1068–1070.

Rader, L.E., and Grimaldi, F.S., 1961, Chemical analysis for selected minor elements in Pierre Shale: Geological Survey Professional Paper 391-A.:

Rassam, H., Ghamizi, M., Benaissa, H., Clewing, C., and Albrech, C., 2021, The fingernail clams (Bivalvia: Veneroida: Sphaeriidae) of Morocco: Diversity, distributionand conservation status: Biodiversity Data Journal, v. 9, doi:10.3897/BDJ.9.E73346.

Rathbun, M.J., 1926, Crustacea, *in* Wade, B. ed., The fauna of the Ripley Formation of Coon Creek, Tennessee, United States Geological Survey Professional Paper 137, p. 184–191.

Rathbun, M.J., 1935, Fossil Crustacea of the Atlantic and Gulf Coastal Plain: Geological Society of America Special Papers,.

Récluz, C.A., 1844, Monographie du genre Narica: Revue Zoologique, par la Société Cuvierienne, v. 7.

Reeside, J.B.Jr., 1927, Cephalopods from the lower part of the Cody shale of Oregon Basin, Wyoming: U.S. Geological Survey Professional Paper 150A, p. 1–19.

Reeside, J.B., 1928, New Cretaceous mollusks from Colorado and: Journal of the Washington Academy of Sciences, v. 18, p. 306–313.

Rehfeld, U., and Ernst, G., 1998, Hydrozoan build-ups of Millepora irregularis sp. nov. and fungiid coral meadows of Cunnolites Alloiteau (Anthozoa) - palaeoecological and palaeoceanographical implications for the Upper Cretaceous of north Cantabria (northern Spain): Facies, v. 39, p. 125–138.

Reiskind, J., 1975, Marine concretionary faunas of the uppermost Bearpaw Shale (Maestrichtian) in eastern Montana and southwestern Saskatchewan: Geological Association of Canada Special Paper, v. 13, p. 235–252.

Riccardi, A.C., 1983, Scaphitids from the Upper Campanian–Lower Maastrichtian Bearpaw Formation of the Western Interior of Canada: Geological Survey of Canada Bulletin 354, p. 1–51.

Richards, H.G. et al., 1962a, The Cretaceous Fossils of New Jersey: Geological Survey of New Jersey Bulletin 61 Part II, p. 1–345.

Richards, H.G. et al., 1962b, The Cretaceous Fossils of New Jersey: Geological Survey of New Jersey Bulletin 61 Part II, p. 1–345.

Richards, H.G. et al., 1958a, The Cretaceous Fossils of New Jersey. Part 1. Porifera, Coelenterata, Annelida, Echinoidea, Brachipoda, and Pelecypoda: Trenton, N J, State of New Jersey Department of Conservation and Economic Development, 1–266 p.

Richards, H.G., Garner, H.F., Howell, B.F., Jeletzky, J.A., Miller, A.K., Miller Jr., H.W., Ramsdell, R.C., Richards, H.G., Reeside Jr., J.B., and Wells, J.W., 1958b, The Cretaceous Fossils of New Jersey Part 1.:

Richards, H.G., Garner, H.F., Howell, B.F., Jeletzky, J.A., Miller, A.K., Miller Jr., H.W., Ramsdell, R.C., Richards, H.G., Reeside Jr., J.B., and Wells, J.W., 1958c, The Cretaceous Fossils of New Jersey Part 1.:

Richards, H.G., and Shapiro, E., 1963, Delaware Geological Survey Report of Investigations No. 7: An invertebrate macrofauna from the Upper Cretaceous of Delaware.:

Rindsberg, A.K., 2000, Upper Cretaceous Bivalvia of Alabama, published by the author: http:fly.hiwaay.net/ ~dwills/fossils/uppercre.html,.

Rioux, R.L., 1958, Geology of the Spence-Kane Area, Big Horn County, Wyoming [PhD Thesis]: University of Illinois, 1–183 p.

Robinson, C.S., Mapel, W.J., and Cobban, W.A., 1959, Pierre Shale along western and northern flanks of Black Hills, Wyoming and Montana: Bulletin of the American Association of Petroleum Geologists, v. 43, p. 101–123, http://pubs.geoscienceworld.org/aapgbull/article-pdf/43/1/101/4379769/aapg_1959_0043_0001_0101.pdf.

Roehler, H.W., 1979, Geology and Mineral Resources of the Mud Springs Ranch Quadrangle, Sweetwater County, Wyoming: Geological Survey Professional Paper, v. 1065–C, p. C1–C35.

Roemer, F., 1852, Die Kreidebildungen von Texas und ihre organischen Einschlüsse: Bonn, Adolph Marcus, 1–100 p.

Ros-Franch, S., Marquez-Aliaga, A., and Damborenea, S.E., 2014, Comprehensive Database on Induan (Lower Triassic) to Sinemurian (Lower Jurassic) Marine Bivalve Genera and Their Paleobiogeographic Record: Paleontological Contributions, v. 8, p. 1–219.

Rosso, A., 2002, Amphiblestrom Gray, 1848 (Bryozoa Cheilostomatida) from the Atlantic-Mediterranean area, with description of a new species: Journal of Natural History, v. 36, p. 1489–1508.

Roth, Z., and Hartman, J.H., 1998, A Probable Cerion (Gastropoda: Pulmonata) from Uppermost Cretaceous Hell Creek Formation, Garfield County, Montana.:

Russell, L.S., 1964, Cretaceous non-marine faunas of northwestern North America: Royal Ontario Museum, Life Sciences, Contribution No. 61, p. 1–24.

Russell, L.S., 1935, Fauna of the Upper Milk River beds, Southern Alberta: Transactions of the Royal Society of Canada, Series 3, v. 29, p. 115–128.

Russell, L.S., 1926, Mollusca of the Paskapoo Formation in Alberta: Transactions of the Royal Society of Canada, third series. Section IV - Geological Sciences, v. 20, p. 207–220.

Russell, L.S., 1976, Pelecypods of the Hell Creek Formation (Upper Cretaceous) of Garfield County, Montana: Canadian Journal of Earth Science, v. 13, p. 365–388.

Russell, L.S., and Landes, R.W., 1940, Geology of the Southern Alberta Plains. Part II. Paleontology of the marine formations of the Montana Group: Geological Survey of Canada Memoir, v. 221.

Sari, B., Kandemir, R., Özer, S., Walaszczyk, I., Görmüş, M., Demircan, H., and Yilmaz, C., 2014, Upper Campanian calciclastic turbidite sequences from the Hacimehmet area (eastern Pontides, NE

Turkey): Integrated biostratigraphy and microfacies analysis: Acta Geologica Polonica, v. 64, p. 393–418, doi:10.2478/agp-2014-0022.

Sava, L.A., 2007, The molluscan and brachiopod fauna of the Late Cretaceous Pierre Shale (Baculites compressus/ Baculites cuneatus biozones) near Kremmling, Colorado: University of South Florida, http://scholarcommons.usf.edu/etdhttp://scholarcommons.usf.edu/etd/2354.

Schneider, J.A., 1995a, Phylogeny of the Cardiidae (Mollusca, Bivalvia): Protocardiinae, Laevicardiinae, Lahilliinae, Tulongocardiinae subfam. n. and Pleuriocardiinae subfam. n.: Zoologica Scripta, v. 24, p. 321–346.

Schneider, J.A., 1995b, Phylogeny of the Cardiidae (Mollusca, Bivalvia): Protocardiinae, Laevicardiinae, Lahilliinae, Tulongocardiinae subfam. n. and Pleuriocardiinae subfam. n.: Zoologica Scripta, v. 24, p. 321–346.

Schneider, S., Jäger, M., Kroh, A., Mitterer, A., Niebuhr, B., Vodrážka, R., Wilmsen, M., Wood, C.J., and Zágoršek, K., 2013, Silicified sea life - Macrofauna and palaeoecology of the Neuburg Kieselerde Member (Cenomanian to Lower Turonian Wellheim Formation, Bavaria, southern Germany): Acta Geologica Polonica, v. 63, p. 555–610, doi:10.2478/agp-2013-0025.

Schweitzer, C.E., Feldmann, R.M., Fam, J., Hessin, W.A., Hetrick, S.W., Nyborg, T.G., and Ross, R.L.M., 2003, Cretaceous and Eocene Decapod crustaceans from Southern Vancouver Island, British Columbia, Canada: National Research Council of Canada.

Schweitzer, C., Feldmann, R., Garassino, A., Karasawa, H., and Schwigert, G., 2010, Systematic list of fossil decapod crustacean species: Crustacean Monographs, v. 10, p. 1–222.

Schweizerbart, E., 1894, New Yearbook of Mineralogy, Geology and Paleontology, Part 2:

Scott, R.W., 1978, Paleobiology of Comanchean (Cretaceous) Cardiids (Cardiinae), North America:, https://www.jstor.org/stable/1303907.

Scott, G.R., and Cobban, W.A., 1964, Stratigraphy of the Niobrara Formation at Pueblo, Colorado: Geological Survey Professional Paper 454-L, p. L1–L30.

Scott, G.R., Cobban, W.A., and Merewether, E.A., 1986a, New Mexico Bureau of Land Mines and Mineral Resources, Bulletin 115: Stratigraphy of the Upper Cretaceous Niobrara Formation in the Raton basin, New Mexico.:

Scott, G.R., Cobban, W.A., and Merewether, E.A., 1986b, New Mexico Bureau of Mines & Mineral Resources Stratigraphy of the Upper Cretaceous Niobrara Formation in the Raton basin, New Mexico: New Mexico Bureau of Mines & Mineral Resources Bulletin 115, p. 1–34.

Sealey, P.L., and Lucas, S.G., 2022, Baculites Baculus Meek and Hayden, 1861 (Earliest Maastrichtian) from the Uppermost Pierre Shale in the Raton basin of Northeastern New Mexico and its Significance, *in* New Mexico Geological Society, p. 73–80, doi:10.56577/ffc-70.73.

Sealey, P.L., and Lucas, S.G., 2018a, Campanian ammonites and other mollusks from the Lewis Shale, eastern San Juan Basin, Rio Arriba County, New Mexico: New Mexico Museum of Natural History and Science Bulletin, v. 79, p. 603–642.

Sealey, P.L., and Lucas, S.G., 2018b, Campanian ammonites and other mollusks from the Lewis Shale, Eastern San Juan Basin, Rio Arriba County, New Mexico: New Mexico Museum of Natural History and Science, p. 603–642.

Sealey, P.L., and Lucas, S.G., 2011, Upper Cretaceous (Turonian) ammonites from the Carlile Member and Reference are of the Juana Lopez Member of the Mancos Shale, Eastern side of the San Juan Basin, La Ventana, Sandoval County, New Mexico: New Mexico Museum of Natural History and Science, Bulletin, v. 53, p. 37\0-379.

Sepkoski Jr., J.J., 2002, A Compendium of Fossil Marine Animal Genera: Bulletins of American Paleontology,.

Sessa, J.A., Bralower, T.J., Patzkowsky, M.E., Handley, J.C., and Ivany, L.C., 2012, Environmental and biological controls on the diversity and ecology of Late Cretaceous through early Paleogene marine ecosystems in the U.S. Gulf Coastal Plain: Paleobiology, v. 38, p. 218–239, doi:10.5061/dryad.6kp1m6r2.

Shaw, N.G., 1967, Cheilostomata of the Gulfian Cretaceous of Southwestern Arkansas. [Historical Dissertations and Theses]: Louisiana State University, https://digitalcommons.lsu.edu/gradschool_disstheses.

Shigeta, Y., and Izukura, M., 2018, Discovery of the middle Campanian (Late Cretaceous) “Soya Fauna” ammonoids in the Hidaka area: The Bulletin of the Hobetsu Museum, p. 11–25.

Sidwell, R., 1932, New Species from the Colorado Group, Cretaceous, in South Central Wyoming.:

Simon, E., and Owen, E.F., 2001, A first step in the revision of the genus Cretirhynchia Pettitt, 1950: Bulletin de l’Institut Royal des Sciences Naturelles de Belgique, Sciences de la Terre, v. 71, p. 53–118.

Simons, F.S., 1964, Geology of the Klondyke Quadrangle, Graham and Pinal Counties, Arizona: Geological Survey Professional Paper, v. 461, p. 1–173.

Simpson, C.T., 1900, Synopsis of the Naiades, or pearly fresh-water mussels: Proceedings U.S. National Museum, v. XXII, p. 501–998.

Skawina, A., and Dzik, J., 2011, Umbonal musculature and relationships of the late triassic filibranch unionoid bivalves: Zoological Journal of the Linnean Society, v. 163, p. 863–883, doi:10.1111/j.1096-3642.2011.00728.x.

Skelton, P.W., 2013, Rudist classification for the revised Bivalvia volumes of the ’Treatise on Invertebrate Paleontology:, http://www.paleotax.de/rudists/intro.htm/Taxonom.

Smith, A.G., Sohl, N.F., and Yochelson, E.L., 1968, New Upper Cretaceous Amphineura (Mollusca): Geological Survey Professional Paper 593-G.:

Sohl, N.F., 1960, Archeogastropoda, Mesogastropoda and Stratigraphy of the Ripley Owl Creek, and Prairie Bluff Formations: Geological Survey Professional Paper 331-A, p. 151.

Sohl, N.F., 1964a, Gastropods from the Coffee Sand (Upper Cretaceous) of Mississippi: Geological Survey Professional Paper 331-C, p. 1–394.

Sohl, N.F., 1964b, Neogastropoda, Opisthobranchia and Basommatophora from the Ripley, Owl Creek, and Prairie Bluff Formations: Geological Survey Professional Paper 331-B, p. 1–344.

Sohl, N.F., 1963, New Gastropod Genera from the Late Upper Cretaceous of the East Gulf Coastal Plain:, https://about.jstor.org/terms.

Sohl, N.F., 1967, Upper Cretaceous Gastropods From the Pierre Shale at Red Bird, Wyoming: Geological Survey Professional Paper 393-B, p. 1–46.

Sohl, N.F., and Kauffman, E.G., 1964, Giant Upper Cretaceous Oysters from the Gulf Coast and Caribbean: Geological Survey Professional Paper 483-H, p. 1–22.

Sohl, N.F., and Koch, C.F., 1984, Upper Cretaceous (Maestrichtian) Larger Invertebrate Fossils form the Haustator bilara Assemblage Zone in the West Gulf Coastal Plain: Open File Report 84-687.:

Sohl, N.F., and Koch, C.F., 1983, Upper Cretaceous (Maestrichtian) Mollusca from the Haustator bilira Assemblage Zone in the East Gulf Coastal Plain: Open File Report 83-451.:

Sohl, N.F., and Kollmann, H.A., 1985, Cretaceous Actaeonellid Gastropods from the Western Hemisphere: United States Geological Survey Professional Paper, v. 1304, p. 1–104.

Sowerby, J. de C., 1829, The mineral conchology of Great Britain; or coloured figures and descriptions of those remains of testaceous animals or shells, which have been preserved at various times and depths in the Earth: London, privately published, v. VI.

Spath, L.F., 1953, The Upper Cretaceous cephalopod fauna of Graham Land: Falkland Islands Dependencies Survey Scientific Reports, v. 3, p. 1–60.

Speden, I.G., 1967, Revision of Syncyclonema (Upper Cretaceous) and comparison with other small pectinid bivalves and Entolium: Bulletin of the Peabody Museum of Natural History, v. 110, p. 1–36, https://bioone.org/.

Speden, I.G., 1970, The Type Fox Hills Formation, Cretaceous (Maestrichtian), South Dakota. Part 2. Systematics of the Bivalvia: Peabody Museum of Natural History, Bulletin, v. 33, p. 1–222.

Stanton, T.W., 1909, The age and stratigraphic relations of the “Ceratops Beds” of Wyoming and Montana: Proceedings of the Washington Academy of Sciences, v. 11, p. 239–293.

Stanton, T.W., 1893, The Colorado formation and its invertebrate fauna: Bulletin of the United States Geological Survey No. 106, p. 1–189.

Stanton, T.W., 1921, The fauna of the Cannonball Marine Member of the Lance Formation: Professional Paper US Geological Survey, v. 128, p. 1–64.

Stefanini, G., 1939, Molluschi del Giuralias della Somalia, Gasteropodi e Lamellibranchi: Palaeontographia Italica, v. 32, p. 103–270.

Steiner, G., and Kabat, A.R., 2004a, Catalog of species-group names of Recent and fossil Scaphopoda (Mollusca): Zoosystema, v. 26, p. 549–726.

Steiner, G., and Kabat, A.R., 2004b, Catalog of species-group names of Recent and fossil Scaphopoda (Mollusca): Zoosystema, v. 26, p. 549–726.

Steiner, G., Kabat, A.R., and R, K.A., 2001, Catalogue of supraspecific taxa of Scaphopoda (Mollusca):, http://www.mnhn.fr/publication/zoosyst/z01n3a.

Stephenson, L.W., 1954, Additions to the Fauna of the Raritan Formation (Cenomanian) of New Jersey: Geological Survey Professional Paper 264-B.:

Stephenson, L.W., 1955a, Basal Eagle Ford Fauna (Cenomanian) in Johnson and Tarrant Counties Texas: Geological Survey Professional Paper 274-C, p. 1–67.

Stephenson, L.W., 1914, Cretaceous Deposits of the Eastern Gulf Region and Species of Exogyra from the Eastern Gulf Region and the Carolinas: United States Geological Survey Professional Paper 81, p. 1–77.

Stephenson, L.W., 1952, Larger invertebrate fossils of the Woodbine Formation (Cenomanian) of Texas: Geologic Survey Professional Paper 242, p. 1–226.

Stephenson, L.W., 1955b, Owl Creek (Upper Cretaceous) fossils from Crowleys Ridge, Southeastern Missouri: Shorter Contributions to Geology: Geological Survey Professional Paper 274, p. 97–140.

Stephenson, L.W., 1941, The larger invertebrate fossils of the Navarro Group of Texas: University of Texas Publication number 4101, p. 1–641.

Stephenson, Ll.W., 1923, Volume V: The Cretaceous Formations of North Carolina: Part 1. Invertebrate fossils of the Upper Cretaceous formations.:

Stephenson, L.W., and Monroe, W.H., 1940, The Upper Cretaceous deposits: Mississippi State Geological Survey Bulletin 40, p. 1–296.

Summesberger, H., 1979, Eine obersantone Ammonitenfauna aus dem Becken von Gosau (Oberösterreich): Annalen des Naturhistorischen Museums in Wien, v. Serie A, p. 109–176.

Summesberger, H., Kennedy, W.J., and Skoumal, P., 2017, On late santonian ammonites from the hofergraben member (Gosau group, upper cretaceous, Austria): Austrian Journal of Earth Sciences, v. 110, p. 120–132, doi:10.17738/ajes.2017.0009.

Taylor, D.W., 1975, Early Tertiary Mollusks from the Powder River Basin, Wyoming-Montana, and adjacent regions: United States Geological Survey Open File Report Part II, p. 217–331.

Taylor, J., and Glover, E., 2021, Biology, evolution and generic review of the chemosymbiotic bivalve family Lucinidae: London, The Ray Society, 1–151 p.

Taylor, P.D., and Mckinney, F.K., 2006, Bryozoa from the Campanian and Maastrichtian of the Atlan-tic and Gulf Coastal Plains, United States: Scripta Geologica, v. 132, p. 1–346.

Toots, H., and Cutler, J.F., 1962, Bryozoa from the "Mesaverde" Formation (Upper Cretaceous) of southeastern Wyoming: Journal of Paleontology, v. 36, p. 81–86.

Tourtelot, H., and Cobban, W.A., 1968, Stratigraphic significance and petrology of phosphate nodules at base of Niobrara Formation, East Flank of Black Hills, South Dakota: Geological Survey Professional Paper 594-L, p. L1–L22.

Troger, K.A., 1969, Inoceramus incurvatus n. sp. aus dem Unter-Coniac der Subherzynen Kreidemulde: Freiberger Forschungshefte, v. C256, p. 65–71.

Troger, K.-A., 1967, Zur Paläontologie, Biostratigraphie und faziellen Ausbildung der unteren Oberkreide (Cenoman bis Turon); Teil I, Paläontologie und Biostratigraphie der Inoceramen des Cenoman bis Turons Mitteleuropas: Abhandlungen des Staatlichen Museums für Mineralogie und Geologie zu Dresden, v. 12, p. 13–208.

Turkay, M., 2001, Decapoda, *in* Costello, M.J. ed., European register of marine species: a check-list of the marine species in Europe and a bibliography of guides to their identification, Collection Patrimoines Naturels, v. 50, p. 284–292.

Turner, R.D., 1966, A Survey and Illustrated Catalogue of Teredinidae (Mollusca: Bivalvia): Cambridge, Museum of Comparative Zoology.

Twenhofel, W., 1924, State Geological Survey of Kansas Bulletin 9: The Geology and Invertebrate Paleontology of the Comanchean and "Dakota" Formations of Kansas:

Vega, F.J., Garassino, A., and Zapata Jaime, R., 2013, Enoploclytia tepeyacensis n. sp. (Crustacea, Decapoda, Erymidae) from the Cretaceous (Campanian) of Coahuila, NE Mexico: la Sociedad Geológica Mexicana, v. 65, p. 207–211, doi:10.2307/24921216.

Veldsman, S.G., 2017, Taxonomic reclassification of the genus Marginella Lamarck, 1799 and description of new subgenera (Neogastropoda: Marginellidae): Visaya, v. supple, p. 5–46.

Vokes, H.E., 1980, Genera of the Bivalvia: a systematic and bibliographic Catalogue: Genera of the Bivalvia: a systematic and bibliographic Catalogue,.

Vokes, H.E., 1990, Genera of the Bivalvia: A systematic and bibliographic catalogue - Addenda and Errata: Tulane Studies in Geology and Paleontology, v. 23, p. 97–120.

Vokes, H.E., 1945, Protodonax, a New Cretaceous Molluscan Genus: Journal of Paleontology, v. 19, p. 295–308, https://about.jstor.org/terms.

Vredenburg, E., 1925a, Description of Mollusca from the post-Eocene Tertiary formation of north-western India: Cephalopoda, Opisthobranchiata, Siphonostomata: Memoirs of the Geological Survey of India, v. 50, p. 1–350.

Vredenburg, E., 1925b, Description of Mollusca from the post-Eocene Tertiary formation of north-western India: Cephalopoda, Opisthobranchiata, Siphonostomata: Memoirs of the Geological Survey of India, v. 50, p. 1–350.

Wade, B., 1917, New and Little Known Gastropoda from the Upper Cretaceous of Tennessee: Proceedings of the Academy of Natural Sciences of Philadelphia, v. 69, p. 280–304.

Wade, B., 1926, The Fauna of the Ripley Formation on Coon Creek, Tennessee: Professional Paper 137.:

Waggoner, B.M., and Langer, M.R., 1993a, A new hydroid from the Upper Cretaceous of Mississippi: Palaontologische Zeitschrift, v. 67, p. 253–259.

Waggoner, B.M., and Langer, M.R., 1993b, A new hydroid from the Upper Cretaceous of Mississippi: Palaontologische Zeitschrift, v. 67, p. 253–259.

Wagreich, M., Summesberger, H., and Kroh, A., 2010, Late Santonian bioevents in the Schattau section, Gosau Group of Austria - implications for the Santonian-Campanian boundary stratigraphy: Cretaceous Research, v. 31, p. 181–191, doi:10.1016/j.cretres.2009.10.003.

Walaszczyk, I., 2004, Inoceramids and inoceramid biostratigraphy of the Upper Campanian to basal Maastrichtian of the Middle Vistula River section, central Poland.:

Walaszczyk, I., and Cobban, W.A., 2000, Inoceramid faunas and biostratigraphy of the Upper Turonian-Lower Coniacian of the Western Interior of the United States: Special Papers in Paleontology, v. 64.

Walaszczyk, I., and Cobban, W.A., 2006, Palaeontology and biostratigraphy of the Middle-Upper Coniacian and Santonian inoceramids of the US Western Interior: Acta Geologica Polonica, v. 56, p. 241–348, https://www.researchgate.net/publication/286250061.

Walaszczyk, I., Cobban, W.A., and Harries, P.J., 2001, Inoceramids and inoceramid biostratigraphy of the Campanian and Maastrichtian of the United States Western Interior Basin: Revue de Paleobiologie, v. 20, p. 117–234.

Walaszczyk, I., Kennedy, W.J., and Paranjape, A.R., 2018, Inoceramids and associated ammonite faunas from the uppermost Turonian−lower Coniacian (Upper Cretaceous) of the Anaipadi–Saradamangalam region of the Cauvery Basin, south-east India: Acta Geologica Polonica, v. 68, p. 663–687, doi:10.1515/agp-2018-0036.

Walaszczyk, I., Odin, G.S., and Dhondt, A. v, 2002, Inoceramids from the Upper Campanian and Lower Maastrichtian of the Tercis section (SW France), the Global Stratotype Section and Point for the Campanian – Maastrichtian boundary; taxonomy, biostratigraphy and correlation potential: Acta Geological Polonica, v. 52, p. 269–305, https://www.researchgate.net/publication/290195909.

Walaszczyk, I., Plint, A.G., and Kennedy, W.J., 2016, Biostratigraphy and inoceramus survival across the Cenomanian-Turonian (Cretaceous) boundary in the Ram River section, Alberta, Canada: Acta Geologica Polonica, v. 66, p. 715–728, doi:10.1515/agp-2016-0039.

Walaszczyk, I., Plint, A.G., and Landman, N.H., 2017, Chapter 2: Inoceramid Bivalves from the Coniacian and Basal Santonian (Upper Cretaceous) of the Western Canada Foreland Basin, *in* Bulletin of the American Museum of Natural History, American Museum of Natural History Library, v. 2017-June, p. 53–103, doi:10.1206/0003-0090-414.1.4.

Walaszczyk, I., and Troger, K.-A., 1996, The species Inoceramus frechi (Bivalvia, Cretaceous); its characteristics, formal status, and stratigraphical position: Paläontologische Zeitschrift, v. 70, p. 393–404.

Waller, T.R., 2001, Dhondtichlamys, a new name for Microchlamys Sobetski, 1977 (Mollusca : Bivalvia : Pectinidae), preoccupied by Microchlamys Cockerell, 1911 (Rhizopoda : Arcellinida): Proceedings of the Biological Society of Washington, v. 114, p. 858–860.

Ward, L.W., 1992, Molluscan biostratigraphy of the Miocene Middle Atlantic Coastal Plain of North America: Martinsville, Virginia Museum of Natural History, v. 2, 1–159 p.

Warren, P.S., 1931, Invertebrate paleontology of southern plains of Alberta: AAPG Bulletin, v. 15, p. 1283–1291, http://pubs.geoscienceworld.org/aapgbull/article-pdf/15/10/1283/4358974/aapg_1931_0015_0010_1283.pdf?casa_token=EGa1gj_svMMAAAAA:4MH0dwakbCLp.

Warren, P.S., 1930, New species of fossils from Smoky River and Dunvegan formations, Alberta: Research Council of Alberta Geological Survey Report 21, p. 57–68.

Warren, P.S., 1934, Palaeontology of the Bearpaw Formation: Transactions of the Royal Society of Canada, v. 28, p. 81–97.

Weller, S., 1907, A Report on the Cretaceous Paleontology of New Jersey: Trenton, NJ, MacCrellish and Quigley, v. IV, 1–871 p.

Wells, J.W., 1947, Coral Studies, Part III. Three new Cretaceous corals from Texas and Alabama, Part IV. A new species of Phyllangia from the Florida Miocene, Part V. A new Coenocyathus from Florida: Bulletins of American Paleontology, v. 31, p. 162–176.

Wells, J.W., 1933, Corals of the Cretaceous of the Atlantic and Gulf Coastal Plains and Western Interior of the United States: Bulletins of American Paleontology, v. 18, p. 85–288.

Whetstone, K.N., and Collins, J.S.H., 1982, Fossil Crabs (Crustacea: Decapoda) from the Upper Cretaceous Eutaw Formation of Alabama: Journal of Paleontology, v. 56, p. 1218–1222, https://www.jstor.org/stable/1304579.

Whetstone, K.N., and Teichert, C., 1978, A New Genus of Nautiloid Cephalopods from the Mooreville Formation (Cretaceous) of Alabama: Journal of Paleontology, v. 52, p. 440–443, https://about.jstor.org/terms.

White, C.A., 1884, A Review of the Fossil Ostreidae of North America; and a comparison of the fossil with the living forms:

White, C.A., 1883, A review of the non-marine fossil Mollusca of North America : Washington, Geological Survey.

White, C.A., 1889, Bulletin of the United States Geological Survey No. 51: On Invertebrate Fossils from the Pacific Coast.:

White, C.A., 1876, Report upon the invertebrate fossils collected in portions of Nevada, Utah, Colorado, New Mexico, and Arizona: Washington, D.C., 1–219 p.

White, C.A., 1905, The ancestral origin of the North American Unionidae, or fresh-water mussels: Smithsonian Miscellaneous Collections, v. 48, p. 75–88.

Whiteaves, J.F., 1885, Contributions to Canadian Palaeontology: Part 1. Report on the Invertebrata of the Laramie and Cretaceous Rocks of the Vicinity of the Bow and Belly Rivers and adjacent localities in the North-West Territory.:

Whiteaves, J.F., 1879, Mesozoic Fossils: Part II. On the fossils of the Cretaceous Rocks of Vancouver and adjacent Islands in the Strait of Georgia: v. 1, 1–120 p.

Whitfield, R.P., 1982, Gasteropoda and Cephalopoda of the Raritan Clays and Greensand Marls of New Jersey: Monographs of the United States Geological Survey, v. XVIII, p. 1–295.

Whitfield, R.P., 1907, Notice of an American species of the genus Hoploparia McCoy, from the Cretaceous of Montana: Bulletin of the American Museum of Natural History, v. 23, p. 459–461.

Whitfield, R.P., 1877, Preliminary report on the paleontology of the Black Hills: United States Geological and Geographic Survey of the Rocky Mountain Region (Powell), p. 1–49.

Wick, S.L., 2021, Paleontological inventory of Paleozoic, Late Mesozoic, and Cenozoic plant, invertebrate, and vertebrate fossil species from Big Bend National Park, Texas, USA - over a century of paleontological discovery: Zitteliana, v. 95, p. 95–134, doi:10.3897/zitteliana.95.73026.

Williams, J.D., Bogan, A.E., Butler, R.S., Cummings, K.S., Garner, J.T., Harris, J.L., Johnson, N.A., and Watters, G.T., 2019, A Revised List of the Freshwater Mussels (Mollusca: Bivalvia: Unionida) of the United States and Canada: Freshwater Mollusk Biology and Conservation, v. 20, p. 33, doi:10.31931/fmbc.v20i2.2017.33-58.

Williams, J.D., Bogan, A.E., Butler, R.S., Cummings, K.S., Garner, J.T., Harris, J.L., Johnson, N.A., and Watters, G.T., 2017, A Revised List of the Freshwater Mussels (Mollusca: Bivalvia: Unionida) of the United States and Canada: Freshwater Mollusk Biology and Conservation, v. 20, p. 33, doi:10.31931/fmbc.v20i2.2017.33-58.

Williston, S.W., 1898a, The University Geological Survey of Kansas, Paleontology Part I. Upper Cretaceous: Topeka, J.S. Parks, State Printer, v. 4, 1–588 p.

Williston, S.W., 1898b, The University Geological Survey of Kansas, Paleontology Part I. Upper Cretaceous: Topeka, J.S. Parks, State Printer, v. 4, 1–588 p.

Wilson, M.A., and Taylor, P.D., 2012, Palaeoecology, preservation and taxonomy of encrusting ctenostome bryozoans inhabiting ammonite body chambers in the late cretaceous pierre shale of wyoming and South Dakota, USA: Cretaceous ctenostome bryozoans of wyoming and South Dakota, *in* Lecture Notes in Earth System Sciences, Springer International Publishing, v. 143, p. 419–433, doi:10.1007/978-3-642-16411-8_28.

Wingard, G.L., 1993, A Detailed Taxonomy of Upper Cretaceous and Lower Tertiary Crassatellidae in the Eastern United States- An Example of the Nature of Extinction at the Boundary: U.S. Geological Survey Professional Paper 1535, p. 1–131.

Winston, J.E., and Maturo Jr., F.J., 2009, Winston and Maturo 2009 GulfofMexicoOriginWatersandBiota: Texas A&M University Press, v. 1, p. 1147–1164.

Wolleben, J.A., 1977, Paleontology of the Difunta Group (Upper Cretaceous-Tertiary) in Northern Mexico:, https://about.jstor.org/terms.

Yen, T.-C., 1951, Fresh-Water Mollusks of Cretaceous Age From Montana and Wyoming: Geological Survey Professional Paper 233-A, p. 1–20.

Yen, T.-C., 1954, Nonmarine mollusks of Late Cretaceous age from Wyoming, Utah, and Colorado: Geological Survey Professional Paper, v. 254-B, p. 45–124.

Zaborski, P.M., 1990, Some Upper Cretaceous ammonites from southern Nigeria: Journal of African Earth Sciences, v. 10, p. 565–581.

Zullo, V.A., Russell, E.E., and Mellen, F.F., 1987, Brachylepas Woodward and Virgiscalpellum Withers (Cirripedia) from the Upper Cretaceous of Arkansas:, https://www.jstor.org/stable/1305135.

Appendix A-3. R Code used for Analysis

The following scripts were used in R to perform analyses.

*Code for Comparing Generic Richness and Occurrence Number (Sampling Bias)*

```
library(ggplot2)
library(ggmap)
library(tidyverse)
library(sf)
library(mapview)
library(raster)
library(rgdal)
library(dismo)
library(XML)
library(maps)
library(mFD)
library(reshape2)
library(vegan)
library(dplyr)
library(tidyr)

setwd("C:/Users/ceara/Documents/Province Project/Occurrence Data/VettedCombinedDB")

# read in the data table (includes the name of grid cell based on 60km grid)
data <- read.csv("R_Config_for_EN_OUTPUT/Vetted_Substage_Joined_60.csv", header=TRUE)

#### RECONFIGURE ENTIRE DB (same basic code as used to configure into EDENETWORKS
        format) ####

colnames(data)

#convert data to simple dataframes that can be used in EDENETWORKS analyses
simplify_data <- function(data){
  data1 <- data.frame(data)
  data2 <- data1[,c(1:2,8:9,19,25:29)]
  return(data2)
}

trunk_data <- simplify_data(data)
ncol(trunk_data)
colnames(trunk_data)
nrow(trunk_data)

#create collumn collating the grid cell and the age
trunk_data$Abrev_Age <-with(trunk_data, ifelse(Substage_from_Zone.Mbr == "CAM (low)",
                                    'LC', ifelse(Substage_from_Zone.Mbr == "CAM (mid)",
                                            'MC', ifelse(Substage_from_Zone.Mbr == "CAM (up)",
                                                    'UC', ifelse(Substage_from_Zone.Mbr == "MAA (low)",
                                                            'LM', 'UM' )))))
head(trunk_data)
```

```
nrow(trunk_data)

# replace the grid cell name column with the grid cell name and abbreviated age collated
trunk_data$PageName <- paste(trunk_data$PageName,trunk_data$Abrev_Age,sep="")
head(trunk_data)
colnames(trunk_data)
nrow(trunk_data)

# Add the faunal province information
FP_list <- read.csv("FaunalProv_AttributesList.csv")
FP_list <- FP_list[,1:3] # shorted to just the Province and Grid name columns
nrow(FP_list)
unique(FP_list$Province)

FP_list$Color <- with(FP_list, ifelse(Province == "NIP",
                              'mediumblue', ifelse(Province == "SIP",
                                           'gold', ifelse(Province == "CIP",
                                                       'darkolivegreen', ifelse(Province == "GCP",
                                                                             'chocolate', ifelse(Province == "EC",
                                                                                              'darkmagenta', 'grey' ))))))

FP_list$ColorAge <-with(FP_list, ifelse(Age == "LC",
                                 'darkblue', ifelse(Age == "MC",
                                             'olivedrab', ifelse(Age == "UC",
                                                         'yellow', ifelse(Age == "LM",
                                                                     'darkorange', 'darkred' )))))

# Check for difference in the two datasets
pn <- trunk_data[,10]
fp <- FP_list[,2]
setdiff(pn, fp)
#merge the two (table join by Grid cell name)
trunk_data <- merge(trunk_data,FP_list,by="PageName")
colnames(trunk_data)
head(trunk_data)
nrow(trunk_data) #check that there are the same # occ still

# subset just the spp-level occ data into
trunk_data_spp <- subset(trunk_data, Updated_Sp != "")
head(trunk_data_spp)
nrow(trunk_data_spp)

colnames(trunk_data)

# Get the number of fossil occurrence present for each grid cell
trunk_gen_occnum <- as.data.frame(table(trunk_data$PageName))

trunk_spp_occnum <- as.data.frame(table(trunk_data_spp$PageName))

trunk_gen_occnum <- data.frame(trunk_gen_occnum[,-1], row.names = trunk_gen_occnum[,1])
```

```
trunk_spp_occnum <- data.frame(trunk_spp_occnum[,-1], row.names = trunk_spp_occnum[,1])

# function to transform into an abundance matrix of locations with genus names
substg_gen_grid <- function(data){
  data1 <- as.data.frame(subset(data)[,c(1,4)]) # subset out just the genus name and location information
  colnames(data1) <- c("PageName","Genus_Name") # Give the columns names
  data2 <- dcast(data1, PageName~Genus_Name, length) # transform into a pres-abs matrix based on Grid
          cells
}

trunk_matrix_gen <- substg_gen_grid(trunk_data)

# function to transform into an abundance matrix of locations with species names
substg_spp_grid <- function(data){
  data1 <- as.data.frame(subset(data)[,c(1,8)]) # subset out just the species name and location information
  colnames(data1) <- c("PageName","Species_Name") # Give the columns names
  data2 <- dcast(data1, PageName~Species_Name, length) # transform into a pres-abs matrix based on
          Grid cells
}

trunk_matrix_spp <- substg_spp_grid(trunk_data_spp)

# function to configure the matrix into the correct format for EDENETWORKS and add the grid cell info
config_matrix <- function(data){
  #base1 <- (merge(data, lookup, by = 'LocationKey',all.X=TRUE, all.y=FALSE)) # only use if trying to
          make column of latlong using unique latlong key
  x <- nrow(data)
  data$SampleKey <-seq(1:x)
  a <- ncol(data)
  b <- a - 1
  base2 <- data[,c(1,a,2:b)] # This part would need to change if using lat/long as unique ID
  base3 <- data.frame(base2[,-1], row.names = base2[,1]) # make the first column with pagename the
          index
}

trunk_matrix_gen_final <- config_matrix(trunk_matrix_gen)
head(trunk_matrix_gen_final)
trunk_matrix_gen_final[4,]

trunk_matrix_spp_final <- config_matrix(trunk_matrix_spp)

#MATRIX OF TOTAL DB BASED ON SPP NAME (For figuring out faunal provinces... not for final
          analysis)
#trunk_matrix_TOTAL_final <- config_matrix(trunk_matrix_TOTAL)

pres_ab <- function(data){
  numeric_cols <- vapply(data, is.numeric, logical(1)) # make all values 1 if not 0
  data[numeric_cols] <- as.integer(data[numeric_cols] != 0)
  data
}
```

```
trunk_matrix_gen_final_pres <- pres_ab(trunk_matrix_gen_final)

trunk_matrix_spp_final_pres <- pres_ab(trunk_matrix_spp_final)

# get the richness per grid cell count by summing across the rows
#For genera
ncol(trunk_matrix_gen_final_pres)
numb_gen <- data.frame(rowSums(trunk_matrix_gen_final_pres[,2:575])) # sum rows to get total # gen
trunk_matrix_gen_final_pres_pre <- trunk_matrix_gen_final_pres # create new matrix name to use
trunk_matrix_gen_final_pres_pre$nGen <- numb_gen # Add column of total occ number
trunk_gen_rich <- trunk_matrix_gen_final_pres_pre[,c(576)]
colnames(trunk_gen_rich)

#For species
ncol(trunk_matrix_spp_final_pres)
numb_spp <- data.frame(rowSums(trunk_matrix_spp_final_pres[,2:1114])) # sum rows to get total # occ
trunk_matrix_spp_final_pres_pre <- trunk_matrix_spp_final # create new matrix name to use
trunk_matrix_spp_final_pres_pre$nSpp <- numb_spp # Add column of total occ number
trunk_spp_rich <- trunk_matrix_spp_final_pres_pre[,c(1115)]
colnames(trunk_spp_rich)

# Combine the tables

genus_data <- cbind(trunk_gen_occnum,trunk_gen_rich)

colnames(genus_data) <- c("Occ count","Genus Richness")

species_data <- cbind(trunk_spp_occnum,trunk_spp_rich)

colnames(species_data) <- c("Occ count","Species Richness")

# Look at the correlation and distributions of the genus-level data

pdf(file = "DB_Map_Visualize/genus_histograms.pdf")

hist(genus_data[,1],xlab='Genus Occurrence Count',main='Genus Occurrence Frequency')
hist(genus_data[,2],xlab='Genus Richness',main='Genus Richness Frequency')

dev.off()

# perform Shapiro test for normality
shapiro.test(genus_data[,1]) # very much not normal
shapiro.test(genus_data[,2]) # very much not normal

# Q-Q plots to look at normality

pdf(file = "DB_Map_Visualize/genus_QQplots.pdf")

ggqqplot(genus_data[,1], ylab = "Genus Occurrence Count")
ggqqplot(genus_data[,2], ylab = "Genus Richness")
```

```
dev.off()

# conducnt spearmans rho correlation (non-parametric because the data is highly skewed)
cor(genus_data$`Occ count`, genus_data$`Genus Richness`,  method = "spearman", use =
        "complete.obs")
# result: 0.9327709 >> highly correlated

cor.test(genus_data$`Occ count`, genus_data$`Genus Richness`,  method = "spearman",exact=FALSE)
  # p-value is highly significant: < 2.2e-16

   ## Correlation is highly significant

# Look at a correlation plot of the data
library("ggpubr")

pdf(file = "DB_Map_Visualize/genus_scatter_trend.pdf")

ggscatter(genus_data, x = "Occ count", y = "Genus Richness",
        add = "reg.line", conf.int = TRUE,
        cor.coef = TRUE, cor.method = "spearman",
        xlab = "Number of Occurrences", ylab = "Genus Richness")

dev.off()

# create linear model for the correlation
lm.model = lm(genus_data$'Occ count' ~ genus_data$'Genus Richness', data = genus_data)

summary(lm.model)

# create a log transformed model
lm_log.model = lm(log1p(genus_data$'Occ count') ~ log1p(genus_data$'Genus Richness'), data =
        genus_data)

summary(lm_log.model)
str(summary(lm_log.model))
names(summary(lm_log.model))

# Results show that there is a very high correlation between sampling effort (number of fossil occurrences
        collected)
# and the genus richness observed.

pdf(file = "DB_Map_Visualize/genus_logtransformed_data.pdf")

# perform log10 transformation on both sets of data and graph
ggplot(data = genus_data, aes(x = genus_data$'Occ count', y = genus_data$'Genus Richness')) +
  geom_point() +
  scale_x_log10() + scale_y_log10() +
  xlab("Log10 transformed Number of Occurrences") +
  ylab("Log10 transformed Genus Richness") +
  ggtitle("Log Transfromed Richness vs. Occurrence Count") +
  geom_smooth(method=lm, level=0.99) +
```

```
  annotate('text',x = 10, y = 100,label=paste('R-squared:',summary(lm_log.model)$r.squared))

dev.off()

# Look at the correlation and distributions of the species-level data

pdf(file = "DB_Map_Visualize/species_histograms.pdf")

hist(species_data[,1],xlab='Species Occurrence Count',main='Species Occurrence Frequency')
hist(species_data[,2],xlab='Species Richness',main='Species Richness Frequency')

dev.off()

# perform Shapiro test for normality
shapiro.test(species_data[,1]) # very much not normal
shapiro.test(species_data[,2]) # very much not normal

# Q-Q plots to look at normality

pdf(file = "DB_Map_Visualize/species_QQplots.pdf")

ggqqplot(species_data[,1], ylab = "Species Occurrence Count")
ggqqplot(species_data[,2], ylab = "Speices Richness")

dev.off()

# conducnt spearmans rho correlation (non-parametric because the data is highly skewed)
cor(species_data$`Occ count`, species_data$`Species Richness`,  method = "spearman", use =
        "complete.obs")
# result: 0.9353668

cor.test(species_data$`Occ count`, species_data$`Species Richness`,  method =
        "spearman",exact=FALSE)
# p-value is highly significant: < 2.2e-16

## Correlation is highly significant


# Look at a correlation plot of the data
library("ggpubr")

pdf(file = "DB_Map_Visualize/species_scatter_trend.pdf")

ggscatter(species_data, x = "Occ count", y = "Species Richness",
      add = "reg.line", conf.int = TRUE,
      cor.coef = TRUE, cor.method = "spearman",
      xlab = "Number of Occurrences", ylab = "Species Richness")

dev.off()

# create linear model for the correlation
```

```
lm.model_spp = lm(species_data$'Occ count' ~ species_data$'Species Richness', data = species_data)

summary(lm.model_spp)

# create a log transformed model
lm_log.model_spp = lm(log1p(species_data$'Occ count') ~ log1p(species_data$'Species Richness'), data
        = species_data)

summary(lm_log.model_spp)

# Results show that there is a very high correlation between sampling effort (number of fossil occurrences
        collected)
# and the species richness observed.

pdf(file = "DB_Map_Visualize/species_logtransformed_data.pdf")

# perform log10 transformation on both sets of data and graph
ggplot(data = species_data, aes(x = species_data$'Occ count', y = species_data$'Species Richness')) +
  geom_point() +
  scale_x_log10() + scale_y_log10() +
  xlab("Log10 transformed Number of Occurrences") +
  ylab("Log10 transformed Species Richness") +
  ggtitle("Log Transfromed Richness vs. Occurrence Count") +
  geom_smooth(method=lm, level=0.99) +
  annotate('text',x = 10, y = 100,label=paste('R-squared:',summary(lm_log.model_spp)$r.squared))

dev.off()
```

*Code to configure Occurrence Database into format for EDENetworks Analysis*

```
## Load necessary packages
library(ggplot2)
library(ggmap)
library(tidyverse)
library(sf)
library(mapview)
library(raster)
library(rgdal)
library(dismo)
library(XML)
library(maps)
library(mFD)
library(reshape2)
library(vegan)
library(dplyr)
library(tidyr)

# set working directory
setwd("C:/Users/ceara/Documents/Province Project/Occurrence Data/VettedCombinedDB")


##### RECONFIGURE THE FULL DB FILE BASED ON LOCATION AND GRID INFO ####

# read in the csv file of the DB that has only the substage-level occ data
data <- read.csv('Vetted_Combined_Genus-Spp-lvl_SubstageOnly_sansACP-lowLAT.csv')
# add column with the lat long information as a unique location ID
data$LocationKey <- paste(data$Lat,",",data$Lon)
# add a column with the genus and spp names concatenated
data$SppName <- paste(data$Updated_Genus," ",data$Updated_Sp)
# add lat and long columns
data$Latitude <- paste(data$Lat)
data$Longitude <- paste(data$Lon)
head(data)
nrow(data)
latlon <- data[,3:4]
head(latlon)
nrow(unique(latlon))
nrow(data.frame(unique(data$LocationKey)))

# Read in grids made in ArcGIS pro
grid_60km <- st_read("SpatialGrids/grid_60km.shp")
plot(grid_60km$geometry,xlim=c(-118,-
          68),ylim=c(28,53),col="blue",pch=19,xlab="Longitude",ylab="Latitude")
points(data$Long,data$Lat,col="red",pch=19,xlab="Longitude",ylab="Latitude") # add occ pts

#transform database into shapefile
sf_data <- st_as_sf(data, coords = c("Long", "Lat"),  crs = 4326)
```

```
# function to spatially join DB with grids
spatial_join <- function(shapefile_data,grid){
  data1 <- st_join(shapefile_data,left=TRUE,grid["PageName"])
  return(data1)
}

# spatially join Grid cell Page Names to data table
join60_data <- spatial_join(sf_data,grid_60km)
join60_data <- as.data.frame(join60_data)
colnames(join60_data)

nrow(data.frame(unique(join60_data$Updated_Genus)))
nrow(data.frame(unique(join60_data$SppName)))

# add column name to last column for geometry
names(join60_data)[length(names(join60_data))]<-"geometry2"
colnames(join60_data)
nrow(join60_data)


# write csv file with joined info
write.csv(join60_data,file="R_Config_for_EN_OUTPUT/Vetted_Substage_Joined_60.csv")


#### Get summary information about the dataset as a whole ####
class_genus <- join60_data[,c(14,8)]
class_species <- subset(join60_data, Updated_Sp != "")
class_species <- class_species[,c(14,26)]

nb_gen_class <-count(class_genus) # get the frequency of a genus with class ID
nb_spp_class <-count(class_species) # get the frequency of a species with class ID
nb_gen <- nrow(data.frame(unique(class_genus$Updated_Genus))) # get the number of unique genera
nb_spp <- nrow(data.frame(unique(class_species$SppName))) # get the number of unique species
nb_classes <- nrow(data.frame(unique(class_genus$HigherTax_Class))) # get number of occ for each
        unique class-level
nb_occ <- nrow(join60_data) # get number of fosil occurrences

write.csv(unique(class_genus),"R_Config_for_EN_OUTPUT/class_unique_genus.csv",row.names=FAL
        SE)

write.csv(unique(class_species),"R_Config_for_EN_OUTPUT/class_unique_species.csv",row.names=FA
        LSE)

#### RECONFIGURE ENTIRE DB TO USE IN EDENETWORKS ####

#convert data to simple dataframes that can be used in EDENETWORKS analyses
simplify_data <- function(data){
  data1 <- data.frame(data)
  data2 <- data1[,c(1:2,8:9,19,25:29)]
  return(data2)
}
```

```
trunk60_data <- simplify_data(join60_data)
ncol(trunk60_data)
colnames(trunk60_data)
nrow(trunk60_data)

#create collumn collating the grid cell and the age
trunk60_data$Abrev_Age <-with(trunk60_data, ifelse(Substage_from_Zone.Mbr == "CAM (low)",
                                'LC', ifelse(Substage_from_Zone.Mbr == "CAM (mid)",
                                'MC', ifelse(Substage_from_Zone.Mbr == "CAM (up)",
                                'UC', ifelse(Substage_from_Zone.Mbr == "MAA (low)",
                                'LM', 'UM' )))))
head(trunk60_data)
nrow(trunk60_data)

# replace the grid cell name column with the grid cell name and abbreviated age collated
trunk60_data$PageName <- paste(trunk60_data$PageName,trunk60_data$Abrev_Age,sep="")
head(trunk60_data)
colnames(trunk60_data)
nrow(trunk60_data)

# Add the fuanal province information
FP_list <- read.csv("FaunalProv_AttributesList.csv")
FP_list <- FP_list[,1:3] # shorted to just the Province and Grid name columns
nrow(FP_list)
unique(FP_list$Province)

FP_list$Color <- with(FP_list, ifelse(Province == "NIP",
                                'mediumblue', ifelse(Province == "SIP",
                                'gold', ifelse(Province == "CIP",
                                'darkolivegreen', ifelse(Province == "GCP",
                                'chocolate', ifelse(Province == "EC",
                                'darkmagenta', 'grey' ))))))

FP_list$ColorAge <-with(FP_list, ifelse(Age == "LC",
                                'darkblue', ifelse(Age == "MC",
                                'olivedrab', ifelse(Age == "UC",
                                'yellow', ifelse(Age == "LM",
                                'darkorange', 'darkred' )))))

# Check for difference in the two datasets
pn <- trunk60_data[,10]
fp <- FP_list[,2]
setdiff(pn, fp)
#merge the two (table join by Grid cell name)
trunk60_data <- merge(trunk60_data,FP_list,by="PageName")
colnames(trunk60_data)
head(trunk60_data)
nrow(trunk60_data) #check that there are the same # occ still

# subset just the spp-level occ data into
```

```
trunk60_data_spp <- subset(trunk60_data, Updated_Sp != "")
head(trunk60_data_spp)
nrow(trunk60_data_spp)

colnames(trunk60_data)

# function to transform into an abundance matrix of locations with genus names
substg_gen_grid <- function(data){
  data1 <- as.data.frame(subset(data)[,c(1,4)]) # subset out just the genus name and location information
  colnames(data1) <- c("PageName","Genus_Name") # Give the columns names
  data2 <- dcast(data1, PageName~Genus_Name, length) # transform into a pres-abs matrix based on Grid
          cells
}

trunk60_matrix_gen <- substg_gen_grid(trunk60_data)
head(trunk60_matrix_gen)

# function to transform into an abundance matrix of locations with species names
substg_spp_grid <- function(data){
  data1 <- as.data.frame(subset(data)[,c(1,8)]) # subset out just the species name and location information
  colnames(data1) <- c("PageName","Species_Name") # Give the columns names
  data2 <- dcast(data1, PageName~Species_Name, length) # transform into a pres-abs matrix based on
          Grid cells
}

trunk60_matrix_spp <- substg_spp_grid(trunk60_data_spp)
head(trunk60_matrix_spp)

#GET MATRIX OF ALL DB BASED ON SPP NAME (For figuring out faunal provinces... not for final
          analysis)
#trunk60_matrix_TOTAL <- substg_spp_grid(trunk60_data)

# function to configure the matrix into the correct format for EDENETWORKS and add the grid cell info
config_matrix <- function(data){
  #base1 <- (merge(data, lookup, by = 'LocationKey',all.X=TRUE, all.y=FALSE)) # only use if trying to
          make column of latlong using unique latlong key
  x <- nrow(data)
  data$SampleKey <-seq(1:x)
  a <- ncol(data)
  b <- a - 1
  base2 <- data[,c(1,a,2:b)] # This part would need to change if using lat/long as unique ID
  base3 <- data.frame(base2[,-1], row.names = base2[,1]) # make the first column with pagename the
          index
}

trunk60_matrix_gen_final <- config_matrix(trunk60_matrix_gen)

trunk60_matrix_spp_final <- config_matrix(trunk60_matrix_spp)

#MATRIX OF TOTAL DB BASED ON SPP NAME (For figuring out faunal provinces... not for final
          analysis)
```

```
#trunk60_matrix_TOTAL_final <- config_matrix(trunk60_matrix_TOTAL)

pres_ab <- function(data){
  numeric_cols <- vapply(data, is.numeric, logical(1)) # make all values 1 if not 0
  data[numeric_cols] <- as.integer(data[numeric_cols] != 0)
  data
}

trunk60_matrix_gen_final <- pres_ab(trunk60_matrix_gen_final)

trunk60_matrix_spp_final <- pres_ab(trunk60_matrix_spp_final)

#MATRIX OF TOTAL DB BASED ON SPP NAME (For figuring out faunal provinces... not for final
        analysis)
#trunk60_matrix_TOTAL_final <- pres_ab(trunk60_matrix_TOTAL_final)


# write csv and txt files of the matrix for the whole dataset
write.csv(trunk60_matrix_gen_final,file=
        "R_Config_for_EN_OUTPUT/CompleteDB/trunk60_matrix_gen_final.csv")
#write.table(trunk60_matrix_gen_final,file=
        "R_Config_for_EN_OUTPUT/trunk60_matrix_gen_final.txt",col.names = TRUE)


write.csv(trunk60_matrix_spp_final,file=
        "R_Config_for_EN_OUTPUT/CompleteDB/trunk60_matrix_spp_final.csv")

# (For figuring out faunal provinces... not for final analysis)
#write.csv(trunk60_matrix_TOTAL_final,file=
        "R_Config_for_EN_OUTPUT/trunk60_matrix_TOTAL_final.csv")

## Create matrix with grid cells containing less than 3 occ removed

#For genera
ncol(trunk60_matrix_gen_final)
numb_occ_gen <- data.frame(rowSums(trunk60_matrix_gen_final[,2:575])) # sum rows to get total # occ
trunk60_matrix_gen_final_pre <- trunk60_matrix_gen_final # create new matrix name to use
trunk60_matrix_gen_final_pre$nOccGen <- numb_occ_gen # Add column of total occ number
colnames(trunk60_matrix_gen_final_pre)
trunk60_matrix_gen_great3 <- subset(trunk60_matrix_gen_final_pre, nOccGen >= 3) # subset out all grid
        cells with greater 3 occ
ncol(trunk60_matrix_gen_great3)
trunk60_matrix_gen_great3 <- trunk60_matrix_gen_great3[,1:575] # remove occ number column
colnames(trunk60_matrix_gen_great3)

# export new csv and txt files of the dataset with grids containing <3 occ removed
write.csv(trunk60_matrix_gen_great3,file=
        "R_Config_for_EN_OUTPUT/CompleteDB/trunk60_matrix_gen_great3.csv")
# write.table(trunk60_matrix_gen_great3,file=
        "R_Config_for_EN_OUTPUT/trunk60_matrix_gen_great3.txt",col.names = TRUE)
```

```
# for species
ncol(trunk60_matrix_spp_final)
numb_occ_spp <- data.frame(rowSums(trunk60_matrix_spp_final[,2:575])) # sum rows to get total # occ
trunk60_matrix_spp_final_pre <- trunk60_matrix_spp_final # create new matrix name to use
trunk60_matrix_spp_final_pre$nOccSpp <- numb_occ_spp # Add column of total occ number
colnames(trunk60_matrix_spp_final_pre)
trunk60_matrix_spp_great3 <- subset(trunk60_matrix_spp_final_pre, nOccSpp >= 3) # subset out all grid
        cells with greater 3 occ
ncol(trunk60_matrix_spp_great3)
trunk60_matrix_spp_great3 <- trunk60_matrix_spp_great3[,1:1114] # remove occ number column
colnames(trunk60_matrix_spp_great3)

# export new csv and txt files of the dataset with grids containing <3 occ removed
write.csv(trunk60_matrix_spp_great3,file=
        "R_Config_for_EN_OUTPUT/CompleteDB/trunk60_matrix_spp_great3.csv")
# write.table(trunk60_matrix_gen_great3,file=
        "R_Config_for_EN_OUTPUT/trunk60_matrix_gen_great3.txt",col.names = TRUE)

#### Get attribute files for EDENetwork analysis ####

# function for creating an attribute table of averaged coor for each occupied grid cell (attribute data can be
        added to this function)
coor_atrib <- function(data){
  loc <- data[,c(1,9:10)]
  loc$Latitude <- as.numeric(as.character(loc$Latitude))# get just lat/long and grid cell names to get
        average locations
  loc$Longitude <- as.numeric(as.character(loc$Longitude))# get just lat/long and grid cell names to get
        average locations
  avg_lat <- aggregate( Latitude ~ PageName, loc, mean )# create a matrix of lat lon avg for grid cells
  avg_lon <- aggregate( Longitude ~ PageName, loc, mean )# create a matrix of lat lon avg for grid cells
  avg_loc <- merge(avg_lat,avg_lon)
  colnames(avg_loc) <- c("node_label","y","x")
  avg_loc
}

colnames(trunk60_data)
nrow(trunk60_data)

# Get the attribute table of just the locations for genera level
trunk60_gen_avg_loc <- coor_atrib(trunk60_data)
head(trunk60_gen_avg_loc)
nrow(trunk60_gen_avg_loc)

# Get list of all grid cells with greater/less than 2 dev of median occ count?? Trying different methods
colnames(trunk60_matrix_gen_final_pre)
lastcol <- ncol(trunk60_matrix_gen_final_pre)
grid_gen_occ_count <- trunk60_matrix_gen_final_pre[,c(lastcol)]
colnames(grid_gen_occ_count) <- "Numb_Occ"  # get a list of the total # occ in each grid cell
head(grid_gen_occ_count)
summary(grid_gen_occ_count) # check the mean, median, etc for the list
typeof(grid_gen_occ_count)
```

```
grid_gen_occ_count_calc <- as.numeric(unlist(grid_gen_occ_count)) # convert to numeric
hist(grid_gen_occ_count_calc) # values are highly right skewed... most have less than 20 occ
med_ab_dev <- mad((grid_gen_occ_count_calc)) # calculate the median absolute deviation of the data
   ### The MAD is too high, even without being multiplied by 2 like SD, can't really go below median
         with it
sd(grid_gen_occ_count_calc)
   ### The SD is too high too... can't subtract SD from mean either to get low lying values
quantile(grid_gen_occ_count_calc) # get quantile values
   ### Quantiles show that 1st quantile is 2. Might be worth using but very low... try median?

## Going to try to just give color attributes based on below median and below 2nd quantile but will color
         based on bins... no <3 occ
grid_gen_occ_count$OccColor <- with(grid_gen_occ_count, ifelse(Numb_Occ <= 3,
            'black', ifelse(Numb_Occ <= 5,'grey', ifelse(Numb_Occ <= 10,
             'blue',ifelse(Numb_Occ <= 15,'springgreen',
             ifelse(Numb_Occ <= 20,'gold',ifelse(Numb_Occ <= 25,'indianred','purple')))))))
grid_gen_occ_count <- cbind(node_label = rownames(grid_gen_occ_count), grid_gen_occ_count) #
         make the index the first row (Grid cell names)
rownames(grid_gen_occ_count) <- 1:nrow(grid_gen_occ_count) # create a new index

#Add a column of age and province to the attributes table
nrow(trunk60_data)
atrib_gen <- unique(trunk60_data[,c(1,12:15)])
colnames(atrib_gen) <- c("node_label","Province","Age","ProvColor","AgeColor")
nrow(atrib_gen)
head(atrib_gen)

#merge with table of less than median and 2nd quantile value colors
trunk60_gen_atrib <- merge(grid_gen_occ_count,atrib_gen,by="node_label")

#Add age level info to the loc table
loc_age_merge <- trunk60_gen_atrib[,c(1,5)]
head(loc_age_merge)
trunk60_gen_avg_loc <- merge(trunk60_gen_avg_loc,loc_age_merge,by="node_label")

write.csv(trunk60_gen_avg_loc,file="R_Config_for_EN_OUTPUT/CompleteDB/trunk60_gen_avg_loc.c
         sv",row.names=FALSE)

write.csv(trunk60_gen_atrib,file="R_Config_for_EN_OUTPUT/CompleteDB/trunk60_gen_atrib.csv",ro
         w.names=FALSE)

# Get the attribute table of just the locations for species level
trunk60_spp_avg_loc <- coor_atrib(trunk60_data_spp)
head(trunk60_spp_avg_loc)
nrow(trunk60_spp_avg_loc)

# Get list of all grid cells with greater/less than 2 dev of median occ count?? Trying different methods
colnames(trunk60_matrix_spp_final_pre)
lastcol <- ncol(trunk60_matrix_spp_final_pre)
grid_spp_occ_count <- trunk60_matrix_spp_final_pre[,c(lastcol)]
colnames(grid_spp_occ_count) <- "Numb_Occ"  # get a list of the total # occ in each grid cell
```

```
head(grid_spp_occ_count)
summary(grid_spp_occ_count) # check the mean, median, etc for the list

## Going to try to just give color attributes based on below median and below 2nd quantile
grid_spp_occ_count$OccColor <-with(grid_spp_occ_count, ifelse(Numb_Occ <= 3,
                        'black', ifelse(Numb_Occ <= 5,'grey', ifelse(Numb_Occ <= 10,
                        'blue',ifelse(Numb_Occ <= 15,'springgreen',
                        ifelse(Numb_Occ <= 20,'gold',ifelse(Numb_Occ <= 25,'indianred','purple')))))))

grid_spp_occ_count <- cbind(node_label = rownames(grid_spp_occ_count), grid_spp_occ_count) # make
    the index the first row (Grid cell names)
rownames(grid_spp_occ_count) <- 1:nrow(grid_spp_occ_count) # create a new index

#Add a column of age and province to the attributes table
colnames(trunk60_data_spp)
atrib_spp <- unique(trunk60_data_spp[,c(1,12:15)])
colnames(atrib_spp) <- c("node_label","Province","Age","ProvColor","AgeColor")
head(atrib_spp)

# Merge the two tables together
trunk60_spp_atrib <- merge(grid_spp_occ_count,atrib_spp, by="node_label",all.y=TRUE)
head(trunk60_spp_atrib)

#Add age level info to the loc table
loc_age_merge <- trunk60_spp_atrib[,c(1,5)]
head(loc_age_merge)
trunk60_spp_avg_loc <- merge(trunk60_spp_avg_loc,loc_age_merge,by="node_label")

write.csv(trunk60_spp_avg_loc,file="R_Config_for_EN_OUTPUT/CompleteDB/trunk60_spp_avg_loc.c
    sv",row.names=FALSE)

write.csv(trunk60_spp_atrib,file="R_Config_for_EN_OUTPUT/CompleteDB/trunk60_spp_atrib.csv",ro
    w.names=FALSE)

### Make a attributes table for the matrix with grids containing <3 occ removed (this one is messier)

# Reconfigure the matrix without <3 occ grid cells into a simple list of grid names
trunk60_matrix_gen_great3_GridNames <- cbind(node_label = rownames(trunk60_matrix_gen_great3),
    trunk60_matrix_gen_great3) # make the index the first row (Grid cell names)
rownames(trunk60_matrix_gen_great3_GridNames) <- 1:nrow(trunk60_matrix_gen_great3) # create a
    new index
trunk60_matrix_gen_great3_GridNames <- data.frame(trunk60_matrix_gen_great3_GridNames[,1]) #
    remove all rows except the grid cell names
colnames(trunk60_matrix_gen_great3_GridNames) <- "node_label"
nrow(trunk60_matrix_gen_great3_GridNames)
head(trunk60_matrix_gen_great3_GridNames)

# merge the simplified list of occupied grid cells from the matrix which has had <3 occ grids removed
    with the full atrib table
trunk60_matrix_gen_great3_atrib <-
    merge(trunk60_matrix_gen_great3_GridNames,trunk60_gen_atrib,all.x=TRUE)
```

```
nrow(trunk60_matrix_gen_great3_atrib)
head(trunk60_matrix_gen_great3_atrib)

trunk60_matrix_gen_great3_avgloc <-
        merge(trunk60_matrix_gen_great3_GridNames,trunk60_gen_avg_loc,all.x=TRUE)

write.csv(trunk60_matrix_gen_great3_atrib,file="R_Config_for_EN_OUTPUT/CompleteDB/trunk60_m
        atrix_gen_great3_atrib.csv",row.names=FALSE)

write.csv(trunk60_matrix_gen_great3_avgloc,file="R_Config_for_EN_OUTPUT/CompleteDB/trunk60_
        matrix_gen_great3_avgloc.csv",row.names=FALSE)

# Reconfigure the matrix without <3 occ grid cells into a simple list of grid names for species
trunk60_matrix_spp_great3_GridNames <- cbind(node_label = rownames(trunk60_matrix_spp_great3),
        trunk60_matrix_spp_great3) # make the index the first row (Grid cell names)
rownames(trunk60_matrix_spp_great3_GridNames) <- 1:nrow(trunk60_matrix_spp_great3) # create a
        new index
trunk60_matrix_spp_great3_GridNames <- data.frame(trunk60_matrix_spp_great3_GridNames[,1]) #
        remove all rows except the grid cell names
colnames(trunk60_matrix_spp_great3_GridNames) <- "node_label"
nrow(trunk60_matrix_spp_great3_GridNames)
head(trunk60_matrix_spp_great3_GridNames)

# merge the simplified list of occupied grid cells from the matrix which has had <3 occ grids removed
        with the full atrib table
trunk60_matrix_spp_great3_atrib <-
        merge(trunk60_matrix_spp_great3_GridNames,trunk60_spp_atrib,all.x=TRUE)
nrow(trunk60_matrix_spp_great3_atrib)
head(trunk60_matrix_spp_great3_atrib)

trunk60_matrix_spp_great3_avgloc <-
        merge(trunk60_matrix_spp_great3_GridNames,trunk60_spp_avg_loc,all.x=TRUE)

write.csv(trunk60_matrix_spp_great3_atrib,file="R_Config_for_EN_OUTPUT/CompleteDB/trunk60_ma
        trix_spp_great3_atrib.csv",row.names=FALSE)

write.csv(trunk60_matrix_spp_great3_avgloc,file="R_Config_for_EN_OUTPUT/CompleteDB/trunk60_
        matrix_spp_great3_avg_loc.csv",row.names=FALSE)

#### RECONFIGURE INTO SIMPLIFIED SUBSTAGES TO USE IN EDENETWORKS ####

   ## SUBSETTING INTO SUBSTAGE BINS ##

# Subset the genus-level data based on substages
CAM_low <- subset(trunk60_data,Substage_from_Zone.Mbr == "CAM (low)")
CAM_mid <- subset(trunk60_data,Substage_from_Zone.Mbr == "CAM (mid)")
CAM_up <- subset(trunk60_data,Substage_from_Zone.Mbr == "CAM (up)")
MAA_low <- subset(trunk60_data,Substage_from_Zone.Mbr == "MAA (low)")
MAA_up <- subset(trunk60_data,Substage_from_Zone.Mbr == "MAA (up)")

# check the number of genus-level occ for each substage
```

```
nrow(CAM_low) # 1866
nrow(CAM_mid) # 8403
nrow(CAM_up) # 8564
nrow(MAA_low) # 6480
nrow(MAA_up) # 7855

# write csv of subset genus-level data
write.csv(CAM_low, file= "R_Config_for_EN_OUTPUT/Low CAM/CAM_low_genus_all.csv")
write.csv(CAM_mid, file= "R_Config_for_EN_OUTPUT/Mid CAM/CAM_mid_genus_all.csv")
write.csv(CAM_up, file= "R_Config_for_EN_OUTPUT/Up CAM/CAM_up_genus_all.csv")
write.csv(MAA_low, file= "R_Config_for_EN_OUTPUT/Low MAA/MAA_low_genus_all.csv")
write.csv(MAA_up, file= "R_Config_for_EN_OUTPUT/Up MAA/MAA_up_genus_all.csv")

# Subset the spp-level data based on substages
CAM_low_spp <- subset(trunk60_data_spp,Substage_from_Zone.Mbr == "CAM (low)")
CAM_mid_spp <- subset(trunk60_data_spp,Substage_from_Zone.Mbr == "CAM (mid)")
CAM_up_spp <- subset(trunk60_data_spp,Substage_from_Zone.Mbr == "CAM (up)")
MAA_low_spp <- subset(trunk60_data_spp,Substage_from_Zone.Mbr == "MAA (low)")
MAA_up_spp <- subset(trunk60_data_spp,Substage_from_Zone.Mbr == "MAA (up)")

# Check number of spp occ in each substage (have not removed duplicates)
nrow(CAM_low_spp) # 896
nrow(CAM_mid_spp) # 4856
nrow(CAM_up_spp) # 5282
nrow(MAA_low_spp) # 4012
nrow(MAA_up_spp) # 4866

# write csv of subset spp-level data
write.csv(CAM_low_spp, file= "R_Config_for_EN_OUTPUT/Low CAM/CAM_low_species_all.csv")
write.csv(CAM_mid_spp, file= "R_Config_for_EN_OUTPUT/Mid CAM/CAM_mid_species_all.csv")
write.csv(CAM_up_spp, file= "R_Config_for_EN_OUTPUT/Up CAM/CAM_up_species_all.csv")
write.csv(MAA_low_spp, file= "R_Config_for_EN_OUTPUT/Low MAA/MAA_low_species_all.csv")
write.csv(MAA_up_spp, file= "R_Config_for_EN_OUTPUT/Up MAA/MAA_up_species_all.csv")

#### CREATE MATRIX OF GENUS LEVEL OCC INFORMATION FOR EDENETWORKS
        ANALYSIS ####

  # Original attempted made unique lat/long into the ID locations but that doesn't work, so instead
  # I'm going to make the GridCell name the unique location and give a collection ID that is
  # meaningless just so the configuration matches what it's supposed to in EDENETWORKS. Any code
  # that isn't used anymore based on that original attempt has been left but commented out.

# create lookup table of unique location ID (lat,long) to use as "sites" in EN
# lookup <- seq(1:6000)

# function to transform into an abundance matrix of locations with genus names
substg_gen_grid <- function(data){
  data1 <- as.data.frame(subset(data)[,c(1,4)]) # subset out just the genus name and location information
  colnames(data1) <- c("PageName","Genus_Name") # Give the columns names
  data2 <- dcast(data1, PageName~Genus_Name, length) # transform into a pres-abs matrix based on Grid
        cells
```

```
}

colnames(CAM_low)

CAM_low_matrix_gen <- substg_gen_grid(CAM_low)
CAM_mid_matrix_gen <- substg_gen_grid(CAM_mid)
CAM_up_matrix_gen <- substg_gen_grid(CAM_up)
MAA_low_matrix_gen <- substg_gen_grid(MAA_low)
MAA_up_matrix_gen <- substg_gen_grid(MAA_up)

# find number of genera in each substage matrix
ncol(CAM_low_matrix_gen) # 148
ncol(CAM_mid_matrix_gen) # 188
ncol(CAM_up_matrix_gen) # 344
ncol(MAA_low_matrix_gen) # 305
ncol(MAA_up_matrix_gen) # 331

# function to configure the matrix into the correct format for EDENETWORKS and add the grid cell info
config_matrix <- function(data){
  #base1 <- (merge(data, lookup, by = 'LocationKey',all.X=TRUE, all.y=FALSE)) # only use if trying to
          make column of latlong using unique latlong key
  x <- nrow(data)
  data$SampleKey <-seq(1:x)
  a <- ncol(data)
  b <- a - 1
  base2 <- data[,c(1,a,2:b)] # This part would need to change if using lat/long as unique ID
  base3 <- data.frame(base2[,-1], row.names = base2[,1]) # make the first column with pagename the
          index
}

# create matrix in correct config with grid cell names include as first
CAM_low_matrix_gen_final <- config_matrix(CAM_low_matrix_gen)
CAM_mid_matrix_gen_final <- config_matrix(CAM_mid_matrix_gen)
CAM_up_matrix_gen_final <- config_matrix(CAM_up_matrix_gen)
MAA_low_matrix_gen_final <- config_matrix(MAA_low_matrix_gen)
MAA_up_matrix_gen_final <- config_matrix(MAA_up_matrix_gen)

#function to change from abundance to pres-abs matrix
pres_ab <- function(data){
  numeric_cols <- vapply(data, is.numeric, logical(1))
  data[numeric_cols] <- as.integer(data[numeric_cols] != 0)
  data
}

CAM_low_matrix_gen_final <- pres_ab(CAM_low_matrix_gen_final)
CAM_mid_matrix_gen_final <- pres_ab(CAM_mid_matrix_gen_final)
CAM_up_matrix_gen_final <- pres_ab(CAM_up_matrix_gen_final)
MAA_low_matrix_gen_final <- pres_ab(MAA_low_matrix_gen_final)
MAA_up_matrix_gen_final <- pres_ab(MAA_up_matrix_gen_final)

# check the number of rows (# of unique locations w/ genera) for each substage
```

```
nrow(CAM_low_matrix_gen_final) # 122
nrow(CAM_mid_matrix_gen_final) # 161
nrow(CAM_up_matrix_gen_final) # 198
nrow(MAA_low_matrix_gen_final) # 118
nrow(MAA_up_matrix_gen_final) # 75

# check number of genera in each substage matrix
ncol(CAM_low_matrix_gen_final) # 148
ncol(CAM_mid_matrix_gen_final) # 188
ncol(CAM_up_matrix_gen_final) # 344
ncol(MAA_low_matrix_gen_final) # 305
ncol(MAA_up_matrix_gen_final) # 331

   # The number of unique locations is significantly larger for the Mid and Up CAM

write.csv(CAM_low_matrix_gen_final,file= "R_Config_for_EN_OUTPUT/Low
        CAM/CAM_low_matrix_gen_final.csv")
write.csv(CAM_mid_matrix_gen_final,file= "R_Config_for_EN_OUTPUT/Mid
        CAM/CAM_mid_matrix_gen_final.csv")
write.csv(CAM_up_matrix_gen_final,file= "R_Config_for_EN_OUTPUT/Up
        CAM/CAM_up_matrix_gen_final.csv")
write.csv(MAA_low_matrix_gen_final,file= "R_Config_for_EN_OUTPUT/Low
        MAA/MAA_low_matrix_gen_final.csv")
write.csv(MAA_up_matrix_gen_final,file= "R_Config_for_EN_OUTPUT/Up
        MAA/MAA_up_matrix_gen_final.csv")

## Create matrix with grid cells containing less than 3 occ removed for each individual substage

#Lower CAM: For genera
ncol(CAM_low_matrix_gen_final)
CAM_low_numb_occ_gen <- data.frame(rowSums(CAM_low_matrix_gen_final[,2:148])) # sum rows to
        get total # occ
CAM_low_matrix_gen_final_pre <- CAM_low_matrix_gen_final # create new matrix name to use
CAM_low_matrix_gen_final_pre$nOccGen <- CAM_low_numb_occ_gen # Add column of total occ
        number
colnames(CAM_low_matrix_gen_final_pre)
CAM_low_matrix_gen_great3 <- subset(CAM_low_matrix_gen_final_pre, nOccGen >= 3) # subset out
        all grid cells with greater 3 occ
ncol(CAM_low_matrix_gen_great3)
CAM_low_matrix_gen_great3 <- CAM_low_matrix_gen_great3[,1:148] # remove occ number column
colnames(CAM_low_matrix_gen_great3)

#Middle CAM: For genera
ncol(CAM_mid_matrix_gen_final)
CAM_mid_numb_occ_gen <- data.frame(rowSums(CAM_mid_matrix_gen_final[,2:188])) # sum rows to
        get total # occ
CAM_mid_matrix_gen_final_pre <- CAM_mid_matrix_gen_final # create new matrix name to use
CAM_mid_matrix_gen_final_pre$nOccGen <- CAM_mid_numb_occ_gen # Add column of total occ
        number
colnames(CAM_mid_matrix_gen_final_pre)
```

```
CAM_mid_matrix_gen_great3 <- subset(CAM_mid_matrix_gen_final_pre, nOccGen >= 3) # subset out
        all grid cells with greater 3 occ
ncol(CAM_mid_matrix_gen_great3)
CAM_mid_matrix_gen_great3 <- CAM_mid_matrix_gen_great3[,1:188] # remove occ number column
colnames(CAM_mid_matrix_gen_great3)

#Upper CAM: For genera
ncol(CAM_up_matrix_gen_final)
CAM_up_numb_occ_gen <- data.frame(rowSums(CAM_up_matrix_gen_final[,2:344])) # sum rows to
        get total # occ
CAM_up_matrix_gen_final_pre <- CAM_up_matrix_gen_final # create new matrix name to use
CAM_up_matrix_gen_final_pre$nOccGen <- CAM_up_numb_occ_gen # Add column of total occ
        number
colnames(CAM_up_matrix_gen_final_pre)
CAM_up_matrix_gen_great3 <- subset(CAM_up_matrix_gen_final_pre, nOccGen >= 3) # subset out all
        grid cells with greater 3 occ
ncol(CAM_up_matrix_gen_great3)
CAM_up_matrix_gen_great3 <- CAM_up_matrix_gen_great3[,1:344] # remove occ number column
colnames(CAM_up_matrix_gen_great3)

#Lower MAA: For genera
ncol(MAA_low_matrix_gen_final)
MAA_low_numb_occ_gen <- data.frame(rowSums(MAA_low_matrix_gen_final[,2:305])) # sum rows to
        get total # occ
MAA_low_matrix_gen_final_pre <- MAA_low_matrix_gen_final # create new matrix name to use
MAA_low_matrix_gen_final_pre$nOccGen <- MAA_low_numb_occ_gen # Add column of total occ
        number
colnames(MAA_low_matrix_gen_final_pre)
MAA_low_matrix_gen_great3 <- subset(MAA_low_matrix_gen_final_pre, nOccGen >= 3) # subset out
        all grid cells with greater 3 occ
ncol(MAA_low_matrix_gen_great3)
MAA_low_matrix_gen_great3 <- MAA_low_matrix_gen_great3[,1:305] # remove occ number column
colnames(MAA_low_matrix_gen_great3)

#Upper MAA: For genera
ncol(MAA_up_matrix_gen_final)
MAA_up_numb_occ_gen <- data.frame(rowSums(MAA_up_matrix_gen_final[,2:331])) # sum rows to
        get total # occ
MAA_up_matrix_gen_final_pre <- MAA_up_matrix_gen_final # create new matrix name to use
MAA_up_matrix_gen_final_pre$nOccGen <- MAA_up_numb_occ_gen # Add column of total occ
        number
colnames(MAA_up_matrix_gen_final_pre)
MAA_up_matrix_gen_great3 <- subset(MAA_up_matrix_gen_final_pre, nOccGen >= 3) # subset out all
        grid cells with greater 3 occ
ncol(MAA_up_matrix_gen_great3)
MAA_up_matrix_gen_great3 <- MAA_up_matrix_gen_great3[,1:331] # remove occ number column
colnames(MAA_up_matrix_gen_great3)

# export new csv files of the dataset with grids containing <3 occ removed
write.csv(CAM_low_matrix_gen_great3,file= "R_Config_for_EN_OUTPUT/Low
        CAM/CAM_low_matrix_gen_great3.csv")
```

```
write.csv(CAM_mid_matrix_gen_great3,file= "R_Config_for_EN_OUTPUT/Mid
        CAM/CAM_mid_matrix_gen_great3.csv")
write.csv(CAM_up_matrix_gen_great3,file= "R_Config_for_EN_OUTPUT/Up
        CAM/CAM_up_matrix_gen_great3.csv")
write.csv(MAA_low_matrix_gen_great3,file= "R_Config_for_EN_OUTPUT/Low
        MAA/MAA_low_matrix_gen_great3.csv")
write.csv(MAA_up_matrix_gen_great3,file= "R_Config_for_EN_OUTPUT/Up
        MAA/MAA_up_matrix_gen_great3.csv")

#### Get attribute files for EDENetwork analysis for each Substage ####

# subset out the different substages for genera
CAM_low_gen_atrib <- subset(trunk60_gen_atrib, Age == "LC")
CAM_mid_gen_atrib <- subset(trunk60_gen_atrib, Age == "MC")
CAM_up_gen_atrib <- subset(trunk60_gen_atrib, Age == "UC")
MAA_low_gen_atrib <- subset(trunk60_gen_atrib, Age == "LM")
MAA_up_gen_atrib <- subset(trunk60_gen_atrib, Age == "UM")

#write csv files of attributes tables for substages at genus level
write.csv(CAM_low_gen_atrib,file="R_Config_for_EN_OUTPUT/Low
        CAM/CAM_low_gen_atrib.csv",row.names=FALSE)
write.csv(CAM_mid_gen_atrib,file="R_Config_for_EN_OUTPUT/Mid
        CAM/CAM_mid_gen_atrib.csv",row.names=FALSE)
write.csv(CAM_up_gen_atrib,file="R_Config_for_EN_OUTPUT/Up
        CAM/CAM_up_gen_atrib.csv",row.names=FALSE)
write.csv(MAA_low_gen_atrib,file="R_Config_for_EN_OUTPUT/Low
        MAA/MAA_low_gen_atrib.csv",row.names=FALSE)
write.csv(MAA_up_gen_atrib,file="R_Config_for_EN_OUTPUT/Up
        MAA/MAA_up_gen_atrib.csv",row.names=FALSE)

### Make a attributes table for the matrix with grids containing <3 occ removed for substages

# subset out the different substages for genera of greater than 3 occ
CAM_low_gen_great3_atrib <- subset(trunk60_matrix_gen_great3_atrib, Age == "LC")
CAM_mid_gen_great3_atrib <- subset(trunk60_matrix_gen_great3_atrib, Age == "MC")
CAM_up_gen_great3_atrib <- subset(trunk60_matrix_gen_great3_atrib, Age == "UC")
MAA_low_gen_great3_atrib <- subset(trunk60_matrix_gen_great3_atrib, Age == "LM")
MAA_up_gen_great3_atrib <- subset(trunk60_matrix_gen_great3_atrib, Age == "UM")

#write csv files of attributes tables for substages at genus level
write.csv(CAM_low_gen_great3_atrib,file="R_Config_for_EN_OUTPUT/Low
        CAM/CAM_low_gen_great3_atrib.csv",row.names=FALSE)
write.csv(CAM_mid_gen_great3_atrib,file="R_Config_for_EN_OUTPUT/Mid
        CAM/CAM_mid_gen_great3_atrib.csv",row.names=FALSE)
write.csv(CAM_up_gen_great3_atrib,file="R_Config_for_EN_OUTPUT/Up
        CAM/CAM_up_gen_great3_atrib.csv",row.names=FALSE)
write.csv(MAA_low_gen_great3_atrib,file="R_Config_for_EN_OUTPUT/Low
        MAA/MAA_low_gen_great3_atrib.csv",row.names=FALSE)
write.csv(MAA_up_gen_great3_atrib,file="R_Config_for_EN_OUTPUT/Up
        MAA/MAA_up_gen_great3_atrib.csv",row.names=FALSE)
```

```
#### Get average coordinates files for EDENetwork analysis for each Substage ####

# subset out the different substages for genera
CAM_low_gen_avg_loc <- subset(trunk60_gen_avg_loc, Age == "LC")
CAM_mid_gen_avg_loc <- subset(trunk60_gen_avg_loc, Age == "MC")
CAM_up_gen_avg_loc <- subset(trunk60_gen_avg_loc, Age == "UC")
MAA_low_gen_avg_loc <- subset(trunk60_gen_avg_loc, Age == "LM")
MAA_up_gen_avg_loc <- subset(trunk60_gen_avg_loc, Age == "UM")

#write csv files of attributes tables for substages at genus level
write.csv(CAM_low_gen_avg_loc,file="R_Config_for_EN_OUTPUT/Low
        CAM/CAM_low_gen_avg_loc.csv",row.names=FALSE)
write.csv(CAM_mid_gen_avg_loc,file="R_Config_for_EN_OUTPUT/Mid
        CAM/CAM_mid_gen_avg_loc.csv",row.names=FALSE)
write.csv(CAM_up_gen_avg_loc,file="R_Config_for_EN_OUTPUT/Up
        CAM/CAM_up_gen_avg_loc.csv",row.names=FALSE)
write.csv(MAA_low_gen_avg_loc,file="R_Config_for_EN_OUTPUT/Low
        MAA/MAA_low_gen_avg_loc.csv",row.names=FALSE)
write.csv(MAA_up_gen_avg_loc,file="R_Config_for_EN_OUTPUT/Up
        MAA/MAA_up_gen_avg_loc.csv",row.names=FALSE)

### Make a average coordinates table for the matrix with grids containing <3 occ removed for substages

# subset out the different substages for genera of greater than 3 occ
CAM_low_gen_great3_avgloc <- subset(trunk60_matrix_gen_great3_avgloc, Age == "LC")
CAM_mid_gen_great3_avgloc <- subset(trunk60_matrix_gen_great3_avgloc, Age == "MC")
CAM_up_gen_great3_avgloc <- subset(trunk60_matrix_gen_great3_avgloc, Age == "UC")
MAA_low_gen_great3_avgloc <- subset(trunk60_matrix_gen_great3_avgloc, Age == "LM")
MAA_up_gen_great3_avgloc <- subset(trunk60_matrix_gen_great3_avgloc, Age == "UM")

#write csv files of attributes tables for substages at genus level
write.csv(CAM_low_gen_great3_avgloc,file="R_Config_for_EN_OUTPUT/Low
        CAM/CAM_low_gen_great3_avgloc.csv",row.names=FALSE)
write.csv(CAM_mid_gen_great3_avgloc,file="R_Config_for_EN_OUTPUT/Mid
        CAM/CAM_mid_gen_great3_avgloc.csv",row.names=FALSE)
write.csv(CAM_up_gen_great3_avgloc,file="R_Config_for_EN_OUTPUT/Up
        CAM/CAM_up_gen_great3_avgloc.csv",row.names=FALSE)
write.csv(MAA_low_gen_great3_avgloc,file="R_Config_for_EN_OUTPUT/Low
        MAA/MAA_low_gen_great3_avgloc.csv",row.names=FALSE)
write.csv(MAA_up_gen_great3_avgloc,file="R_Config_for_EN_OUTPUT/Up
        MAA/MAA_up_gen_great3_avgloc.csv",row.names=FALSE)

#### CREATE MATRIX OF SPECIES LEVEL OCC INFORMATION FOR EDENETWORKS
        ANALYSIS ####

# create lookup table of unique location ID (lat,long) to use as "sites" in EN
#lookup <- unique(join60_data[,c(25,27)])

# function to transform into a pres-abs matrix of locations with genus names
substg_spp_grid <- function(data){
  data1 <- as.data.frame(subset(data)[,c(1,8)]) # subset out just the genus name and location information
```

```
  colnames(data1) <- c("PageName","Species_Name") # Give the columns names
  data2 <- dcast(data1, PageName~Species_Name, length) # transform into a pres-abs matrix

}

colnames(CAM_low_spp)

CAM_low_matrix_spp <- substg_spp_grid(CAM_low_spp)
CAM_mid_matrix_spp <- substg_spp_grid(CAM_mid_spp)
CAM_up_matrix_spp <- substg_spp_grid(CAM_up_spp)
MAA_low_matrix_spp <- substg_spp_grid(MAA_low_spp)
MAA_up_matrix_spp <- substg_spp_grid(MAA_up_spp)

# find number of spp in each substage matrix
ncol(CAM_low_matrix_spp) # 135
ncol(CAM_mid_matrix_spp) # 211
ncol(CAM_up_matrix_spp) # 586
ncol(MAA_low_matrix_spp) # 381
ncol(MAA_up_matrix_spp) # 452

# function to configure the SPP-level matrix into the correct format for EDENETWORKS and add the
        grid cell info
config_matrix_spp <- function(data){
  # base1 <- (merge(data, lookup, by = 'LocationKey',all.X=TRUE, all.y=FALSE))
  x <- nrow(data)
  data$SampleKey <-seq(1:x)
  a <- ncol(data)
  b <- a - 1
  base2 <- data[,c(1,a,2:b)] # This part would need to change if using lat/long as unique ID
  base3 <- data.frame(base2[,-1], row.names = base2[,1]) # make the first column with pagename the
        index
}

# create matrix in correct config with grid cell names include as first
CAM_low_matrix_spp_final <- config_matrix_spp(CAM_low_matrix_spp)
CAM_mid_matrix_spp_final <- config_matrix_spp(CAM_mid_matrix_spp)
CAM_up_matrix_spp_final <- config_matrix_spp(CAM_up_matrix_spp)
MAA_low_matrix_spp_final <- config_matrix_spp(MAA_low_matrix_spp)
MAA_up_matrix_spp_final <- config_matrix_spp(MAA_up_matrix_spp)

head(CAM_low_matrix_spp_final)

#function to change from abundance to pres-abs matrix
pres_ab <- function(data){
  numeric_cols <- vapply(data, is.numeric, logical(1))
  data[numeric_cols] <- as.integer(data[numeric_cols] != 0)
  data
}

# create pres-abs matrix from the abundance matrix
CAM_low_matrix_spp_final <- pres_ab(CAM_low_matrix_spp_final)
```

```
CAM_mid_matrix_spp_final <- pres_ab(CAM_mid_matrix_spp_final)
CAM_up_matrix_spp_final <- pres_ab(CAM_up_matrix_spp_final)
MAA_low_matrix_spp_final <- pres_ab(MAA_low_matrix_spp_final)
MAA_up_matrix_spp_final <- pres_ab(MAA_up_matrix_spp_final)

# check number of spp in each substage matrix
ncol(CAM_low_matrix_spp) # 135
ncol(CAM_mid_matrix_spp) # 211
ncol(CAM_up_matrix_spp) # 586
ncol(MAA_low_matrix_spp) # 381
ncol(MAA_up_matrix_spp) # 452

# check the number of rows (# of unique locations w/ species) for each substage
nrow(CAM_low_matrix_spp_final) # 107
nrow(CAM_mid_matrix_spp_final) # 156
nrow(CAM_up_matrix_spp_final) # 188
nrow(MAA_low_matrix_spp_final) # 116
nrow(MAA_up_matrix_spp_final) # 73

# The number of unique locations is significantly larger for the Mid and Up CAM

# write csv of matrices for each substage
write.csv(CAM_low_matrix_spp_final,file= "R_Config_for_EN_OUTPUT/Low
        CAM/CAM_low_matrix_spp_final.csv")
write.csv(CAM_mid_matrix_spp_final,file= "R_Config_for_EN_OUTPUT/Mid
        CAM/CAM_mid_matrix_spp_final.csv")
write.csv(CAM_up_matrix_spp_final,file= "R_Config_for_EN_OUTPUT/Up
        CAM/CAM_up_matrix_spp_final.csv")
write.csv(MAA_low_matrix_spp_final,file= "R_Config_for_EN_OUTPUT/Low
        MAA/MAA_low_matrix_spp_final.csv")
write.csv(MAA_up_matrix_spp_final,file= "R_Config_for_EN_OUTPUT/Up
        MAA/MAA_up_matrix_spp_final.csv")

## Create matrix with grid cells containing less than 3 occ removed for each individual substage

#Lower CAM: For species
ncol(CAM_low_matrix_spp_final)
CAM_low_numb_occ_spp <- data.frame(rowSums(CAM_low_matrix_spp_final[,2:135])) # sum rows to
        get total # occ
CAM_low_matrix_spp_final_pre <- CAM_low_matrix_spp_final # create new matrix name to use
CAM_low_matrix_spp_final_pre$nOccGen <- CAM_low_numb_occ_spp # Add column of total occ
        number
colnames(CAM_low_matrix_spp_final_pre)
CAM_low_matrix_spp_great3 <- subset(CAM_low_matrix_spp_final_pre, nOccGen >= 3) # subset out
        all grid cells with greater 3 occ
ncol(CAM_low_matrix_spp_great3)
CAM_low_matrix_spp_great3 <- CAM_low_matrix_spp_great3[,1:135] # remove occ number column
colnames(CAM_low_matrix_spp_great3)

#Middle CAM: For species
ncol(CAM_mid_matrix_spp_final)
```

```
CAM_mid_numb_occ_spp <- data.frame(rowSums(CAM_mid_matrix_spp_final[,2:211])) # sum rows to
        get total # occ
CAM_mid_matrix_spp_final_pre <- CAM_mid_matrix_spp_final # create new matrix name to use
CAM_mid_matrix_spp_final_pre$nOccGen <- CAM_mid_numb_occ_spp # Add column of total occ
        number
colnames(CAM_mid_matrix_spp_final_pre)
CAM_mid_matrix_spp_great3 <- subset(CAM_mid_matrix_spp_final_pre, nOccGen >= 3) # subset out
        all grid cells with greater 3 occ
ncol(CAM_mid_matrix_spp_great3)
CAM_mid_matrix_spp_great3 <- CAM_mid_matrix_spp_great3[,1:211] # remove occ number column
colnames(CAM_mid_matrix_spp_great3)

#Upper CAM: For species
ncol(CAM_up_matrix_spp_final)
CAM_up_numb_occ_spp <- data.frame(rowSums(CAM_up_matrix_spp_final[,2:586])) # sum rows to
        get total # occ
CAM_up_matrix_spp_final_pre <- CAM_up_matrix_spp_final # create new matrix name to use
CAM_up_matrix_spp_final_pre$nOccGen <- CAM_up_numb_occ_spp # Add column of total occ
        number
colnames(CAM_up_matrix_spp_final_pre)
CAM_up_matrix_spp_great3 <- subset(CAM_up_matrix_spp_final_pre, nOccGen >= 3) # subset out all
        grid cells with greater 3 occ
ncol(CAM_up_matrix_spp_great3)
CAM_up_matrix_spp_great3 <- CAM_up_matrix_spp_great3[,1:586] # remove occ number column
colnames(CAM_up_matrix_spp_great3)

#Lower MAA: For species
ncol(MAA_low_matrix_spp_final)
MAA_low_numb_occ_spp <- data.frame(rowSums(MAA_low_matrix_spp_final[,2:381])) # sum rows to
        get total # occ
MAA_low_matrix_spp_final_pre <- MAA_low_matrix_spp_final # create new matrix name to use
MAA_low_matrix_spp_final_pre$nOccGen <- MAA_low_numb_occ_spp # Add column of total occ
        number
colnames(MAA_low_matrix_spp_final_pre)
MAA_low_matrix_spp_great3 <- subset(MAA_low_matrix_spp_final_pre, nOccGen >= 3) # subset out
        all grid cells with greater 3 occ
ncol(MAA_low_matrix_spp_great3)
MAA_low_matrix_spp_great3 <- MAA_low_matrix_spp_great3[,1:382] # remove occ number column
colnames(MAA_low_matrix_spp_great3)

#Upper MAA: For species
ncol(MAA_up_matrix_spp_final)
MAA_up_numb_occ_spp <- data.frame(rowSums(MAA_up_matrix_spp_final[,2:452])) # sum rows to
        get total # occ
MAA_up_matrix_spp_final_pre <- MAA_up_matrix_spp_final # create new matrix name to use
MAA_up_matrix_spp_final_pre$nOccGen <- MAA_up_numb_occ_spp # Add column of total occ
        number
colnames(MAA_up_matrix_spp_final_pre)
MAA_up_matrix_spp_great3 <- subset(MAA_up_matrix_spp_final_pre, nOccGen >= 3) # subset out all
        grid cells with greater 3 occ
ncol(MAA_up_matrix_spp_great3)
```

```
MAA_up_matrix_spp_great3 <- MAA_up_matrix_spp_great3[,1:452] # remove occ number column
colnames(MAA_up_matrix_spp_great3)


# export new csv of the dataset with grids containing <3 occ removed for each substage and spp
write.csv(CAM_low_matrix_spp_great3,file= "R_Config_for_EN_OUTPUT/Low
        CAM/CAM_low_matrix_spp_great3.csv")
write.csv(CAM_mid_matrix_spp_great3,file= "R_Config_for_EN_OUTPUT/Mid
        CAM/CAM_mid_matrix_spp_great3.csv")
write.csv(CAM_up_matrix_spp_great3,file= "R_Config_for_EN_OUTPUT/Up
        CAM/CAM_up_matrix_spp_great3.csv")
write.csv(MAA_low_matrix_spp_great3,file= "R_Config_for_EN_OUTPUT/Low
        MAA/MAA_low_matrix_spp_great3.csv")
write.csv(MAA_up_matrix_spp_great3,file= "R_Config_for_EN_OUTPUT/Up
        MAA/MAA_up_matrix_spp_great3.csv")

#### Get attribute files for EDENetwork analysis for each Substage ####

# subset out the different substages for species
CAM_low_spp_atrib <- subset(trunk60_spp_atrib, Age == "LC")
CAM_mid_spp_atrib <- subset(trunk60_spp_atrib, Age == "MC")
CAM_up_spp_atrib <- subset(trunk60_spp_atrib, Age == "UC")
MAA_low_spp_atrib <- subset(trunk60_spp_atrib, Age == "LM")
MAA_up_spp_atrib <- subset(trunk60_spp_atrib, Age == "UM")

#write csv files of attributes tables for substages at species level
write.csv(CAM_low_spp_atrib,file="R_Config_for_EN_OUTPUT/Low
        CAM/CAM_low_spp_atrib.csv",row.names=FALSE)
write.csv(CAM_mid_spp_atrib,file="R_Config_for_EN_OUTPUT/Mid
        CAM/CAM_mid_spp_atrib.csv",row.names=FALSE)
write.csv(CAM_up_spp_atrib,file="R_Config_for_EN_OUTPUT/Up
        CAM/CAM_up_spp_atrib.csv",row.names=FALSE)
write.csv(MAA_low_spp_atrib,file="R_Config_for_EN_OUTPUT/Low
        MAA/MAA_low_spp_atrib.csv",row.names=FALSE)
write.csv(MAA_up_spp_atrib,file="R_Config_for_EN_OUTPUT/Up
        MAA/MAA_up_spp_atrib.csv",row.names=FALSE)

### Make attributes tables of substages with only >3 occ

# subset out the different substages for species
CAM_low_spp_great3_atrib <- subset(trunk60_matrix_spp_great3_atrib, Age == "LC")
CAM_mid_spp_great3_atrib <- subset(trunk60_matrix_spp_great3_atrib, Age == "MC")
CAM_up_spp_great3_atrib <- subset(trunk60_matrix_spp_great3_atrib, Age == "UC")
MAA_low_spp_great3_atrib <- subset(trunk60_matrix_spp_great3_atrib, Age == "LM")
MAA_up_spp_great3_atrib <- subset(trunk60_matrix_spp_great3_atrib, Age == "UM")

#write csv files of attributes tables for substages at species level
write.csv(CAM_low_spp_great3_atrib,file="R_Config_for_EN_OUTPUT/Low
        CAM/CAM_low_spp_great3_atrib.csv",row.names=FALSE)
write.csv(CAM_mid_spp_great3_atrib,file="R_Config_for_EN_OUTPUT/Mid
        CAM/CAM_mid_spp_great3_atrib.csv",row.names=FALSE)
```

```
write.csv(CAM_up_spp_great3_atrib,file="R_Config_for_EN_OUTPUT/Up
        CAM/CAM_up_spp_great3_atrib.csv",row.names=FALSE)
write.csv(MAA_low_spp_great3_atrib,file="R_Config_for_EN_OUTPUT/Low
        MAA/MAA_low_spp_great3_atrib.csv",row.names=FALSE)
write.csv(MAA_up_spp_great3_atrib,file="R_Config_for_EN_OUTPUT/Up
        MAA/MAA_up_spp_great3_atrib.csv",row.names=FALSE)

#### Get average coordinates files for EDENetwork analysis for each Substage species level ####

# subset out the different substages for genera
CAM_low_spp_avg_loc <- subset(trunk60_spp_avg_loc, Age == "LC")
CAM_mid_spp_avg_loc <- subset(trunk60_spp_avg_loc, Age == "MC")
CAM_up_spp_avg_loc <- subset(trunk60_spp_avg_loc, Age == "UC")
MAA_low_spp_avg_loc <- subset(trunk60_spp_avg_loc, Age == "LM")
MAA_up_spp_avg_loc <- subset(trunk60_spp_avg_loc, Age == "UM")

#write csv files of attributes tables for substages at genus level
write.csv(CAM_low_spp_avg_loc,file="R_Config_for_EN_OUTPUT/Low
        CAM/CAM_low_spp_avg_loc.csv",row.names=FALSE)
write.csv(CAM_mid_spp_avg_loc,file="R_Config_for_EN_OUTPUT/Mid
        CAM/CAM_mid_spp_avg_loc.csv",row.names=FALSE)
write.csv(CAM_up_spp_avg_loc,file="R_Config_for_EN_OUTPUT/Up
        CAM/CAM_up_spp_avg_loc.csv",row.names=FALSE)
write.csv(MAA_low_spp_avg_loc,file="R_Config_for_EN_OUTPUT/Low
        MAA/MAA_low_spp_avg_loc.csv",row.names=FALSE)
write.csv(MAA_up_spp_avg_loc,file="R_Config_for_EN_OUTPUT/Up
        MAA/MAA_up_gen_spp_loc.csv",row.names=FALSE)

### Make a average coordinates table for the matrix with grids containing <3 occ removed for substages
        species level
# subset out the different substages for genera of greater than 3 occ
CAM_low_spp_great3_avgloc <- subset(trunk60_matrix_spp_great3_avgloc, Age == "LC")
CAM_mid_spp_great3_avgloc <- subset(trunk60_matrix_spp_great3_avgloc, Age == "MC")
CAM_up_spp_great3_avgloc <- subset(trunk60_matrix_spp_great3_avgloc, Age == "UC")
MAA_low_spp_great3_avgloc <- subset(trunk60_matrix_spp_great3_avgloc, Age == "LM")
MAA_up_spp_great3_avgloc <- subset(trunk60_matrix_spp_great3_avgloc, Age == "UM")

#write csv files of attributes tables for substages at genus level
write.csv(CAM_low_spp_great3_avgloc,file="R_Config_for_EN_OUTPUT/Low
        CAM/CAM_low_spp_great3_avgloc.csv",row.names=FALSE)
write.csv(CAM_mid_spp_great3_avgloc,file="R_Config_for_EN_OUTPUT/Mid
        CAM/CAM_mid_spp_great3_avgloc.csv",row.names=FALSE)
write.csv(CAM_up_spp_great3_avgloc,file="R_Config_for_EN_OUTPUT/Up
        CAM/CAM_up_spp_great3_avgloc.csv",row.names=FALSE)
write.csv(MAA_low_spp_great3_avgloc,file="R_Config_for_EN_OUTPUT/Low
        MAA/MAA_low_spp_great3_avgloc.csv",row.names=FALSE)
write.csv(MAA_up_spp_great3_avgloc,file="R_Config_for_EN_OUTPUT/Up
        MAA/MAA_up_spp_great3_avgloc.csv",row.names=FALSE)

#### CREATE SUMMARY TABLE OF THE NUMBER OF SPP/GEN AND UNIQUE LOCATIONS
        IN EACH SUBSTAGE ####
```

```
Substage <- c("Low CAM","Mid CAM", "Up CAM", "Low MAA", "Up MAA")

UGenLoc <- c(nrow(CAM_low_matrix_gen_final),nrow(CAM_mid_matrix_gen_final),
             nrow(CAM_up_matrix_gen_final),nrow(MAA_low_matrix_gen_final),
             nrow(MAA_up_matrix_gen_final))

GenGrid <- c(nrow(data.frame(unique(CAM_low_matrix_gen_final$PageName))),
      nrow(data.frame(unique(CAM_mid_matrix_gen_final$PageName))),
      nrow(data.frame(unique(CAM_up_matrix_gen_final$PageName))),
      nrow(data.frame(unique(MAA_low_matrix_gen_final$PageName))),
      nrow(data.frame(unique(MAA_up_matrix_gen_final$PageName))))

GenOcc <- c(nrow(CAM_low),
      nrow(CAM_mid),
      nrow(CAM_up),
      nrow(MAA_low),
      nrow(MAA_up))

Genera <- (c(ncol(CAM_low_matrix_gen_final),ncol(CAM_mid_matrix_gen_final),
      ncol(CAM_up_matrix_gen_final),ncol(MAA_low_matrix_gen_final),
      ncol(MAA_up_matrix_gen_final))) -2

USppLoc <- c(nrow(CAM_low_matrix_spp_final),nrow(CAM_mid_matrix_spp_final),
             nrow(CAM_up_matrix_spp_final),nrow(MAA_low_matrix_spp_final),
             nrow(MAA_up_matrix_spp_final))

SppGrid <- c(nrow(data.frame(unique(CAM_low_matrix_spp_final$PageName))),
      nrow(data.frame(unique(CAM_mid_matrix_spp_final$PageName))),
      nrow(data.frame(unique(CAM_up_matrix_spp_final$PageName))),
      nrow(data.frame(unique(MAA_low_matrix_spp_final$PageName))),
      nrow(data.frame(unique(MAA_up_matrix_spp_final$PageName))))

SppOcc <- c(nrow(CAM_low_spp),
   nrow(CAM_mid_spp),
   nrow(CAM_up_spp),
   nrow(MAA_low_spp),
   nrow(MAA_up_spp))

Species <- (c(ncol(CAM_low_matrix_spp_final),ncol(CAM_mid_matrix_spp_final),
      ncol(CAM_up_matrix_spp_final),ncol(MAA_low_matrix_spp_final),
      ncol(MAA_up_matrix_spp_final))) -2

summary <- as.data.frame(cbind(Substage, GenOcc, UGenLoc, GenGrid,  Genera, SppOcc, USppLoc,
                    SppGrid,
                    Species))
colnames(summary) <- c("Substage","TotGenOcc","UniqueGenLocations", "GenGridPixels", "#Genera",
              "TotSppOcc","UniqueSppLocations","SppGridPixels","#Species")
summary
write.csv(summary, file = "R_Config_for_EN_OUTPUT/SubstageSummary.csv")
```

# The output shows that there are significantly more unique locations in the Mid and Up CAM
# but not a much higher number of occupied grid cells. There are far fewer total gen/spp
# occ points in the Low CAM, but not a much lower number of unique spp and genera relative
# to other substages. The Up MAA has the fewest total grid cells occupied, despite having
# a realtively high number of total spp/gen occ points and unique spp/gen.

*Code for Conducting Network Randomization Comparisons*

```
setwd("C:/Users/ceara/Documents/Province Project/Occurrence Data/VettedCombinedDB")


#### Lower Campanian ####
# read in the network information data from the pajek file

library(igraph)
library(tnet)

low_cam_percol_net <- read.graph('EDENetwork_outputs/Low
        CAM/CAM_low_matrix_gen_great3_thresholded_at_061.net', format = "pajek")

typeof(low_cam_percol_net)

cc_low_cam_percol_net <- transitivity(low_cam_percol_net)

# convert to a i,j,w weighted edgelist (NOTE: Could also do this by getting distance matrix and then
                        # converting to NW using graph.adjacency() then convert using get.data.frame())

### Get the network data into the right format for tnet to calculate random nws (need the nodes to be
        numeric, not grid names)
low_cam_percol_net2 <- cbind(get.edgelist(low_cam_percol_net, names=FALSE),
        E(low_cam_percol_net)$weight)

# run this since the network is undirected, duplicates ties to justify that there is no direction
if(!is.directed(low_cam_percol_net))
  low_cam_percol_net2 <- symmetrise_w(low_cam_percol_net2)

# make sure conforms to tnet format
low_cam_percol_net2 <- as.tnet(low_cam_percol_net2, type="weighted one-mode tnet")

### Convert from tnet format to igraph (this should remove any zero weights, which essentially indicates
        identical connection but only 2 present...)
# Create igraph object
low_cam_percol_net3 <- tnet_igraph(low_cam_percol_net2,type="weighted one-mode tnet")

# Get summary statistics
summary(low_cam_percol_net3)

# check that the cc is still close
cc_low_cam_percol_net2 <- transitivity(low_cam_percol_net3)
     ### loss of the 2 identical links maybe changed, but not much...

## Plot the network

# Get layout of observed network and plot it... need to look into this more...
layout <- layout.fruchterman.reingold(low_cam_percol_net3)
```

```
plot(low_cam_percol_net3, main="Observed Network", layout=layout,
        vertex.label=1:length(low_cam_percol_net3), edge.width=E(low_cam_percol_net3)$weight,
        edge.label=E(low_cam_percol_net3)$weight)

# Create random network from original network

# create function to reshuffle by links (reshuffles weighted links while maintaining negree)
fun <- function(x) {
  rg_reshuffling_w(x, option="links")
}

# set number of times to replicate
n = 1000

# run analysis to create array with different randomized networks
low_cam_percol_1000rand <- replicate(n, fun(low_cam_percol_net2), simplify=FALSE)

# turn into igraph objects
low_cam_percol_1000rand_igraph <- lapply(low_cam_percol_1000rand, function(x)
        tnet_igraph((x),type="weighted one-mode tnet"))

# run cc analysis
cc_low_cam_percol_1000rand <- lapply(low_cam_percol_1000rand_igraph,function(x) transitivity(x))

# get random networks cc average
cc_avg_low_cam_percol_1000rand <- mean(unlist(cc_low_cam_percol_1000rand))

# Calculate standad deviation of the random dist
cc_sd_low_cam_percol_1000rand <- sd(unlist(cc_low_cam_percol_1000rand))

# look at random nw cc distributions
cc_dist_low_cam_percol_1000rand <- (unlist(cc_low_cam_percol_1000rand))
hist(cc_dist_low_cam_percol_1000rand)

### Run statistical test on the real cc using random networks

# H0: cc[real] = mean of cc[rand]

# H1: cc[real] != mean of cc[rand]

# alpha = 0.05

t.test(cc_low_cam_percol_net2, mu = cc_avg_low_cam_percol_1000rand, alternative = 'less')

#### Middle Campanian ####

mid_cam_percol_net <- read.graph('EDENetwork_outputs/Mid
        CAM/CAM_mid_matrix_gen_great3_thresholded_at_057.net', format = "pajek")

typeof(mid_cam_percol_net)
```

```
cc_mid_cam_percol_net <- transitivity(mid_cam_percol_net)

# convert to a i,j,w weighted edgelist (NOTE: Could also do this by getting distance matrix and then
# converting to NW using graph.adjacency() then convert using get.data.frame())

### Get the network data into the right format for tnet to calculate random nws (need the nodes to be
        numeric, not grid names)
mid_cam_percol_net2 <- cbind(get.edgelist(mid_cam_percol_net, names=FALSE),
        E(mid_cam_percol_net)$weight)

# run this since the network is undirected, duplicates ties to justify that there is no direction
if(!is.directed(mid_cam_percol_net))
  mid_cam_percol_net2 <- symmetrise_w(mid_cam_percol_net2)

# make sure conforms to tnet format
mid_cam_percol_net2 <- as.tnet(mid_cam_percol_net2, type="weighted one-mode tnet")

### Convert from tnet format to igraph (this should remove any zero weights, which essentially indicates
        identical connection but only 2 present...)
# Create igraph object
mid_cam_percol_net3 <- tnet_igraph(mid_cam_percol_net2,type="weighted one-mode tnet")

# Get summary statistics
summary(mid_cam_percol_net3)

# check that the cc is still close
cc_mid_cam_percol_net2 <- transitivity(mid_cam_percol_net3)
### exactly the same!

## Plot the network

# Get layout of observed network and plot it... need to look into this more...
layout <- layout.fruchterman.reingold(mid_cam_percol_net3)

plot(mid_cam_percol_net3, main="Observed Network", layout=layout,
        vertex.label=1:length(mid_cam_percol_net3), edge.width=E(mid_cam_percol_net3)$weight,
        edge.label=E(mid_cam_percol_net3)$weight)

# Create random network from original network

# create function to reshuffle by links (reshuffles weighted links while maintaining negree)
fun <- function(x) {
  rg_reshuffling_w(x, option="links")
}

# set number of times to replicate
n = 1000

# run analysis to create array with different randomized networks
mid_cam_percol_1000rand <- replicate(n, fun(mid_cam_percol_net2), simplify=FALSE)
```

```
# turn into igraph objects
mid_cam_percol_1000rand_igraph <- lapply(mid_cam_percol_1000rand, function(x)
        tnet_igraph((x),type="weighted one-mode tnet"))

# run cc analysis
cc_mid_cam_percol_1000rand <- lapply(mid_cam_percol_1000rand_igraph,function(x) transitivity(x))

# get random networks cc average
cc_avg_mid_cam_percol_1000rand <- mean(unlist(cc_mid_cam_percol_1000rand))

# Calculate standad deviation of the random dist
cc_sd_mid_cam_percol_1000rand <- sd(unlist(cc_mid_cam_percol_1000rand))

# look at random nw cc distributions
cc_dist_mid_cam_percol_1000rand <- (unlist(cc_mid_cam_percol_1000rand))
hist(cc_dist_mid_cam_percol_1000rand)



### Run statistical test on the real cc using random networks

# H0: cc[real] = mean of cc[rand]

# H1: cc[real] != mean of cc[rand]

# alpha = 0.05

t.test(cc_mid_cam_percol_net2, mu = cc_avg_mid_cam_percol_1000rand, alternative = 'less')

#### Middle Campanian ####

mid_cam_percol_net <- read.graph('EDENetwork_outputs/Mid
        CAM/CAM_mid_matrix_gen_great3_thresholded_at_057.net', format = "pajek")

typeof(mid_cam_percol_net)

cc_mid_cam_percol_net <- transitivity(mid_cam_percol_net)

# convert to a i,j,w weighted edgelist (NOTE: Could also do this by getting distance matrix and then
# converting to NW using graph.adjacency() then convert using get.data.frame())

### Get the network data into the right format for tnet to calculate random nws (need the nodes to be
        numeric, not grid names)
mid_cam_percol_net2 <- cbind(get.edgelist(mid_cam_percol_net, names=FALSE),
        E(mid_cam_percol_net)$weight)

# run this since the network is undirected, duplicates ties to justify that there is no direction
if(!is.directed(mid_cam_percol_net))
  mid_cam_percol_net2 <- symmetrise_w(mid_cam_percol_net2)
```

```
# make sure conforms to tnet format
mid_cam_percol_net2 <- as.tnet(mid_cam_percol_net2, type="weighted one-mode tnet")


### Convert from tnet format to igraph (this should remove any zero weights, which essentially indicates
        identical connection but only 2 present...)
# Create igraph object
mid_cam_percol_net3 <- tnet_igraph(mid_cam_percol_net2,type="weighted one-mode tnet")

# Get summary statistics
summary(mid_cam_percol_net3)

# check that the cc is still close
cc_mid_cam_percol_net2 <- transitivity(mid_cam_percol_net3)
### exactly the same!

## Plot the network

# Get layout of observed network and plot it... need to look into this more...
layout <- layout.fruchterman.reingold(mid_cam_percol_net3)

plot(mid_cam_percol_net3, main="Observed Network", layout=layout,
        vertex.label=1:length(mid_cam_percol_net3), edge.width=E(mid_cam_percol_net3)$weight,
        edge.label=E(mid_cam_percol_net3)$weight)

# Create random network from original network

# create function to reshuffle by links (reshuffles weighted links while maintaining negree)
fun <- function(x) {
  rg_reshuffling_w(x, option="links")
}

# set number of times to replicate
n = 1000

# run analysis to create array with different randomized networks
mid_cam_percol_1000rand <- replicate(n, fun(mid_cam_percol_net2), simplify=FALSE)

# turn into igraph objects
mid_cam_percol_1000rand_igraph <- lapply(mid_cam_percol_1000rand, function(x)
        tnet_igraph((x),type="weighted one-mode tnet"))

# run cc analysis
cc_mid_cam_percol_1000rand <- lapply(mid_cam_percol_1000rand_igraph,function(x) transitivity(x))

# get random networks cc average
cc_avg_mid_cam_percol_1000rand <- mean(unlist(cc_mid_cam_percol_1000rand))

# Calculate standad deviation of the random dist
cc_sd_mid_cam_percol_1000rand <- sd(unlist(cc_mid_cam_percol_1000rand))
```

```
# look at random nw cc distributions
cc_dist_mid_cam_percol_1000rand <- (unlist(cc_mid_cam_percol_1000rand))
hist(cc_dist_mid_cam_percol_1000rand)

### Run statistical test on the real cc using random networks

# H0: cc[real] = mean of cc[rand]

# H1: cc[real] != mean of cc[rand]

# alpha = 0.05

t.test(cc_mid_cam_percol_net2, mu = cc_avg_mid_cam_percol_1000rand, alternative = 'less')

#### Upper Campanian ####

up_cam_percol_net <- read.graph('EDENetwork_outputs/Up
        CAM/CAM_up_matrix_gen_great3_thresholded_at_050.net', format = "pajek")

typeof(up_cam_percol_net)

cc_up_cam_percol_net <- transitivity(up_cam_percol_net)

# convert to a i,j,w weighted edgelist (NOTE: Could also do this by getting distance matrix and then
# converting to NW using graph.adjacency() then convert using get.data.frame())

### Get the network data into the right format for tnet to calculate random nws (need the nodes to be
        numeric, not grid names)
up_cam_percol_net2 <- cbind(get.edgelist(up_cam_percol_net, names=FALSE),
        E(up_cam_percol_net)$weight)

# run this since the network is undirected, duplicates ties to justify that there is no direction
if(!is.directed(up_cam_percol_net))
  up_cam_percol_net2 <- symmetrise_w(up_cam_percol_net2)

# make sure conforms to tnet format
up_cam_percol_net2 <- as.tnet(up_cam_percol_net2, type="weighted one-mode tnet")

### Convert from tnet format to igraph (this should remove any zero weights, which essentially indicates
        identical connection but only 2 present...)
# Create igraph object
up_cam_percol_net3 <- tnet_igraph(up_cam_percol_net2,type="weighted one-mode tnet")

# Get summary statistics
summary(up_cam_percol_net3)

# check that the cc is still close
cc_up_cam_percol_net2 <- transitivity(up_cam_percol_net3)
### slightly lower, loss of 3 identical connections (wieghts = 0) but not very different

## Plot the network
```

```
# Get layout of observed network and plot it... need to look into this more...
layout <- layout.fruchterman.reingold(up_cam_percol_net3)

plot(up_cam_percol_net3, main="Observed Network", layout=layout,
        vertex.label=1:length(up_cam_percol_net3), edge.width=E(up_cam_percol_net3)$weight,
        edge.label=E(up_cam_percol_net3)$weight)

# Create random network from original network

# create function to reshuffle by links (reshuffles weighted links while maintaining negree)
fun <- function(x) {
  rg_reshuffling_w(x, option="links")
}

# set number of times to replicate
n = 1000

# run analysis to create array with different randomized networks
up_cam_percol_1000rand <- replicate(n, fun(up_cam_percol_net2), simplify=FALSE)

# turn into igraph objects
up_cam_percol_1000rand_igraph <- lapply(up_cam_percol_1000rand, function(x)
        tnet_igraph((x),type="weighted one-mode tnet"))

# run cc analysis
cc_up_cam_percol_1000rand <- lapply(up_cam_percol_1000rand_igraph,function(x) transitivity(x))

# get random networks cc average
cc_avg_up_cam_percol_1000rand <- mean(unlist(cc_up_cam_percol_1000rand))

# Calculate standad deviation of the random dist
cc_sd_up_cam_percol_1000rand <- sd(unlist(cc_up_cam_percol_1000rand))

# look at random nw cc distributions
cc_dist_up_cam_percol_1000rand <- (unlist(cc_up_cam_percol_1000rand))
hist(cc_dist_up_cam_percol_1000rand)

### Run statistical test on the real cc using random networks

# H0: cc[real] = mean of cc[rand]

# H1: cc[real] != mean of cc[rand]

# alpha = 0.05

t.test(cc_up_cam_percol_net2, mu = cc_avg_up_cam_percol_1000rand, alternative = 'less')

#### Lower Maastrichtian ####
```

```
low_maa_percol_net <- read.graph('EDENetwork_outputs/Low
        MAA/maa_low_matrix_gen_great3_thresholded_at_080.net', format = "pajek")

typeof(low_maa_percol_net)


cc_low_maa_percol_net <- transitivity(low_maa_percol_net)

# convert to a i,j,w weighted edgelist (NOTE: Could also do this by getting distance matrix and then
# converting to NW using graph.adjacency() then convert using get.data.frame())


### Get the network data into the right format for tnet to calculate random nws (need the nodes to be
        numeric, not grid names)
low_maa_percol_net2 <- cbind(get.edgelist(low_maa_percol_net, names=FALSE),
        E(low_maa_percol_net)$weight)

# run this since the network is undirected, duplicates ties to justify that there is no direction
if(!is.directed(low_maa_percol_net))
  low_maa_percol_net2 <- symmetrise_w(low_maa_percol_net2)

# make sure conforms to tnet format
low_maa_percol_net2 <- as.tnet(low_maa_percol_net2, type="weighted one-mode tnet")


### Convert from tnet format to igraph (this should remove any zero weights, which essentially indicates
        identical connection but only 2 present...)
# Create igraph object
low_maa_percol_net3 <- tnet_igraph(low_maa_percol_net2,type="weighted one-mode tnet")

# Get summary statistics
summary(low_maa_percol_net3)

# check that the cc is still close
cc_low_maa_percol_net2 <- transitivity(low_maa_percol_net3)
### slightly lower, loss of 1 identical connections (wieghts = 0) but not very different

## Plot the network

# Get layout of observed network and plot it... need to look into this more...
layout <- layout.fruchterman.reingold(low_maa_percol_net3)

plot(low_maa_percol_net3, main="Observed Network", layout=layout,
        vertex.label=1:length(low_maa_percol_net3), edge.width=E(low_maa_percol_net3)$weight,
        edge.label=E(low_maa_percol_net3)$weight)

# Create random network from original network

# create function to reshuffle by links (reshuffles weighted links while maintaining negree)
fun <- function(x) {
  rg_reshuffling_w(x, option="links")
```

```
}

# set number of times to replicate
n = 1000

# run analysis to create array with different randomized networks
low_maa_percol_1000rand <- replicate(n, fun(low_maa_percol_net2), simplify=FALSE)

# turn into igraph objects
low_maa_percol_1000rand_igraph <- lapply(low_maa_percol_1000rand, function(x)
        tnet_igraph((x),type="weighted one-mode tnet"))

# run cc analysis
cc_low_maa_percol_1000rand <- lapply(low_maa_percol_1000rand_igraph,function(x) transitivity(x))

# get random networks cc average
cc_avg_low_maa_percol_1000rand <- mean(unlist(cc_low_maa_percol_1000rand))

# Calculate standad deviation of the random dist
cc_sd_low_maa_percol_1000rand <- sd(unlist(cc_low_maa_percol_1000rand))

# look at random nw cc distributions
cc_dist_low_maa_percol_1000rand <- (unlist(cc_low_maa_percol_1000rand))
hist(cc_dist_low_maa_percol_1000rand)

### Run statistical test on the real cc using random networks

# H0: cc[real] = mean of cc[rand]

# H1: cc[real] != mean of cc[rand]

# alpha = 0.05

t.test(cc_low_maa_percol_net2, mu = cc_avg_low_maa_percol_1000rand, alternative = 'less')

#### Upper Maastrichtian ####

up_maa_percol_net <- read.graph('EDENetwork_outputs/Up
        MAA/maa_up_matrix_gen_great3_thresholded_at_086.net', format = "pajek")

typeof(up_maa_percol_net)

cc_up_maa_percol_net <- transitivity(up_maa_percol_net)

# convert to a i,j,w weighted edgelist (NOTE: Could also do this by getting distance matrix and then
# converting to NW using graph.adjacency() then convert using get.data.frame())

### Get the network data into the right format for tnet to calculate random nws (need the nodes to be
        numeric, not grid names)
up_maa_percol_net2 <- cbind(get.edgelist(up_maa_percol_net, names=FALSE),
        E(up_maa_percol_net)$weight)
```

```
# run this since the network is undirected, duplicates ties to justify that there is no direction
if(!is.directed(up_maa_percol_net))
  up_maa_percol_net2 <- symmetrise_w(up_maa_percol_net2)

# make sure conforms to tnet format
up_maa_percol_net2 <- as.tnet(up_maa_percol_net2, type="weighted one-mode tnet")


### Convert from tnet format to igraph (this should remove any zero weights, which essentially indicates
        identical connection but only 2 present...)
# Create igraph object
up_maa_percol_net3 <- tnet_igraph(up_maa_percol_net2,type="weighted one-mode tnet")

# Get summary statistics
summary(up_maa_percol_net3)

# check that the cc is still close
cc_up_maa_percol_net2 <- transitivity(up_maa_percol_net3)
### slightly lower, loss of 1 identical connections (wieghts = 0) but not very different

## Plot the network

# Get layout of observed network and plot it... need to look into this more...
layout <- layout.fruchterman.reingold(up_maa_percol_net3)

plot(up_maa_percol_net3, main="Observed Network", layout=layout,
        vertex.label=1:length(up_maa_percol_net3), edge.width=E(up_maa_percol_net3)$weight,
        edge.label=E(up_maa_percol_net3)$weight)

# Create random network from original network

# create function to reshuffle by links (reshuffles weighted links while maintaining negree)
fun <- function(x) {
  rg_reshuffling_w(x, option="links")
}

# set number of times to replicate
n = 1000

# run analysis to create array with different randomized networks
up_maa_percol_1000rand <- replicate(n, fun(up_maa_percol_net2), simplify=FALSE)

# turn into igraph objects
up_maa_percol_1000rand_igraph <- lapply(up_maa_percol_1000rand, function(x)
        tnet_igraph((x),type="weighted one-mode tnet"))

# run cc analysis
cc_up_maa_percol_1000rand <- lapply(up_maa_percol_1000rand_igraph,function(x) transitivity(x))

# get random networks cc average
```

```
cc_avg_up_maa_percol_1000rand <- mean(unlist(cc_up_maa_percol_1000rand))

# Calculate standad deviation of the random dist
cc_sd_up_maa_percol_1000rand <- sd(unlist(cc_up_maa_percol_1000rand))

# look at random nw cc distributions
cc_dist_up_maa_percol_1000rand <- (unlist(cc_up_maa_percol_1000rand))
hist(cc_dist_up_maa_percol_1000rand)

### Run statistical test on the real cc using random networks

# H0: cc[real] = mean of cc[rand]

# H1: cc[real] != mean of cc[rand]

# alpha = 0.05


t.test(cc_up_maa_percol_net2, mu = cc_avg_up_maa_percol_1000rand, alternative = 'less')

#### Complete DB ####


comp_db_percol_net <-
        read.graph('EDENetwork_outputs/CompleteDB/trunk60_matrix_gen_great3_thresholded_at_076
        .net', format = "pajek")

typeof(comp_db_percol_net)

cc_comp_db_percol_net <- transitivity(comp_db_percol_net)

# convert to a i,j,w weighted edgelist (NOTE: Could also do this by getting distance matrix and then
# converting to NW using graph.adjacency() then convert using get.data.frame())

### Get the network data into the right format for tnet to calculate random nws (need the nodes to be
        numeric, not grid names)
comp_db_percol_net2 <- cbind(get.edgelist(comp_db_percol_net, names=FALSE),
        E(comp_db_percol_net)$weight)

# run this since the network is undirected, dcomplicates ties to justify that there is no direction
if(!is.directed(comp_db_percol_net))
  comp_db_percol_net2 <- symmetrise_w(comp_db_percol_net2)

# make sure conforms to tnet format
comp_db_percol_net2 <- as.tnet(comp_db_percol_net2, type="weighted one-mode tnet")

### Convert from tnet format to igraph (this should remove any zero weights, which essentially indicates
        identical connection but only 2 present...)
# Create igraph object
comp_db_percol_net3 <- tnet_igraph(comp_db_percol_net2,type="weighted one-mode tnet")
```

```
# Get summary statistics
summary(comp_db_percol_net3)

# check that the cc is still close
cc_comp_db_percol_net2 <- transitivity(comp_db_percol_net3)
### slightly lower, loss of 16 identical connections (wieghts = 0) but not very different

## Plot the network

# Get layout of observed network and plot it... need to look into this more...
layout <- layout.fruchterman.reingold(comp_db_percol_net3)

plot(comp_db_percol_net3, main="Observed Network", layout=layout,
        vertex.label=1:length(comp_db_percol_net3), edge.width=E(comp_db_percol_net3)$weight,
        edge.label=E(comp_db_percol_net3)$weight)

# Create random network from original network

# create function to reshuffle by links (reshuffles weighted links while maintaining negree)
fun <- function(x) {
  rg_reshuffling_w(x, option="links")
}

# set number of times to replicate
n = 1000

# run analysis to create array with different randomized networks
comp_db_percol_1000rand <- replicate(n, fun(comp_db_percol_net2), simplify=FALSE)

# turn into igraph objects
comp_db_percol_1000rand_igraph <- lapply(comp_db_percol_1000rand, function(x)
        tnet_igraph((x),type="weighted one-mode tnet"))

# run cc analysis
cc_comp_db_percol_1000rand <- lapply(comp_db_percol_1000rand_igraph,function(x) transitivity(x))

# get random networks cc average
cc_avg_comp_db_percol_1000rand <- mean(unlist(cc_comp_db_percol_1000rand))

# Calculate standad deviation of the random dist
cc_sd_comp_db_percol_1000rand <- sd(unlist(cc_comp_db_percol_1000rand))

# look at random nw cc distributions
cc_dist_comp_db_percol_1000rand <- (unlist(cc_comp_db_percol_1000rand))
hist(cc_dist_comp_db_percol_1000rand)

### Run statistical test on the real cc using random networks

# H0: cc[real] = mean of cc[rand]

# H1: cc[real] != mean of cc[rand]
```

```
# alpha = 0.05


t.test(cc_comp_db_percol_net2, mu = cc_avg_comp_db_percol_1000rand, alternative = 'less')

### Take randomized results (mean and sd) and put in table for comparison with real cc

rand_cc_means <-
        c(cc_avg_low_cam_percol_1000rand,cc_avg_mid_cam_percol_1000rand,cc_avg_up_cam_perco
        l_1000rand,cc_avg_low_maa_percol_1000rand,
             cc_avg_up_maa_percol_1000rand,cc_avg_comp_db_percol_1000rand)

rand_cc_sd <-
        c(cc_sd_low_cam_percol_1000rand,cc_sd_mid_cam_percol_1000rand,cc_sd_up_cam_percol_10
        00rand,cc_sd_low_maa_percol_1000rand,
            cc_sd_up_maa_percol_1000rand,cc_sd_comp_db_percol_1000rand)

rand_cc_1sd <- (rand_cc_sd) + rand_cc_means

rand_cc_2sd <- (rand_cc_sd *2) + rand_cc_means

rand_cc_3sd <- (rand_cc_sd *3) + + rand_cc_means

cc_means <-
        c(cc_low_cam_percol_net2,cc_mid_cam_percol_net2,cc_up_cam_percol_net2,cc_low_maa_perc
        ol_net2,
         cc_up_maa_percol_net2,cc_comp_db_percol_net2)

cc_table <- cbind(cc_means,rand_cc_means,rand_cc_1sd,rand_cc_2sd,rand_cc_3sd)
colnames(cc_table) <- c("Networks <CC>","Mean Rand. <CC>","1 S.D. Rand. <CC>","2 S.D. Rand.
        <CC>","3 S.D. Rand. <CC>")
rownames(cc_table) <- c("Lower Camp.", "Middle Camp.", "Upper Camp.", "Lower Maastr.", "Upper
        Maastr.", "Complete Database")


cc_table

write.csv(cc_table,file="NWR_Outputs/rnw_dist_results_table_1000.csv")

x <- 1:5
set.seed(49)
mat1 <- do.call(rbind,lapply(1:10,function(y) sample(x,3)))

lapply(1:10,sample(x,2))

x <- seq(1,10)

r <- array(dim = c(10,3,1))

for (i in 1:1){
```

```
  r[i] <- rg_reshuffling_w(low_cam_061_net2, option="links")
  print(r)
}

set.seed(1)
gen_mat <- function(x) matrix(c(1, 1, 1, x + rnorm(1)), nrow = 2)

gen_mat(5)

n <- 10

# give 1 to gen_mat n-times
lapply(rep(1, n), gen_mat)

try <- replicate(1,r)

try[[1]]

ri <- tnet_igraph(r,type="weighted one-mode tnet")
r1 <- tnet_igraph(r11,type="weighted one-mode tnet")
r2 <- tnet_igraph(r22,type="weighted one-mode tnet")
r3 <- tnet_igraph(r33,type="weighted one-mode tnet")
r4 <- tnet_igraph(r44,type="weighted one-mode tnet")
r5 <- tnet_igraph(r55,type="weighted one-mode tnet")


transitivity(ri)
transitivity(r1)
transitivity(r2)
transitivity(r3)
transitivity(r4)
transitivity(r5)

gg <- get.data.frame(g)

sum(gg_final$weight != 0)
nrow(gg)


gg_names <- data.frame(unique(gg[,1]))
colnames(gg_names) <- "names"
gg_names2 <- data.frame(unique(gg[,2]))
colnames(gg_names2) <- "names"
gg_names_list <- unique(rbind(gg_names,gg_names2))

gg_names_list <- cbind(newColName = rownames(gg_names_list), gg_names_list)
rownames(gg_names_list) <- 1:nrow(gg_names_list)
gg_names_list$to <- gg_names_list[,2]
colnames(gg_names_list) <- c("ID","from","to")

gg_mergefrom <- merge(gg,gg_names_list,by='from',all.x=TRUE)
```

```
colnames(gg_mergefrom) <- c("from","to",'weight','ID','to_y')
gg_mergeto <- merge(gg_mergefrom,gg_names_list,by='to',all.x=TRUE)

head(gg_mergeto)

gg_final <- gg_mergeto[,c(4,6,3)]

h <- subset(gg_final, weight != 0)
colnames(h) <- c("i","j","w")

rnw1 <- rg_reshuffling_w(h, option="links")

rnw1_igraph <- tnet_igraph(rnw,type="weighted one-mode tnet",directed =FALSE)

betweenness(rnw_igraph)

# Look at distribution of weights in the network
low_cam_dist <- read.table("EDENetwork_outputs/Low CAM/low_CAM_distanceMatrix.txt", header =
        FALSE)

low_cam_names <- read.table("EDENetwork_outputs/Low
        CAM/low_CAM_distanceMatrix_names.txt",header = FALSE)
summary(low_cam_names)
colnames(low_cam_names) <- "Node"
nrow(low_cam_names)
nrow(low_cam_dist)

colnames(low_cam_dist) <- c(low_cam_names$Node)
rownames(low_cam_dist) <- c(low_cam_names$Node)


low_cam_dist_long <- melt(low_cam_dist)

#make into a numeric table
try <- matrix(data = NA, nrow = dim(low_cam_dist)[1], ncol = dim(low_cam_dist)[2])

colnames(try) <- rownames(low_cam_dist)
rownames(try) <- colnames(low_cam_dist)

for (i in 1:dim(low_cam_dist)[2]) {
  try[,i] <- c(as.numeric(low_cam_dist[[i]]))
}

try
hist(try)

### Node weights not normally distributed, so probably definitely need to resample links from the
        existing networks, rather than creating them from a dist
```

*Code for Manually Comparing Potential Network Subcomponents*

```
## Code to assess the number and attributes of network components, or clusters of interconnected nodes, in the substage networks.
## Notes are included throughout which clarify results and interpretations.

setwd("C:/Users/ceara/Documents/Province Project/Occurrence Data/VettedCombinedDB")

#### Lower Campanian ####

# read in text files for different thresholds
cam_low_0.61th <- readLines("EDENetwork_outputs/Low CAM/CAM_low_matrix_gen_great3_thresholded_at_061.txt")
cam_low_0.55th <- readLines("EDENetwork_outputs/Low CAM/CAM_low_matrix_gen_great3_thresholded_at_055.txt")
cam_low_0.50th <- readLines("EDENetwork_outputs/Low CAM/CAM_low_matrix_gen_great3_thresholded_at_050.txt")
cam_low_0.45th <- readLines("EDENetwork_outputs/Low CAM/CAM_low_matrix_gen_great3_thresholded_at_045.txt")
cam_low_0.40th <- readLines("EDENetwork_outputs/Low CAM/CAM_low_matrix_gen_great3_thresholded_at_040.txt")

# Check which th have components consisting of multiple nodes

cam_low_0.61th # one major component, all others less than 2 nodes each
cam_low_0.55th # one major component, all others less than 2 nodes each
cam_low_0.50th # one major component, all others less than 2 nodes each
cam_low_0.45th # one major component, all others less than 2 nodes each
cam_low_0.40th  # one major component, all others less than 2 nodes each

## Do not need to run further visualization of these components, all in WI for LOW CAM

#### Middle Campanian ####

# read in text files for different thresholds
cam_mid_0.57th <- readLines("EDENetwork_outputs/Mid CAM/CAM_mid_matrix_gen_great3_thresholded_at_057.txt")
cam_mid_0.50th <- readLines("EDENetwork_outputs/Mid CAM/CAM_mid_matrix_gen_great3_thresholded_at_050.txt")
cam_mid_0.45th <- readLines("EDENetwork_outputs/Mid CAM/CAM_mid_matrix_gen_great3_thresholded_at_045.txt")
cam_mid_0.40th <- readLines("EDENetwork_outputs/Mid CAM/CAM_mid_matrix_gen_great3_thresholded_at_040.txt")

# Check which th have components consisting of multiple nodes

cam_mid_0.57th # one major component, all others less than 2 nodes each
cam_mid_0.50th # one major component, one minor with 6, all others less than 3 nodes each
```

cam_mid_0.45th # one major component, one minor with 6, one with 4, all others less than 3 nodes each
cam_mid_0.40th  # one major component, one minor with 5, all others less than 3 nodes each

    ## Should map to check location of minor components within WI

#### Upper Campanian ####

# read in text files for different thresholds
cam_up_0.50th <- readLines("EDENetwork_outputs/Up CAM/CAM_up_matrix_gen_great3_thresholded_at_050.txt")
cam_up_0.45th <- readLines("EDENetwork_outputs/Up CAM/CAM_up_matrix_gen_great3_thresholded_at_045.txt")
cam_up_0.40th <- readLines("EDENetwork_outputs/Up CAM/CAM_up_matrix_gen_great3_thresholded_at_040.txt")

# Check which th have components consisting of multiple nodes

cam_up_0.50th # one major component, all others less than 3 nodes each
cam_up_0.45th # one major component, all others less than 3 nodes each
cam_up_0.40th  # one major component, all others less than 1 nodes each

## Do not need to run further visualization of these components

#### Lower Maastrichtian ####

# read in text files for different thresholds
maa_low_0.80th <- readLines("EDENetwork_outputs/Low MAA/MAA_low_matrix_gen_great3_thresholded_at_080.txt")
maa_low_0.75th <- readLines("EDENetwork_outputs/Low MAA/MAA_low_matrix_gen_great3_thresholded_at_075.txt")
maa_low_0.70th <- readLines("EDENetwork_outputs/Low MAA/MAA_low_matrix_gen_great3_thresholded_at_070.txt")
maa_low_0.65th <- readLines("EDENetwork_outputs/Low MAA/MAA_low_matrix_gen_great3_thresholded_at_065.txt")
maa_low_0.60th <- readLines("EDENetwork_outputs/Low MAA/MAA_low_matrix_gen_great3_thresholded_at_060.txt")
maa_low_0.55th <- readLines("EDENetwork_outputs/Low MAA/MAA_low_matrix_gen_great3_thresholded_at_055.txt")
maa_low_0.50th <- readLines("EDENetwork_outputs/Low MAA/MAA_low_matrix_gen_great3_thresholded_at_050.txt")
maa_low_0.45th <- readLines("EDENetwork_outputs/Low MAA/MAA_low_matrix_gen_great3_thresholded_at_045.txt")
maa_low_0.40th <- readLines("EDENetwork_outputs/Low MAA/MAA_low_matrix_gen_great3_thresholded_at_040.txt")

# Check which th have components consisting of multiple nodes

maa_low_0.80th # one major component, all others less than 1 nodes each
maa_low_0.75th # one major component, one with 7, all others less than 1 nodes each
maa_low_0.70th # one major component, one with 7, all others less than 1 nodes each
maa_low_0.65th # one major component, one with 4, all others less than 2 nodes each
maa_low_0.60th # one major component, one with 4, all others less than 2 nodes each
maa_low_0.55th # one major component, 2 with 7, 1 with 4, all others less than 2 nodes each
maa_low_0.50th # one major component, one with 5, all others less than 3 nodes each
maa_low_0.45th # one component with 5, one with 4, all others less than 2 nodes each
maa_low_0.40th # all component less than 3 nodes each

## Map out 0.75, 0.60, and 0.50 to look at component locations

#### Upper Maastrichtian ####

# read in text files for different thresholds
maa_up_0.86th <- readLines("EDENetwork_outputs/Up
MAA/MAA_up_matrix_gen_great3_thresholded_at_086.txt")
maa_up_0.80th <- readLines("EDENetwork_outputs/Up
MAA/MAA_up_matrix_gen_great3_thresholded_at_080.txt")
maa_up_0.75th <- readLines("EDENetwork_outputs/Up
MAA/MAA_up_matrix_gen_great3_thresholded_at_075.txt")
maa_up_0.70th <- readLines("EDENetwork_outputs/Up
MAA/MAA_up_matrix_gen_great3_thresholded_at_070.txt")
maa_up_0.65th <- readLines("EDENetwork_outputs/Up
MAA/MAA_up_matrix_gen_great3_thresholded_at_065.txt")
maa_up_0.60th <- readLines("EDENetwork_outputs/Up
MAA/MAA_up_matrix_gen_great3_thresholded_at_060.txt")
maa_up_0.55th <- readLines("EDENetwork_outputs/Up
MAA/MAA_up_matrix_gen_great3_thresholded_at_055.txt")
maa_up_0.50th <- readLines("EDENetwork_outputs/Up
MAA/MAA_up_matrix_gen_great3_thresholded_at_050.txt")
maa_up_0.45th <- readLines("EDENetwork_outputs/Up
MAA/MAA_up_matrix_gen_great3_thresholded_at_045.txt")
maa_up_0.40th <- readLines("EDENetwork_outputs/Up
MAA/MAA_up_matrix_gen_great3_thresholded_at_040.txt")

# Check which th have components consisting of multiple nodes

maa_up_0.86th # one major component, all others less than 2 nodes
maa_up_0.80th # one major component, 2ndary component with 13 nodes, all others less than 1 nodes each
maa_up_0.75th # one major component, 2ndary component with 13 nodes, all others less than 1 nodes each
maa_up_0.70th # one major component, 2ndary component with 13 nodes, all others less than 1 nodes each
maa_up_0.65th # one major component, 2ndary component with 12 nodes, all others less than 1 nodes each

```
maa_up_0.60th # one major component, 2ndary component with 12 nodes, all others less than 1 nodes each
maa_up_0.55th # one major component, 2ndary component with 10 nodes, all others less than 1 nodes each
maa_up_0.50th # one major component, 2ndary component with 8 nodes, all others less than 1 nodes each
maa_up_0.45th # all components less than 3 nodes each
maa_up_0.40th # all component less than 3 nodes each

#### Middle Campanian Components for 0.50 to 0.40 thresholds

# Extract component info and map to visualize:

typeof(cam_mid_0.50th) ## Contains: one major component, one minor with 6, all others less than 3 nodes each >> NEED TO LOOK AT FIRST 2 COMPONENTS

# split the character string into lists of component information
cam_mid_0.50th_split <- split(cam_mid_0.50th, cumsum( grepl("^---", cam_mid_0.50th)))
cam_mid_0.50th_split$"0"

# create individual list objects of each relevant component
cam_mid_0.50th_split_1st <- cam_mid_0.50th_split$"0"
cam_mid_0.50th_split_2nd <- cam_mid_0.50th_split$"1"

typeof(cam_mid_0.50th_split_2nd)

# get length of each component (number of rows)
nrow_1st <- length(cam_mid_0.50th_split_1st)
nrow_2nd <- length(cam_mid_0.50th_split_2nd)

# create an empty matrix to put the first component in (use length of rows for dimensions, remember all components after 1st have extra leading row)
cam_mid_0.50th_split_1st_matrix <- matrix(data = NA, nrow=nrow_1st-2,ncol=11)
cam_mid_0.50th_split_2nd_matrix <- matrix(data = NA, nrow=nrow_2nd-3,ncol=11)

# split the string information from the first component in the list by "\t" and unlist to put values into vectors

# 1st component
for (i in 3:length(cam_mid_0.50th_split_1st)){
  cam_mid_0.50th_split_1st_matrix[i-2,] <- unlist(strsplit(cam_mid_0.50th_split_1st[i], "\t"))
}

# 2nd component
for (i in 4:length(cam_mid_0.50th_split_2nd)){
  cam_mid_0.50th_split_2nd_matrix[i-3,] <- unlist(strsplit(cam_mid_0.50th_split_2nd[i], "\t"))
}
```

```
# add column names to the matrices
colnames(cam_mid_0.50th_split_1st_matrix) <- c("Node","Degree",'Clustering', 'Province', 'Numb_Occ',
'Age', 'betweenness', 'location', 'OccColor', 'AgeColor', 'ProvColor')
colnames(cam_mid_0.50th_split_2nd_matrix) <- c("Node","Degree",'Clustering', 'Province',
'Numb_Occ', 'Age', 'betweenness', 'location', 'OccColor', 'AgeColor', 'ProvColor')

# make the matrices into data frames
cam_mid_0.50th_split_1st_df <- as.data.frame(cam_mid_0.50th_split_1st_matrix)
cam_mid_0.50th_split_2nd_df <- as.data.frame(cam_mid_0.50th_split_2nd_matrix)

typeof(cam_mid_0.50th_split_1st_df)

# read in the lat/long file for the substage
cam_mid_loc <- read.csv('R_Config_for_EN_OUTPUT/Mid CAM/CAM_mid_gen_great3_avgloc.csv')
colnames(cam_mid_loc) <- c("Node","y","x","Age")


# merge the dfs together to get the locaiton information with the components
cam_mid_0.50th_split_1st_merge <- merge(cam_mid_0.50th_split_1st_df,cam_mid_loc,by="Node")
cam_mid_0.50th_split_2nd_merge <- merge(cam_mid_0.50th_split_2nd_df,cam_mid_loc,by="Node")

# simplify to just the x,y coord for each component
cam_mid_0.50th_split_1st_loc <- cam_mid_0.50th_split_1st_merge[,c(12,13)]
cam_mid_0.50th_split_2nd_loc <- cam_mid_0.50th_split_2nd_merge[,c(12,13)]

# plot the map of the different components

library("ggplot2")
theme_set(theme_bw())
library("sf")

library("rnaturalearth")
library("rnaturalearthdata")
library("ggspatial")

world <- ne_countries(scale = "medium", returnclass = "sf") # get sf info for whole world from
rnaturalearth
class(world)

# Plotting the map
ggplot(data = world) +
  geom_sf(size=0.1) +
  xlab("Longitude") + ylab("Latitude") +
  annotation_scale(location = "bl", width_hint = 0.5) +
  annotation_north_arrow(location = "bl", which_north = "true",
                         height = unit(0.4,"in"), width = unit(0.3,"in"),
                         pad_x = unit(0.1, "in"), pad_y = unit(0.3, "in"),
```

```
                    style = north_arrow_minimal) +
  annotate(geom = "text", x = -90, y = 26, label = "Gulf of Mexico",
           fontface = "italic", color = "grey22", size = 3) +
  theme(panel.grid.major = element_line(color = "grey",
                                        size = 0.3,
                                        linetype = 2)) +
  geom_point(data = cam_mid_0.50th_split_1st_loc, aes(x = x, y = y), size = 3,
             shape = 21, fill = "darkred") +
  geom_point(data = cam_mid_0.50th_split_2nd_loc, aes(x = x, y = y), size = 3,
             shape = 21, fill = "lightblue") +
  coord_sf(xlim = c(-120, -75), ylim = c(20, 60), expand = FALSE)
ggsave("ComponentAnalysis/cam_mid_0.50th.pdf")

# Second TH level:

### Contains: one major component, one minor with 6, one with 4, all others less than 3 nodes each >>>
Plot 3 Components

# split the character string into lists of component information
cam_mid_0.45th_split <- split(cam_mid_0.45th, cumsum( grepl("^---", cam_mid_0.45th)))
cam_mid_0.45th_split$"0"

# create individual list objects of each relevant component
cam_mid_0.45th_split_1st <- cam_mid_0.45th_split$"0"
cam_mid_0.45th_split_2nd <- cam_mid_0.45th_split$"1"
cam_mid_0.45th_split_3rd <- cam_mid_0.45th_split$"2"

typeof(cam_mid_0.45th_split_2nd)

# get length of each component (number of rows)
nrow_1st <- length(cam_mid_0.45th_split_1st)
nrow_2nd <- length(cam_mid_0.45th_split_2nd)
nrow_3rd <- length(cam_mid_0.45th_split_3rd)

# create an empty matrix to put the first component in (use length of rows for dimentions, remember all
components after 1st have extra leading row)
cam_mid_0.45th_split_1st_matrix <- matrix(data = NA, nrow=nrow_1st-2,ncol=11)
cam_mid_0.45th_split_2nd_matrix <- matrix(data = NA, nrow=nrow_2nd-3,ncol=11)
cam_mid_0.45th_split_3rd_matrix <- matrix(data = NA, nrow=nrow_3rd-3,ncol=11)

# split the string information from the first component in the list by "\t" and unlist to put values into
vectors

# 1st component
for (i in 3:length(cam_mid_0.45th_split_1st)){
  cam_mid_0.45th_split_1st_matrix[i-2,] <- unlist(strsplit(cam_mid_0.45th_split_1st[i], "\t"))
}
```

```
# 2nd component
for (i in 4:length(cam_mid_0.45th_split_2nd)){
  cam_mid_0.45th_split_2nd_matrix[i-3,] <- unlist(strsplit(cam_mid_0.45th_split_2nd[i], "\t"))
}

# 3rd component
for (i in 4:length(cam_mid_0.45th_split_3rd)){
  cam_mid_0.45th_split_3rd_matrix[i-3,] <- unlist(strsplit(cam_mid_0.45th_split_3rd[i], "\t"))
}

# add column names to the matrices
colnames(cam_mid_0.45th_split_1st_matrix) <- c("Node","Degree",'Clustering', 'Province', 'Numb_Occ',
'Age', 'betweenness', 'location', 'OccColor', 'AgeColor', 'ProvColor')
colnames(cam_mid_0.45th_split_2nd_matrix) <- c("Node","Degree",'Clustering', 'Province',
'Numb_Occ', 'Age', 'betweenness', 'location', 'OccColor', 'AgeColor', 'ProvColor')
colnames(cam_mid_0.45th_split_3rd_matrix) <- c("Node","Degree",'Clustering', 'Province', 'Numb_Occ',
'Age', 'betweenness', 'location', 'OccColor', 'AgeColor', 'ProvColor')

# make the matrices into data frames
cam_mid_0.45th_split_1st_df <- as.data.frame(cam_mid_0.45th_split_1st_matrix)
cam_mid_0.45th_split_2nd_df <- as.data.frame(cam_mid_0.45th_split_2nd_matrix)
cam_mid_0.45th_split_3rd_df <- as.data.frame(cam_mid_0.45th_split_3rd_matrix)

typeof(cam_mid_0.45th_split_1st_df)

# read in the lat/long file for the substage
cam_mid_loc <- read.csv('R_Config_for_EN_OUTPUT/Mid CAM/CAM_mid_gen_great3_avgloc.csv')
colnames(cam_mid_loc) <- c("Node","y","x","Age")

# merge the dfs together to get the locaiton information with the components
cam_mid_0.45th_split_1st_merge <- merge(cam_mid_0.45th_split_1st_df,cam_mid_loc,by="Node")
cam_mid_0.45th_split_2nd_merge <- merge(cam_mid_0.45th_split_2nd_df,cam_mid_loc,by="Node")
cam_mid_0.45th_split_3rd_merge <- merge(cam_mid_0.45th_split_3rd_df,cam_mid_loc,by="Node")

# simplify to just the x,y coord for each component
cam_mid_0.45th_split_1st_loc <- cam_mid_0.45th_split_1st_merge[,c(12,13)]
cam_mid_0.45th_split_2nd_loc <- cam_mid_0.45th_split_2nd_merge[,c(12,13)]
cam_mid_0.45th_split_3rd_loc <- cam_mid_0.45th_split_3rd_merge[,c(12,13)]

# plot the map of the different components

library("ggplot2")
theme_set(theme_bw())
library("sf")

library("rnaturalearth")
library("rnaturalearthdata")
library("ggspatial")
```

```
world <- ne_countries(scale = "medium", returnclass = "sf") # get sf info for whole world from
rnaturalearth
class(world)

# Plotting the map
ggplot(data = world) +
  geom_sf(size=0.1) +
  xlab("Longitude") + ylab("Latitude") +
  annotation_scale(location = "bl", width_hint = 0.5) +
  annotation_north_arrow(location = "bl", which_north = "true",
                         height = unit(0.4,"in"), width = unit(0.3,"in"),
                         pad_x = unit(0.1, "in"), pad_y = unit(0.3, "in"),
                         style = north_arrow_minimal) +
  annotate(geom = "text", x = -90, y = 26, label = "Gulf of Mexico",
           fontface = "italic", color = "grey22", size = 3) +
  theme(panel.grid.major = element_line(color = "grey",
                                        size = 0.3,
                                        linetype = 2)) +
  geom_point(data = cam_mid_0.45th_split_1st_loc, aes(x = x, y = y), size = 3,
             shape = 21, fill = "darkred") +
  geom_point(data = cam_mid_0.45th_split_2nd_loc, aes(x = x, y = y), size = 3,
             shape = 21, fill = "lightblue") +
  geom_point(data = cam_mid_0.45th_split_3rd_loc, aes(x = x, y = y), size = 3,
             shape = 21, fill = "limegreen") +
  coord_sf(xlim = c(-120, -75), ylim = c(20, 60), expand = FALSE)
ggsave("ComponentAnalysis/cam_mid_0.45th.pdf")

### NOTE: though the 2nd component points appear to overlap in the same geographic region as in the
0.50 TH level in both maps, they do not
   # represent the same nodes. So, it isn't a consistent component through time... Weird that it overlaps
so well spatially but based on different nodes.
   # The 3rd component here also doesn't correspond with the 2nd component from the 0.50 TH.

# Third TH level:

### Contains: one major component, one minor with 5, all others less than 3 nodes each >>> NEED TO
PLOT 2 COMPONENTS ONLY

# split the character string into lists of component information
cam_mid_0.40th_split <- split(cam_mid_0.40th, cumsum( grepl("^---", cam_mid_0.40th)))
cam_mid_0.40th_split$"0"

# create individual list objects of each relevant component
cam_mid_0.40th_split_1st <- cam_mid_0.40th_split$"0"
cam_mid_0.40th_split_2nd <- cam_mid_0.40th_split$"1"

typeof(cam_mid_0.40th_split_2nd)
```

```
# get lenght of each component (number of rows)
nrow_1st <- length(cam_mid_0.40th_split_1st)
nrow_2nd <- length(cam_mid_0.40th_split_2nd)

# create an empty matrix to put the first component in (use length of rows for dimentions, remember all components after 1st have extra leading row)
cam_mid_0.40th_split_1st_matrix <- matrix(data = NA, nrow=nrow_1st-2,ncol=11)
cam_mid_0.40th_split_2nd_matrix <- matrix(data = NA, nrow=nrow_2nd-3,ncol=11)

# split the string information from the first component in the list by "\t" and unlist to put values into vectors

# 1st component
for (i in 3:length(cam_mid_0.40th_split_1st)){
  cam_mid_0.40th_split_1st_matrix[i-2,] <- unlist(strsplit(cam_mid_0.40th_split_1st[i], "\t"))
}

# 2nd component
for (i in 4:length(cam_mid_0.40th_split_2nd)){
  cam_mid_0.40th_split_2nd_matrix[i-3,] <- unlist(strsplit(cam_mid_0.40th_split_2nd[i], "\t"))
}

# add column names to the matrices
colnames(cam_mid_0.40th_split_1st_matrix) <- c("Node","Degree",'Clustering', 'Province', 'Numb_Occ', 'Age', 'betweenness', 'location', 'OccColor', 'AgeColor', 'ProvColor')
colnames(cam_mid_0.40th_split_2nd_matrix) <- c("Node","Degree",'Clustering', 'Province', 'Numb_Occ', 'Age', 'betweenness', 'location', 'OccColor', 'AgeColor', 'ProvColor')

# make the matrices into data frames
cam_mid_0.40th_split_1st_df <- as.data.frame(cam_mid_0.40th_split_1st_matrix)
cam_mid_0.40th_split_2nd_df <- as.data.frame(cam_mid_0.40th_split_2nd_matrix) # REPRESENTS SAME NODES AS THE 2ND COMPONENT IN THE 0.45 TH (but one less)

typeof(cam_mid_0.40th_split_1st_df)

# read in the lat/long file for the substage
cam_mid_loc <- read.csv('R_Config_for_EN_OUTPUT/Mid CAM/CAM_mid_gen_great3_avgloc.csv')
colnames(cam_mid_loc) <- c("Node","y","x","Age")

# merge the dfs together to get the locaiton information with the components
cam_mid_0.40th_split_1st_merge <- merge(cam_mid_0.40th_split_1st_df,cam_mid_loc,by="Node")
cam_mid_0.40th_split_2nd_merge <- merge(cam_mid_0.40th_split_2nd_df,cam_mid_loc,by="Node")

# simplify to just the x,y coord for each component
cam_mid_0.40th_split_1st_loc <- cam_mid_0.40th_split_1st_merge[,c(12,13)]
cam_mid_0.40th_split_2nd_loc <- cam_mid_0.40th_split_2nd_merge[,c(12,13)]
```

```
# plot the map of the different components

library("ggplot2")
theme_set(theme_bw())
library("sf")

library("rnaturalearth")
library("rnaturalearthdata")
library("ggspatial")

world <- ne_countries(scale = "medium", returnclass = "sf") # get sf info for whole world from
rnaturalearth
class(world)

# Plotting the map
ggplot(data = world) +
  geom_sf(size=0.1) +
  xlab("Longitude") + ylab("Latitude") +
  annotation_scale(location = "bl", width_hint = 0.5) +
  annotation_north_arrow(location = "bl", which_north = "true",
                         height = unit(0.4,"in"), width = unit(0.3,"in"),
                         pad_x = unit(0.1, "in"), pad_y = unit(0.3, "in"),
                         style = north_arrow_minimal) +
  annotate(geom = "text", x = -90, y = 26, label = "Gulf of Mexico",
           fontface = "italic", color = "grey22", size = 3) +
  theme(panel.grid.major = element_line(color = "grey",
                                        size = 0.3,
                                        linetype = 2)) +
  geom_point(data = cam_mid_0.40th_split_1st_loc, aes(x = x, y = y), size = 3,
             shape = 21, fill = "darkred") +
  geom_point(data = cam_mid_0.40th_split_2nd_loc, aes(x = x, y = y), size = 3,
             shape = 21, fill = "lightblue") +
  coord_sf(xlim = c(-120, -75), ylim = c(20, 60), expand = FALSE)
ggsave("ComponentAnalysis/cam_mid_0.40th.pdf")

### NOTE: Secondary component in the 0.40 TH same as that of the 0.45 TH, so consistent between the
two. Again, not the same as the 2nd Component from
   # the 0.50 TH, however.

#### Lower Maastrichtian Components for 0.75, 0.60, and 0.50 thresholds

# Extract component info and map to visualize:

typeof(maa_low_0.50th)

### Contains: one major component, one with 5, all others less than 3 nodes each >>> PLOT 2
COMPONENTS
```

```
# split the character string into lists of component information
maa_low_0.50th_split <- split(maa_low_0.50th, cumsum( grepl("^---", maa_low_0.50th)))
maa_low_0.50th_split$"0"

# create individual list objects of each relevant component
maa_low_0.50th_split_1st <- maa_low_0.50th_split$"0"
maa_low_0.50th_split_2nd <- maa_low_0.50th_split$"1"

typeof(maa_low_0.50th_split_2nd)

# get lenght of each component (number of rows)
nrow_1st <- length(maa_low_0.50th_split_1st)
nrow_2nd <- length(maa_low_0.50th_split_2nd)

# create an empty matrix to put the first component in (use length of rows for dimentions, remember all components after 1st have extra leading row)
maa_low_0.50th_split_1st_matrix <- matrix(data = NA, nrow=nrow_1st-2,ncol=11)
maa_low_0.50th_split_2nd_matrix <- matrix(data = NA, nrow=nrow_2nd-3,ncol=11)

# split the string information from the first component in the list by "\t" and unlist to put values into vectors

# 1st component
for (i in 3:length(maa_low_0.50th_split_1st)){
  maa_low_0.50th_split_1st_matrix[i-2,] <- unlist(strsplit(maa_low_0.50th_split_1st[i], "\t"))
}

# 2nd component
for (i in 4:length(maa_low_0.50th_split_2nd)){
  maa_low_0.50th_split_2nd_matrix[i-3,] <- unlist(strsplit(maa_low_0.50th_split_2nd[i], "\t"))
}

# add column names to the matrices
colnames(maa_low_0.50th_split_1st_matrix) <- c("Node","Degree",'Clustering', 'Province', 'Numb_Occ', 'Age', 'betweenness', 'location', 'OccColor', 'AgeColor', 'ProvColor')
colnames(maa_low_0.50th_split_2nd_matrix) <- c("Node","Degree",'Clustering', 'Province', 'Numb_Occ', 'Age', 'betweenness', 'location', 'OccColor', 'AgeColor', 'ProvColor')

# make the matrices into data frames
maa_low_0.50th_split_1st_df <- as.data.frame(maa_low_0.50th_split_1st_matrix)
maa_low_0.50th_split_2nd_df <- as.data.frame(maa_low_0.50th_split_2nd_matrix)

typeof(maa_low_0.50th_split_1st_df)

# read in the lat/long file for the substage
maa_low_loc <- read.csv('R_Config_for_EN_OUTPUT/low maa/maa_low_gen_great3_avgloc.csv')
colnames(maa_low_loc) <- c("Node","y","x","Age")
```

```
# merge the dfs together to get the locaiton information with the components
maa_low_0.50th_split_1st_merge <- merge(maa_low_0.50th_split_1st_df,maa_low_loc,by="Node")
maa_low_0.50th_split_2nd_merge <- merge(maa_low_0.50th_split_2nd_df,maa_low_loc,by="Node")

# simplify to just the x,y coord for each component
maa_low_0.50th_split_1st_loc <- maa_low_0.50th_split_1st_merge[,c(12,13)]
maa_low_0.50th_split_2nd_loc <- maa_low_0.50th_split_2nd_merge[,c(12,13)]

# plot the map of the different components

library("ggplot2")
theme_set(theme_bw())
library("sf")

library("rnaturalearth")
library("rnaturalearthdata")
library("ggspatial")

world <- ne_countries(scale = "medium", returnclass = "sf") # get sf info for whole world from
rnaturalearth
class(world)

# Plotting the map
ggplot(data = world) +
 geom_sf(size=0.1) +
 xlab("Longitude") + ylab("Latitude") +
 annotation_scale(location = "bl", width_hint = 0.5) +
 annotation_north_arrow(location = "bl", which_north = "true",
                height = unit(0.4,"in"), width = unit(0.3,"in"),
                pad_x = unit(0.1, "in"), pad_y = unit(0.3, "in"),
                style = north_arrow_minimal) +
 annotate(geom = "text", x = -90, y = 26, label = "Gulf of Mexico",
       fontface = "italic", color = "grey22", size = 3) +
 theme(panel.grid.major = element_line(color = "grey",
                           size = 0.3,
                           linetype = 2)) +
 geom_point(data = maa_low_0.50th_split_1st_loc, aes(x = x, y = y), size = 3,
        shape = 21, fill = "darkred") +
 geom_point(data = maa_low_0.50th_split_2nd_loc, aes(x = x, y = y), size = 3,
        shape = 21, fill = "lightblue") +
 coord_sf(xlim = c(-120, -75), ylim = c(20, 60), expand = FALSE)
ggsave("ComponentAnalysis/maa_low_0.50th.pdf")

# Second TH level:

### Contains: one major component, one with 4, all others less than 2 nodes each >>> PLOT 2
COMPONENTS
```

```
# split the character string into lists of component information
maa_low_0.60th_split <- split(maa_low_0.60th, cumsum( grepl("^---", maa_low_0.60th)))
maa_low_0.60th_split$"0"

# create individual list objects of each relevant component
maa_low_0.60th_split_1st <- maa_low_0.60th_split$"0"
maa_low_0.60th_split_2nd <- maa_low_0.60th_split$"1"

typeof(maa_low_0.60th_split_2nd)

# get lenght of each component (number of rows)
nrow_1st <- length(maa_low_0.60th_split_1st)
nrow_2nd <- length(maa_low_0.60th_split_2nd)


# create an empty matrix to put the first component in (use length of rows for dimensions, remember all components after 1st have extra leading row)
maa_low_0.60th_split_1st_matrix <- matrix(data = NA, nrow=nrow_1st-2,ncol=11)
maa_low_0.60th_split_2nd_matrix <- matrix(data = NA, nrow=nrow_2nd-3,ncol=11)

# split the string information from the first component in the list by "\t" and unlist to put values into vectors

# 1st component
for (i in 3:length(maa_low_0.60th_split_1st)){
  maa_low_0.60th_split_1st_matrix[i-2,] <- unlist(strsplit(maa_low_0.60th_split_1st[i], "\t"))
}

# 2nd component
for (i in 4:length(maa_low_0.60th_split_2nd)){
  maa_low_0.60th_split_2nd_matrix[i-3,] <- unlist(strsplit(maa_low_0.60th_split_2nd[i], "\t"))
}

# add column names to the matrices
colnames(maa_low_0.60th_split_1st_matrix) <- c("Node","Degree",'Clustering', 'Province', 'Numb_Occ', 'Age', 'betweenness', 'location', 'OccColor', 'AgeColor', 'ProvColor')
colnames(maa_low_0.60th_split_2nd_matrix) <- c("Node","Degree",'Clustering', 'Province', 'Numb_Occ', 'Age', 'betweenness', 'location', 'OccColor', 'AgeColor', 'ProvColor')

# make the matrices into data frames
maa_low_0.60th_split_1st_df <- as.data.frame(maa_low_0.60th_split_1st_matrix)
maa_low_0.60th_split_2nd_df <- as.data.frame(maa_low_0.60th_split_2nd_matrix) # NOTE: Not the same component as the 2nd component for the 0.50 TH level

typeof(maa_low_0.60th_split_1st_df)

# read in the lat/long file for the substage
maa_low_loc <- read.csv('R_Config_for_EN_OUTPUT/low maa/maa_low_gen_great3_avgloc.csv')
```

```
colnames(maa_low_loc) <- c("Node","y","x","Age")

# merge the dfs together to get the locaiton information with the components
maa_low_0.60th_split_1st_merge <- merge(maa_low_0.60th_split_1st_df,maa_low_loc,by="Node")
maa_low_0.60th_split_2nd_merge <- merge(maa_low_0.60th_split_2nd_df,maa_low_loc,by="Node")

# simplify to just the x,y coord for each component
maa_low_0.60th_split_1st_loc <- maa_low_0.60th_split_1st_merge[,c(12,13)]
maa_low_0.60th_split_2nd_loc <- maa_low_0.60th_split_2nd_merge[,c(12,13)]

# plot the map of the different components

library("ggplot2")
theme_set(theme_bw())
library("sf")

library("rnaturalearth")
library("rnaturalearthdata")
library("ggspatial")

world <- ne_countries(scale = "medium", returnclass = "sf") # get sf info for whole world from
rnaturalearth
class(world)

# Plotting the map
ggplot(data = world) +
  geom_sf(size=0.1) +
  xlab("Longitude") + ylab("Latitude") +
  annotation_scale(location = "bl", width_hint = 0.5) +
  annotation_north_arrow(location = "bl", which_north = "true",
                         height = unit(0.4,"in"), width = unit(0.3,"in"),
                         pad_x = unit(0.1, "in"), pad_y = unit(0.3, "in"),
                         style = north_arrow_minimal) +
  annotate(geom = "text", x = -90, y = 26, label = "Gulf of Mexico",
           fontface = "italic", color = "grey22", size = 3) +
  theme(panel.grid.major = element_line(color = "grey",
                                        size = 0.3,
                                        linetype = 2)) +
  geom_point(data = maa_low_0.60th_split_1st_loc, aes(x = x, y = y), size = 3,
             shape = 21, fill = "darkred") +
  geom_point(data = maa_low_0.60th_split_2nd_loc, aes(x = x, y = y), size = 3,
             shape = 21, fill = "lightblue") +
  coord_sf(xlim = c(-120, -75), ylim = c(20, 60), expand = FALSE)
ggsave("ComponentAnalysis/maa_low_0.60th.pdf")

# Third TH level:
```

```
### Contains: one major component, one with 7, all others less than 1 nodes each >>> PLOT 2 COMPONENTS

# split the character string into lists of component information
maa_low_0.75th_split <- split(maa_low_0.75th, cumsum( grepl("^---", maa_low_0.75th)))
maa_low_0.75th_split$"0"

# create individual list objects of each relevant component
maa_low_0.75th_split_1st <- maa_low_0.75th_split$"0"
maa_low_0.75th_split_2nd <- maa_low_0.75th_split$"1"

typeof(maa_low_0.75th_split_2nd)

# get lenght of each component (number of rows)
nrow_1st <- length(maa_low_0.75th_split_1st)
nrow_2nd <- length(maa_low_0.75th_split_2nd)

# create an empty matrix to put the first component in (use length of rows for dimensions, remember all components after 1st have extra leading row)
maa_low_0.75th_split_1st_matrix <- matrix(data = NA, nrow=nrow_1st-2,ncol=11)
maa_low_0.75th_split_2nd_matrix <- matrix(data = NA, nrow=nrow_2nd-3,ncol=11)

# split the string information from the first component in the list by "\t" and unlist to put values into vectors

# 1st component
for (i in 3:length(maa_low_0.75th_split_1st)){
  maa_low_0.75th_split_1st_matrix[i-2,] <- unlist(strsplit(maa_low_0.75th_split_1st[i], "\t"))
}

# 2nd component
for (i in 4:length(maa_low_0.75th_split_2nd)){
  maa_low_0.75th_split_2nd_matrix[i-3,] <- unlist(strsplit(maa_low_0.75th_split_2nd[i], "\t"))
}

# add column names to the matrices
colnames(maa_low_0.75th_split_1st_matrix) <- c("Node","Degree",'Clustering', 'Province', 'Numb_Occ', 'Age', 'betweenness', 'location', 'OccColor', 'AgeColor', 'ProvColor')
colnames(maa_low_0.75th_split_2nd_matrix) <- c("Node","Degree",'Clustering', 'Province', 'Numb_Occ', 'Age', 'betweenness', 'location', 'OccColor', 'AgeColor', 'ProvColor')

# make the matrices into data frames
maa_low_0.75th_split_1st_df <- as.data.frame(maa_low_0.75th_split_1st_matrix)
maa_low_0.75th_split_2nd_df <- as.data.frame(maa_low_0.75th_split_2nd_matrix) # NOTE: Same 2nd component as the 0.60 TH level but with more nodes

typeof(maa_low_0.75th_split_1st_df)
```

```
# read in the lat/long file for the substage
maa_low_loc <- read.csv('R_Config_for_EN_OUTPUT/low maa/maa_low_gen_great3_avgloc.csv')
colnames(maa_low_loc) <- c("Node","y","x","Age")

# merge the dfs together to get the locaiton information with the components
maa_low_0.75th_split_1st_merge <- merge(maa_low_0.75th_split_1st_df,maa_low_loc,by="Node")
maa_low_0.75th_split_2nd_merge <- merge(maa_low_0.75th_split_2nd_df,maa_low_loc,by="Node")

# simplify to just the x,y coord for each component
maa_low_0.75th_split_1st_loc <- maa_low_0.75th_split_1st_merge[,c(12,13)]
maa_low_0.75th_split_2nd_loc <- maa_low_0.75th_split_2nd_merge[,c(12,13)]

# plot the map of the different components

library("ggplot2")
theme_set(theme_bw())
library("sf")

library("rnaturalearth")
library("rnaturalearthdata")
library("ggspatial")

world <- ne_countries(scale = "medium", returnclass = "sf") # get sf info for whole world from
rnaturalearth
class(world)

# Plotting the map
ggplot(data = world) +
 geom_sf(size=0.1) +
 xlab("Longitude") + ylab("Latitude") +
 annotation_scale(location = "bl", width_hint = 0.5) +
 annotation_north_arrow(location = "bl", which_north = "true",
                height = unit(0.4,"in"), width = unit(0.3,"in"),
                pad_x = unit(0.1, "in"), pad_y = unit(0.3, "in"),
                style = north_arrow_minimal) +
 annotate(geom = "text", x = -90, y = 26, label = "Gulf of Mexico",
       fontface = "italic", color = "grey22", size = 3) +
 theme(panel.grid.major = element_line(color = "grey",
                           size = 0.3,
                           linetype = 2)) +
 geom_point(data = maa_low_0.75th_split_1st_loc, aes(x = x, y = y), size = 3,
        shape = 21, fill = "darkred") +
 geom_point(data = maa_low_0.75th_split_2nd_loc, aes(x = x, y = y), size = 3,
        shape = 21, fill = "lightblue") +
 coord_sf(xlim = c(-120, -75), ylim = c(20, 60), expand = FALSE)
ggsave("ComponentAnalysis/maa_low_0.75th.pdf")
```

### Maps show that the 2nd component for TH levels 0.75-0.60 are the same, located in the eastern side of the GCP. The secondary componenet in the 0.50 TH,

# however, is not the same and occupies the middle WIS (assumably it is a result of the WIS major component breaking down at lower THs). At TH below the

# 0.60 TH, it appears that the GCP breaks down compoletely to less than 4 connected node components.

*Code for Calculating Average Network Link Weights across Distance Bins*

```
setwd("C:/Users/ceara/Documents/Province Project/Occurrence Data/VettedCombinedDB")


# Calculalte geographic distance for paleo-coords between 5deg lats
x0 <- c(-100.0,35.0)
y0 <- c(-100.0,40.0)

x <- c(-100.0,40.0)
y <- c(-100.0,45.0)

x1 <- c(-100.0,45.0)
y1 <- c(-100.0,50.0)

x2 <- c(-100.0,50.0)
y2 <- c(-100.0,55.0)

distGeo(x0,y0, a=6378137, f=1/298.257223563)
distGeo(x,y, a=6378137, f=1/298.257223563)
distGeo(x1,y1, a=6378137, f=1/298.257223563)
distGeo(x2,y2, a=6378137, f=1/298.257223563)

555*2

library(geosphere) # load geosphere package

# load the chronosphere package to run this code. This should allow for reconstructions of plat and plong:
# https://www.evolv-ed.net/post/chronosphere-paleomap/chronosphere-paleomap/
library(chronosphere)
citation("geosphere")

##### Lower Campanian #####

# read in the lat/long file for the substage
cam_low_loc <- read.csv('R_Config_for_EN_OUTPUT/Low CAM/CAM_low_gen_great3_avgloc.csv')

# get just the lat long locations from the table
cam_low_simple_loc <- cam_low_loc[,c(3,2)]

    # NOTE: Remember that these are average locations based on the fossil occ found in each grid cell, and therefore already approsimations
    # which may be biased by not converting the fossil locations to paleocoordinates originally...

# Get the map information for this substage (age chosen is very approximate for this substage because limited options)
cam_low_maps <- reconstruct("plates", age=80) # get paleo plate info
cam_low_coast <- reconstruct("coastlines", age=80) # get coastline info

# Reconstruct paleocoord using age of rough age of substage (approximate, will use the same age for M Cam because limited options...)
```

```
cam_low_ploc <- reconstruct(cam_low_simple_loc, age=c(80))

# get min and max paleolong
cam_low_minlong <- min(cam_low_ploc[,1]) - 10
cam_low_maxlong <- max(cam_low_ploc[,1]) + 10

# get min and max paleolat
cam_low_minlat <- min(cam_low_ploc[,2]) - 15
cam_low_maxlat <- max(cam_low_ploc[,2]) + 15

# plot the maps and points in paleo positions to get an idea about their layout
par(mfrow = c(1,1))
plot(cam_low_maps, col = 'grey', border = NA, xlim = c(cam_low_minlong,cam_low_maxlong), ylim =
c(cam_low_minlat,cam_low_maxlat))
lines(cam_low_coast, col = "black", xlim = c(cam_low_minlong,cam_low_maxlong), ylim =
c(cam_low_maxlat,cam_low_minlat))
points(cam_low_ploc, pch=3, col="red")

# get number of unique locations to use for matrix dimensions
low_cam_count <- nrow(cam_low_loc)

# create empty dataframe to put the geographic distances into (use the names of the diffrent grid cells as
row and column names)
low_cam_geodist_matrix <- data.frame(matrix(ncol = low_cam_count, nrow = low_cam_count))
colnames(low_cam_geodist_matrix) <- cam_low_loc$node_label # give the matrix column names based
on original grid names
rownames(low_cam_geodist_matrix) <- cam_low_loc$node_label # give the matrix row names based on
original grid names

# run for loop to get distances using geodesic distance calc
for (i in 1:nrow(cam_low_ploc)){
  low_cam_geodist_matrix[i,] <- distGeo(cam_low_simple_loc[i,],cam_low_simple_loc, a=6378137,
f=1/298.257223563)
}

     # results are in meters of distance!!!

# check the dimentions and summary info for the matrix
ncol(low_cam_geodist_matrix)
head(low_cam_geodist_matrix)

max(low_cam_geodist_matrix)
min(low_cam_geodist_matrix)
mean(as.matrix(low_cam_geodist_matrix), na.rm = TRUE)
median(as.matrix(low_cam_geodist_matrix), na.rm = TRUE)

# make the matrix into a lower triangle (to make following calc less cumbersome and remove duplicates)
low_cam_geodist_matrix[lower.tri(low_cam_geodist_matrix, diag = TRUE)] <- NA
low_cam_geodist_matrix
```

```
# create a sequence of numbers to use for making bins (these bins were chosen to try to capture variation in values while minimizing bin number)
seq <- c(0,50000, 100000, 200000, 400000, 600000, 1000000, 1500000, 2000000, 2500000, 3000000, 3500000)

# read in the Jaccard dist matrix created using EDENetworks
low_cam_dist <- read.table("EDENetwork_outputs/Low CAM/low_CAM_distanceMatrix.txt", header = FALSE) # jaccard values matrix

low_cam_names <- read.table("EDENetwork_outputs/Low CAM/low_CAM_distanceMatrix_names.txt",header = FALSE) # node names vector
colnames(low_cam_names) <- "Node"

# give the rows and columns names based on the node names file
colnames(low_cam_dist) <- c(low_cam_names$Node)
rownames(low_cam_dist) <- c(low_cam_names$Node)

# check that the two matrices have an identical arrangement
a <- colnames(low_cam_geodist_matrix)
b <- colnames(low_cam_dist)

a == b # They do, yay!

#### Calculate the average and mean jaccard weights for the different bins

## 0-50 km ##

# get index positions of all values in geographic distance matrix within 0-50 km
low_cam_geodist_0_50 = which(low_cam_geodist_matrix <50000, arr.ind = TRUE)
# extract corresponding values from the jaccard distance matrix
low_cam_geodist_0_50_jaccard <- low_cam_dist[c(low_cam_geodist_0_50[,1]),c(low_cam_geodist_0_50[,2])]

# get the average and median jaccard weight values for the 0-50 km bin
low_cam_geodist_0_50_avg <- mean(as.matrix(low_cam_geodist_0_50_jaccard), na.rm = TRUE)
low_cam_geodist_0_50_med <- median(as.matrix(low_cam_geodist_0_50_jaccard), na.rm = TRUE)

## 50-100 km ##

# get index positions of all values in geographic distance matrix within 50-100 km
low_cam_geodist_50_100 = which(low_cam_geodist_matrix >50000 & low_cam_geodist_matrix < 100000, arr.ind = TRUE)
# extract corresponding values from the jaccard distance matrix
low_cam_geodist_50_100_jaccard <- low_cam_dist[c(low_cam_geodist_50_100[,1]),c(low_cam_geodist_50_100[,2])]

# get the average and median jaccard weight values for the 50-100 km bin
low_cam_geodist_50_100_avg <- mean(as.matrix(low_cam_geodist_50_100_jaccard), na.rm = TRUE)
low_cam_geodist_50_100_med <- median(as.matrix(low_cam_geodist_50_100_jaccard), na.rm = TRUE)
```

```
## 100-200 km ##

# get index positions of all values in geographic distance matrix within 100-200 km
low_cam_geodist_100_200 = which(low_cam_geodist_matrix >100000 & low_cam_geodist_matrix <
200000, arr.ind = TRUE)
# extract corresponding values from the jaccard distance matrix
low_cam_geodist_100_200_jaccard <-
low_cam_dist[c(low_cam_geodist_100_200[,1]),c(low_cam_geodist_100_200[,2])]

# get the average and median jaccard weight values for the 100-200 km bin
low_cam_geodist_100_200_avg <- mean(as.matrix(low_cam_geodist_100_200_jaccard), na.rm = TRUE)
low_cam_geodist_100_200_med <- median(as.matrix(low_cam_geodist_100_200_jaccard), na.rm =
TRUE)

## 200-400 km ##

# get index positions of all values in geographic distance matrix within 200-400 km
low_cam_geodist_200_400 = which(low_cam_geodist_matrix >200000 & low_cam_geodist_matrix <
400000, arr.ind = TRUE)
# extract corresponding values from the jaccard distance matrix
low_cam_geodist_200_400_jaccard <-
low_cam_dist[c(low_cam_geodist_200_400[,1]),c(low_cam_geodist_200_400[,2])]

# get the average and median jaccard weight values for the 200-400 km bin
low_cam_geodist_200_400_avg <- mean(as.matrix(low_cam_geodist_200_400_jaccard), na.rm = TRUE)
low_cam_geodist_200_400_med <- median(as.matrix(low_cam_geodist_200_400_jaccard), na.rm =
TRUE)

## 400-600 km ##

# get index positions of all values in geographic distance matrix within 400-600 km
low_cam_geodist_400_600 = which(low_cam_geodist_matrix >400000 & low_cam_geodist_matrix <
600000, arr.ind = TRUE)
# extract corresponding values from the jaccard distance matrix
low_cam_geodist_400_600_jaccard <-
low_cam_dist[c(low_cam_geodist_400_600[,1]),c(low_cam_geodist_400_600[,2])]

# get the average and median jaccard weight values for the 400-600 km bin
low_cam_geodist_400_600_avg <- mean(as.matrix(low_cam_geodist_400_600_jaccard), na.rm = TRUE)
low_cam_geodist_400_600_med <- median(as.matrix(low_cam_geodist_400_600_jaccard), na.rm =
TRUE)

## 600-1000 km ##

# get index positions of all values in geographic distance matrix within 600-1000 km
low_cam_geodist_600_1000 = which(low_cam_geodist_matrix >600000 & low_cam_geodist_matrix <
1000000, arr.ind = TRUE)
# extract corresponding values from the jaccard distance matrix
low_cam_geodist_600_1000_jaccard <-
low_cam_dist[c(low_cam_geodist_600_1000[,1]),c(low_cam_geodist_600_1000[,2])]
```

```
# get the average and median jaccard weight values for the 600-1000 km bin
low_cam_geodist_600_1000_avg <- mean(as.matrix(low_cam_geodist_600_1000_jaccard), na.rm =
TRUE)
low_cam_geodist_600_1000_med <- median(as.matrix(low_cam_geodist_600_1000_jaccard), na.rm =
TRUE)

## 1000-1500 km ##

# get index positions of all values in geographic distance matrix within 1000-1500 km
low_cam_geodist_1000_1500 = which(low_cam_geodist_matrix >1000000 & low_cam_geodist_matrix
< 1500000, arr.ind = TRUE)
# extract corresponding values from the jaccard distance matrix
low_cam_geodist_1000_1500_jaccard <-
low_cam_dist[c(low_cam_geodist_1000_1500[,1]),c(low_cam_geodist_1000_1500[,2])]

# get the average and median jaccard weight values for the 1000-1500 km bin
low_cam_geodist_1000_1500_avg <- mean(as.matrix(low_cam_geodist_1000_1500_jaccard), na.rm =
TRUE)
low_cam_geodist_1000_1500_med <- median(as.matrix(low_cam_geodist_1000_1500_jaccard), na.rm =
TRUE)

## 1500-2000 km ##

# get index positions of all values in geographic distance matrix within 1500-2000 km
low_cam_geodist_1500_2000 = which(low_cam_geodist_matrix >1500000 & low_cam_geodist_matrix
< 2000000, arr.ind = TRUE)
# extract corresponding values from the jaccard distance matrix
low_cam_geodist_1500_2000_jaccard <-
low_cam_dist[c(low_cam_geodist_1500_2000[,1]),c(low_cam_geodist_1500_2000[,2])]

# get the average and median jaccard weight values for the 1500-2000 km bin
low_cam_geodist_1500_2000_avg <- mean(as.matrix(low_cam_geodist_1500_2000_jaccard), na.rm =
TRUE)
low_cam_geodist_1500_2000_med <- median(as.matrix(low_cam_geodist_1500_2000_jaccard), na.rm =
TRUE)

## 2000-2500 km ##

# get index positions of all values in geographic distance matrix within 2000-2500 km
low_cam_geodist_2000_2500 = which(low_cam_geodist_matrix >2000000 & low_cam_geodist_matrix
< 2500000, arr.ind = TRUE)
# extract corresponding values from the jaccard distance matrix
low_cam_geodist_2000_2500_jaccard <-
low_cam_dist[c(low_cam_geodist_2000_2500[,1]),c(low_cam_geodist_2000_2500[,2])]

# get the average and median jaccard weight values for the 2000-2500 km bin
low_cam_geodist_2000_2500_avg <- mean(as.matrix(low_cam_geodist_2000_2500_jaccard), na.rm =
TRUE)
low_cam_geodist_2000_2500_med <- median(as.matrix(low_cam_geodist_2000_2500_jaccard), na.rm =
TRUE)
```

```
## 2500-3000 km ##

# get index positions of all values in geographic distance matrix within 2500-2500 km
low_cam_geodist_2500_3000 = which(low_cam_geodist_matrix >2500000 & low_cam_geodist_matrix
< 3000000, arr.ind = TRUE)
# extract corresponding values from the jaccard distance matrix
low_cam_geodist_2500_3000_jaccard <-
low_cam_dist[c(low_cam_geodist_2500_3000[,1]),c(low_cam_geodist_2500_3000[,2])]

# get the average and median jaccard weight values for the 2500-2500 km bin
low_cam_geodist_2500_3000_avg <- mean(as.matrix(low_cam_geodist_2500_3000_jaccard), na.rm =
TRUE)
low_cam_geodist_2500_3000_med <- median(as.matrix(low_cam_geodist_2500_3000_jaccard), na.rm =
TRUE)

## 3000-3500 km ##

# get index positions of all values in geographic distance matrix within 3000-3500 km
low_cam_geodist_3000_3500 = which(low_cam_geodist_matrix >3000000, arr.ind = TRUE)
# extract corresponding values from the jaccard distance matrix
low_cam_geodist_3000_3500_jaccard <-
low_cam_dist[c(low_cam_geodist_3000_3500[,1]),c(low_cam_geodist_3000_3500[,2])]

# get the average and median jaccard weight values for the 3000-3500 km bin
low_cam_geodist_3000_3500_avg <- mean(as.matrix(low_cam_geodist_3000_3500_jaccard), na.rm =
TRUE)
low_cam_geodist_3000_3500_med <- median(as.matrix(low_cam_geodist_3000_3500_jaccard), na.rm =
TRUE)

low_cam_freq_geodist_bins <- c(nrow(low_cam_geodist_0_50),
                  nrow(low_cam_geodist_50_100),
                  nrow(low_cam_geodist_100_200),
                  nrow(low_cam_geodist_200_400),
                  nrow(low_cam_geodist_400_600),
                  nrow(low_cam_geodist_600_1000),
                  nrow(low_cam_geodist_1000_1500),
                  nrow(low_cam_geodist_1500_2000),
                  nrow(low_cam_geodist_2000_2500),
                  nrow(low_cam_geodist_2500_3000),
                  nrow(low_cam_geodist_3000_3500))

low_cam_geodist_avg_bins <- c(low_cam_geodist_0_50_avg,
                  low_cam_geodist_50_100_avg,
                  low_cam_geodist_100_200_avg,
                  low_cam_geodist_200_400_avg,
                  low_cam_geodist_400_600_avg,
                  low_cam_geodist_600_1000_avg,
                  low_cam_geodist_1000_1500_avg,
                  low_cam_geodist_1500_2000_avg,
                  low_cam_geodist_2000_2500_avg,
                  low_cam_geodist_2500_3000_avg,
```

```
                low_cam_geodist_3000_3500_avg)


low_cam_geodist_med_bins <- c(low_cam_geodist_0_50_med,
                low_cam_geodist_50_100_med,
                low_cam_geodist_100_200_med,
                low_cam_geodist_200_400_med,
                low_cam_geodist_400_600_med,
                low_cam_geodist_600_1000_med,
                low_cam_geodist_1000_1500_med,
                low_cam_geodist_1500_2000_med,
                low_cam_geodist_2000_2500_med,
                low_cam_geodist_2500_3000_med,
                low_cam_geodist_3000_3500_med)

low_cam_geodist_bins_table <-
rbind(low_cam_geodist_avg_bins,low_cam_geodist_med_bins,low_cam_freq_geodist_bins)

colnames(low_cam_geodist_bins_table) <- c("0-50 km","50-100 km","100-200 km","200-400 km","400-
600 km","600-1000 km",
                        "1000-1500 km","1500-2000 km","2000-2500 km","2500-3000 km","3000-3500
km")
rownames(low_cam_geodist_bins_table) <- c("Mean Jaccard Weight","Median Jaccard Weight","Numb.
Sites in Bin")

low_cam_geodist_bins_table

write.csv(low_cam_geodist_bins_table, file = 'WeightAnalysis/low_cam_geodist_bins_table.csv')

### PLot the lines #####

plot(low_cam_geodist_bins_table[1,],col="blue",pch=16,axes=FALSE, ann=FALSE)
axis(1, at=1:11, lab=c("0-50 ","50-100","100-200","200-400","400-600","600-1000",
                "1000-1500","1500-2000","2000-2500","2500-3000","3000-3500"))
axis(2, at=c(0.7,0.75,0.80,0.85,0.90,0.95,1))
box()
title(xlab="Distance Bins (km)")
title(ylab="Bray–Curtis Dissimilarity Index Weight ")
title(main="Average Weight by Geographicc Distance")

##### Middle Campanian #####

# read in the lat/long file for the substage
cam_mid_loc <- read.csv('R_Config_for_EN_OUTPUT/Mid CAM/CAM_mid_gen_great3_avgloc.csv')

# get just the lat long locations from the table
cam_mid_simple_loc <- cam_mid_loc[,c(3,2)]

# NOTE: Remember that these are average locations based on the fossil occ found in each grid cell, and
therefore already approsimations
# which may be biased by not converting the fossil locations to paleocoordinates originally...
```

```
# Get the map information for this substage (age chosen is very approximate for this substage because limited options)
cam_mid_maps <- reconstruct("plates", age=80) # get paleo plate info
cam_mid_coast <- reconstruct("coastlines", age=80) # get coastline info

# Reconstruct paleocoord using age of rough age of substage (approximate, will use the same age for M Cam because limited options...)
cam_mid_ploc <- reconstruct(cam_mid_simple_loc, age=c(80))

# get min and max paleolong
cam_mid_minlong <- min(cam_mid_ploc[,1]) - 10
cam_mid_maxlong <- max(cam_mid_ploc[,1]) + 10

# get min and max paleolat
cam_mid_minlat <- min(cam_mid_ploc[,2]) - 15
cam_mid_maxlat <- max(cam_mid_ploc[,2]) + 15

# plot the maps and points in paleo positions to get an idea about their layout
par(mfrow = c(1,1))
plot(cam_mid_maps, col = 'grey', border = NA, xlim = c(cam_mid_minlong,cam_mid_maxlong), ylim = c(cam_mid_minlat,cam_mid_maxlat))
lines(cam_mid_coast, col = "black", xlim = c(cam_mid_minlong,cam_mid_maxlong), ylim = c(cam_mid_maxlat,cam_mid_minlat))
points(cam_mid_ploc, pch=3, col="red")

# get number of unique locations to use for matrix dimensions
mid_cam_count <- nrow(cam_mid_loc)

# create empty dataframe to put the geographic distances into (use the names of the diffrent grid cells as row and column names)
mid_cam_geodist_matrix <- data.frame(matrix(ncol = mid_cam_count, nrow = mid_cam_count))
colnames(mid_cam_geodist_matrix) <- cam_mid_loc$node_label # give the matrix column names based on original grid names
rownames(mid_cam_geodist_matrix) <- cam_mid_loc$node_label # give the matrix row names based on original grid names

# run for loop to get distances using geodesic distance calc
for (i in 1:nrow(cam_mid_ploc)){
  mid_cam_geodist_matrix[i,] <- distGeo(cam_mid_simple_loc[i,],cam_mid_simple_loc, a=6378137, f=1/298.257223563)
}

# results are in meters of distance!!!

# check the dimentions and summary info for the matrix
ncol(mid_cam_geodist_matrix)
head(mid_cam_geodist_matrix)

max(mid_cam_geodist_matrix)
min(mid_cam_geodist_matrix)
```

```
mean(as.matrix(mid_cam_geodist_matrix), na.rm = TRUE)
median(as.matrix(mid_cam_geodist_matrix), na.rm = TRUE)

# make the matrix into a mider triangle (to make folmiding calc less cumbersome and remove duplicates)
mid_cam_geodist_matrix[lower.tri(mid_cam_geodist_matrix, diag = TRUE)] <- NA
mid_cam_geodist_matrix

# create a sequence of numbers to use for making bins (these bins were chosen to try to capture variation
in values while minimizing bin number)
seq <- c(0,50000, 100000, 200000, 400000, 600000, 1000000, 1500000, 2000000, 2500000, 3000000,
3500000)

# read in the Jaccard dist matrix created using EDENetworks
mid_cam_dist <- read.table("EDENetwork_outputs/mid CAM/CAM_mid_Matrix_distance.txt", header =
FALSE) # jaccard values matrix

mid_cam_names <- read.table("EDENetwork_outputs/mid
CAM/CAM_mid_Matrix_distance_names.txt",header = FALSE) # node names vector
colnames(mid_cam_names) <- "Node"

# give the rows and columns names based on the node names file
colnames(mid_cam_dist) <- c(mid_cam_names$Node)
rownames(mid_cam_dist) <- c(mid_cam_names$Node)

# check that the two matrices have an identical arrangement
a <- colnames(mid_cam_geodist_matrix)
b <- colnames(mid_cam_dist)

a == b # They do, yay!

#### Calculate the average and mean jaccard weights for the different bins

## 0-50 km ##

# get index positions of all values in geographic distance matrix within 0-50 km
mid_cam_geodist_0_50 = which(mid_cam_geodist_matrix <50000, arr.ind = TRUE)
# extract corresponding values from the jaccard distance matrix
mid_cam_geodist_0_50_jaccard <-
mid_cam_dist[c(mid_cam_geodist_0_50[,1]),c(mid_cam_geodist_0_50[,2])]

# get the average and median jaccard weight values for the 0-50 km bin
mid_cam_geodist_0_50_avg <- mean(as.matrix(mid_cam_geodist_0_50_jaccard), na.rm = TRUE)
mid_cam_geodist_0_50_med <- median(as.matrix(mid_cam_geodist_0_50_jaccard), na.rm = TRUE)

## 50-100 km ##

# get index positions of all values in geographic distance matrix within 50-100 km
mid_cam_geodist_50_100 = which(mid_cam_geodist_matrix >50000 & mid_cam_geodist_matrix <
100000, arr.ind = TRUE)
# extract corresponding values from the jaccard distance matrix
```

```
mid_cam_geodist_50_100_jaccard <-
mid_cam_dist[c(mid_cam_geodist_50_100[,1]),c(mid_cam_geodist_50_100[,2])]

# get the average and median jaccard weight values for the 50-100 km bin
mid_cam_geodist_50_100_avg <- mean(as.matrix(mid_cam_geodist_50_100_jaccard), na.rm = TRUE)
mid_cam_geodist_50_100_med <- median(as.matrix(mid_cam_geodist_50_100_jaccard), na.rm =
TRUE)

## 100-200 km ##

# get index positions of all values in geographic distance matrix within 100-200 km
mid_cam_geodist_100_200 = which(mid_cam_geodist_matrix >100000 & mid_cam_geodist_matrix <
200000, arr.ind = TRUE)
# extract corresponding values from the jaccard distance matrix
mid_cam_geodist_100_200_jaccard <-
mid_cam_dist[c(mid_cam_geodist_100_200[,1]),c(mid_cam_geodist_100_200[,2])]


# get the average and median jaccard weight values for the 100-200 km bin
mid_cam_geodist_100_200_avg <- mean(as.matrix(mid_cam_geodist_100_200_jaccard), na.rm =
TRUE)
mid_cam_geodist_100_200_med <- median(as.matrix(mid_cam_geodist_100_200_jaccard), na.rm =
TRUE)

## 200-400 km ##

# get index positions of all values in geographic distance matrix within 200-400 km
mid_cam_geodist_200_400 = which(mid_cam_geodist_matrix >200000 & mid_cam_geodist_matrix <
400000, arr.ind = TRUE)
# extract corresponding values from the jaccard distance matrix
mid_cam_geodist_200_400_jaccard <-
mid_cam_dist[c(mid_cam_geodist_200_400[,1]),c(mid_cam_geodist_200_400[,2])]


# get the average and median jaccard weight values for the 200-400 km bin
mid_cam_geodist_200_400_avg <- mean(as.matrix(mid_cam_geodist_200_400_jaccard), na.rm =
TRUE)
mid_cam_geodist_200_400_med <- median(as.matrix(mid_cam_geodist_200_400_jaccard), na.rm =
TRUE)

## 400-600 km ##

# get index positions of all values in geographic distance matrix within 400-600 km
mid_cam_geodist_400_600 = which(mid_cam_geodist_matrix >400000 & mid_cam_geodist_matrix <
600000, arr.ind = TRUE)
# extract corresponding values from the jaccard distance matrix
mid_cam_geodist_400_600_jaccard <-
mid_cam_dist[c(mid_cam_geodist_400_600[,1]),c(mid_cam_geodist_400_600[,2])]


# get the average and median jaccard weight values for the 400-600 km bin
```

```
mid_cam_geodist_400_600_avg <- mean(as.matrix(mid_cam_geodist_400_600_jaccard), na.rm =
TRUE)
mid_cam_geodist_400_600_med <- median(as.matrix(mid_cam_geodist_400_600_jaccard), na.rm =
TRUE)

## 600-1000 km ##

# get index positions of all values in geographic distance matrix within 600-1000 km
mid_cam_geodist_600_1000 = which(mid_cam_geodist_matrix >600000 & mid_cam_geodist_matrix <
1000000, arr.ind = TRUE)
# extract corresponding values from the jaccard distance matrix
mid_cam_geodist_600_1000_jaccard <-
mid_cam_dist[c(mid_cam_geodist_600_1000[,1]),c(mid_cam_geodist_600_1000[,2])]

# get the average and median jaccard weight values for the 600-1000 km bin
mid_cam_geodist_600_1000_avg <- mean(as.matrix(mid_cam_geodist_600_1000_jaccard), na.rm =
TRUE)
mid_cam_geodist_600_1000_med <- median(as.matrix(mid_cam_geodist_600_1000_jaccard), na.rm =
TRUE)




## 1000-1500 km ##

# get index positions of all values in geographic distance matrix within 1000-1500 km
mid_cam_geodist_1000_1500 = which(mid_cam_geodist_matrix >1000000 & mid_cam_geodist_matrix
< 1500000, arr.ind = TRUE)
# extract corresponding values from the jaccard distance matrix
mid_cam_geodist_1000_1500_jaccard <-
mid_cam_dist[c(mid_cam_geodist_1000_1500[,1]),c(mid_cam_geodist_1000_1500[,2])]


# get the average and median jaccard weight values for the 1000-1500 km bin
mid_cam_geodist_1000_1500_avg <- mean(as.matrix(mid_cam_geodist_1000_1500_jaccard), na.rm =
TRUE)
mid_cam_geodist_1000_1500_med <- median(as.matrix(mid_cam_geodist_1000_1500_jaccard), na.rm =
TRUE)

## 1500-2000 km ##

# get index positions of all values in geographic distance matrix within 1500-2000 km
mid_cam_geodist_1500_2000 = which(mid_cam_geodist_matrix >1500000 & mid_cam_geodist_matrix
< 2000000, arr.ind = TRUE)
# extract corresponding values from the jaccard distance matrix
mid_cam_geodist_1500_2000_jaccard <-
mid_cam_dist[c(mid_cam_geodist_1500_2000[,1]),c(mid_cam_geodist_1500_2000[,2])]


# get the average and median jaccard weight values for the 1500-2000 km bin
```

```
mid_cam_geodist_1500_2000_avg <- mean(as.matrix(mid_cam_geodist_1500_2000_jaccard), na.rm =
TRUE)
mid_cam_geodist_1500_2000_med <- median(as.matrix(mid_cam_geodist_1500_2000_jaccard), na.rm =
TRUE)

## 2000-2500 km ##

# get index positions of all values in geographic distance matrix within 2000-2500 km
mid_cam_geodist_2000_2500 = which(mid_cam_geodist_matrix >2000000 & mid_cam_geodist_matrix
< 2500000, arr.ind = TRUE)
# extract corresponding values from the jaccard distance matrix
mid_cam_geodist_2000_2500_jaccard <-
mid_cam_dist[c(mid_cam_geodist_2000_2500[,1]),c(mid_cam_geodist_2000_2500[,2])]

# get the average and median jaccard weight values for the 2000-2500 km bin
mid_cam_geodist_2000_2500_avg <- mean(as.matrix(mid_cam_geodist_2000_2500_jaccard), na.rm =
TRUE)
mid_cam_geodist_2000_2500_med <- median(as.matrix(mid_cam_geodist_2000_2500_jaccard), na.rm =
TRUE)

## 2500-3000 km ##

# get index positions of all values in geographic distance matrix within 2500-2500 km
mid_cam_geodist_2500_3000 = which(mid_cam_geodist_matrix >2500000 & mid_cam_geodist_matrix
< 3000000, arr.ind = TRUE)
# extract corresponding values from the jaccard distance matrix
mid_cam_geodist_2500_3000_jaccard <-
mid_cam_dist[c(mid_cam_geodist_2500_3000[,1]),c(mid_cam_geodist_2500_3000[,2])]


# get the average and median jaccard weight values for the 2500-2500 km bin
mid_cam_geodist_2500_3000_avg <- mean(as.matrix(mid_cam_geodist_2500_3000_jaccard), na.rm =
TRUE)
mid_cam_geodist_2500_3000_med <- median(as.matrix(mid_cam_geodist_2500_3000_jaccard), na.rm =
TRUE)

## 3000-3500 km ##

# get index positions of all values in geographic distance matrix within 3000-3500 km
mid_cam_geodist_3000_3500 = which(mid_cam_geodist_matrix >3000000, arr.ind = TRUE)
# extract corresponding values from the jaccard distance matrix
mid_cam_geodist_3000_3500_jaccard <-
mid_cam_dist[c(mid_cam_geodist_3000_3500[,1]),c(mid_cam_geodist_3000_3500[,2])]


# get the average and median jaccard weight values for the 3000-3500 km bin
mid_cam_geodist_3000_3500_avg <- mean(as.matrix(mid_cam_geodist_3000_3500_jaccard), na.rm =
TRUE)
mid_cam_geodist_3000_3500_med <- median(as.matrix(mid_cam_geodist_3000_3500_jaccard), na.rm =
TRUE)
```

```
mid_cam_freq_geodist_bins <- c(nrow(mid_cam_geodist_0_50),
                        nrow(mid_cam_geodist_50_100),
                        nrow(mid_cam_geodist_100_200),
                        nrow(mid_cam_geodist_200_400),
                        nrow(mid_cam_geodist_400_600),
                        nrow(mid_cam_geodist_600_1000),
                        nrow(mid_cam_geodist_1000_1500),
                        nrow(mid_cam_geodist_1500_2000),
                        nrow(mid_cam_geodist_2000_2500),
                        nrow(mid_cam_geodist_2500_3000),
                        nrow(mid_cam_geodist_3000_3500))

mid_cam_geodist_avg_bins <- c(mid_cam_geodist_0_50_avg,
                       mid_cam_geodist_50_100_avg,
                       mid_cam_geodist_100_200_avg,
                       mid_cam_geodist_200_400_avg,
                       mid_cam_geodist_400_600_avg,
                       mid_cam_geodist_600_1000_avg,
                       mid_cam_geodist_1000_1500_avg,
                       mid_cam_geodist_1500_2000_avg,
                       mid_cam_geodist_2000_2500_avg,
                       mid_cam_geodist_2500_3000_avg,
                       mid_cam_geodist_3000_3500_avg)

mid_cam_geodist_med_bins <- c(mid_cam_geodist_0_50_med,
                       mid_cam_geodist_50_100_med,
                       mid_cam_geodist_100_200_med,
                       mid_cam_geodist_200_400_med,
                       mid_cam_geodist_400_600_med,
                       mid_cam_geodist_600_1000_med,
                       mid_cam_geodist_1000_1500_med,
                       mid_cam_geodist_1500_2000_med,
                       mid_cam_geodist_2000_2500_med,
                       mid_cam_geodist_2500_3000_med,
                       mid_cam_geodist_3000_3500_med)

mid_cam_geodist_bins_table <-
rbind(mid_cam_geodist_avg_bins,mid_cam_geodist_med_bins,mid_cam_freq_geodist_bins)

colnames(mid_cam_geodist_bins_table) <- c("0-50 km","50-100 km","100-200 km","200-400 km","400-
600 km","600-1000 km",
                              "1000-1500 km","1500-2000 km","2000-2500 km","2500-3000 km","3000-
3500 km")
rownames(mid_cam_geodist_bins_table) <- c("Mean Jaccard Weight","Median Jaccard Weight","Numb.
Sites in Bin")

mid_cam_geodist_bins_table

write.csv(mid_cam_geodist_bins_table, file = 'WeightAnalysis/mid_cam_geodist_bins_table.csv')

### PLot the lines #####
```

```
plot(low_cam_geodist_bins_table[1,],col="blue",pch=16,axes=FALSE, ann=FALSE)
points(mid_cam_geodist_bins_table[1,],col="red",pch=17)
axis(1, at=1:11, lab=c("0-50 ","50-100","100-200","200-400","400-600","600-1000",
                       "1000-1500","1500-2000","2000-2500","2500-3000","3000-3500"))
axis(2, at=c(0.7,0.75,0.80,0.85,0.90,0.95,1))
box()
title(xlab="Distance Bins (km)")
title(ylab="Bray–Curtis Dissimilarity Index Weight ")
title(main="Average Weight by Geographicc Distance")

##### Upper Campanian #####

# read in the lat/long file for the substage
cam_up_loc <- read.csv('R_Config_for_EN_OUTPUT/up CAM/CAM_up_gen_great3_avgloc.csv')


# get just the lat long locations from the table
cam_up_simple_loc <- cam_up_loc[,c(3,2)]

# NOTE: Remember that these are average locations based on the fossil occ found in each grid cell, and
therefore already approsimations
# which may be biased by not converting the fossil locations to paleocoordinates originally...

# Get the map information for this substage (age chosen is very approximate for this substage because
limited options)
cam_up_maps <- reconstruct("plates", age=75) # get paleo plate info
cam_up_coast <- reconstruct("coastlines", age=75) # get coastline info

# Reconstruct paleocoord using age of rough age of substage (approximate)
cam_up_ploc <- reconstruct(cam_up_simple_loc, age=c(75))

# get min and max paleolong
cam_up_minlong <- min(cam_up_ploc[,1]) - 10
cam_up_maxlong <- max(cam_up_ploc[,1]) + 10

# get min and max paleolat
cam_up_minlat <- min(cam_up_ploc[,2]) - 15
cam_up_maxlat <- max(cam_up_ploc[,2]) + 15

# plot the maps and points in paleo positions to get an idea about their layout
par(mfrow = c(1,1))
plot(cam_up_maps, col = 'grey', border = NA, xlim = c(cam_up_minlong,cam_up_maxlong), ylim =
c(cam_up_minlat,cam_up_maxlat))
lines(cam_up_coast, col = "black", xlim = c(cam_up_minlong,cam_up_maxlong), ylim =
c(cam_up_maxlat,cam_up_minlat))
points(cam_up_ploc, pch=3, col="red")

# get number of unique locations to use for matrix dimensions
up_cam_count <- nrow(cam_up_loc)
```

```
# create empty dataframe to put the geographic distances into (use the names of the diffrent grid cells as row and column names)
up_cam_geodist_matrix <- data.frame(matrix(ncol = up_cam_count, nrow = up_cam_count))
colnames(up_cam_geodist_matrix) <- cam_up_loc$node_label # give the matrix column names based on original grid names
rownames(up_cam_geodist_matrix) <- cam_up_loc$node_label # give the matrix row names based on original grid names

# run for loop to get distances using geodesic distance calc
for (i in 1:nrow(cam_up_ploc)){
  up_cam_geodist_matrix[i,] <- distGeo(cam_up_simple_loc[i,],cam_up_simple_loc, a=6378137, f=1/298.257223563)
}

# results are in meters of distance!!!

# check the dimentions and summary info for the matrix
ncol(up_cam_geodist_matrix)
head(up_cam_geodist_matrix)

max(up_cam_geodist_matrix)
min(up_cam_geodist_matrix)
mean(as.matrix(up_cam_geodist_matrix), na.rm = TRUE)
median(as.matrix(up_cam_geodist_matrix), na.rm = TRUE)

# make the matrix into a uper triangle (to make foluping calc less cumbersome and remove duplicates)
up_cam_geodist_matrix[lower.tri(up_cam_geodist_matrix, diag = TRUE)] <- NA
up_cam_geodist_matrix

# create a sequence of numbers to use for making bins (these bins were chosen to try to capture variation in values while minimizing bin number)
seq <- c(0,50000, 100000, 200000, 400000, 600000, 1000000, 1500000, 2000000, 2500000, 3000000, 3500000)

# read in the Jaccard dist matrix created using EDENetworks
up_cam_dist <- read.table("EDENetwork_outputs/Up CAM/UCam_distance.txt", header = FALSE) # jaccard values matrix

up_cam_names <- read.table("EDENetwork_outputs/Up CAM/UCam_distance_names.txt",header = FALSE) # node names vector
colnames(up_cam_names) <- "Node"

# give the rows and columns names based on the node names file
colnames(up_cam_dist) <- c(up_cam_names$Node)
rownames(up_cam_dist) <- c(up_cam_names$Node)

# check that the two matrices have an identical arrangement
a <- colnames(up_cam_geodist_matrix)
b <- colnames(up_cam_dist)

a == b # They do, yay!
```

```
#### Calculate the average and mean jaccard weights for the different bins

## 0-50 km ##

# get index positions of all values in geographic distance matrix within 0-50 km
up_cam_geodist_0_50 = which(up_cam_geodist_matrix <50000, arr.ind = TRUE)
# extract corresponding values from the jaccard distance matrix
up_cam_geodist_0_50_jaccard <-
up_cam_dist[c(up_cam_geodist_0_50[,1]),c(up_cam_geodist_0_50[,2])]

# get the average and median jaccard weight values for the 0-50 km bin
up_cam_geodist_0_50_avg <- mean(as.matrix(up_cam_geodist_0_50_jaccard), na.rm = TRUE)
up_cam_geodist_0_50_med <- median(as.matrix(up_cam_geodist_0_50_jaccard), na.rm = TRUE)

## 50-100 km ##

# get index positions of all values in geographic distance matrix within 50-100 km
up_cam_geodist_50_100 = which(up_cam_geodist_matrix >50000 & up_cam_geodist_matrix < 100000,
arr.ind = TRUE)
# extract corresponding values from the jaccard distance matrix
up_cam_geodist_50_100_jaccard <-
up_cam_dist[c(up_cam_geodist_50_100[,1]),c(up_cam_geodist_50_100[,2])]


# get the average and median jaccard weight values for the 50-100 km bin
up_cam_geodist_50_100_avg <- mean(as.matrix(up_cam_geodist_50_100_jaccard), na.rm = TRUE)
up_cam_geodist_50_100_med <- median(as.matrix(up_cam_geodist_50_100_jaccard), na.rm = TRUE)

## 100-200 km ##

# get index positions of all values in geographic distance matrix within 100-200 km
up_cam_geodist_100_200 = which(up_cam_geodist_matrix >100000 & up_cam_geodist_matrix <
200000, arr.ind = TRUE)
# extract corresponding values from the jaccard distance matrix
up_cam_geodist_100_200_jaccard <-
up_cam_dist[c(up_cam_geodist_100_200[,1]),c(up_cam_geodist_100_200[,2])]

# get the average and median jaccard weight values for the 100-200 km bin
up_cam_geodist_100_200_avg <- mean(as.matrix(up_cam_geodist_100_200_jaccard), na.rm = TRUE)
up_cam_geodist_100_200_med <- median(as.matrix(up_cam_geodist_100_200_jaccard), na.rm =
TRUE)

## 200-400 km ##

# get index positions of all values in geographic distance matrix within 200-400 km
up_cam_geodist_200_400 = which(up_cam_geodist_matrix >200000 & up_cam_geodist_matrix <
400000, arr.ind = TRUE)
# extract corresponding values from the jaccard distance matrix
up_cam_geodist_200_400_jaccard <-
up_cam_dist[c(up_cam_geodist_200_400[,1]),c(up_cam_geodist_200_400[,2])]
```

```
# get the average and median jaccard weight values for the 200-400 km bin
up_cam_geodist_200_400_avg <- mean(as.matrix(up_cam_geodist_200_400_jaccard), na.rm = TRUE)
up_cam_geodist_200_400_med <- median(as.matrix(up_cam_geodist_200_400_jaccard), na.rm =
TRUE)

## 400-600 km ##

# get index positions of all values in geographic distance matrix within 400-600 km
up_cam_geodist_400_600 = which(up_cam_geodist_matrix >400000 & up_cam_geodist_matrix <
600000, arr.ind = TRUE)
# extract corresponding values from the jaccard distance matrix
up_cam_geodist_400_600_jaccard <-
up_cam_dist[c(up_cam_geodist_400_600[,1]),c(up_cam_geodist_400_600[,2])]

# get the average and median jaccard weight values for the 400-600 km bin
up_cam_geodist_400_600_avg <- mean(as.matrix(up_cam_geodist_400_600_jaccard), na.rm = TRUE)
up_cam_geodist_400_600_med <- median(as.matrix(up_cam_geodist_400_600_jaccard), na.rm =
TRUE)

## 600-1000 km ##

# get index positions of all values in geographic distance matrix within 600-1000 km
up_cam_geodist_600_1000 = which(up_cam_geodist_matrix >600000 & up_cam_geodist_matrix <
1000000, arr.ind = TRUE)
# extract corresponding values from the jaccard distance matrix
up_cam_geodist_600_1000_jaccard <-
up_cam_dist[c(up_cam_geodist_600_1000[,1]),c(up_cam_geodist_600_1000[,2])]

# get the average and median jaccard weight values for the 600-1000 km bin
up_cam_geodist_600_1000_avg <- mean(as.matrix(up_cam_geodist_600_1000_jaccard), na.rm = TRUE)
up_cam_geodist_600_1000_med <- median(as.matrix(up_cam_geodist_600_1000_jaccard), na.rm =
TRUE)

## 1000-1500 km ##

# get index positions of all values in geographic distance matrix within 1000-1500 km
up_cam_geodist_1000_1500 = which(up_cam_geodist_matrix >1000000 & up_cam_geodist_matrix <
1500000, arr.ind = TRUE)
# extract corresponding values from the jaccard distance matrix
up_cam_geodist_1000_1500_jaccard <-
up_cam_dist[c(up_cam_geodist_1000_1500[,1]),c(up_cam_geodist_1000_1500[,2])]

# get the average and median jaccard weight values for the 1000-1500 km bin
up_cam_geodist_1000_1500_avg <- mean(as.matrix(up_cam_geodist_1000_1500_jaccard), na.rm =
TRUE)
up_cam_geodist_1000_1500_med <- median(as.matrix(up_cam_geodist_1000_1500_jaccard), na.rm =
TRUE)

## 1500-2000 km ##
```

```
# get index positions of all values in geographic distance matrix within 1500-2000 km
up_cam_geodist_1500_2000 = which(up_cam_geodist_matrix >1500000 & up_cam_geodist_matrix <
2000000, arr.ind = TRUE)
# extract corresponding values from the jaccard distance matrix
up_cam_geodist_1500_2000_jaccard <-
up_cam_dist[c(up_cam_geodist_1500_2000[,1]),c(up_cam_geodist_1500_2000[,2])]

# get the average and median jaccard weight values for the 1500-2000 km bin
up_cam_geodist_1500_2000_avg <- mean(as.matrix(up_cam_geodist_1500_2000_jaccard), na.rm =
TRUE)
up_cam_geodist_1500_2000_med <- median(as.matrix(up_cam_geodist_1500_2000_jaccard), na.rm =
TRUE)

## 2000-2500 km ##

# get index positions of all values in geographic distance matrix within 2000-2500 km
up_cam_geodist_2000_2500 = which(up_cam_geodist_matrix >2000000 & up_cam_geodist_matrix <
2500000, arr.ind = TRUE)
# extract corresponding values from the jaccard distance matrix
up_cam_geodist_2000_2500_jaccard <-
up_cam_dist[c(up_cam_geodist_2000_2500[,1]),c(up_cam_geodist_2000_2500[,2])]

# get the average and median jaccard weight values for the 2000-2500 km bin
up_cam_geodist_2000_2500_avg <- mean(as.matrix(up_cam_geodist_2000_2500_jaccard), na.rm =
TRUE)
up_cam_geodist_2000_2500_med <- median(as.matrix(up_cam_geodist_2000_2500_jaccard), na.rm =
TRUE)

## 2500-3000 km ##

# get index positions of all values in geographic distance matrix within 2500-2500 km
up_cam_geodist_2500_3000 = which(up_cam_geodist_matrix >2500000 & up_cam_geodist_matrix <
3000000, arr.ind = TRUE)
# extract corresponding values from the jaccard distance matrix
up_cam_geodist_2500_3000_jaccard <-
up_cam_dist[c(up_cam_geodist_2500_3000[,1]),c(up_cam_geodist_2500_3000[,2])]

# get the average and median jaccard weight values for the 2500-2500 km bin
up_cam_geodist_2500_3000_avg <- mean(as.matrix(up_cam_geodist_2500_3000_jaccard), na.rm =
TRUE)
up_cam_geodist_2500_3000_med <- median(as.matrix(up_cam_geodist_2500_3000_jaccard), na.rm =
TRUE)

## 3000-3500 km ##

# get index positions of all values in geographic distance matrix within 3000-3500 km
up_cam_geodist_3000_3500 = which(up_cam_geodist_matrix >3000000, arr.ind = TRUE)
# extract corresponding values from the jaccard distance matrix
up_cam_geodist_3000_3500_jaccard <-
up_cam_dist[c(up_cam_geodist_3000_3500[,1]),c(up_cam_geodist_3000_3500[,2])]
```

```
# get the average and median jaccard weight values for the 3000-3500 km bin
up_cam_geodist_3000_3500_avg <- mean(as.matrix(up_cam_geodist_3000_3500_jaccard), na.rm =
TRUE)
up_cam_geodist_3000_3500_med <- median(as.matrix(up_cam_geodist_3000_3500_jaccard), na.rm =
TRUE)

up_cam_freq_geodist_bins <- c(nrow(up_cam_geodist_0_50),
                         nrow(up_cam_geodist_50_100),
                         nrow(up_cam_geodist_100_200),
                         nrow(up_cam_geodist_200_400),
                         nrow(up_cam_geodist_400_600),
                         nrow(up_cam_geodist_600_1000),
                         nrow(up_cam_geodist_1000_1500),
                         nrow(up_cam_geodist_1500_2000),
                         nrow(up_cam_geodist_2000_2500),
                         nrow(up_cam_geodist_2500_3000),
                         nrow(up_cam_geodist_3000_3500))

up_cam_geodist_avg_bins <- c(up_cam_geodist_0_50_avg,
                        up_cam_geodist_50_100_avg,
                        up_cam_geodist_100_200_avg,
                        up_cam_geodist_200_400_avg,
                        up_cam_geodist_400_600_avg,
                        up_cam_geodist_600_1000_avg,
                        up_cam_geodist_1000_1500_avg,
                        up_cam_geodist_1500_2000_avg,
                        up_cam_geodist_2000_2500_avg,
                        up_cam_geodist_2500_3000_avg,
                        up_cam_geodist_3000_3500_avg)

up_cam_geodist_med_bins <- c(up_cam_geodist_0_50_med,
                        up_cam_geodist_50_100_med,
                        up_cam_geodist_100_200_med,
                        up_cam_geodist_200_400_med,
                        up_cam_geodist_400_600_med,
                        up_cam_geodist_600_1000_med,
                        up_cam_geodist_1000_1500_med,
                        up_cam_geodist_1500_2000_med,
                        up_cam_geodist_2000_2500_med,
                        up_cam_geodist_2500_3000_med,
                        up_cam_geodist_3000_3500_med)


up_cam_geodist_bins_table <-
rbind(up_cam_geodist_avg_bins,up_cam_geodist_med_bins,up_cam_freq_geodist_bins)

colnames(up_cam_geodist_bins_table) <- c("0-50 km","50-100 km","100-200 km","200-400 km","400-
600 km","600-1000 km",
                                         "1000-1500 km","1500-2000 km","2000-2500 km","2500-3000 km","3000-
3500 km")
```

```
rownames(up_cam_geodist_bins_table) <- c("Mean Jaccard Weight","Median Jaccard Weight","Numb. Sites in Bin")

up_cam_geodist_bins_table

write.csv(up_cam_geodist_bins_table, file = 'WeightAnalysis/up_cam_geodist_bins_table.csv')

### Plot the lines #####

plot(low_cam_geodist_bins_table[1,],col="blue",pch=16,axes=FALSE, ann=FALSE)
points(mid_cam_geodist_bins_table[1,],col="green",pch=17)
points(up_cam_geodist_bins_table[1,],col="gold",pch=18)
axis(1, at=1:11, lab=c("0-50 ","50-100","100-200","200-400","400-600","600-1000",
                       "1000-1500","1500-2000","2000-2500","2500-3000","3000-3500"))
axis(2, at=c(0.7,0.75,0.80,0.85,0.90,0.95,1))
box()
title(xlab="Distance Bins (km)")
title(ylab="Bray–Curtis Dissimilarity Index Weight ")
title(main="Average Weight by Geographicc Distance")

##### Lower Maastrichtian #####

# read in the lat/long file for the substage
maa_low_loc <- read.csv('R_Config_for_EN_OUTPUT/low maa/maa_low_gen_great3_avgloc.csv')


# get just the lat long locations from the table
maa_low_simple_loc <- maa_low_loc[,c(3,2)]

# NOTE: Remember that these are average locations based on the fossil occ found in each grid cell, and therefore already approsimations
# which may be biased by not converting the fossil locations to paleocoordinates originally...

# Get the map information for this substage (age chosen is very approximate for this substage because limited options)
maa_low_maps <- reconstruct("plates", age=70) # get paleo plate info
maa_low_coast <- reconstruct("coastlines", age=70) # get coastline info

# Reconstruct paleocoord using age of rough age of substage (approximate)
maa_low_ploc <- reconstruct(maa_low_simple_loc, age=c(70))

# get min and max paleolong
maa_low_minlong <- min(maa_low_ploc[,1]) - 10
maa_low_maxlong <- max(maa_low_ploc[,1]) + 10

# get min and max paleolat
maa_low_minlat <- min(maa_low_ploc[,2]) - 15
maa_low_maxlat <- max(maa_low_ploc[,2]) + 15

# plot the maps and points in paleo positions to get an idea about their layout
par(mfrow = c(1,1))
```

```
plot(maa_low_maps, col = 'grey', border = NA, xlim = c(maa_low_minlong,maa_low_maxlong), ylim =
c(maa_low_minlat,maa_low_maxlat))
lines(maa_low_coast, col = "black", xlim = c(maa_low_minlong,maa_low_maxlong), ylim =
c(maa_low_maxlat,maa_low_minlat))
points(maa_low_ploc, pch=3, col="red")

# get number of unique locations to use for matrix dimensions
low_maa_count <- nrow(maa_low_loc)

# create empty dataframe to put the geographic distances into (use the names of the diffrent grid cells as
row and column names)
low_maa_geodist_matrix <- data.frame(matrix(ncol = low_maa_count, nrow = low_maa_count))
colnames(low_maa_geodist_matrix) <- maa_low_loc$node_label # give the matrix column names based
on original grid names
rownames(low_maa_geodist_matrix) <- maa_low_loc$node_label # give the matrix row names based on
original grid names

# run for loop to get distances using geodesic distance calc
for (i in 1:nrow(maa_low_ploc)){
  low_maa_geodist_matrix[i,] <- distGeo(maa_low_simple_loc[i,],maa_low_simple_loc, a=6378137,
f=1/298.257223563)
}

# results are in meters of distance!!!

# check the dimentions and summary info for the matrix
ncol(low_maa_geodist_matrix)
head(low_maa_geodist_matrix)

max(low_maa_geodist_matrix)
min(low_maa_geodist_matrix)
mean(as.matrix(low_maa_geodist_matrix), na.rm = TRUE)
median(as.matrix(low_maa_geodist_matrix), na.rm = TRUE)

# make the matrix into a lower triangle (to make following calc less cumbersome and remove dlowlicates)
low_maa_geodist_matrix[lower.tri(low_maa_geodist_matrix, diag = TRUE)] <- NA
low_maa_geodist_matrix

# create a sequence of numbers to use for making bins (these bins were chosen to try to capture variation
in values while minimizing bin number)
seq <- c(0,50000, 100000, 200000, 400000, 600000, 1000000, 1500000, 2000000, 2500000, 3000000,
3500000)

# read in the Jaccard dist matrix created using EDENetworks
low_maa_dist <- read.table("EDENetwork_outputs/Low MAA/Low_Maa_distance.txt", header =
FALSE) # jaccard values matrix

low_maa_names <- read.table("EDENetwork_outputs/Low MAA/Low_Maa_distance_names.txt",header
= FALSE) # node names vector
colnames(low_maa_names) <- "Node"
```

```
# give the rows and columns names based on the node names file
colnames(low_maa_dist) <- c(low_maa_names$Node)
rownames(low_maa_dist) <- c(low_maa_names$Node)

# check that the two matrices have an identical arrangement
a <- colnames(low_maa_geodist_matrix)
b <- colnames(low_maa_dist)

a == b # They do, yay!

#### Calculate the average and mean jaccard weights for the different bins

## 0-50 km ##

# get index positions of all values in geographic distance matrix within 0-50 km
low_maa_geodist_0_50 = which(low_maa_geodist_matrix <50000, arr.ind = TRUE)
# extract corresponding values from the jaccard distance matrix
low_maa_geodist_0_50_jaccard <-
low_maa_dist[c(low_maa_geodist_0_50[,1]),c(low_maa_geodist_0_50[,2])]

# get the average and median jaccard weight values for the 0-50 km bin
low_maa_geodist_0_50_avg <- mean(as.matrix(low_maa_geodist_0_50_jaccard), na.rm = TRUE)
low_maa_geodist_0_50_med <- median(as.matrix(low_maa_geodist_0_50_jaccard), na.rm = TRUE)

## 50-100 km ##

# get index positions of all values in geographic distance matrix within 50-100 km
low_maa_geodist_50_100 = which(low_maa_geodist_matrix >50000 & low_maa_geodist_matrix <
100000, arr.ind = TRUE)
# extract corresponding values from the jaccard distance matrix
low_maa_geodist_50_100_jaccard <-
low_maa_dist[c(low_maa_geodist_50_100[,1]),c(low_maa_geodist_50_100[,2])]

# get the average and median jaccard weight values for the 50-100 km bin
low_maa_geodist_50_100_avg <- mean(as.matrix(low_maa_geodist_50_100_jaccard), na.rm = TRUE)
low_maa_geodist_50_100_med <- median(as.matrix(low_maa_geodist_50_100_jaccard), na.rm =
TRUE)

## 100-200 km ##

# get index positions of all values in geographic distance matrix within 100-200 km
low_maa_geodist_100_200 = which(low_maa_geodist_matrix >100000 & low_maa_geodist_matrix <
200000, arr.ind = TRUE)
# extract corresponding values from the jaccard distance matrix
low_maa_geodist_100_200_jaccard <-
low_maa_dist[c(low_maa_geodist_100_200[,1]),c(low_maa_geodist_100_200[,2])]


# get the average and median jaccard weight values for the 100-200 km bin
low_maa_geodist_100_200_avg <- mean(as.matrix(low_maa_geodist_100_200_jaccard), na.rm = TRUE)
```

```
low_maa_geodist_100_200_med <- median(as.matrix(low_maa_geodist_100_200_jaccard), na.rm =
TRUE)

## 200-400 km ##

# get index positions of all values in geographic distance matrix within 200-400 km
low_maa_geodist_200_400 = which(low_maa_geodist_matrix >200000 & low_maa_geodist_matrix <
400000, arr.ind = TRUE)
# extract corresponding values from the jaccard distance matrix
low_maa_geodist_200_400_jaccard <-
low_maa_dist[c(low_maa_geodist_200_400[,1]),c(low_maa_geodist_200_400[,2])]

# get the average and median jaccard weight values for the 200-400 km bin
low_maa_geodist_200_400_avg <- mean(as.matrix(low_maa_geodist_200_400_jaccard), na.rm = TRUE)
low_maa_geodist_200_400_med <- median(as.matrix(low_maa_geodist_200_400_jaccard), na.rm =
TRUE)



## 400-600 km ##

# get index positions of all values in geographic distance matrix within 400-600 km
low_maa_geodist_400_600 = which(low_maa_geodist_matrix >400000 & low_maa_geodist_matrix <
600000, arr.ind = TRUE)
# extract corresponding values from the jaccard distance matrix
low_maa_geodist_400_600_jaccard <-
low_maa_dist[c(low_maa_geodist_400_600[,1]),c(low_maa_geodist_400_600[,2])]

# get the average and median jaccard weight values for the 400-600 km bin
low_maa_geodist_400_600_avg <- mean(as.matrix(low_maa_geodist_400_600_jaccard), na.rm = TRUE)
low_maa_geodist_400_600_med <- median(as.matrix(low_maa_geodist_400_600_jaccard), na.rm =
TRUE)

## 600-1000 km ##

# get index positions of all values in geographic distance matrix within 600-1000 km
low_maa_geodist_600_1000 = which(low_maa_geodist_matrix >600000 & low_maa_geodist_matrix <
1000000, arr.ind = TRUE)
# extract corresponding values from the jaccard distance matrix
low_maa_geodist_600_1000_jaccard <-
low_maa_dist[c(low_maa_geodist_600_1000[,1]),c(low_maa_geodist_600_1000[,2])]

# get the average and median jaccard weight values for the 600-1000 km bin
low_maa_geodist_600_1000_avg <- mean(as.matrix(low_maa_geodist_600_1000_jaccard), na.rm =
TRUE)
low_maa_geodist_600_1000_med <- median(as.matrix(low_maa_geodist_600_1000_jaccard), na.rm =
TRUE)

## 1000-1500 km ##
```

```
# get index positions of all values in geographic distance matrix within 1000-1500 km
low_maa_geodist_1000_1500 = which(low_maa_geodist_matrix >1000000 & low_maa_geodist_matrix
< 1500000, arr.ind = TRUE)
# extract corresponding values from the jaccard distance matrix
low_maa_geodist_1000_1500_jaccard <-
low_maa_dist[c(low_maa_geodist_1000_1500[,1]),c(low_maa_geodist_1000_1500[,2])]

# get the average and median jaccard weight values for the 1000-1500 km bin
low_maa_geodist_1000_1500_avg <- mean(as.matrix(low_maa_geodist_1000_1500_jaccard), na.rm =
TRUE)
low_maa_geodist_1000_1500_med <- median(as.matrix(low_maa_geodist_1000_1500_jaccard), na.rm =
TRUE)

## 1500-2000 km ##

# get index positions of all values in geographic distance matrix within 1500-2000 km
low_maa_geodist_1500_2000 = which(low_maa_geodist_matrix >1500000 & low_maa_geodist_matrix
< 2000000, arr.ind = TRUE)
# extract corresponding values from the jaccard distance matrix
low_maa_geodist_1500_2000_jaccard <-
low_maa_dist[c(low_maa_geodist_1500_2000[,1]),c(low_maa_geodist_1500_2000[,2])]


# get the average and median jaccard weight values for the 1500-2000 km bin
low_maa_geodist_1500_2000_avg <- mean(as.matrix(low_maa_geodist_1500_2000_jaccard), na.rm =
TRUE)
low_maa_geodist_1500_2000_med <- median(as.matrix(low_maa_geodist_1500_2000_jaccard), na.rm =
TRUE)

## 2000-2500 km ##

# get index positions of all values in geographic distance matrix within 2000-2500 km
low_maa_geodist_2000_2500 = which(low_maa_geodist_matrix >2000000 & low_maa_geodist_matrix
< 2500000, arr.ind = TRUE)
# extract corresponding values from the jaccard distance matrix
low_maa_geodist_2000_2500_jaccard <-
low_maa_dist[c(low_maa_geodist_2000_2500[,1]),c(low_maa_geodist_2000_2500[,2])]

# get the average and median jaccard weight values for the 2000-2500 km bin
low_maa_geodist_2000_2500_avg <- mean(as.matrix(low_maa_geodist_2000_2500_jaccard), na.rm =
TRUE)
low_maa_geodist_2000_2500_med <- median(as.matrix(low_maa_geodist_2000_2500_jaccard), na.rm =
TRUE)

## 2500-3000 km ##

# get index positions of all values in geographic distance matrix within 2500-2500 km
low_maa_geodist_2500_3000 = which(low_maa_geodist_matrix >2500000 & low_maa_geodist_matrix
< 3000000, arr.ind = TRUE)
# extract corresponding values from the jaccard distance matrix
```

```
low_maa_geodist_2500_3000_jaccard <-
low_maa_dist[c(low_maa_geodist_2500_3000[,1]),c(low_maa_geodist_2500_3000[,2])]


# get the average and median jaccard weight values for the 2500-2500 km bin
low_maa_geodist_2500_3000_avg <- mean(as.matrix(low_maa_geodist_2500_3000_jaccard), na.rm =
TRUE)
low_maa_geodist_2500_3000_med <- median(as.matrix(low_maa_geodist_2500_3000_jaccard), na.rm =
TRUE)

## 3000-3500 km ##

# get index positions of all values in geographic distance matrix within 3000-3500 km
low_maa_geodist_3000_3500 = which(low_maa_geodist_matrix >3000000, arr.ind = TRUE)
# extract corresponding values from the jaccard distance matrix
low_maa_geodist_3000_3500_jaccard <-
low_maa_dist[c(low_maa_geodist_3000_3500[,1]),c(low_maa_geodist_3000_3500[,2])]

# get the average and median jaccard weight values for the 3000-3500 km bin
low_maa_geodist_3000_3500_avg <- mean(as.matrix(low_maa_geodist_3000_3500_jaccard), na.rm =
TRUE)
low_maa_geodist_3000_3500_med <- median(as.matrix(low_maa_geodist_3000_3500_jaccard), na.rm =
TRUE)

low_maa_freq_geodist_bins <- c(nrow(low_maa_geodist_0_50),
                        nrow(low_maa_geodist_50_100),
                        nrow(low_maa_geodist_100_200),
                        nrow(low_maa_geodist_200_400),
                        nrow(low_maa_geodist_400_600),
                        nrow(low_maa_geodist_600_1000),
                        nrow(low_maa_geodist_1000_1500),
                        nrow(low_maa_geodist_1500_2000),
                        nrow(low_maa_geodist_2000_2500),
                        nrow(low_maa_geodist_2500_3000),
                        nrow(low_maa_geodist_3000_3500))

low_maa_geodist_avg_bins <- c(low_maa_geodist_0_50_avg,
                       low_maa_geodist_50_100_avg,
                       low_maa_geodist_100_200_avg,
                       low_maa_geodist_200_400_avg,
                       low_maa_geodist_400_600_avg,
                       low_maa_geodist_600_1000_avg,
                       low_maa_geodist_1000_1500_avg,
                       low_maa_geodist_1500_2000_avg,
                       low_maa_geodist_2000_2500_avg,
                       low_maa_geodist_2500_3000_avg,
                       low_maa_geodist_3000_3500_avg)

low_maa_geodist_med_bins <- c(low_maa_geodist_0_50_med,
                       low_maa_geodist_50_100_med,
                       low_maa_geodist_100_200_med,
```

```
                    low_maa_geodist_200_400_med,
                    low_maa_geodist_400_600_med,
                    low_maa_geodist_600_1000_med,
                    low_maa_geodist_1000_1500_med,
                    low_maa_geodist_1500_2000_med,
                    low_maa_geodist_2000_2500_med,
                    low_maa_geodist_2500_3000_med,
                    low_maa_geodist_3000_3500_med)

low_maa_geodist_bins_table <-
rbind(low_maa_geodist_avg_bins,low_maa_geodist_med_bins,low_maa_freq_geodist_bins)

colnames(low_maa_geodist_bins_table) <- c("0-50 km","50-100 km","100-200 km","200-400 km","400-
600 km","600-1000 km",
                                "1000-1500 km","1500-2000 km","2000-2500 km","2500-3000 km","3000-
3500 km")
rownames(low_maa_geodist_bins_table) <- c("Mean Jaccard Weight","Median Jaccard Weight","Numb.
Sites in Bin")

low_maa_geodist_bins_table

write.csv(low_maa_geodist_bins_table, file = 'WeightAnalysis/low_maa_geodist_bins_table.csv')

### Plot the lines #####

plot(low_cam_geodist_bins_table[1,],col="blue",pch=16,axes=FALSE, ann=FALSE)
points(mid_cam_geodist_bins_table[1,],col="green",pch=17)
points(up_cam_geodist_bins_table[1,],col="gold",pch=18)
points(low_maa_geodist_bins_table[1,],col="orange",pch=19)
axis(1, at=1:11, lab=c("0-50 ","50-100","100-200","200-400","400-600","600-1000",
                    "1000-1500","1500-2000","2000-2500","2500-3000","3000-3500"))
axis(2, at=c(0.7,0.75,0.80,0.85,0.90,0.95,1))
box()
title(xlab="Distance Bins (km)")
title(ylab="Bray–Curtis Dissimilarity Index Weight ")
title(main="Average Weight by Geographicc Distance")

##### Upper Maastrichtian #####

# read in the lat/long file for the substage
maa_up_loc <- read.csv('R_Config_for_EN_OUTPUT/up maa/maa_up_gen_great3_avgloc.csv')

# get just the lat long locations from the table
maa_up_simple_loc <- maa_up_loc[,c(3,2)]

# NOTE: Remember that these are average locations based on the fossil occ found in each grid cell, and
therefore already approsimations
# which may be biased by not converting the fossil locations to paleocoordinates originally...

# Get the map information for this substage (age chosen is very approximate for this substage because
limited options)
```

```
maa_up_maps <- reconstruct("plates", age=70) # get paleo plate info
maa_up_coast <- reconstruct("coastlines", age=70) # get coastline info

# Reconstruct paleocoord using age of rough age of substage (approximate)
maa_up_ploc <- reconstruct(maa_up_simple_loc, age=c(70))

# get min and max paleolong
maa_up_minlong <- min(maa_up_ploc[,1]) - 10
maa_up_maxlong <- max(maa_up_ploc[,1]) + 10

# get min and max paleolat
maa_up_minlat <- min(maa_up_ploc[,2]) - 15
maa_up_maxlat <- max(maa_up_ploc[,2]) + 15

# plot the maps and points in paleo positions to get an idea about their layout
par(mfrow = c(1,1))
plot(maa_up_maps, col = 'grey', border = NA, xlim = c(maa_up_minlong,maa_up_maxlong), ylim =
c(maa_up_minlat,maa_up_maxlat))
lines(maa_up_coast, col = "black", xlim = c(maa_up_minlong,maa_up_maxlong), ylim =
c(maa_up_maxlat,maa_up_minlat))
points(maa_up_ploc, pch=3, col="red")

# get number of unique locations to use for matrix dimensions
up_maa_count <- nrow(maa_up_loc)

# create empty dataframe to put the geographic distances into (use the names of the diffrent grid cells as
row and column names)
up_maa_geodist_matrix <- data.frame(matrix(ncol = up_maa_count, nrow = up_maa_count))
colnames(up_maa_geodist_matrix) <- maa_up_loc$node_label # give the matrix column names based on
original grid names
rownames(up_maa_geodist_matrix) <- maa_up_loc$node_label # give the matrix row names based on
original grid names

# run for loop to get distances using geodesic distance calc
for (i in 1:nrow(maa_up_ploc)){
  up_maa_geodist_matrix[i,] <- distGeo(maa_up_simple_loc[i,],maa_up_simple_loc, a=6378137,
f=1/298.257223563)
}

# results are in meters of distance!!!

# check the dimentions and summary info for the matrix
ncol(up_maa_geodist_matrix)
head(up_maa_geodist_matrix)

max(up_maa_geodist_matrix)
min(up_maa_geodist_matrix)
mean(as.matrix(up_maa_geodist_matrix), na.rm = TRUE)
median(as.matrix(up_maa_geodist_matrix), na.rm = TRUE)

# make the matrix into a uper triangle (to make foluping calc less cumbersome and remove duplicates)
```

```
up_maa_geodist_matrix[lower.tri(up_maa_geodist_matrix, diag = TRUE)] <- NA
up_maa_geodist_matrix

# create a sequence of numbers to use for making bins (these bins were chosen to try to capture variation in values while minimizing bin number)
seq <- c(0,50000, 100000, 200000, 400000, 600000, 1000000, 1500000, 2000000, 2500000, 3000000, 3500000)

# read in the Jaccard dist matrix created using EDENetworks
up_maa_dist <- read.table("EDENetwork_outputs/Up MAA/UMaa_distance.txt", header = FALSE) # jaccard values matrix

up_maa_names <- read.table("EDENetwork_outputs/Up MAA/UMaa_distance_names.txt",header = FALSE) # node names vector
colnames(up_maa_names) <- "Node"

# give the rows and columns names based on the node names file
colnames(up_maa_dist) <- c(up_maa_names$Node)
rownames(up_maa_dist) <- c(up_maa_names$Node)

# check that the two matrices have an identical arrangement
a <- colnames(up_maa_geodist_matrix)
b <- colnames(up_maa_dist)

a == b # They do, yay!

#### Calculate the average and mean jaccard weights for the different bins

## 0-50 km ##

# get index positions of all values in geographic distance matrix within 0-50 km
up_maa_geodist_0_50 = which(up_maa_geodist_matrix <50000, arr.ind = TRUE)
# extract corresponding values from the jaccard distance matrix
up_maa_geodist_0_50_jaccard <-
up_maa_dist[c(up_maa_geodist_0_50[,1]),c(up_maa_geodist_0_50[,2])]

# get the average and median jaccard weight values for the 0-50 km bin
up_maa_geodist_0_50_avg <- mean(as.matrix(up_maa_geodist_0_50_jaccard), na.rm = TRUE)
up_maa_geodist_0_50_med <- median(as.matrix(up_maa_geodist_0_50_jaccard), na.rm = TRUE)

## 50-100 km ##

# get index positions of all values in geographic distance matrix within 50-100 km
up_maa_geodist_50_100 = which(up_maa_geodist_matrix >50000 & up_maa_geodist_matrix < 100000, arr.ind = TRUE)
# extract corresponding values from the jaccard distance matrix
up_maa_geodist_50_100_jaccard <-
up_maa_dist[c(up_maa_geodist_50_100[,1]),c(up_maa_geodist_50_100[,2])]


# get the average and median jaccard weight values for the 50-100 km bin
```

```
up_maa_geodist_50_100_avg <- mean(as.matrix(up_maa_geodist_50_100_jaccard), na.rm = TRUE)
up_maa_geodist_50_100_med <- median(as.matrix(up_maa_geodist_50_100_jaccard), na.rm = TRUE)

## 100-200 km ##

# get index positions of all values in geographic distance matrix within 100-200 km
up_maa_geodist_100_200 = which(up_maa_geodist_matrix >100000 & up_maa_geodist_matrix <
200000, arr.ind = TRUE)
# extract corresponding values from the jaccard distance matrix
up_maa_geodist_100_200_jaccard <-
up_maa_dist[c(up_maa_geodist_100_200[,1]),c(up_maa_geodist_100_200[,2])]

# get the average and median jaccard weight values for the 100-200 km bin
up_maa_geodist_100_200_avg <- mean(as.matrix(up_maa_geodist_100_200_jaccard), na.rm = TRUE)
up_maa_geodist_100_200_med <- median(as.matrix(up_maa_geodist_100_200_jaccard), na.rm =
TRUE)

## 200-400 km ##

# get index positions of all values in geographic distance matrix within 200-400 km
up_maa_geodist_200_400 = which(up_maa_geodist_matrix >200000 & up_maa_geodist_matrix <
400000, arr.ind = TRUE)
# extract corresponding values from the jaccard distance matrix
up_maa_geodist_200_400_jaccard <-
up_maa_dist[c(up_maa_geodist_200_400[,1]),c(up_maa_geodist_200_400[,2])]

# get the average and median jaccard weight values for the 200-400 km bin
up_maa_geodist_200_400_avg <- mean(as.matrix(up_maa_geodist_200_400_jaccard), na.rm = TRUE)
up_maa_geodist_200_400_med <- median(as.matrix(up_maa_geodist_200_400_jaccard), na.rm =
TRUE)

## 400-600 km ##

# get index positions of all values in geographic distance matrix within 400-600 km
up_maa_geodist_400_600 = which(up_maa_geodist_matrix >400000 & up_maa_geodist_matrix <
600000, arr.ind = TRUE)
# extract corresponding values from the jaccard distance matrix
up_maa_geodist_400_600_jaccard <-
up_maa_dist[c(up_maa_geodist_400_600[,1]),c(up_maa_geodist_400_600[,2])]

# get the average and median jaccard weight values for the 400-600 km bin
up_maa_geodist_400_600_avg <- mean(as.matrix(up_maa_geodist_400_600_jaccard), na.rm = TRUE)
up_maa_geodist_400_600_med <- median(as.matrix(up_maa_geodist_400_600_jaccard), na.rm =
TRUE)

## 600-1000 km ##

# get index positions of all values in geographic distance matrix within 600-1000 km
up_maa_geodist_600_1000 = which(up_maa_geodist_matrix >600000 & up_maa_geodist_matrix <
1000000, arr.ind = TRUE)
# extract corresponding values from the jaccard distance matrix
```

```
up_maa_geodist_600_1000_jaccard <-
up_maa_dist[c(up_maa_geodist_600_1000[,1]),c(up_maa_geodist_600_1000[,2])]

# get the average and median jaccard weight values for the 600-1000 km bin
up_maa_geodist_600_1000_avg <- mean(as.matrix(up_maa_geodist_600_1000_jaccard), na.rm = TRUE)
up_maa_geodist_600_1000_med <- median(as.matrix(up_maa_geodist_600_1000_jaccard), na.rm =
TRUE)

## 1000-1500 km ##

# get index positions of all values in geographic distance matrix within 1000-1500 km
up_maa_geodist_1000_1500 = which(up_maa_geodist_matrix >1000000 & up_maa_geodist_matrix <
1500000, arr.ind = TRUE)
# extract corresponding values from the jaccard distance matrix
up_maa_geodist_1000_1500_jaccard <-
up_maa_dist[c(up_maa_geodist_1000_1500[,1]),c(up_maa_geodist_1000_1500[,2])]


# get the average and median jaccard weight values for the 1000-1500 km bin
up_maa_geodist_1000_1500_avg <- mean(as.matrix(up_maa_geodist_1000_1500_jaccard), na.rm =
TRUE)
up_maa_geodist_1000_1500_med <- median(as.matrix(up_maa_geodist_1000_1500_jaccard), na.rm =
TRUE)

## 1500-2000 km ##

# get index positions of all values in geographic distance matrix within 1500-2000 km
up_maa_geodist_1500_2000 = which(up_maa_geodist_matrix >1500000 & up_maa_geodist_matrix <
2000000, arr.ind = TRUE)
# extract corresponding values from the jaccard distance matrix
up_maa_geodist_1500_2000_jaccard <-
up_maa_dist[c(up_maa_geodist_1500_2000[,1]),c(up_maa_geodist_1500_2000[,2])]


# get the average and median jaccard weight values for the 1500-2000 km bin
up_maa_geodist_1500_2000_avg <- mean(as.matrix(up_maa_geodist_1500_2000_jaccard), na.rm =
TRUE)
up_maa_geodist_1500_2000_med <- median(as.matrix(up_maa_geodist_1500_2000_jaccard), na.rm =
TRUE)

## 2000-2500 km ##

# get index positions of all values in geographic distance matrix within 2000-2500 km
up_maa_geodist_2000_2500 = which(up_maa_geodist_matrix >2000000 & up_maa_geodist_matrix <
2500000, arr.ind = TRUE)
# extract corresponding values from the jaccard distance matrix
up_maa_geodist_2000_2500_jaccard <-
up_maa_dist[c(up_maa_geodist_2000_2500[,1]),c(up_maa_geodist_2000_2500[,2])]

# get the average and median jaccard weight values for the 2000-2500 km bin
```

```
up_maa_geodist_2000_2500_avg <- mean(as.matrix(up_maa_geodist_2000_2500_jaccard), na.rm =
TRUE)
up_maa_geodist_2000_2500_med <- median(as.matrix(up_maa_geodist_2000_2500_jaccard), na.rm =
TRUE)

## 2500-3000 km ##

# get index positions of all values in geographic distance matrix within 2500-2500 km
up_maa_geodist_2500_3000 = which(up_maa_geodist_matrix >2500000 & up_maa_geodist_matrix <
3000000, arr.ind = TRUE)
# extract corresponding values from the jaccard distance matrix
up_maa_geodist_2500_3000_jaccard <-
up_maa_dist[c(up_maa_geodist_2500_3000[,1]),c(up_maa_geodist_2500_3000[,2])]


# get the average and median jaccard weight values for the 2500-2500 km bin
up_maa_geodist_2500_3000_avg <- mean(as.matrix(up_maa_geodist_2500_3000_jaccard), na.rm =
TRUE)
up_maa_geodist_2500_3000_med <- median(as.matrix(up_maa_geodist_2500_3000_jaccard), na.rm =
TRUE)

## 3000-3500 km ##

# get index positions of all values in geographic distance matrix within 3000-3500 km
up_maa_geodist_3000_3500 = which(up_maa_geodist_matrix >3000000, arr.ind = TRUE)
# extract corresponding values from the jaccard distance matrix
up_maa_geodist_3000_3500_jaccard <-
up_maa_dist[c(up_maa_geodist_3000_3500[,1]),c(up_maa_geodist_3000_3500[,2])]

# get the average and median jaccard weight values for the 3000-3500 km bin
up_maa_geodist_3000_3500_avg <- mean(as.matrix(up_maa_geodist_3000_3500_jaccard), na.rm =
TRUE)
up_maa_geodist_3000_3500_med <- median(as.matrix(up_maa_geodist_3000_3500_jaccard), na.rm =
TRUE)

up_maa_freq_geodist_bins <- c(nrow(up_maa_geodist_0_50),
                              nrow(up_maa_geodist_50_100),
                              nrow(up_maa_geodist_100_200),
                              nrow(up_maa_geodist_200_400),
                              nrow(up_maa_geodist_400_600),
                              nrow(up_maa_geodist_600_1000),
                              nrow(up_maa_geodist_1000_1500),
                              nrow(up_maa_geodist_1500_2000),
                              nrow(up_maa_geodist_2000_2500),
                              nrow(up_maa_geodist_2500_3000),
                              nrow(up_maa_geodist_3000_3500))

up_maa_geodist_avg_bins <- c(up_maa_geodist_0_50_avg,
                             up_maa_geodist_50_100_avg,
                             up_maa_geodist_100_200_avg,
                             up_maa_geodist_200_400_avg,
```

```
                    up_maa_geodist_400_600_avg,
                    up_maa_geodist_600_1000_avg,
                    up_maa_geodist_1000_1500_avg,
                    up_maa_geodist_1500_2000_avg,
                    up_maa_geodist_2000_2500_avg,
                    up_maa_geodist_2500_3000_avg,
                    up_maa_geodist_3000_3500_avg)

up_maa_geodist_med_bins <- c(up_maa_geodist_0_50_med,
                    up_maa_geodist_50_100_med,
                    up_maa_geodist_100_200_med,
                    up_maa_geodist_200_400_med,
                    up_maa_geodist_400_600_med,
                    up_maa_geodist_600_1000_med,
                    up_maa_geodist_1000_1500_med,
                    up_maa_geodist_1500_2000_med,
                    up_maa_geodist_2000_2500_med,
                    up_maa_geodist_2500_3000_med,
                    up_maa_geodist_3000_3500_med)

up_maa_geodist_bins_table <-
rbind(up_maa_geodist_avg_bins,up_maa_geodist_med_bins,up_maa_freq_geodist_bins)

colnames(up_maa_geodist_bins_table) <- c("0-50 km","50-100 km","100-200 km","200-400 km","400-
600 km","600-1000 km",
                                "1000-1500 km","1500-2000 km","2000-2500 km","2500-3000 km","3000-
3500 km")
rownames(up_maa_geodist_bins_table) <- c("Mean Jaccard Weight","Median Jaccard Weight","Numb.
Sites in Bin")

up_maa_geodist_bins_table


write.csv(up_maa_geodist_bins_table, file = 'WeightAnalysis/up_maa_geodist_bins_table.csv')

low_cam_geodist_bins_table
mid_cam_geodist_bins_table
up_cam_geodist_bins_table
low_maa_geodist_bins_table


### Plot the lines #####

setwd("C:/Users/ceara/Documents/Province Project/Proposal tables and figures")

pdf(file = "Geographic_Dist_Bins_average_weights.pdf")

# average values plot

plot(low_cam_geodist_bins_table[1,],col="blue",type="o",pch=16,axes=FALSE,
ann=FALSE,ylim=c(0.65,0.99))
```

```
points(mid_cam_geodist_bins_table[1,],col="green",pch=17,type="o")
points(up_cam_geodist_bins_table[1,],col="gold",pch=18,type="o")
points(low_maa_geodist_bins_table[1,],col="orange",pch=19,type="o")
points(up_maa_geodist_bins_table[1,],col="maroon1",pch=15,type="o")
axis(1, at=1:11, lab=c("0-50 ","50-100","100-200","200-400","400-600","600-1000",
                  "1000-1500","1500-2000","2000-2500","2500-3000","3000-3500"))
axis(2, at=c(0.65,0.7,0.75,0.80,0.85,0.90,0.95,0.99))
box()
title(xlab="Distance Bins (km)")
title(ylab="Jaccard Distance Weight")
title(main="Average Weight by Geographic Distance")
legend(x = 8.75, y = 0.75,  legend = c("Low Camp.","Mid Camp.", "Up Camp.", "Low Maastr.","Up
Maastr."), col = c("blue","darkgreen","gold","orange","maroon1"),
      pch = c(16,17,18,19,15))

# median values plot

plot(low_cam_geodist_bins_table[2,],col="blue",type="o",pch=16,axes=FALSE,
ann=FALSE,ylim=c(0.65,0.99))
points(mid_cam_geodist_bins_table[2,],col="green",pch=17,type="o")
points(up_cam_geodist_bins_table[2,],col="gold",pch=18,type="o")
points(low_maa_geodist_bins_table[2,],col="orange",pch=19,type="o")
points(up_maa_geodist_bins_table[2,],col="maroon1",pch=15,type="o")
axis(1, at=1:11, lab=c("0-50 ","50-100","100-200","200-400","400-600","600-1000",
                  "1000-1500","1500-2000","2000-2500","2500-3000","3000-3500"))
axis(2, at=c(0.65,0.7,0.75,0.80,0.85,0.90,0.95,0.99))
box()
title(xlab="Distance Bins (km)")
title(ylab="Jaccard Distance Weight")
title(main="Median Weight by Geographic Distance")
legend(x = 8.75, y = 0.75,  legend = c("Low Camp.","Mid Camp.", "Up Camp.", "Low Maastr.","Up
Maastr."), col = c("blue","darkgreen","gold","orange","maroon1"),
      pch = c(16,17,18,19,15))

dev.off()

pdf(file = "PaleoLocation_Maps_Substages.pdf")

plot(cam_low_maps, col = 'grey', border = NA, xlim = c(cam_low_minlong,maa_up_maxlong), ylim =
c(cam_low_minlat,cam_low_maxlat))
lines(cam_low_coast, col = "black", xlim = c(cam_low_minlong,cam_low_maxlong), ylim =
c(cam_low_maxlat,cam_low_minlat))
points(cam_low_ploc, pch=21, bg = "blue", col = "black")

plot(cam_mid_maps, col = 'grey', border = NA, xlim = c(cam_mid_minlong,cam_mid_maxlong), ylim =
c(cam_mid_minlat,cam_mid_maxlat))
lines(cam_mid_coast, col = "black", xlim =  c(cam_low_minlong,cam_low_maxlong), ylim =
c(cam_low_maxlat,cam_low_minlat))
points(cam_mid_ploc, pch=21, bg = "darkgreen", col="black")
```

```
plot(cam_up_maps, col = 'grey', border = NA, xlim = c(cam_up_minlong,cam_up_maxlong), ylim =
c(cam_up_minlat,cam_up_maxlat))
lines(cam_up_coast, col = "black", xlim =  c(cam_low_minlong,cam_low_maxlong), ylim =
c(cam_low_maxlat,cam_low_minlat))
points(cam_up_ploc, pch=21, bg = "gold", col="black")

plot(maa_low_maps, col = 'grey', border = NA, xlim = c(maa_low_minlong,maa_up_maxlong), ylim =
c(maa_low_minlat,maa_low_maxlat))
lines(maa_low_coast, col = "black", xlim =  c(cam_low_minlong,cam_low_maxlong), ylim =
c(cam_low_maxlat,cam_low_minlat))
points(maa_low_ploc, pch=21, bg = "orange", col="black")

plot(maa_up_maps, col = 'grey', border = NA, xlim = c(maa_up_minlong,maa_up_maxlong), ylim =
c(maa_up_minlat,maa_up_maxlat))
lines(maa_up_coast, col = "black", xlim =  c(cam_low_minlong,cam_low_maxlong), ylim =
c(cam_low_maxlat,cam_low_minlat))
points(maa_up_ploc, pch=21, bg = "maroon1", col="black")

dev.off()
```

# APPENDIX B. Supplementary Materials for "A network analysis approach to understanding trends in functional diversity in the Late Cretaceous of North America" by Purcell and Myers (*in press*)

Appendix B-1. Detailed Methods and Results

## Relationship between FR and GR in Database:

A basic correlation plot of generic vs functional richness was created based on the values present in each grid cell for the raw database in each substage. The plot showed a strong positive trend with fewer points at the higher levels of GR (>100 genera) that potentially begin to plateau (Figure S1). A squared linear regression computed using the lm function in R was found to have the best fit ($R^2 = 0.958$) of the simple models tested, and visually fits the data (Figure S1). Evaluation plots of this regression (Figure S2) suggest that the data itself is non-normal based on the Q-Q plot (further confirmed with a Shapiro-Wilk test in R: p-value < 0.05), has a linear relationship based on the Residual vs. Fitted plot, contains only one significant outlier based on the Residual vs. Leverage plot, and has fairly homogeneous variance except at very low fitted values (i.e., < ~40) based on the Scale-Location plot. Overall, these results suggest that the regression is a good fit for the data, despite the fact that the data is non-normal.

The strong correlation between the two richness variables indicates that the two are closely linked and low values of generic richness will very likely correspond with low functional richness in further analysis. However, at very high generic richness levels (i.e., >100), it is likely that the two will be less strongly correlated. **This may impact interpretation at different scales, given that the number of unique genera at coarser spatial aggregations in the dataset, such as within latitudinal bins or within the entire study region, will be higher than at more localize scales, such as at the level of the node.** It should also be noted that all 60km nodes with >100 genera are representative of the Maastrichtian. This indicates that richness is better represented in the Maastrichtian than in the Campanian, though not all Maastrichtian nodes are as well sampled.

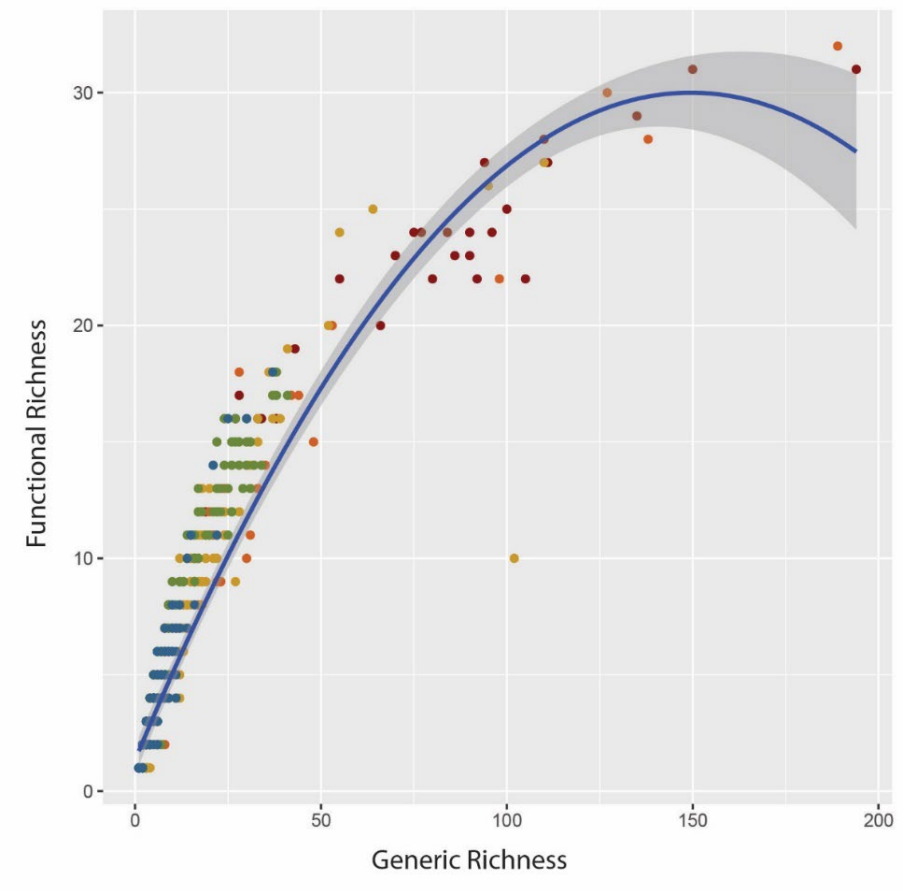



Figure S1. Plot of generic versus functional richness values for each node in each substage. Point color indicates substage interval (blue = Lower Camp., green = Middle Camp., yellow = Upper Camp., orange = Lower Maastr., and red = Upper Maastr.). The blue line represent a squared linear regression line fitted to the data (R2 = 0.958) using the stat_smooth function from the ggplot2 package, with grey shading indicating the 95% confidence interval.

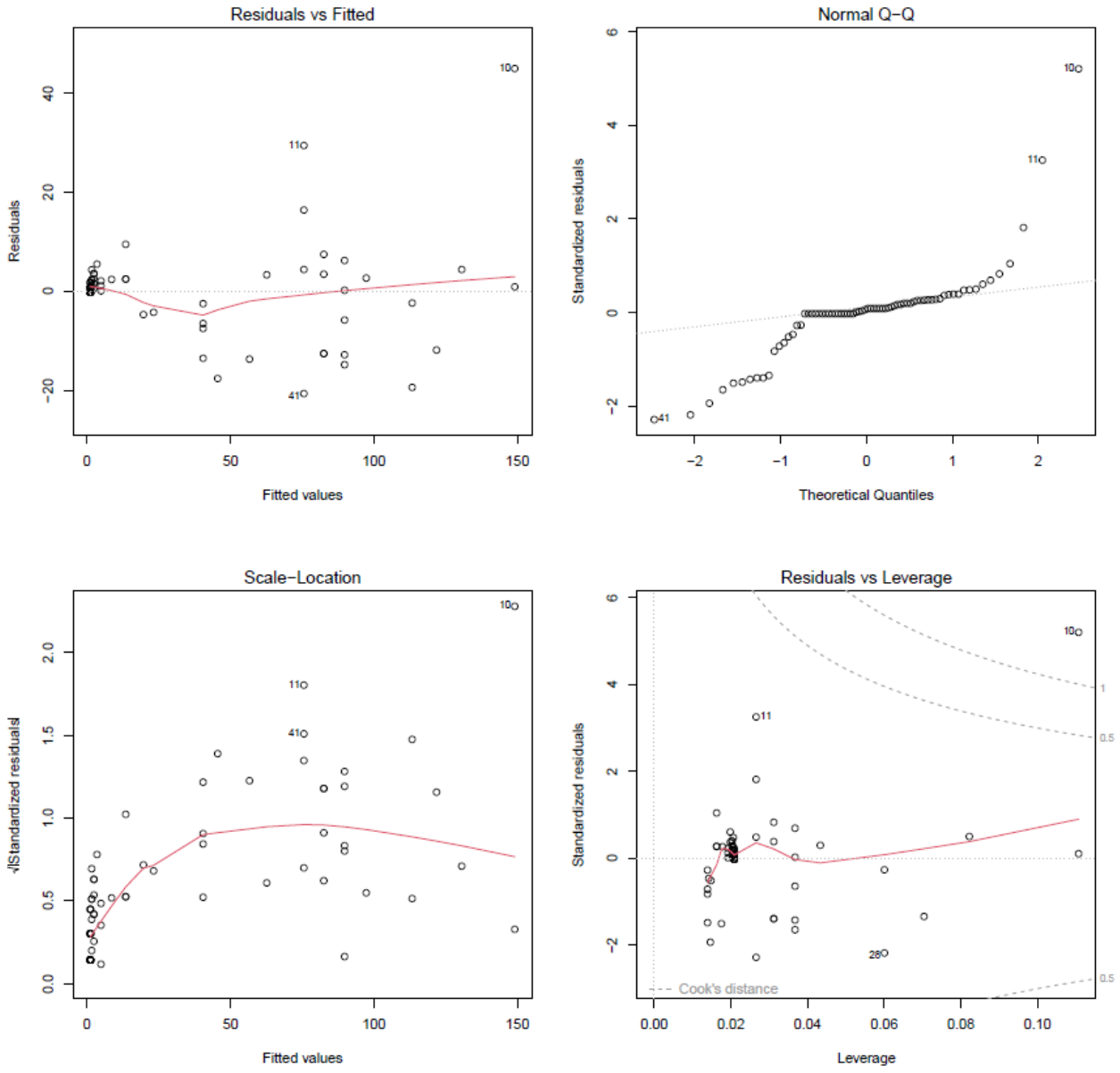



Figure S2. Evaluation plots of the squared linear regression fitted to the data in R using the lm function ($y = 0.095x^2 + 2.64$; $R^2 = 0.949$), including a) fitted values plotted against residual values, b) a normal Q-Q plot, c) fitted values plotted against the square root of the standardized residuals, and d) a plot of Cook's distance in the data, wherein leverage is plotted against the standardized residuals.

Bootstrap results for variation in FR values:

Bootstrap analysis of nodes and latitude bins found strong positive correlation between increasing FR variability and generic richness (Figure S3). Only when a node has a FR value close to the maximum for the dataset (20) does the standard error appear to potentially level off (Figure S3). This analysis tests how variable FR is when the unique genera in a node are subsampled. Given the assumption that genera are perfectly evenly distributed in FEs, FR values would not change, producing no correlation between standard error and FR. However, a positive correlation indicates that genera are not evenly distributed. A squared regression model fits this data distribution well ($R^2 = 0.96$), indicating a depreciation of the correlation at approximately more than 20 FEs.

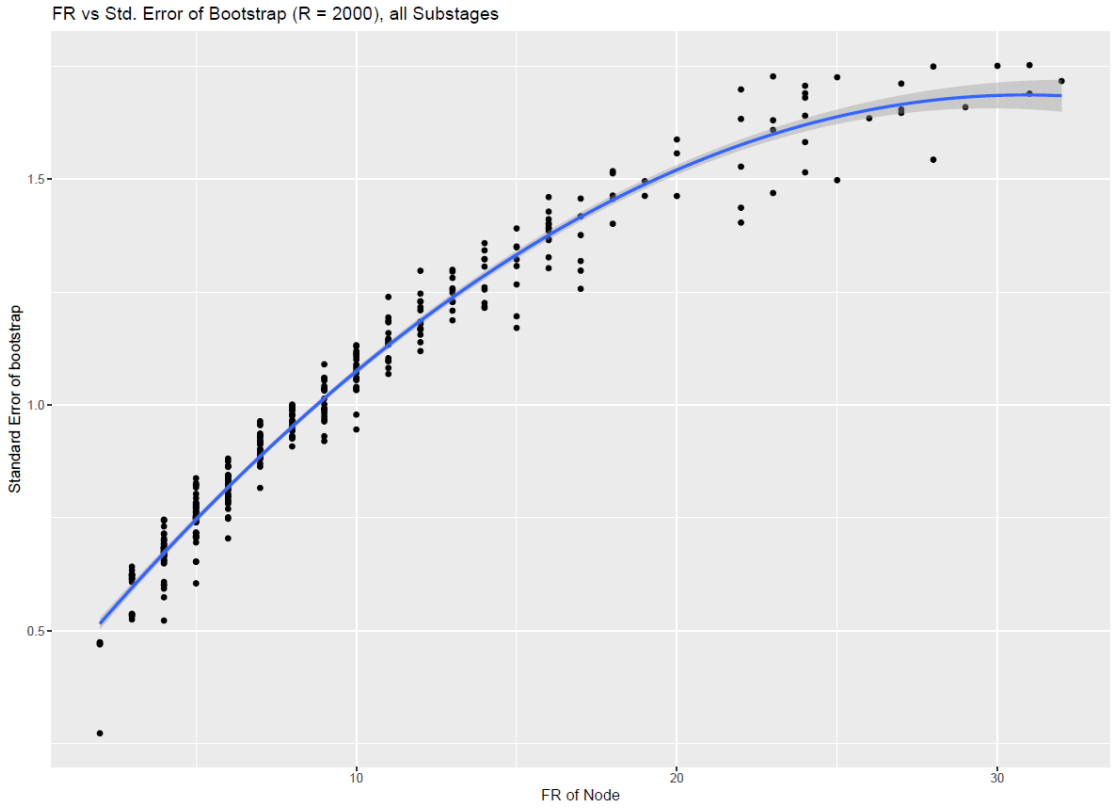



Figure S3. Bootstrap of FR results for nodes in all substagaes. Blue line represents a squared linear regression fitted to the data.

## SQS to determine lower sampling limits

To determine a lower sample size limit that can be assumed to capture the majority of FE in each substage, SQS subsampling (Alroy 2010) was performed on each substage using FEs as taxa and calculating their frequencies in each interval based on the number of unique genera they represent (i.e., no duplicate genera in a node; Table S1). This subsampling procedure indicates that for the Campanian substages, generic richness values greater than 16 and for the Maastrichtian substages generic richness values greater than 18 are likely to capture the majority of common FEs in each substage. This shift in the generic richness level required to adequately subsample the different stages is most likely due to a shift in the dominance of specific FEs and decreasing functional evenness, as is observed at both the paleo-latitudinal bin and substage level. When using only specific Classes of taxa, the limits are much lower, particularly for Cephalopods and Gastropods. Bivalves, which have a high functional diversity (Table 3) have SQS sample size estimates of approximately ten in the Campanian and 11 and 13 in the Lower and Upper Maastrichtian, respectively. Gastropods, also a functionally diverse group, are dominated by a few important functional entities, meaning that their SQS sample size estimate is approximately two in the Campanian and three in the Maastrichtian. The Cephalopods represent only four FEs, and their SQS sample size estimate is three for the Campanian and two for the Maastrichtian.

Table S1. Lower sample size estimates provided by SQS subsampling for each substage:

| Substage | Entire Database |
|---|---|
| Upper Maastr. | 18 |
| Lower Maastr. | 17.9 |
| Upper Camp. | 15.7 |
| Middle Camp. | 15.7 |
| Lower Camp. | 15.3 |

## Note on Shannon Diversity Index:

Initially, the Shannon Diversity Index (SEI) was calculated alongside Simpson's Measure of Evenness (SME), but given that the SEI values followed identical patterns to SME but with values which were not as easy to read and compare, and given that SEI is more susceptible to bias and error (Magurran 2003), it was not included in the final manuscript. However, SEI results are included here in the Supplementary Material for transparency. SEI is calculated based on the following formula:

$$\text{Shannon Equitability Index (H1)} = {-\sum(p_i * \ln(p_i))}\Big/{\ln(S)} \qquad (EQ\ 1)$$

SEI ranges from 0 to 1 with 1 indicating perfect evenness.

In all cases, SME estimates more significant proportional changes in evenness through time, as can be seen here. SME is an evenness metric which is weighted by the abundance or dominance of specific species (or in this case, FE) while SEI is a standardized form of the Shannon index, which quantifies aspects of diversity based on richness (Magurran, 2004). The SEI is therefore more prone to error when sample sizes are low, and not all elements (i.e., FE) are represented in the sample. For this reason, more emphasis is placed on SME in this analysis, though SEI is included for comparison. Though the two metrics of evenness produce such strikingly different proportional changes, they each show shifts in the same direction in all cases.

## Functional Diversity Patterns:

Not all FE are present in all substages but the relative abundance of each is generally very stable through time (Figure 5, Table S4). Of the FE which have continuous presence across the study interval, only two show an increase or decrease in the relative abundance of genera greater than 0.05: facultatively mobile, infaunal suspension feeders and mobile, unattached, epifaunal carnivores, and these only once in each. These FEs are two of the three FEs with highest relative abundance in each substage. The third FE with consistently high relative abundance throughout the study interval is immobile, attached, epifaunal suspension feeders. These three FEs account for at least 8% of all genera in each substage, though which of the three contains the largest relative proportion of genera in each substage does change across the interval. In the Lower Campanian, immobile, attached, epifaunal suspension feeders have the highest relative abundance of genera (16%) and in the Middle Campanian, facultatively mobile, unattached, infaunal suspension feeders have the highest relative abundance of genera (17%). In the Upper Campanian and both Maastrichtian substages, mobile, unattached, epifaunal carnivores have the highest relative abundance (30%, 26%, and 29%, respectively).

Most FEs in each substage contain less than 1% of all genera in each substage, and some FEs are lost or gained across substages (Figure 5, Table S4). Between the Lower and Middle Campanian the total of 30 FE is maintained though there is some shift in the specific FE present. Facultatively mobile, unattached, epifaunal herbivores, which only make up 1% of the relative abundance of genera, are lost and immobile, attached, infaunal suspension feeders, which make up only 3% of the Middle Campanian relative abundance are gained. Between the Middle and Late Campanian, however, no FEs are lost and facultatively mobile, unattached, epifaunal herbivores, making up 2% of the relative abundance of genera, are regained and mobile, unattached, epifaunal, detritivores (1% relative abundance of genera) and immobile, attached, epifaunal, photosymbiotic taxa (1% relative abundance of genera) are gained. Between the Upper Campanian and Lower Maastrichtian, immobile, attached, epifaunal, photosymbiotic taxa are lost but immobile, unattached, boring, suspension feeders (2% relative abundance of genera), facultatively mobile, attached, infaunal, detritivores (1% relative abundance of genera), and immobile,

attached, semi-infaunal, photosymbiots (1% relative abundance of genera) are gained. Finally, all FE are present in the Upper Maastrichtian, meaning that facultatively mobile, unattached, boring suspension feeders (2% relative abundance of genera), and immobile, unattached, semi-infaunal, photosymbiotic taxa (1% relative abundance of genera) are gained and immobile, attached, epifaunal photosymbionts (2% relative abundance of genera) are regained.

In total, eight FE which were not originally present in the Lower Campanian are gained at some point across the study interval, but each has very low relative abundance which do not greatly change across substages (1-3%). Only two FE are "lost" at some point across the study interval, but all are recovered by the Upper Maastrichtian. Again, all FE which appear to be lost have very low relative abundances. These patterns indicate that these FE are probably rare and their loss/gain may be a product of fossilization bias or inadequate sampling. Based on previous analyses which have looked at global patterns of FE (Foster and Twitchet, 2014; Edie et al., 2018), it seems highly unlikely that these FE had gone globally extinct across the study intervals, though they may have been regionally extirpated for a time.

Table S2. Table of the functional entities present in each substage.

| Lower Camp. | Middle Camp. | Upper Camp. | Lower Maastr. | Upper Maastr. |
|---|---|---|---|---|
| IA-E-SF | IA-E-SF | IA-E-SF | IA-E-SF | IA-E-SF |
| FU-I-C | FU-I-C | FU-I-C | FU-I-C | FU-I-C |
| FU-I-CH | FU-I-CH | FU-I-CH | FU-I-CH | FU-I-CH |
| MU-PI-D | MU-PI-D | MU-PI-D | MU-PI-D | MU-PI-D |
| IU-B-SF | IU-B-SF | IU-B-SF | IU-B-SF | IU-B-SF |
| IU-I-SF | IU-I-SF | IU-I-SF | IU-I-SF | IU-I-SF |
| FA-E-SF | FA-E-SF | FA-E-SF | FA-E-SF | FA-E-SF |
| FU-E-SF | FU-E-SF | FU-E-SF | FU-E-SF | FU-E-SF |
| MU-N-SF | MU-N-SF | MU-N-SF | MU-N-SF | MU-N-SF |
| MU-PI-C | MU-PI-C | MU-PI-C | MU-PI-C | MU-PI-C |
| MU-E-H | MU-E-H | MU-E-H | MU-E-H | MU-E-H |
| FU-I-SF | FU-I-SF | FU-I-SF | FU-I-SF | FU-I-SF |
| MU-PI-H | MU-PI-H | MU-PI-H | MU-PI-H | MU-PI-H |
| FA-I-SF | FA-I-SF | FA-I-SF | FA-I-SF | FA-I-SF |
| MU-I-C | MU-I-C | MU-I-C | MU-I-C | MU-I-C |
| MU-I-CH | MU-I-CH | MU-I-CH | MU-I-CH | MU-I-CH |
| FU-E-H | FU-PI-H | FU-E-H | FU-E-H | FU-E-H |
| FU-PI-H | MU-E-O | FU-PI-H | FU-PI-H | FU-PI-H |
| MU-E-O | MU-E-SF | MU-E-O | MU-E-O | MU-E-O |
| MU-E-SF | FA-PI-SF | MU-E-SF | MU-E-SF | MU-E-SF |
| FA-PI-SF | IA-PI-SF | FA-PI-SF | FA-PI-SF | FA-PI-SF |
| IA-PI-SF | MU-E-C | IA-PI-SF | IA-PI-SF | IA-PI-SF |
| MU-E-C | MU-I-SF | MU-E-C | MU-E-C | MU-E-C |
| MU-I-SF | MU-NB-SF | MU-I-SF | MU-I-SF | MU-I-SF |
| MU-NB-SF | IU-E-SF | MU-NB-SF | MU-NB-SF | MU-NB-SF |
| IU-E-SF | MU-N-C | IU-E-SF | IU-E-SF | IU-E-SF |
| MU-N-C | FU-I-D | MU-N-C | MU-N-C | MU-N-C |
| FU-I-D | MU-NB-C | FU-I-D | FU-I-D | FU-I-D |
| MU-NB-C | MU-I-D | MU-NB-C | MU-NB-C | MU-NB-C |
| MU-I-D | IA-I-SF | MU-I-D | MU-I-D | MU-I-D |
| NA | NA | IA-I-SF | IA-I-SF | IA-I-SF |
| NA | NA | MU-E-D | MU-E-D | MU-E-D |
| NA | NA | IA-E-P | IA-B-SF | IA-E-P |
| NA | NA | NA | FA-I-D | IA-B-SF |
| NA | NA | NA | IA-PI-P | FA-I-D |
| NA | NA | NA | NA | IA-PI-P |
| NA | NA | NA | NA | FU-B-SF |
| NA | NA | NA | NA | IU-PI-P |

Summary of Regional Functional vs Taxonomic Diversity change in Substages:

Table S3. Table of proportional change in metrics across substages in the raw database.

| | Proportional Change Across Substages: | | | |
|---|---|---|---|---|
| | **GR** | **FR** | **SME** | **SEI** |
| **Low Camp. - Mid Camp.** | 0.27 | 0.00 | 0.04 | 0.01 |
| **Mid Camp. - Up Camp.** | 0.81 | 0.10 | -0.50 | -0.11 |
| **Up Camp. - Low Maastr.** | -0.09 | 0.06 | 0.11 | 0.02 |
| **Low Maastr. - Up Maastr.** | 0.08 | 0.09 | -0.19 | -0.03 |
| **Average:** | **0.27** | **0.06** | **-0.13** | **-0.02** |
| **Median:** | **0.17** | **0.07** | **-0.07** | **-0.01** |
| **S.D.:** | **0.39** | **0.04** | **0.27** | **0.06** |

Table S4. Relative abundance of each FE in each substage. Highest relative abundance values in each substage are in bold. The three most abundant FEs in all substages are denoted with an asterisk.

| Functional Entities | Lower Camp. | Middle Camp. | Upper Camp. | Lower Maastr. | Upper Maastr. |
|---|---|---|---|---|---|
| *IA-E-SF | 0.161 | 0.115 | 0.097 | 0.120 | 0.111 |
| FU-I-C | 0.021 | 0.016 | 0.009 | 0.010 | 0.009 |
| FU-I-CH | 0.021 | 0.033 | 0.012 | 0.017 | 0.015 |
| MU-PI-D | 0.021 | 0.027 | 0.006 | 0.007 | 0.012 |
| IU-B-SF | 0.021 | 0.005 | 0.018 | 0.023 | 0.022 |
| IU-I-SF | 0.021 | 0.016 | 0.006 | 0.007 | 0.006 |
| FA-E-SF | 0.021 | 0.044 | 0.033 | 0.040 | 0.040 |
| FU-E-SF | 0.021 | 0.016 | 0.021 | 0.030 | 0.028 |
| MU-N-SF | 0.021 | 0.066 | 0.033 | 0.010 | 0.012 |
| MU-PI-C | 0.014 | 0.016 | 0.018 | 0.020 | 0.015 |
| MU-E-H | 0.014 | 0.011 | 0.043 | 0.053 | 0.053 |
| *FU-I-SF | 0.133 | 0.170 | 0.100 | 0.113 | 0.090 |
| MU-PI-H | 0.014 | 0.022 | 0.009 | 0.010 | 0.006 |
| FA-I-SF | 0.014 | 0.011 | 0.018 | 0.020 | 0.015 |
| MU-I-C | 0.007 | 0.005 | 0.003 | 0.010 | 0.009 |
| MU-I-CH | 0.007 | 0.005 | 0.003 | 0.003 | 0.003 |
| FU-E-H | 0.007 | NA | 0.006 | 0.010 | 0.009 |
| FU-PI-H | 0.007 | 0.005 | 0.015 | 0.017 | 0.015 |
| MU-E-O | 0.007 | 0.016 | 0.012 | 0.010 | 0.031 |
| MU-E-SF | 0.007 | 0.011 | 0.012 | 0.017 | 0.015 |
| FA-PI-SF | 0.007 | 0.005 | 0.003 | 0.007 | 0.006 |
| IA-PI-SF | 0.007 | 0.005 | 0.003 | 0.007 | 0.006 |
| *MU-E-C | 0.091 | 0.077 | 0.301 | 0.257 | 0.294 |
| MU-I-SF | 0.007 | 0.016 | 0.009 | 0.013 | 0.009 |
| MU-NB-SF | 0.084 | 0.066 | 0.049 | 0.037 | 0.037 |
| IU-E-SF | 0.063 | 0.038 | 0.024 | 0.013 | 0.012 |
| MU-N-C | 0.049 | 0.044 | 0.024 | 0.013 | 0.006 |
| FU-I-D | 0.049 | 0.038 | 0.024 | 0.033 | 0.028 |
| MU-NB-C | 0.042 | 0.049 | 0.043 | 0.013 | 0.015 |
| MU-I-D | 0.042 | 0.033 | 0.030 | 0.030 | 0.028 |
| IA-I-SF | NA | 0.011 | 0.006 | 0.013 | 0.003 |
| MU-E-D | NA | NA | 0.003 | 0.003 | 0.003 |
| IA-E-P | NA | NA | 0.003 | NA | 0.009 |
| IA-B-SF | NA | NA | NA | 0.007 | 0.006 |
| FA-I-D | NA | NA | NA | 0.003 | 0.003 |
| IA-PI-P | NA | NA | NA | 0.003 | 0.003 |
| FU-B-SF | NA | NA | NA | NA | 0.006 |
| IU-PI-P | NA | NA | NA | NA | 0.003 |

Basin-level Substage Trends:

When data representative of the WIS only are subset and analyzed, only five FEs are lost: facultatively mobile, attached, infaunal deposit feeders (bivalve); facultatively mobile, unattached, boring suspension feeders (bivalve); immobile, attached, boring suspension feeders (bivalve, sponge); immobile, attached, epifaunal photosymbionts (bivalve, coral); immobile, attached, semi-infaunal photosymbionts (bivalve); and immobile unattached, semi-infaunal photosybionts (bivalve). These results show that only one FE with a boring lifestyle (immobile, unattached, boring suspension feeders) are present, represented by bivalves, and no photosymbionts are present once GCP data are removed. Indeed, these same five FEs not present in the WIS alone among the eight FEs that are "gained" through time in the database (Figure 5). **While these FEs are consistently rare in the database, accounting for less than 0.9% of all genera in a given substage, their absence from the WIS suggests that these FEs may indeed not have existed in the WIS during the Campanian and Maastrichtian.**

When considering the WIS alone, generic richness rises through the Campanian to peak in the Upper Campanian, before declining through the Maastrichtian (Table S5; Figure S2). FR parallels this general trend, but proportionally changes much less across the Campanian (only 4% and 7% increase from Lower to Upper Campanian), before falling dramatically through the Maastrichtian (13% and 30% decrease; Table S6). SME decreases first by only 4% from the Lower to Middle Campanian, then by 20% from the Middle to Upper Campanian before rising by first 23% then 54% across the Upper Campanian to Lower Maastrichtian and the Lower to Upper Maastrichtian, respectively. **This suggests that functional evenness fell in the later part of the Campanian in the WIS as generic richness increased, even when FR remained fairly stable, but when generic richness fell in the later substages, evenness increased.** This pattern is similar to the one observed when both the WIS and GCP data are considered together (Table S3, Figure 4), indicating that as generic richness decreases, evenness increases as taxa more evenly distributed within remaining FEs. Declining FR may also contribute to this phenomenon. Indeed, the loss of FEs with low relative abundance of genera most likely contributes to rising evenness, as only those FEs with high redundancy are left.

The GCP region alone experiences a consistent increase in both generic and functional richness throughout the study interval (Table S7; Figure S3). Interestingly, the inverse pattern of rising generic richness coupled with a clear decline in SME across the Middle to Upper Campanian is easily observed in the GCP alone (Table S8). **Again, this indicates that, as generic richness increases, and despite a parallel increase in FR, uneven packing within FEs increases.** Similarly, to when FEs with low relative abundance of genera are lost and evenness increases, most FE that are gained through time in the GCP have very low relative abundances, increasing the uneven distribution of taxa within FEs (Table S10, Figure S5). It should also be noted, however, that shifts in relative abundance of genera in the GCP FEs is more dramatic that those observed in the WIS or overall database (Figure S4 and 5).

Table S5. Table of the generic richness, FR, SME, and SEI values in each substage in the WIS.

| WIS Data | Generic Richness | FR | SME | SEI |
|---|---|---|---|---|
| **Lower Camp.** | 108 | 28 | 0.53 | 0.87 |
| **Mid Camp.** | 151 | 29 | 0.51 | 0.88 |
| **Upper Camp.** | 199 | 31 | 0.41 | 0.84 |
| **Lower Maastr.** | 119 | 27 | 0.51 | 0.87 |
| **Upper Maastr.** | 42 | 19 | 0.78 | 0.91 |
| **Mean** | **123.8** | **26.8** | **0.55** | **0.87** |
| **Median** | **119** | **28** | **0.51** | **0.87** |
| **S.D.** | **57.79** | **4.60** | **0.14** | **0.02** |

Table S6. Table of proportional change in metrics across substages in the WIS.

| WIS Data | Proportional Change Across Substages: | | | |
|---|---|---|---|---|
| | GR | FR | SME | SEI |
| **Low Camp. - Mid Camp.** | 0.40 | 0.04 | -0.04 | 0.01 |
| **Mid Camp. - Up Camp.** | 0.32 | 0.07 | -0.20 | -0.04 |
| **Up Camp. - Low Maastr.** | -0.40 | -0.13 | 0.23 | 0.04 |
| **Low Maastr. - Up Maastr.** | -0.65 | -0.30 | 0.54 | 0.04 |
| **Average:** | **-0.08** | **-0.08** | **0.13** | **0.01** |
| **Median:** | **-0.04** | **-0.05** | **0.10** | **0.02** |
| **S.D.:** | **0.52** | **0.17** | **0.32** | **0.04** |

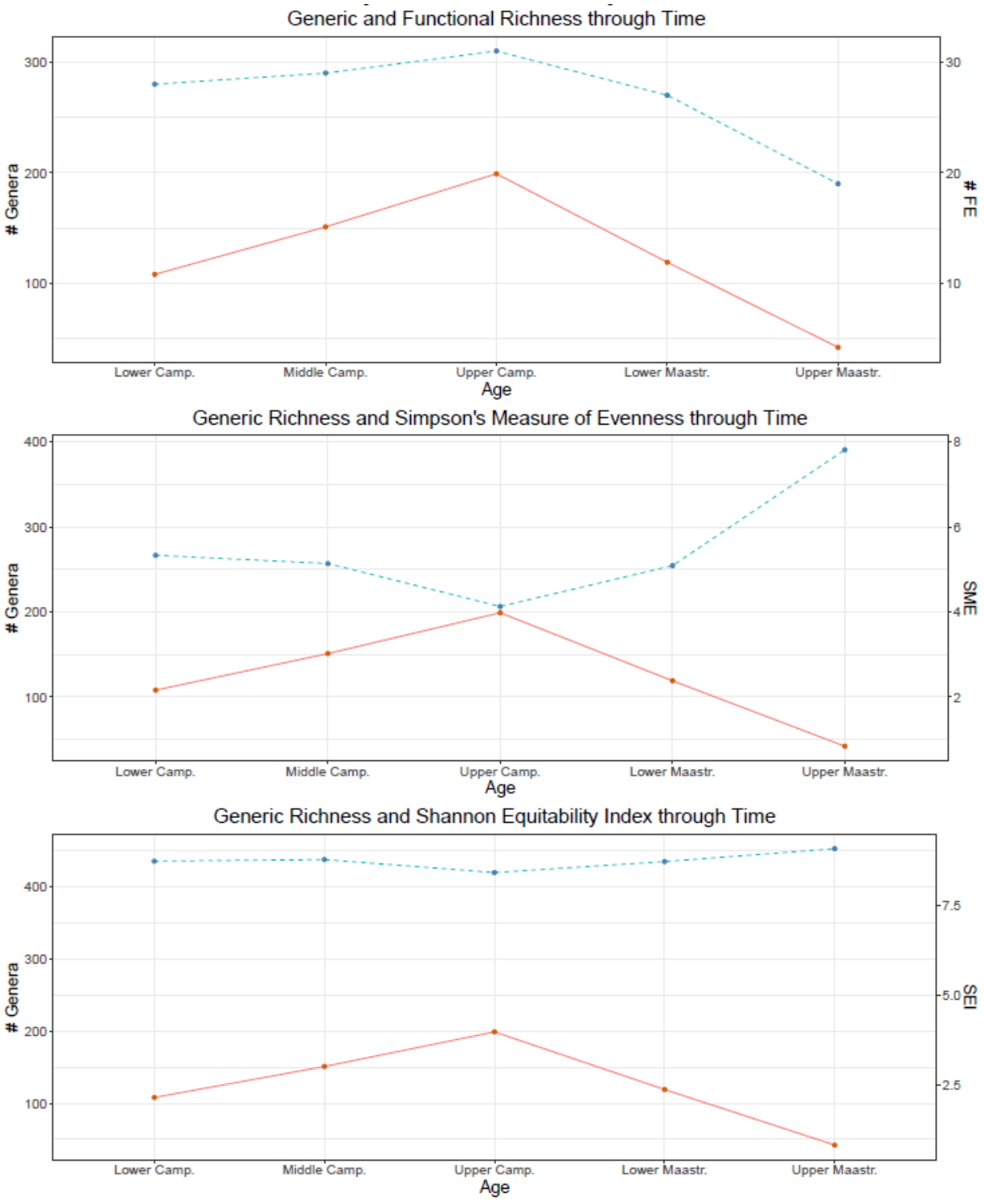



Figure S4. Line plots of changes in functional ecology metrics across substages for WIS. The left y-axis in all plots depicts generic richness values, while the right y-axis presents values for the other metric..

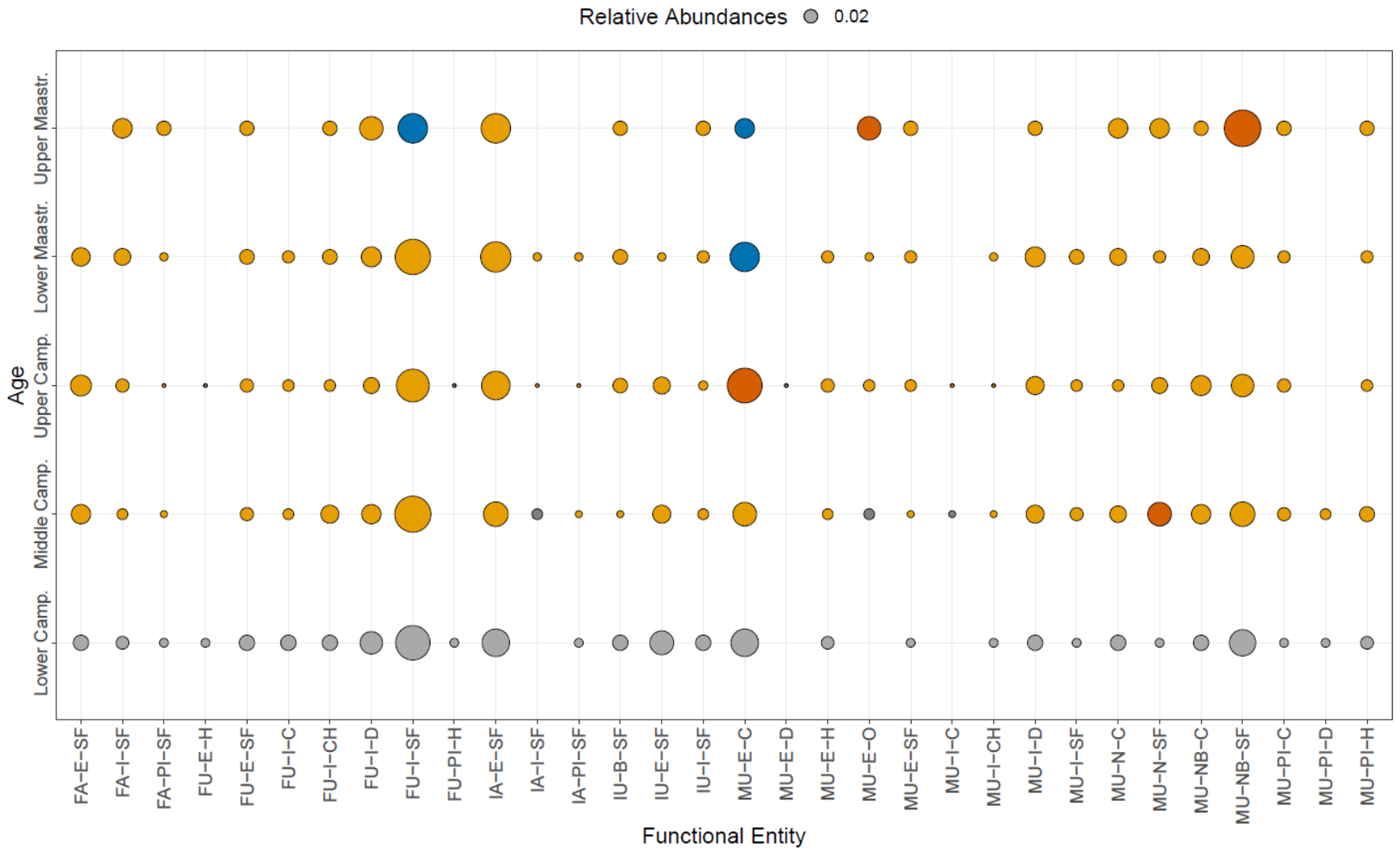



Figure S5. Plot of relative abundance of each FE through time based WIS data alone. Circle size indicates relative abundance of a FE within each substage and color indicates relative abundance increase across subsequent substages greater than 0.05 (orange), decrease greater than 0.05 (blue), and increase/decrease less than 0.05 (yellow). Points are grey if they cannot be compared with a previous substage. See Table 1 for keys to each FE names along the x-axis.

Table S7. Table of the generic richness, FR, SME, and SEI values in each substage in the GCP.

| GCP Data | Generic Richness | FR | SME | SEI |
|---|---|---|---|---|
| **Lower Camp.** | 47 | 14 | 0.61 | 0.87 |
| **Mid Camp.** | 65 | 22 | 0.67 | 0.90 |
| **Upper Camp.** | 213 | 29 | 0.21 | 0.76 |
| **Lower Maastr.** | 248 | 34 | 0.26 | 0.79 |
| **Upper Maastr.** | 309 | 38 | 0.22 | 0.77 |
| **Mean** | **176.4** | **27.4** | **0.39** | **0.82** |
| **Median** | **213** | **29** | **0.26** | **0.79** |
| **S.D.** | **115.33** | **9.58** | **0.23** | **0.06** |

Table S8. Table of proportional change in metrics across substages in the GCP.

| GCP Data | Proportional Change Across Substages: | | | |
|---|---|---|---|---|
| | GR | FR | SME | SEI |
| **Low Camp. - Mid Camp.** | 0.38 | 0.57 | 0.09 | 0.04 |
| **Mid Camp. - Up Camp.** | 2.28 | 0.32 | -0.68 | -0.16 |
| **Up Camp. - Low Maastr.** | 0.16 | 0.17 | 0.23 | 0.04 |
| **Low Maastr. - Up Maastr.** | 0.25 | 0.12 | -0.16 | -0.02 |
| **Average:** | **0.77** | **0.29** | **-0.13** | **-0.03** |
| **Median:** | **0.31** | **0.25** | **-0.04** | **0.01** |
| **S.D.:** | **1.01** | **0.20** | **0.40** | **0.09** |

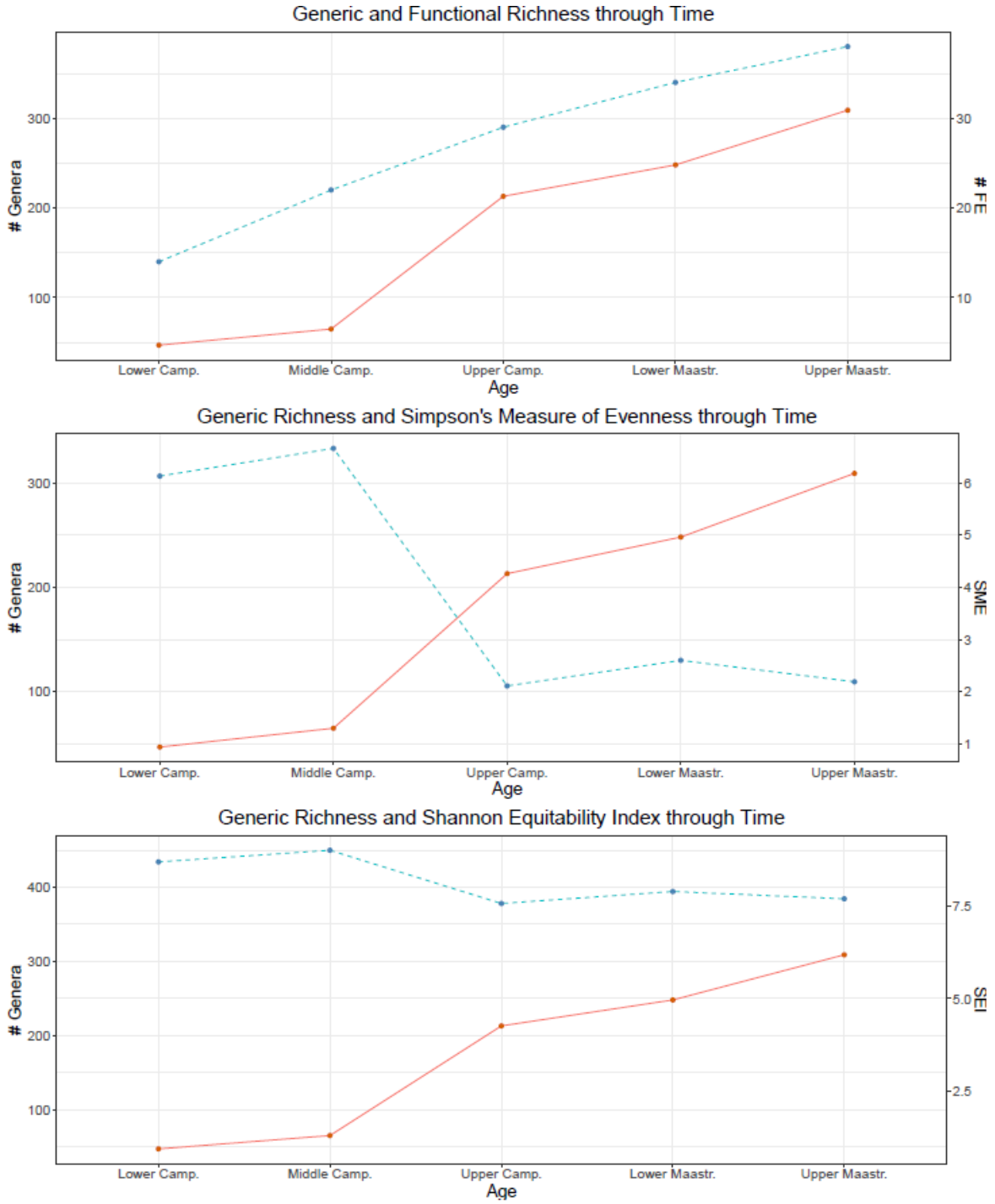



Figure S6. Line plots of changes in functional ecology metrics across substages for GCP. The left y-axis in all plots depicts generic richness values, while the right y-axis presents values for the other metric.

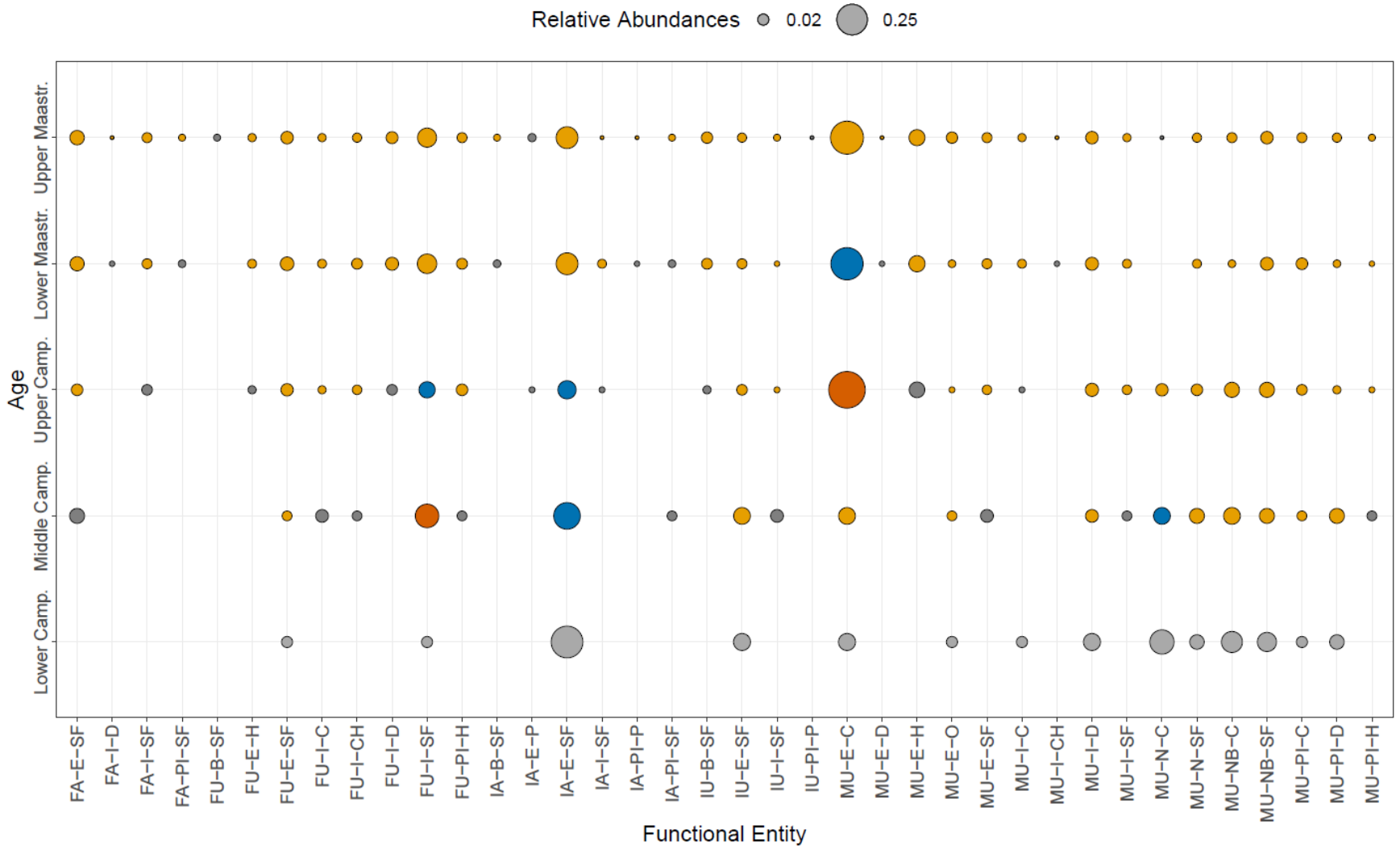



Figure S7. Plot of relative abundance of each FE through time based on GCP data alone. Circle size indicates relative abundance of a FE within each substage and color indicates relative abundance increase across subsequent substages greater than 0.05 (orange), decrease greater than 0.05 (blue), and increase/decrease less than 0.05 (yellow). Points are grey if they cannot be compared with a previous substage. See Table 1 for keys to each FE names along the x-axis.

Table S9. Relative abundance of each FE in each substage for the WIS.

| WIS Data | | | | | |
|---|---|---|---|---|---|
| **Functional Entities** | **Lower Camp.** | **Middle Camp.** | **Upper Camp.** | **Lower Maastr.** | **Upper Maastr.** |
| FA-E-SF | 0.028 | 0.046 | 0.055 | 0.042 | NA |
| FA-I-SF | 0.019 | 0.013 | 0.020 | 0.034 | 0.048 |
| FA-PI-SF | 0.009 | 0.007 | 0.005 | 0.008 | 0.024 |
| FU-E-H | 0.009 | NA | 0.005 | NA | NA |
| FU-E-SF | 0.028 | 0.020 | 0.020 | 0.025 | 0.024 |
| FU-I-C | 0.028 | 0.013 | 0.015 | 0.017 | NA |
| FU-I-CH | 0.028 | 0.040 | 0.015 | 0.025 | 0.024 |
| FU-I-D | 0.065 | 0.046 | 0.030 | 0.050 | 0.071 |
| FU-I-SF | 0.167 | 0.185 | 0.151 | 0.176 | 0.119 |
| FU-PI-H | 0.009 | NA | 0.005 | NA | NA |
| IA-E-SF | 0.102 | 0.079 | 0.111 | 0.126 | 0.119 |
| IA-I-SF | NA | 0.013 | 0.005 | 0.008 | NA |
| IA-PI-SF | 0.009 | 0.007 | 0.005 | 0.008 | NA |
| IU-B-SF | 0.028 | 0.007 | 0.025 | 0.025 | 0.024 |
| IU-E-SF | 0.074 | 0.040 | 0.035 | 0.008 | NA |
| IU-I-SF | 0.028 | 0.013 | 0.010 | 0.017 | 0.024 |
| MU-E-C | 0.102 | 0.073 | 0.171 | 0.118 | 0.048 |
| MU-E-D | NA | NA | 0.005 | NA | NA |
| MU-E-H | 0.019 | 0.013 | 0.020 | 0.017 | NA |
| MU-E-O | NA | 0.013 | 0.015 | 0.008 | 0.071 |
| MU-E-SF | 0.009 | 0.007 | 0.015 | 0.017 | 0.024 |
| MU-I-C | NA | 0.007 | 0.005 | NA | NA |
| MU-I-CH | 0.009 | 0.007 | 0.005 | 0.008 | NA |
| MU-I-D | 0.028 | 0.040 | 0.040 | 0.050 | 0.024 |
| MU-I-SF | 0.009 | 0.020 | 0.015 | 0.025 | NA |
| MU-NB-C | 0.028 | 0.046 | 0.050 | 0.034 | 0.024 |
| MU-NB-SF | 0.093 | 0.079 | 0.065 | 0.067 | 0.190 |
| MU-N-C | 0.028 | 0.033 | 0.015 | 0.034 | 0.048 |
| MU-N-SF | 0.009 | 0.073 | 0.030 | 0.017 | 0.048 |
| MU-PI-C | 0.009 | 0.020 | 0.020 | 0.017 | 0.024 |
| MU-PI-D | 0.009 | 0.013 | NA | NA | NA |
| MU-PI-H | 0.019 | 0.026 | 0.015 | 0.017 | 0.024 |

Table S10. Relative abundance of each FE in each substage for the GCP.

| GCP Data | | | | | |
|---|---|---|---|---|---|
| **Functional Entities** | **Lower Camp.** | **Middle Camp.** | **Upper Camp.** | **Lower Maastr.** | **Upper Maastr.** |
| IA-E-SF | **0.277** | **0.185** | 0.075 | 0.117 | 0.117 |
| MU-PI-C | 0.021 | 0.015 | 0.019 | 0.024 | 0.016 |
| MU-I-C | 0.021 | NA | 0.005 | 0.012 | 0.010 |
| MU-E-O | 0.021 | 0.015 | 0.005 | 0.008 | 0.023 |
| FU-E-SF | 0.021 | 0.015 | 0.028 | 0.036 | 0.029 |
| FU-I-SF | 0.021 | 0.138 | 0.056 | 0.089 | 0.084 |
| MU-N-C | 0.149 | 0.062 | 0.028 | NA | 0.003 |
| MU-NB-C | 0.106 | 0.062 | 0.047 | 0.008 | 0.016 |
| MU-NB-SF | 0.085 | 0.046 | 0.047 | 0.032 | 0.029 |
| MU-E-C | 0.064 | 0.062 | 0.380 | 0.286 | 0.301 |
| MU-I-D | 0.064 | 0.031 | 0.033 | 0.032 | 0.029 |
| IU-E-SF | 0.064 | 0.062 | 0.019 | 0.016 | 0.013 |
| MU-PI-D | 0.043 | 0.046 | 0.009 | 0.008 | 0.013 |
| MU-N-SF | 0.043 | 0.046 | 0.023 | 0.012 | 0.013 |
| FU-I-C | NA | 0.031 | 0.009 | 0.012 | 0.010 |
| IU-I-SF | NA | 0.031 | 0.005 | 0.004 | 0.006 |
| MU-E-SF | NA | 0.031 | 0.014 | 0.016 | 0.016 |
| FU-I-CH | NA | 0.015 | 0.014 | 0.020 | 0.013 |
| FU-PI-H | NA | 0.015 | 0.023 | 0.020 | 0.016 |
| MU-PI-H | NA | 0.015 | 0.005 | 0.004 | 0.006 |
| IA-PI-SF | NA | 0.015 | NA | 0.008 | 0.006 |
| MU-I-SF | NA | 0.015 | 0.014 | 0.012 | 0.010 |
| FA-E-SF | NA | 0.046 | 0.023 | 0.040 | 0.042 |
| FU-I-D | NA | NA | 0.019 | 0.032 | 0.026 |
| FA-I-SF | NA | NA | 0.019 | 0.016 | 0.016 |
| FU-E-H | NA | NA | 0.009 | 0.012 | 0.010 |
| IU-B-SF | NA | NA | 0.009 | 0.020 | 0.023 |
| IA-E-P | NA | NA | 0.005 | NA | 0.010 |
| IA-I-SF | NA | NA | 0.005 | 0.012 | 0.003 |
| MU-E-H | NA | NA | 0.052 | 0.056 | 0.055 |
| IA-B-SF | NA | NA | NA | 0.008 | 0.006 |
| FA-PI-SF | NA | NA | NA | 0.008 | 0.006 |
| MU-I-CH | NA | NA | NA | 0.004 | 0.003 |
| FA-I-D | NA | NA | NA | 0.004 | 0.003 |
| IA-PI-P | NA | NA | NA | 0.004 | 0.003 |
| MU-E-D | NA | NA | NA | 0.004 | 0.003 |
| FU-B-SF | NA | NA | NA | NA | 0.006 |
| IU-PI-P | NA | NA | NA | NA | 0.003 |

Class-level Influences:

Generic richness for cephalopods increases greatly across the Campanian (by 46% and 17%) but declines from the Upper Campanian to Lower Maastrichtian (54%; Table S11; Figure S8). SME is very high within cephalopod FEs, ranging from 0.76 to 1.05, and shifts by no more than 16% across any substage (Table S12). This means that evenness values for cephalopods are consistent, and only change by more than 8% between the Lower and Middle Campanian (16% increase) and Upper Campanian to Lower Maastrichtian (16% decrease; Table S12). These changes are caused by shifts in the distribution of existing genera withing the four FEs (Figure S9). Mobile, unattached, nekton-benthic suspension feeders is consistently the most common FE, but it and all but one other cephalopod FE experiences a rise or decline in relative abundance greater than 5% across these intervals (Table S17). Overall, cephalopods make up a functionally highly stable portion of the database, despite dramatic changes to generic richness.

Bivalves are the most common taxonomic group in the database, including a total of 184 genera and making up a total of 24 of the 38 FEs, meaning that they most likely have the strongest influence on overall patterns in the database. Bivalve generic and functional richness increase across the study interval and SME values are fairly stable (Table S13; Figure S10), indicating that packing withing bivalve FEs is consistent through time (Figure S10). SME only changes more than 7% between the Lower and Middle Campanian, where SME decreases by 23% (Table S14). This decrease in evenness occurs at the same time as a dramatic increase in bivalve generic diversity (33%) and a slight increase in bivalve FR (6%). However, subsequent increases in both generic and functional richness, such as between the Middle and Upper Campanian, are not coupled with a decrease in SME but rather a very slight increase (7%). Across the later substages, SME declines only very slightly (2% and 3%) while generic richness increases slightly (8%) and decreases slightly (5%) and FR increases (5% and 15%). SME is decoupled from both generic and FR, remaining fairly stable in substages after the Middle Campanian, despite changes to the other two metrics.

Facultatively unattached, infaunal suspension feeders are represented by the most genera in each substage, and increase in dominance from the Lower to Middle Campanian while other FEs, including those while high relative abundance in the Lower Campanian, decline in dominance (Figure S11; Table S18). Furthermore, eight FEs are "gained" throughout the study interval among bivalves alone (Figure S11). Gaining FEs in the collective database was assumed to result in declining evenness, but in the case of bivalves, only a single FE is gained between the Lower and Middle Campanian (immobile, attached, infaunal suspension feeders) declines. Overall, these patterns suggest that declining evenness between the Early to Middle Campanian among bivalves is caused by increased packing of genera into the facultatively mobile, unattached infaunal suspension feeder group, the addition of the immobile, attached, infaunal suspension feeders, and loss of genera in other FEs. Throughout the rest of the substages, evenness is highly stable among bivalves despite changes to both generic and functional richness, as the addition of genera is coupled with the addition of FE and no significant packing into any one group.

Gastropods are the second most common class of taxa in the database, making up 184 genera and eleven FEs. They are therefore likely to greatly influence general patterns in the database. Notably, gastropod generic richness increases by 408% across the Middle to Upper Campanian, but it experiences shifts of greater than 14% across all other substages as well (rises 44% from the Lower to Middle Campanian, declines 14% from the Upper Campanian to the Lower Maastrichtian, and rises 17% through the Maastrichtian; Table S15; Figure S12). FR is stable in the Early to Middle Campanian, but rises by 57% between the Middle and Upper Campanian when generic richness greatly increases, before falling between the Upper Campanian and Lower Maastrichtian (18%), and rising in the Maastrichtian (22%; Table S16). Functional evenness in gastropod genera is much lower than that seen in bivalves, except in the Middle Campanian, where gastropod SME is 0.52 and bivalves are only 0.39, and fluctuates strongly. In particular, a sharp decrease in SME from the Middle to Upper Campanian (66%) coincides with the increase to both generic and functional richness.

Mobile, unattached, epifaunal carnivores are consistently the most dominate gastropod FE, taking up at least 50% of all genera in a given substage (Table S19; Figure S13). However, between the Middle and Upper Campanian, this FE experiences a dramatic increase from a relative abundance of 50% to 70% while, at the same time, two of the other most abundant FEs (MU-PI-C and MU-PI-H) experience a decline in relative abundance from 15% and 12% to 3% and 2%, respectively. Furthermore, four FEs are gained between these substages. These factors all contribute to the sharp decrease in evenness between the Middle and Upper Campanian. Throughout the rest of the intervals excepting the Late Maastrichtian, the relative abundance of mobile, unattached, epifaunal carnivores increases or decreases by more than 5% and as FE are lost, gained, and change in relative abundance. Clearly, gastropods are highly variable in generic richness as well as FR and functional evenness relative to bivalves. Their presence, therefore, likely contributes to the greater variability of the database as a whole relative to one including bivalves alone.

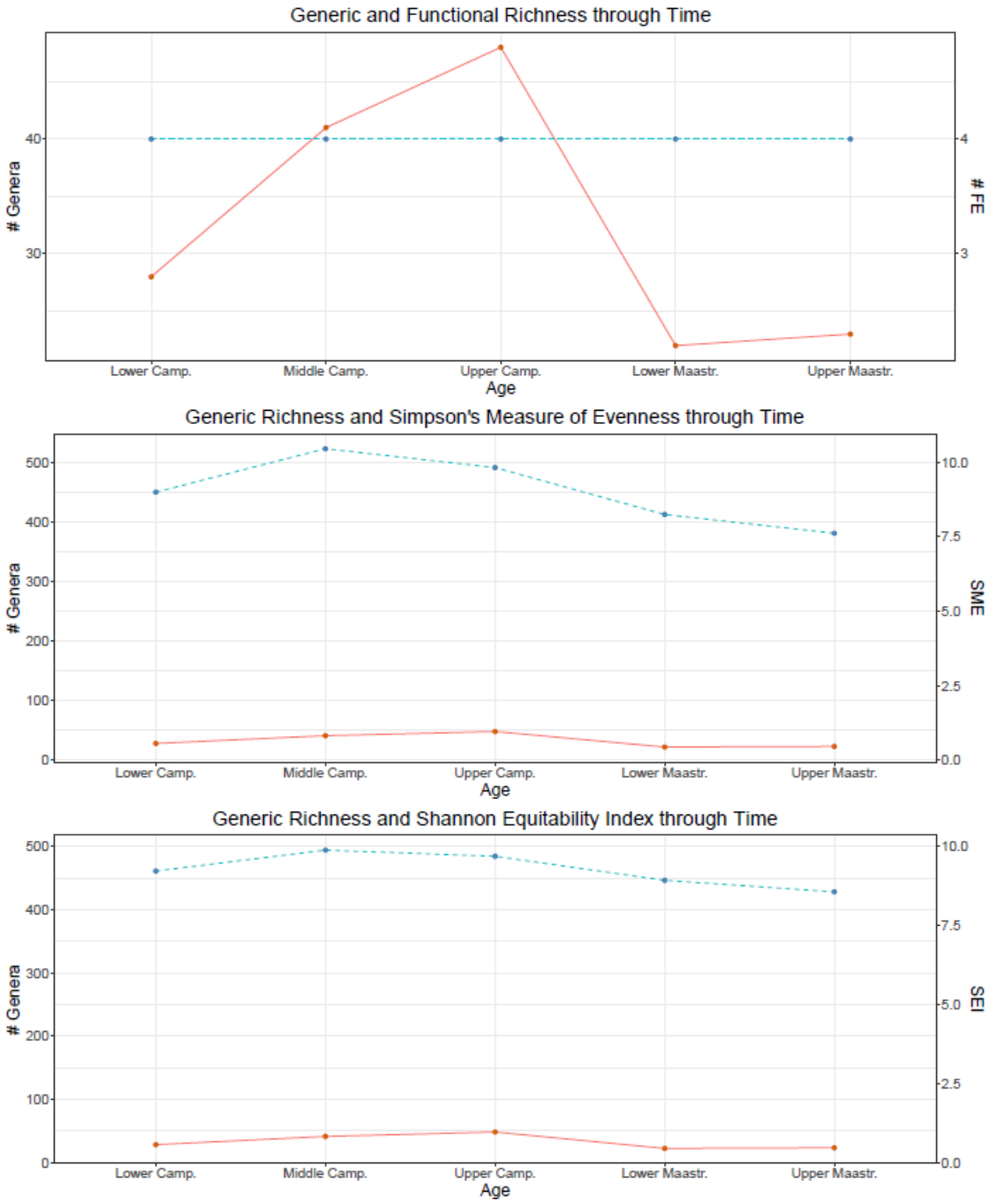



Figure S8. Line plots of changes in functional ecology metrics across substages for cephalopods only. The left y-axis in all plots depicts generic richness values, while the right y-axis presents values for the other metric.

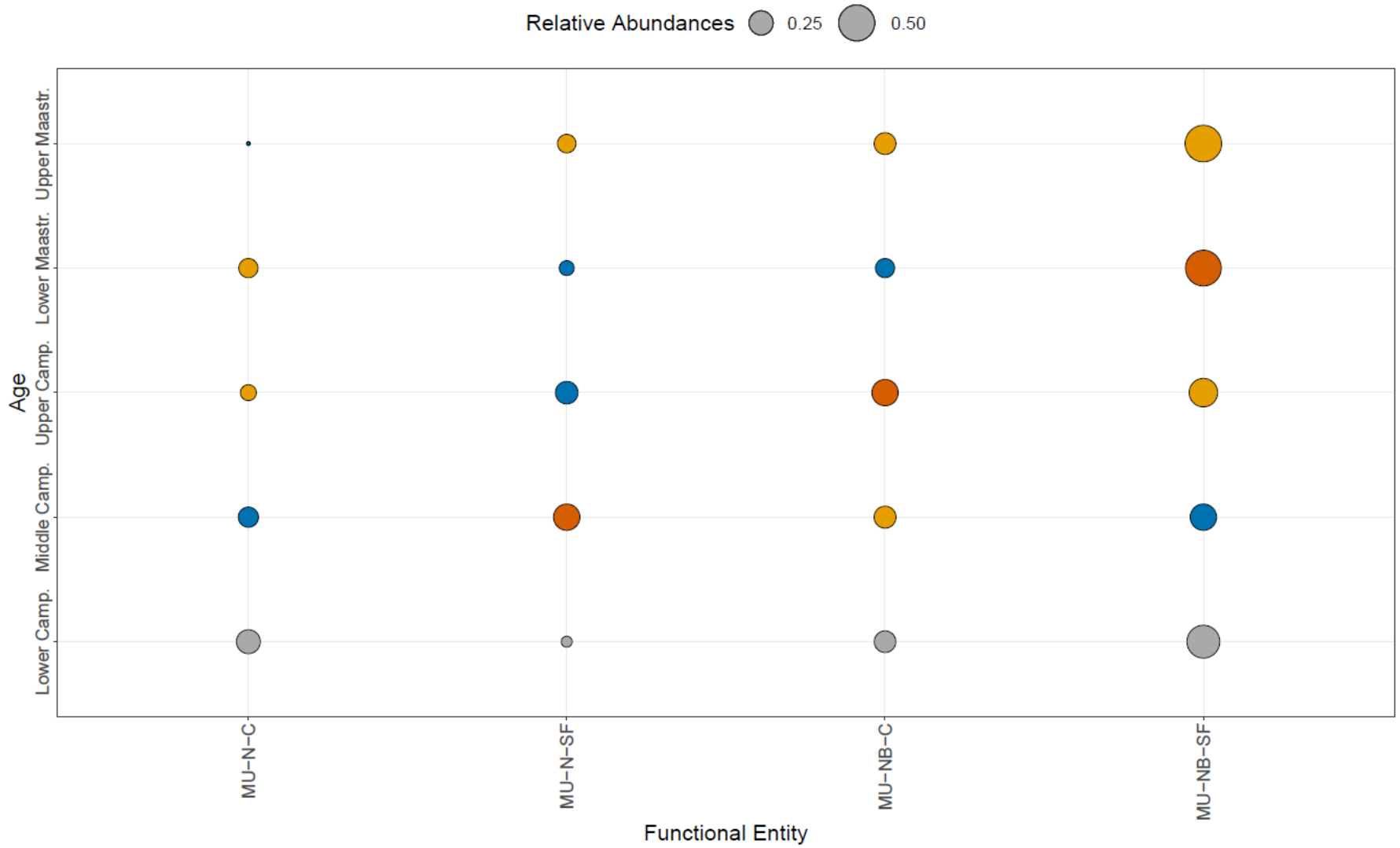



Figure S9. Plot of relative abundance of each FE through time based cephalopod data alone. Circle size indicates relative abundance of a FE within each substage and color indicates relative abundance increase across subsequent substages greater than 0.05 (orange), decrease greater than 0.05 (blue), and increase/decrease less than 0.05 (yellow). Points are grey if they cannot be compared with a previous substage. See Table 1 for keys to each FE names along the x-axis.

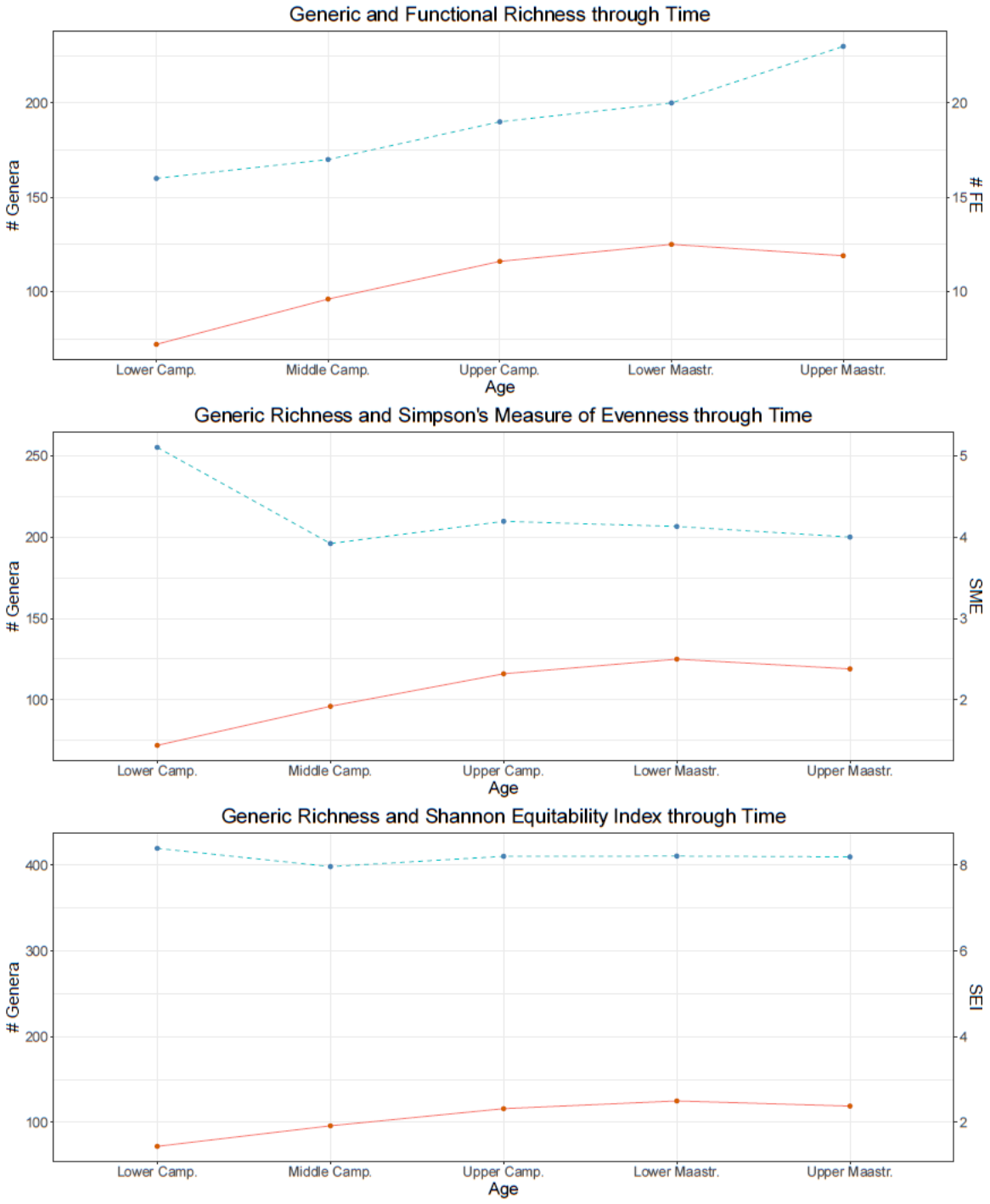



Figure S10. Line plots of changes in functional ecology metrics across substages for bivalves only. The left y-axis in all plots depicts generic richness values, while the right y-axis presents values for the other metric.

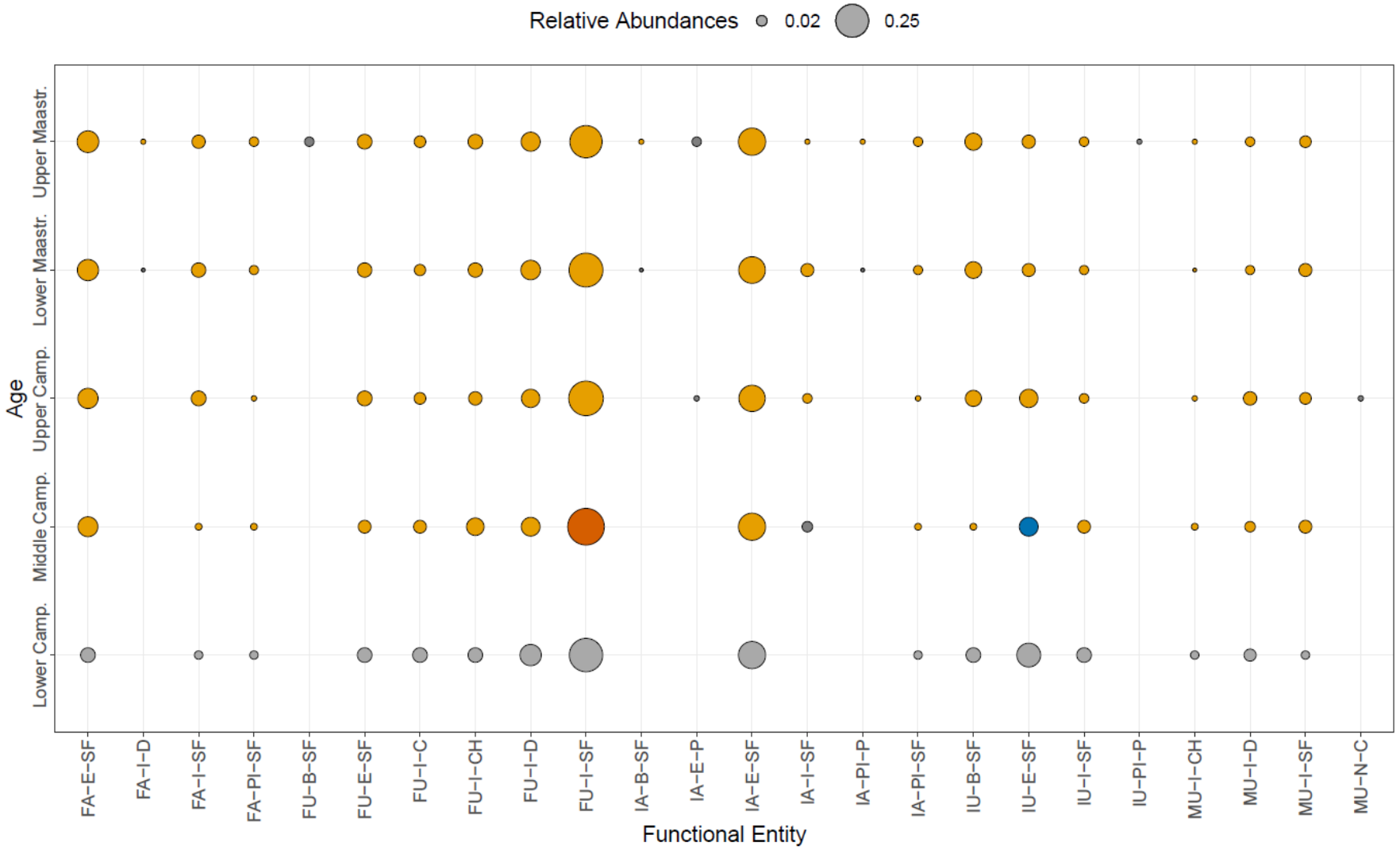



Figure S11. Plot of relative abundance of each FE through time based bivalve data alone. Circle size indicates relative abundance of a FE within each substage and color indicates relative abundance increase across subsequent substages greater than 0.05 (orange), decrease greater than 0.05 (blue), and increase/decrease less than 0.05 (yellow). Points are grey if they cannot be compared with a previous substage. See Table 1 for keys to each FE names along the x-axis.

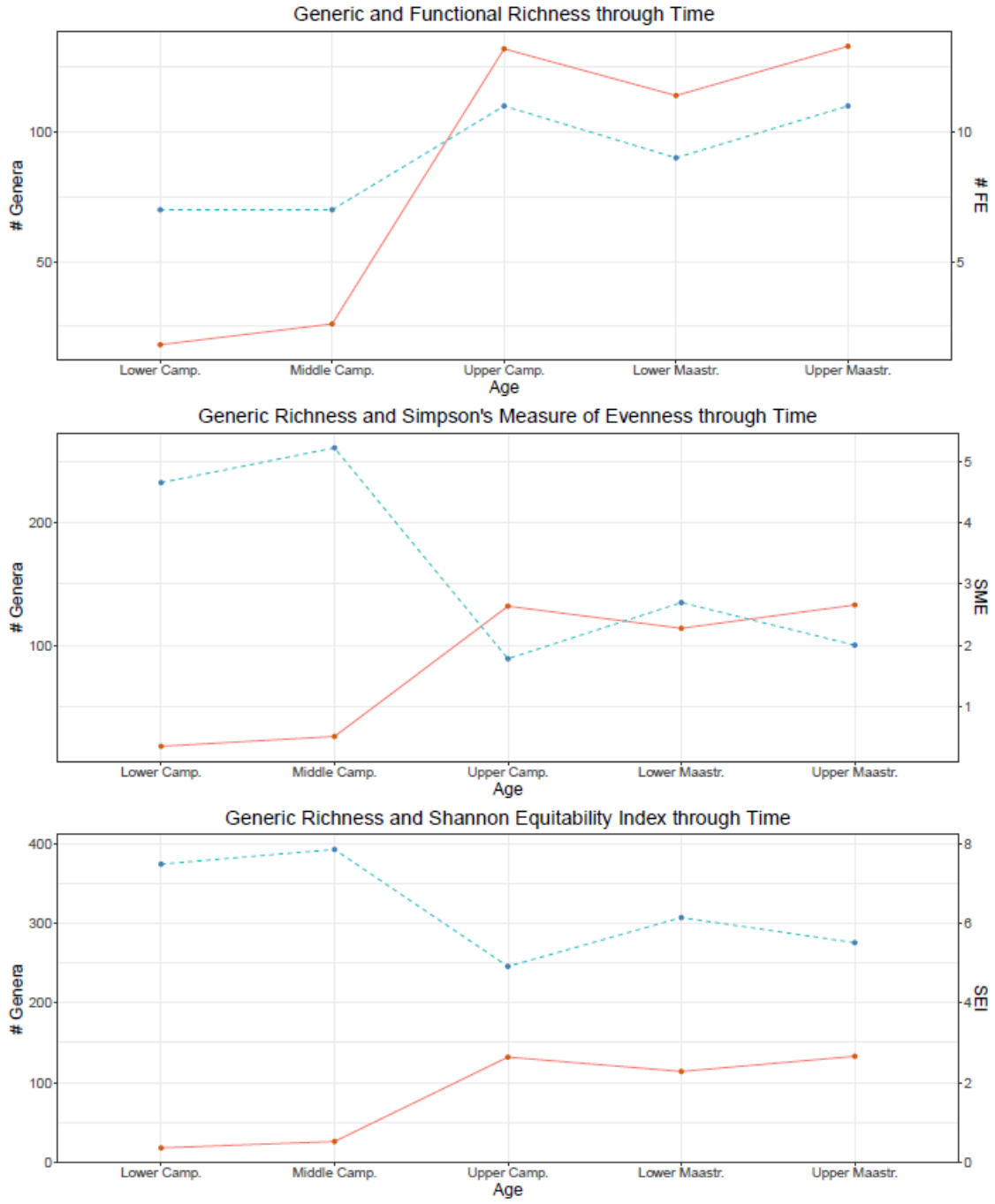



Figure S12. Line plots of changes in functional ecology metrics across substages for gasropods only. The left y-axis in all plots depicts generic richness values, while the right y-axis presents values for the other metric.

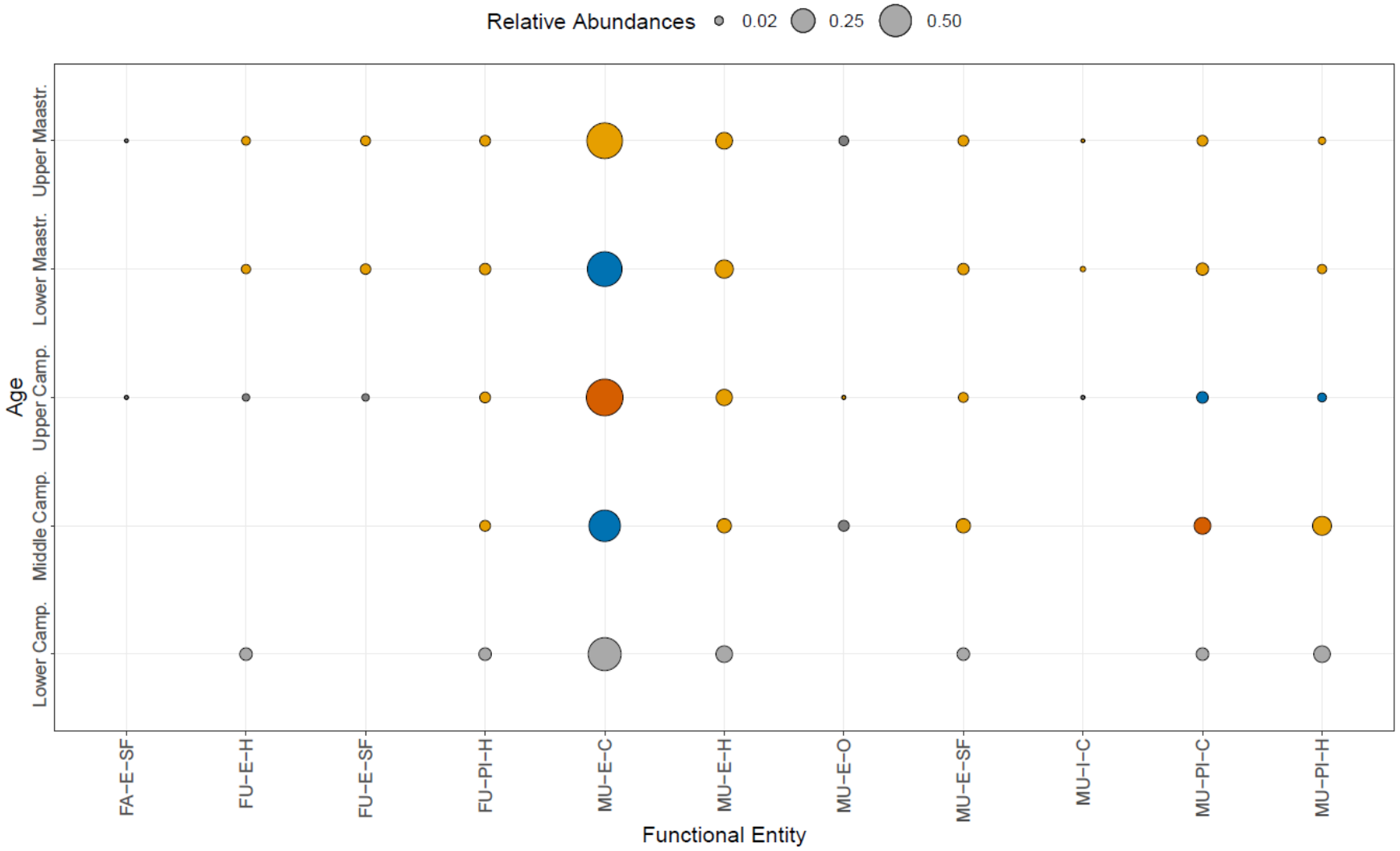



Figure S13. Plot of relative abundance of each FE through time based gastropod data alone. Circle size indicates relative abundance of a FE within each substage and color indicates relative abundance increase across subsequent substages greater than 0.05 (orange), decrease greater than 0.05 (blue), and increase/decrease less than 0.05 (yellow). Points are grey if they cannot be compared with a previous substage. See Table 1 for keys to each FE names along the x-axis.

Table S11. Table of the generic richness, FR, SME, and SEI values in each substage in cephalopods only.

| | Generic Richness | FR | SME | SEI |
|---|---|---|---|---|
| **Lower Camp.** | 28 | 4 | 0.90 | 0.92 |
| **Mid Camp.** | 41 | 4 | 1.05 | 0.99 |
| **Upper Camp.** | 48 | 4 | 0.98 | 0.97 |
| **Lower Maastr.** | 22 | 4 | 0.83 | 0.89 |
| **Upper Maastr.** | 23 | 4 | 0.76 | 0.86 |
| **Mean** | **32.4** | **4** | **0.90** | **0.93** |
| **Median** | **28** | **4** | **0.90** | **0.92** |
| **S.D.** | **11.55** | **0.00** | **0.11** | **0.05** |

Table S12. Table of proportional change in metrics across substages in cephalopods only.

| | Proportional Change Across Substages: | | | |
|---|---|---|---|---|
| | **GR** | **FR** | **SME** | **SEI** |
| **Low Camp.-Mid Camp.** | 0.46 | 0.00 | 0.16 | 0.07 |
| **Mid Camp.-Up Camp.** | 0.17 | 0.00 | -0.06 | -0.02 |
| **Up Camp.-Low Maastr.** | -0.54 | 0.00 | -0.16 | -0.08 |
| **Low Maastr.-Up Maastr.** | 0.05 | 0.00 | -0.08 | -0.04 |
| **Average:** | **0.03** | **0.00** | **-0.03** | **-0.02** |
| **Median:** | **0.11** | **0.00** | **-0.07** | **-0.03** |
| **S.D.:** | **0.42** | **0.00** | **0.14** | **0.06** |

Table S13. Table of the generic richness, FR, SME, and SEI values in each substage in bivalves only.

| | Generic Richness | FR | SME | SEI |
|---|---|---|---|---|
| **Lower Camp.** | 72 | 16 | 0.51 | 0.84 |
| **Mid Camp.** | 96 | 17 | 0.39 | 0.80 |
| **Upper Camp.** | 116 | 19 | 0.42 | 0.82 |
| **Lower Maastr.** | 125 | 20 | 0.41 | 0.82 |
| **Upper Maastr.** | 119 | 23 | 0.40 | 0.82 |
| **Mean** | **105.6** | **19** | **0.43** | **0.82** |
| **Median** | **116** | **19** | **0.41** | **0.82** |
| **S.D.** | **21.71** | **2.74** | **0.05** | **0.02** |

Table S14. Table of proportional change in metrics across substages in bivalves only.

| | Proportional Change Across Substages: | | | |
|---|---|---|---|---|
| | GR | FR | SME | SEI |
| Low Camp.-Mid Camp. | 0.33 | 0.06 | -0.23 | -0.05 |
| Mid Camp.-Up Camp. | 0.21 | 0.12 | 0.07 | 0.03 |
| Up Camp.-Low Maastr. | 0.08 | 0.05 | -0.02 | 0.00 |
| Low Maastr.-Up Maastr. | -0.05 | 0.15 | -0.03 | 0.00 |
| **Average:** | **0.14** | **0.10** | **-0.05** | **-0.01** |
| **Median:** | **0.14** | **0.09** | **-0.02** | **0.00** |
| **S.D.:** | **0.16** | **0.05** | **0.13** | **0.03** |

Table S15. Table of the generic richness, FR, SME, and SEI values in each substage in gastropods only.

| | Generic Richness | FR | SME | SEI |
|---|---|---|---|---|
| Lower Camp. | 18 | 7 | 0.47 | 0.75 |
| Mid Camp. | 26 | 7 | 0.52 | 0.79 |
| Upper Camp. | 132 | 11 | 0.18 | 0.49 |
| Lower Maastr. | 114 | 9 | 0.27 | 0.61 |
| Upper Maastr. | 133 | 11 | 0.20 | 0.55 |
| **Mean** | **84.6** | **9** | **0.33** | **0.64** |
| **Median** | **114** | **9** | **0.27** | **0.61** |
| **S.D.** | **57.71** | **2.00** | **0.16** | **0.13** |

Table S16. Table of proportional change in metrics across substages in gastropods only.

| | Proportional Change Across Substages: | | | |
|---|---|---|---|---|
| | GR | FR | SME | SEI |
| Low Camp.-Mid Camp. | 0.44 | 0.00 | 0.12 | 0.05 |
| Mid Camp.-Up Camp. | 4.08 | 0.57 | -0.66 | -0.37 |
| Up Camp.-Low Maastr. | -0.14 | -0.18 | 0.51 | 0.25 |
| Low Maastr.-Up Maastr. | 0.17 | 0.22 | -0.26 | -0.10 |
| **Average:** | **1.14** | **0.15** | **-0.07** | **-0.04** |
| **Median:** | **0.31** | **0.11** | **-0.07** | **-0.03** |
| **S.D.:** | **1.97** | **0.32** | **0.50** | **0.26** |

Table S17. Relative abundance of each FE in each substage for cephalopods only.

| Functional Entities | Lower Camp. | Middle Camp. | Upper Camp. | Lower Maastr. | Upper Maastr. |
|---|---|---|---|---|---|
| MU-NB-SF | 0.429 | 0.293 | 0.333 | 0.500 | 0.522 |
| MU-N-C | 0.250 | 0.195 | 0.146 | 0.182 | 0.087 |
| MU-NB-C | 0.214 | 0.220 | 0.292 | 0.182 | 0.217 |
| MU-N-SF | 0.107 | 0.293 | 0.229 | 0.136 | 0.174 |

Table S18. Relative abundance of each FE in each substage for bivalves only.

| Functional Entities | Lower Camp. | Middle Camp. | Upper Camp. | Lower Maastr. | Upper Maastr. |
|---|---|---|---|---|---|
| FU-I-SF | 0.264 | 0.323 | 0.284 | 0.272 | 0.244 |
| FU-E-SF | 0.042 | 0.031 | 0.043 | 0.040 | 0.042 |
| MU-I-D | 0.028 | 0.021 | 0.034 | 0.016 | 0.017 |
| MU-I-CH | 0.014 | 0.010 | 0.009 | 0.008 | 0.008 |
| FA-I-SF | 0.014 | 0.010 | 0.043 | 0.040 | 0.034 |
| FA-PI-SF | 0.014 | 0.010 | 0.009 | 0.016 | 0.017 |
| IA-PI-SF | 0.014 | 0.010 | 0.009 | 0.016 | 0.017 |
| MU-I-SF | 0.014 | 0.031 | 0.026 | 0.032 | 0.025 |
| IA-E-SF | 0.167 | 0.167 | 0.155 | 0.160 | 0.168 |
| IU-E-SF | 0.125 | 0.073 | 0.069 | 0.032 | 0.034 |
| FU-I-D | 0.097 | 0.073 | 0.069 | 0.080 | 0.076 |
| FU-I-C | 0.042 | 0.031 | 0.026 | 0.024 | 0.025 |
| FU-I-CH | 0.042 | 0.063 | 0.034 | 0.040 | 0.042 |
| IU-B-SF | 0.042 | 0.010 | 0.052 | 0.056 | 0.059 |
| IU-I-SF | 0.042 | 0.031 | 0.017 | 0.016 | 0.017 |
| FA-E-SF | 0.042 | 0.083 | 0.086 | 0.096 | 0.101 |
| IA-I-SF | NA | 0.021 | 0.017 | 0.032 | 0.008 |
| MU-N-C | NA | NA | 0.009 | NA | NA |
| IA-E-P | NA | NA | 0.009 | NA | 0.017 |
| FA-I-D | NA | NA | NA | 0.008 | 0.008 |
| IA-PI-P | NA | NA | NA | 0.008 | 0.008 |
| IA-B-SF | NA | NA | NA | 0.008 | 0.008 |
| FU-B-SF | NA | NA | NA | NA | 0.017 |
| IU-PI-P | NA | NA | NA | NA | 0.008 |

Table S19. Relative abundance of each FE in each substage for gastropods only.

| Functional Entities | Lower Camp. | Middle Camp. | Upper Camp. | Lower Maastr. | Upper Maastr. |
|---|---|---|---|---|---|
| MU-E-C | 0.556 | 0.500 | 0.705 | 0.623 | 0.662 |
| MU-E-H | 0.111 | 0.077 | 0.106 | 0.140 | 0.113 |
| MU-PI-H | 0.111 | 0.154 | 0.023 | 0.026 | 0.015 |
| MU-PI-C | 0.056 | 0.115 | 0.045 | 0.053 | 0.038 |
| FU-E-H | 0.056 | NA | 0.015 | 0.026 | 0.023 |
| FU-PI-H | 0.056 | 0.038 | 0.038 | 0.044 | 0.038 |
| MU-E-SF | 0.056 | 0.077 | 0.030 | 0.044 | 0.038 |
| MU-E-O | NA | 0.038 | 0.008 | NA | 0.030 |
| FA-E-SF | NA | NA | 0.008 | NA | 0.008 |
| FU-E-SF | NA | NA | 0.015 | 0.035 | 0.030 |
| MU-I-C | NA | NA | 0.008 | 0.009 | 0.008 |

## Summary of Functional vs Taxonomic Diversity change in Latitudinal Bins:

*Metrics in General:*

When the data is converted to paleo-coordinates and separated into seven 5° paleo-latitudinal bins for analysis, almost all bins have continuous occupation across the study interval (Table 4). Only in the Middle Campanian for the 30-35°N bin, and in the Lower and Upper Maastrichtian for the 60-65°N bin is there missing data for comparison. However, only the 30-35 and 50-55 bins have greater than the minimum estimated sample size of unique genera for each interval (see Table S1), though most substages have sufficient generic richness to represent common FEs across all but the northern-most and southern-most paleo-latitudinal bins (Table 4). Generic richness within paleo-latitude bins has greater spread of values through time but FR values have similar spreads, except in the Middle Campanian, where the first and third quartile values are closer to the median than in the other intervals (Figure S14). However, median and mean values of both generic and functional richness show similar trends within latitudinal bins, increasing across the Campanian before decreasing across the Maastrichtian. SME values have the highest spread in the Upper Maastrichtian and mean and median values appear to decrease slightly (i.e., become less even) across the Campanian then increase slightly (become more even) across the Maastrichtian (Figure S14). SEI values follow a similar pattern overall.

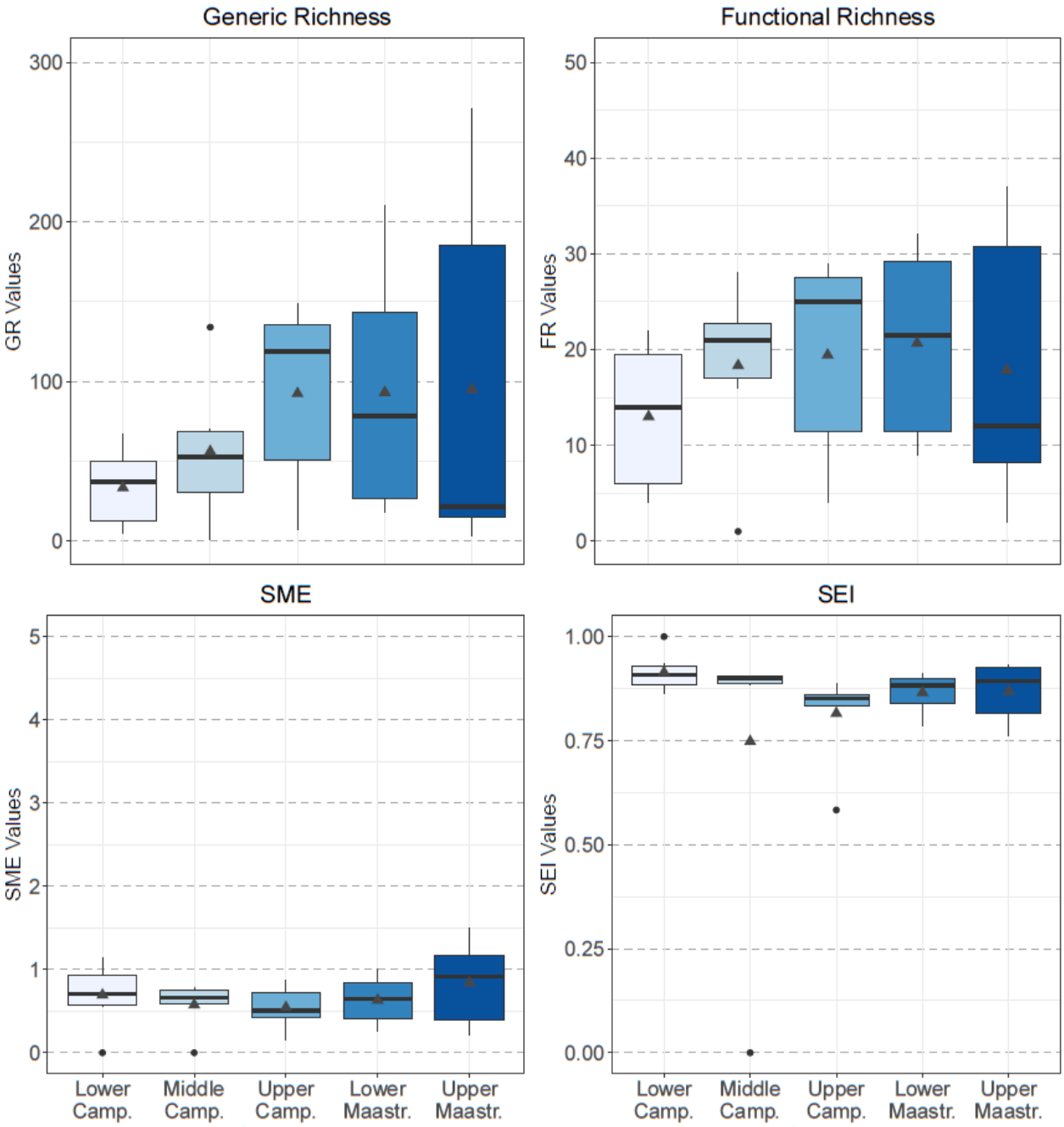



Figure S14. Box and whisker plots of generic richness, FR, SME, and SEI for paleo-latitudinal bins in each substage. Median values are denoted by a horizontal line and means by triangles.

Both generic and functional richness shift across time based on paleo-latitude bins (Table S20, Figure S15). Median and average proportional change in FR between substages, taking into account all paleo-latitudinal bins, indicates that FR increases between the Lower and Middle Campanian (~16% median) and between the Middle to Upper Campanian (22% median), and fell between substages from the Upper Campanian to the Lower Maastrichtian and Lower to Upper Maastrichtian (~24% median and ~3% median, respectively). Only between the Lower and Upper Maastrichtian does the median fall particularly close to zero, suggesting little overall change, but the first and third interquartile range values indicate high variability relative to other substage comparisons.

All box plots showing proportional change values for functional richness cross the zero line, either along the hinge (first or third quartile values) or along the whisker (1.5* interquartile range). This indicates that some paleo-latitude bins experience opposing directions of change in FR, despite the prevailing trends between substages. Proportional change in generic richness follow a similar general trend, except that the spread of values is consistently larger than those observed for FR and proportional changes are more extreme. For example, all proportional changes in generic richness increase between the Middle and Upper Campanian, indicating that generic richness values increase within all paleo-latitude bins between these substages. This once again suggests that generic and functional richness, despite being tightly linked, are at times decoupled, even at this scale. When paleo-latitudinal bins with less than the estimated SQS lower limit of generic richness are removed from calculations, these patterns are generally maintained, but the spread of values is less extreme in all cases (Figure S16). Only in the Upper Campanian to Lower Maastrichtian comparison do statistic value change appreciably to a median value of -43% rather than -21%, indicating stronger loss of FR.

SME values are more variable across substages, decreasing between Campanian substages and rising between the Upper Campanian and Maastrichtian substages (Figure S15). However, only the Middle to Upper Campanian comparison shows unanimous proportional decrease in values. When paleo-latitude bins with less than the estimated SQS generic richness level are removed, the spread of

proportional change values for all substages is diminished and mean and median values for SME fall closer to zero, excluding the Middle to Upper Campanian comparison where no values were removed. This suggest that the removal of generic richness values which may not appropriately represent FEs dampens both negative and positive proportional changes in evenness. Overall, proportional changes in SME indicate that the Lower to Middle Campanian has fairly stable evenness but evenness drops between the Middle to Upper Campanian, then rises from the Upper Campanian to Upper Maastrichtian.

SEI values, unlike SME, are highly stable across all intervals, particularly when paleo-latitude bins with generic richness lower than the SQS estimated requirements are removed (Figure S16). This indicates overall evenness (as opposed to dominance) in the data at the paleo-latitude-bin level of aggregation, however error introduced by lower richness levels for this metric is assumed to be high. Slight shifts in SEI proportional change values do mimic patterns observed with SME, indicating that shifts in the distribution of genera within FEs does occur through time.

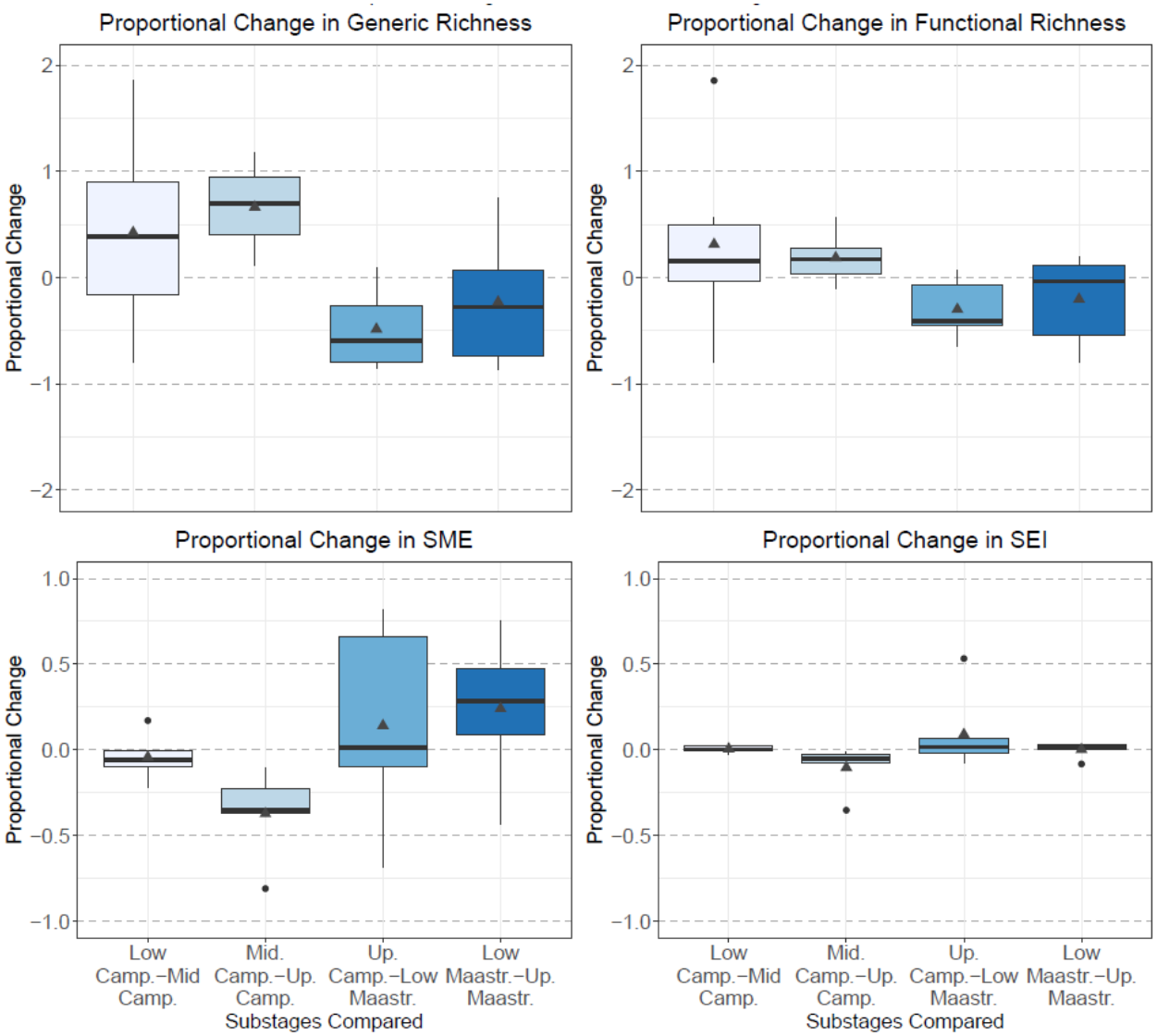



Figure S15. Box and whisker plots of proportional change in paleo-latitude bins between substages for each metric value. Triangles indicate mean values and horizontal lines indicate median values.

Table S20. Summary statistics for proportional change in generic richness, FR, SME, and SEI for paleo-latitude bins across adjacent substages with all values included, regardless of if generic richness for that paleo-latitude bin was less than estimated SQS sample size requirement.

| | | Low Camp. - Mid Camp. | Mid. Camp. - Up. Camp. | Up. Camp. - Low Maastr. | Low Maastr. - Up. Maastr. |
|---|---|---|---|---|---|
| Generic Richness | Mean | 0.43 | 2.07 | 1.93 | -0.23 |
| | Median | 0.39 | 1.70 | -0.43 | -0.27 |
| | SD | 0.94 | 2.10 | 5.92 | 0.62 |
| Functional Richness | Mean | 0.31 | 0.66 | 0.65 | -0.20 |
| | Median | 0.16 | 0.22 | -0.24 | -0.03 |
| | SD | 0.88 | 1.17 | 2.34 | 0.44 |
| SME | Mean | -0.04 | -0.37 | 0.91 | 0.24 |
| | Median | -0.06 | -0.35 | 0.33 | 0.28 |
| | SD | 0.14 | 0.27 | 1.97 | 0.41 |
| SEI | Mean | 0.01 | -0.10 | 0.09 | 0.00 |
| | Median | 0.00 | -0.05 | 0.02 | 0.02 |
| | SD | 0.03 | 0.14 | 0.22 | 0.04 |

Table S21. Summary statistics for proportional change in generic richness, functional richness ($\alpha_f$), and functional evenness (SME) for paleo-latitude bins across adjacent substages. Camp. = Campanian, Maastr. = Maastrichtian. Values were removed from calculations if they did not meet the minimum SQS sample size estimate.

| | | Low Camp. - Mid Camp. | Mid. Camp. - Up. Camp. | Up. Camp. - Low Maastr. | Low Maastr. - Up. Maastr. |
|---|---|---|---|---|---|
| Generic Richness | Mean | 0.38 | 1.29 | -0.48 | -0.23 |
| | Median | 0.39 | 1.18 | -0.60 | -0.27 |
| | SD | 0.54 | 0.93 | 0.39 | 0.62 |
| Functional Richness | Mean | 0.21 | 0.19 | -0.30 | -0.20 |
| | Median | 0.16 | 0.17 | -0.41 | -0.03 |
| | SD | 0.28 | 0.25 | 0.29 | 0.44 |
| SME | Mean | 0.00 | -0.37 | 1.23 | 0.24 |
| | Median | -0.03 | -0.35 | 0.66 | 0.28 |
| | SD | 0.12 | 0.27 | 2.02 | 0.41 |
| SEI | Mean | 0.01 | -0.10 | 0.12 | 0.00 |
| | Median | 0.01 | -0.05 | 0.02 | 0.02 |
| | SD | 0.03 | 0.14 | 0.23 | 0.04 |

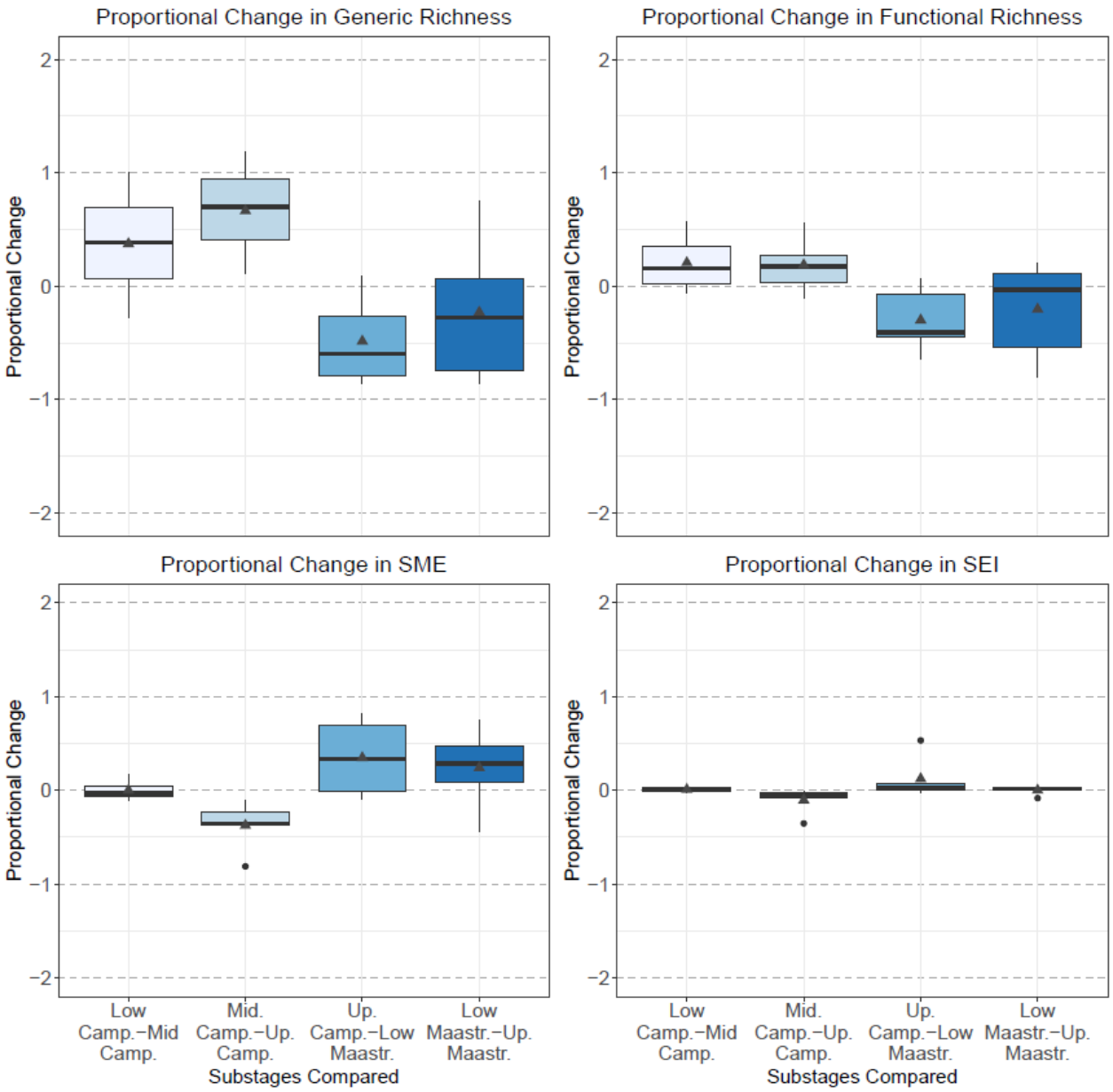



Figure S16. Box and whisker plots of proportional change between substages for each metric value with paleo-latitude bin values corresponding to generic richness less than the required estimate from SQS removed (i.e., <16 unique genera for Campanian substages, and <18 unique genera for Maastrichtian substages). Triangles indicate mean values and horizontal lines indicate median values.

*Proportional Change in Metrics Across Latitude:*

Proportional change across paleo-latitudes within paleo-latitude bins indicates that, in general, substages experience decreasing generic and functional richness across paleo-latitude from south to north, except between the 40-45°N to 45-50°N bins (Figure 4). Both generic and functional richness decrease dramatically moving north across paleo-latitude above the 40-45°N bin. The Upper Maastrichtian is one exception to this pattern, showing an increase of close to ~88% in FR across the 45-50°N and 50-55°N bins. Between the 40-45°N to 45-50°N bins, FR and generic richness increase by more than 40% in all substages (Table S22). Proportional change in generic richness across these bins, while following similar patterns, is not identical. For example, the Middle Campanian experiences a proportional increase of 235% generic richness but only a 40% increase in FR, but the Upper Campanian experiences a proportional increase of generic richness of 16% while FR for this substage only increases by 61%. Once again, the two factors are decoupled though similar.

Across the 35-40°N to 40-45°N transition, both functional and generic richness decrease in all substages while between the 30-35°N and 35-40°N paleo-latitude bins, proportional change in richness is more variable. In this southernmost transition, the Upper and Lower Campanian experience very noticeable increases in both generic and functional richness, while the Lower and Upper Maastrichtian experience a fall of only 6% and 3%, respectively (Table S22). It should be pointed out, that these southern-most bins are dominated by Maastrichtian data, and therefore this relative stability may be influenced by better sampling for the Maastrichtian than for the Campanian substages. Indeed, the more dramatic decreases in proportional richness values in the most northern latitude transitions may also be a product of poor sampling. Removing values which have lower generic richness than the SQS estimated sample size for FR in each substage produces very similar patterns, except that the northernmost and southernmost values are removed for the Campanian substages (Table S23; Figure S17). This suggests that these extraneous values may be caused by poor sampling and strengthens the conclusion that both generic and functional richness decrease in similar proportions across latitude in all substages, except between the 40-45°N to 45-50°N bins where values increase notably.

Shifts in dominance across latitude are also highly variable, though, in general, SME increases in substages across latitudes between 45°N and 60°N. This indicates that evenness increases between latitudes moving north. In the 40-45°N to 45-50°N transition, SEM values decrease in all substages, indicating that evenness drops across latitudes in this range. South of 40°N, SEM show both increase and decrease in different substages, with an extremely noticeable increase between 35-40°N and 40-45°N for the Upper Maastrichtian (Figure 4). When values with fewer than the SQS sample size estimate are removed, these patterns are maintained.

Table S22. Values of proportional change across paleo-latitude bins in each substages

| | 30-35°N to 35-40°N | 35-40°N to 40-45°N | 40-45°N to 45-50°N | 45-50°N to 50-55°N | 50-55°N to 55-60°N | 55-60°N to 60-65°N |
|---|---|---|---|---|---|---|
| | Generic Richness | | | | | |
| **Low Camp.** | 2.73 | -0.66 | 3.79 | -0.12 | -0.37 | -0.86 |
| **Mid Camp.** | NA | -0.38 | 2.35 | -0.48 | -0.61 | -0.96 |
| **Up Camp.** | 9.14 | **-0.09** | **0.16** | -0.20 | -0.27 | -0.92 |
| **Low Maastr.** | **-0.26** | -0.88 | 4.74 | -0.56 | -0.63 | NA |
| **Up Maastr.** | **0.14** | -0.99 | 4.00 | 0.80 | -0.41 | NA |
| | Functional Richness | | | | | |
| **Low Camp.** | 2.50 | -0.50 | 2.14 | 0.00 | -0.23 | -0.71 |
| **Mid Camp.** | NA | -0.09 | 0.40 | -0.18 | -0.30 | -0.94 |
| **Up Camp.** | 4.60 | -0.36 | 0.61 | -0.07 | -0.07 | -0.84 |
| **Low Maastr.** | -0.06 | -0.67 | 1.70 | -0.41 | -0.44 | NA |
| **Up Maastr.** | -0.03 | -0.94 | 3.00 | 0.88 | -0.40 | NA |
| | SME | | | | | |
| **Low Camp.** | -0.50 | 0.76 | -0.44 | 0.25 | 0.23 | -1.00 |
| **Mid Camp.** | NA | 0.17 | -0.28 | 0.18 | 0.18 | -1.00 |
| **Up Camp.** | -0.49 | -0.65 | 2.42 | -0.15 | 0.40 | 0.45 |
| **Low Maastr.** | 0.46 | 1.26 | -0.40 | 0.53 | 0.28 | NA |
| **Up Maastr.** | -0.21 | 5.99 | -0.49 | 0.38 | 0.14 | NA |
| | SEI | | | | | |
| **Low Camp.** | -0.07 | 0.05 | -0.04 | 0.04 | 0.03 | 0.08 |
| **Mid Camp.** | NA | 0.00 | -0.02 | 0.02 | 0.00 | -1.00 |
| **Up Camp.** | 0.00 | -0.32 | 0.48 | -0.03 | 0.06 | -0.06 |
| **Low Maastr.** | 0.06 | 0.08 | -0.03 | 0.04 | 0.01 | NA |
| **Up Maastr.** | -0.05 | 0.21 | -0.05 | 0.07 | 0.00 | NA |

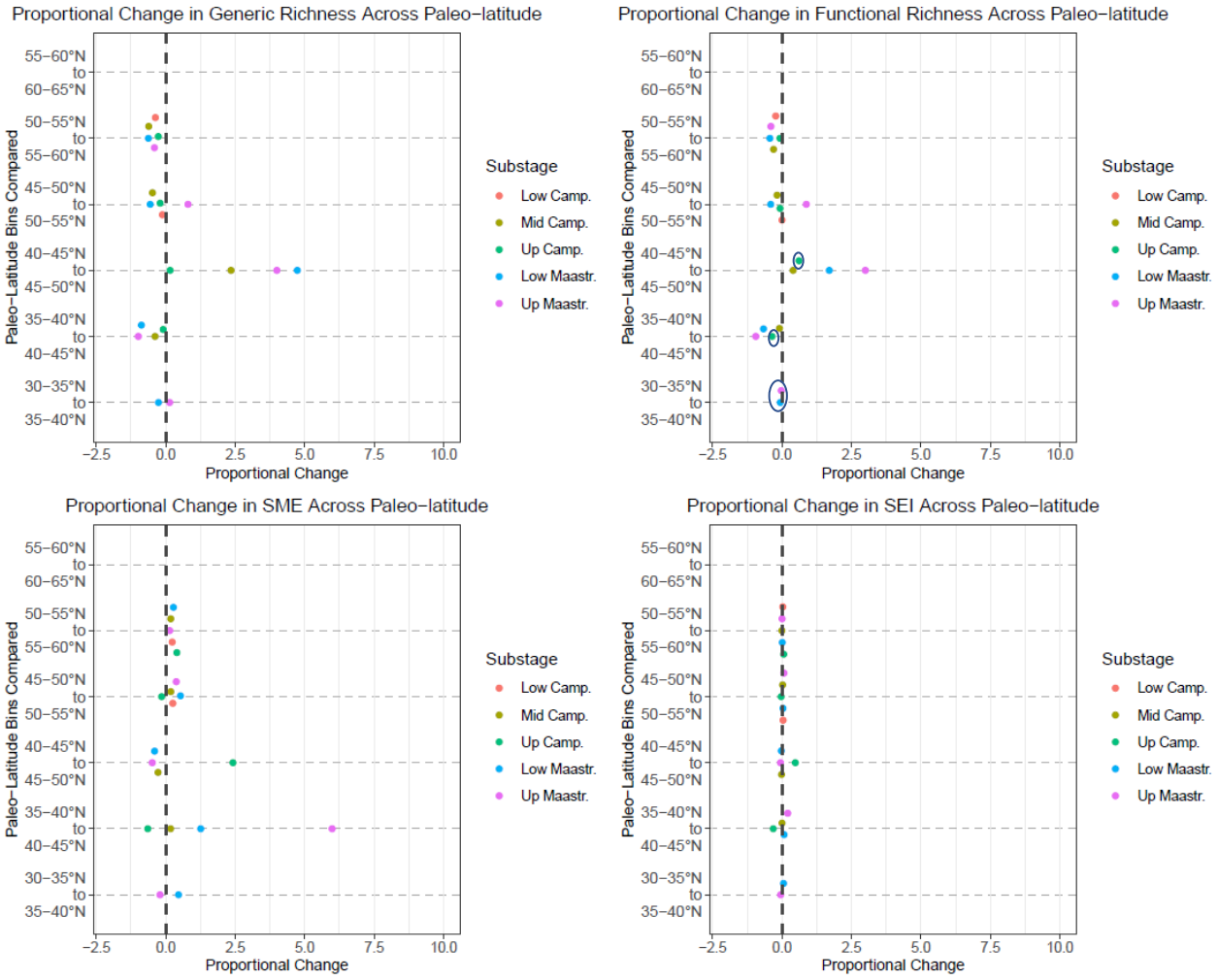



Figure S17. Beeswarm plots of proportional change between paleo-latitude bins for each metric value with paleo-latitude bin values corresponding to generic richness less than the required estimate from SQS removed (i.e., <16 unique genera for Campanian substages, and <18 unique genera for Maastrichtian substages).

Table S23. Values of proportional change across paleo-latitude bins in each substages with values below SQS sample size estimate removed.

| | 30-35°N to 35-40°N | 35-40°N to 40-45°N | 40-45°N to 45-50°N | 45-50°N to 50-55°N | 50-55°N to 55-60°N | 55-60°N to 60-65°N |
|---|---|---|---|---|---|---|
| | Generic Richness | | | | | |
| **Low Camp.** | NA | NA | NA | -0.12 | -0.37 | NA |
| **Mid Camp.** | NA | -0.38 | 2.35 | -0.48 | -0.61 | NA |
| **Up Camp.** | NA | -0.09 | 0.16 | -0.20 | -0.27 | NA |
| **Low Maastr.** | -0.26 | -0.88 | 4.74 | -0.56 | -0.63 | NA |
| **Up Maastr.** | 0.14 | -0.99 | 4.00 | 0.80 | -0.41 | NA |
| | Functional Richness | | | | | |
| **Low Camp.** | NA | NA | NA | 0.00 | -0.23 | NA |
| **Mid Camp.** | NA | -0.09 | 0.40 | -0.18 | -0.30 | NA |
| **Up Camp.** | NA | -0.36 | 0.61 | -0.07 | -0.07 | NA |
| **Low Maastr.** | -0.06 | -0.67 | 1.70 | -0.41 | -0.44 | NA |
| **Up Maastr.** | -0.03 | -0.94 | 3.00 | 0.88 | -0.40 | NA |
| | SME | | | | | |
| **Low Camp.** | NA | NA | NA | 0.25 | 0.23 | NA |
| **Mid Camp.** | NA | 0.17 | -0.28 | 0.18 | 0.18 | NA |
| **Up Camp.** | NA | -0.65 | 2.42 | -0.15 | 0.40 | NA |
| **Low Maastr.** | 0.46 | 1.26 | -0.40 | 0.53 | 0.28 | NA |
| **Up Maastr.** | -0.21 | 5.99 | -0.49 | 0.38 | 0.14 | NA |
| | SEI | | | | | |
| **Low Camp.** | NA | NA | NA | 0.04 | 0.03 | NA |
| **Mid Camp.** | NA | 0.00 | -0.02 | 0.02 | 0.00 | NA |
| **Up Camp.** | NA | -0.32 | 0.48 | -0.03 | 0.06 | NA |
| **Low Maastr.** | 0.06 | 0.08 | -0.03 | 0.04 | 0.01 | NA |
| **Up Maastr.** | -0.05 | 0.21 | -0.05 | 0.07 | 0.00 | NA |

*Nestedness verses Turnover in Functional Entities over Paleolatitude*

Nestedness, or the subsetting of FE assemblages across latitude, was tested using the *betapart* package in R (Baselga and Orme 2012) following Schumm et al. (2019). This package allows users to calculate the pairwise Sørensen dissimilarity between FE presence-absence data, which can be interpreted as a metric of nestedness. Sørensen values of 0 indicate turnover fully accounts for the total dissimilarity and 1 indicates that nestedness accounts for the dissimilarity. Results of nestedness analysis for paleolatitude bins indicate that nestedness is low (Sørensen dissimilarity <0.5) between most paleolatitude bins, and rarely high (Sørensen dissimilarity >0.5) between spatially adjacent bins (Table. S24-S28). Thus, nestedness does not account for dissimilarity well across latitude.

Table S24. Pairwise Sørensen dissimilarity values of FEs between Lower Campanian paleolatitude bins. Values greater than 0.5 are in bold.

| | **30-35°N** | **35-40°N** | **40-45°N** | **45-50°N** | **50-55°N** | **55-60°N** |
|---|---|---|---|---|---|---|
| **35-40°N** | **0.556** | | | | | |
| **40-45°N** | 0.205 | 0.286 | | | | |
| **45-50°N** | **0.692** | 0.175 | **0.517** | | | |
| **50-55°N** | **0.519** | 0.143 | **0.517** | 0.000 | | |
| **55-60°N** | 0.310 | 0.062 | 0.357 | 0.098 | 0.106 | |
| **60-65°N** | 0.028 | 0.284 | 0.033 | **0.630** | **0.630** | 0.436 |

Table S25. Pairwise Sørensen dissimilarity values of FEs between Middle Campanian paleolatitude bins. Values greater than 0.5 are in bold.

| | **35-40°N** | **40-45°N** | **45-50°N** | **50-55°N** | **55-60°N** |
|---|---|---|---|---|---|
| **40-45°N** | 0.038 | | | | |
| **45-50°N** | 0.115 | 0.158 | | | |
| **50-55°N** | 0.018 | 0.066 | 0.098 | | |
| **55-60°N** | 0.109 | 0.097 | 0.256 | 0.168 | |
| **60-65°N** | **0.913** | 0.000 | **0.931** | 0.000 | 0.000 |

Table S26. Pairwise Sørensen dissimilarity values of FEs between Upper Campanian paleolatitude bins. Values greater than 0.5 are in bold.

| | 30-35°N | 35-40°N | 40-45°N | 45-50°N | 50-55°N | 55-60°N |
|---|---|---|---|---|---|---|
| **35-40°N** | **0.697** | | | | | |
| **40-45°N** | 0.339 | 0.205 | | | | |
| **45-50°N** | **0.706** | 0.015 | 0.221 | | | |
| **50-55°N** | **0.550** | 0.015 | 0.189 | 0.034 | | |
| **55-60°N** | **0.533** | 0.050 | 0.154 | 0.071 | 0.034 | |
| **60-65°N** | 0.028 | **0.563** | 0.477 | **0.758** | **0.742** | **0.724** |

Table S27. Pairwise Sørensen dissimilarity values of FEs between Lower Maastrichtian paleolatitude bins. Values greater than 0.5 are in bold.

| | 30-35°N | 35-40°N | 40-45°N | 45-50°N | 50-55°N |
|---|---|---|---|---|---|
| **35-40°N** | 0.030 | | | | |
| **40-45°N** | 0.471 | 0.450 | | | |
| **45-50°N** | 0.082 | 0.045 | 0.459 | | |
| **50-55°N** | 0.313 | 0.266 | 0.208 | 0.256 | |
| **55-60°N** | **0.561** | 0.479 | 0.029 | 0.500 | 0.218 |

Table S28. Pairwise Sørensen dissimilarity values of FEs between Upper Maastrichtian paleolatitude bins. Values greater than 0.5 are in bold.

| | 30-35°N | 35-40°N | 40-45°N | 45-50°N | 50-55°N |
|---|---|---|---|---|---|
| **35-40°N** | 0.013 | | | | |
| **40-45°N** | **0.897** | **0.895** | | | |
| **45-50°N** | **0.644** | **0.636** | **0.600** | | |
| **50-55°N** | 0.423 | 0.412 | **0.765** | 0.228 | |
| **55-60°N** | **0.609** | **0.600** | 0.318 | 0.022 | 0.167 |

Summary of Functional vs Taxonomic Diversity change in 360km2 Nodes across Substages:

At the 360km aggregation scale, generic and functional richness distributions are fairly similar in all substages (Figure S18, Table S29), though the Lower Campanian has a noticeably smaller spread of FR values in the first and third quartiles than later substages and the Upper Maastrichtian has a greater spread of values in this range than other substages. Generic richness outliers are present in most substages, but only the Lower Campanian and Lower Maastrichtian have FR outliers. SME values has small distributions, in all substages with some outliers in the Mid Campanian and Lower Maastrichtian. SEM values have outliers at zero in all cases, except the Middle Campanian, where the first interquartile hinge is at zero. Again, zeros in this metric indicate "perfect evenness" which most commonly occurs when only a single FE exists in that node. Therefore, these values are not considered important in general.

Mean and median FR values for each substage at the 360km node scale are fluctuate across the substages, ranging from 6.21-14.5 and 4-11, respectively. Generic richness mean and median values follow a similar pattern, but means are consistently much larger than median values, ranging from 12.36-56.85 while median values range from only 6-19. Mean and median values for SME are fairly consistent, ranging from 0.59-0.97 and 0.43-0.85, respectively. However, the Lower, and to a lesser degree, Upper Maastrichtian distributions have a noticeable increase in all values (Figure S16), indicating lower overall evenness in these substages compared to others.

Proportional change in 360km node metrics indicate that **FR increases across the Campanian, decreases between the Upper Campanian and Lower Maastrichtian, and is fairly stable overall between the Lower to Upper Maastrichtian** (Figure 5, Table S30). Median FR proportional change values range from only 2% between the Lower and Upper Maastrichtian to -39% (decrease) between the Upper Campanian and Lower Maastrichtian. The distribution of values around these means are fairly consistent, except in the Upper Campanian to Lower Maastrichtian when hinge values are noticeably closer to the median though at least one outlier point is observed in all substages. The very small spread of first and third interquartile values between the Upper Campanian and Lower Maastrichtian suggests

that the median value is highly representative of the data for this comparison. Mean FR values of proportional change are higher than median values, ranging from 98% to 208% increase in FR values between nodes. These high mean values are most likely influenced by outliers, which are present in all substages (Figure 5). Proportional change in generic richness follows a very similar pattern to FR, but mean and median values both indicate a decrease in generic richness between the Lower and Upper Maastrichtian rather than stability in the metric. Again, the two are similar but decoupled. In all cases richness distributions cross the zero line either along the interquartile range or along the whisker (1.5* IQR). This suggests that not all generic or functional richness values within 360km nodes experience only increase or decrease across substages.

SME values indicate that evenness declined through the Campanian (median: -24% and -27%) and rose between the Upper Campanian and Lower Maastrichtian (median: 15%), but was fairly stable between the Lower and Upper Maastrichtian (median: -1%) (Figure 5, Table S30). Again, mean values are frequently higher than median values for this metric, most likely influenced by outliers. All distributions for SME cross the zero line, indicating that evenness change is not unanimously increasing or decreasing in 360km nodes between substages. **Overall, SME values at this scale of analysis suggest declining evenness in the Campanian, then rising evenness in the Upper Campanian to Lower Maastrichtian, followed by relative stability in the Maastrichtian.**

Summary of Functional vs Taxonomic Diversity change in 60km2 Nodes across Substages:

At the 60km aggregation scale, FR distributions are fairly similar in all substages (Figure S19, Table S32), though Upper Maastrichtian has a noticeably larger interquartile range of FR values than earlier substages. Th Upper Maastrichtian generic richness distribution, on the other hand, is more similar in its interquartile range to other substages, except the Middle Campanian where the interquartile range is noticeably larger. Generic richness outliers are present in all substages, and only the Middle Campanian and Upper Maastrichtian lack FR outliers. SME values present interquartile ranges with similar distances from the median value in each substage and outliers in all substages. SEM values either have outliers at

zero in all cases or their first interquartile hinge is positioned at zero. Again, zeros in this metric indicate "perfect evenness" which most commonly occurs when only a single FE exists in that node. Therefore, these values are not considered important and may be misleading.

Median FR values for each substage at the 60km node scale are fairly similar across all the substages, ranging from 3-4, but mean values are more variable, ranging from 3.79-10.11. This is most likely a result of outliers in the distribution. Generic richness mean and median values follow a similar pattern, where medians are fairly similar (3-6) and mean values are much more diverse (5.7-33.5). Again, mean values for generic richness are probably influenced by outliers in the distributions. Median values for SME show a clear shift from higher values in the Campanian, ranging from 0.83-0.99, to lower values in the Maastrichtian (0.37 and 0.38). This suggests a decrease in evenness between the Campanian and Maastrichtian. Mean values of SME follow a similar trend but are more variable and assumed to be more influenced by outlier values.

Proportional change in 60km node metrics indicate that **FR increases across the Lower to Middle Campanian, is stable across the Middle to Upper Campanian, decreases between the Upper Campanian and Lower Maastrichtian, and is fairly stable between the Lower to Upper Maastrichtian** (Figure 5, Table S33). Median FR proportional change values range from 0% between the Middle and Upper Campanian to 90% (increase) between the Lower to Middle Campanian. The upper and lower interquartile range values are farther from the median in both the Lower to Middle Campanian and Lower to Upper Maastrichtian comparisons than they are in the other two comparisons, and all contain numerous outlier values. All distributions cross the zero line, suggesting that not all node values shift in the same direction (increase or decrease). Due to outliers with extremely high proportional increases, mean values are consistently higher than median values of proportional change in FR.

Proportional change in generic richness follows a very similar pattern to FR, though median values indicate an increase in generic richness between the Middle and Upper Campanian rather than stability. Again, the two are similar but decoupled. In all cases richness distributions cross the zero line.

This suggests that not all generic or functional richness values within 60km nodes experience only increase or decrease across substages.

SME values indicate that evenness was stable from the Lower to Middle Campanian (median: -0%), decreased between the Middle and Upper Campanian (median: -21%), increased between the Upper Campanian and Lower Maastrichtian (median: 13%), and decreased between the Lower and Upper Maastrichtian (median: -27%) (Figure 5, Table S33). Mean values are frequently higher (more positive) than median values for this metric, influenced by outliers which have significant increases in evenness through time. All distributions for SME cross the zero line, indicating that evenness change is not unanimously increasing or decreasing in 60km nodes between substages. **Overall, SME values at this scale of analysis suggest first stable then decreasing evenness in the Campanian, then rising evenness in the Upper Campanian to Lower Maastrichtian, followed by evenness decrease in the Maastrichtian.**

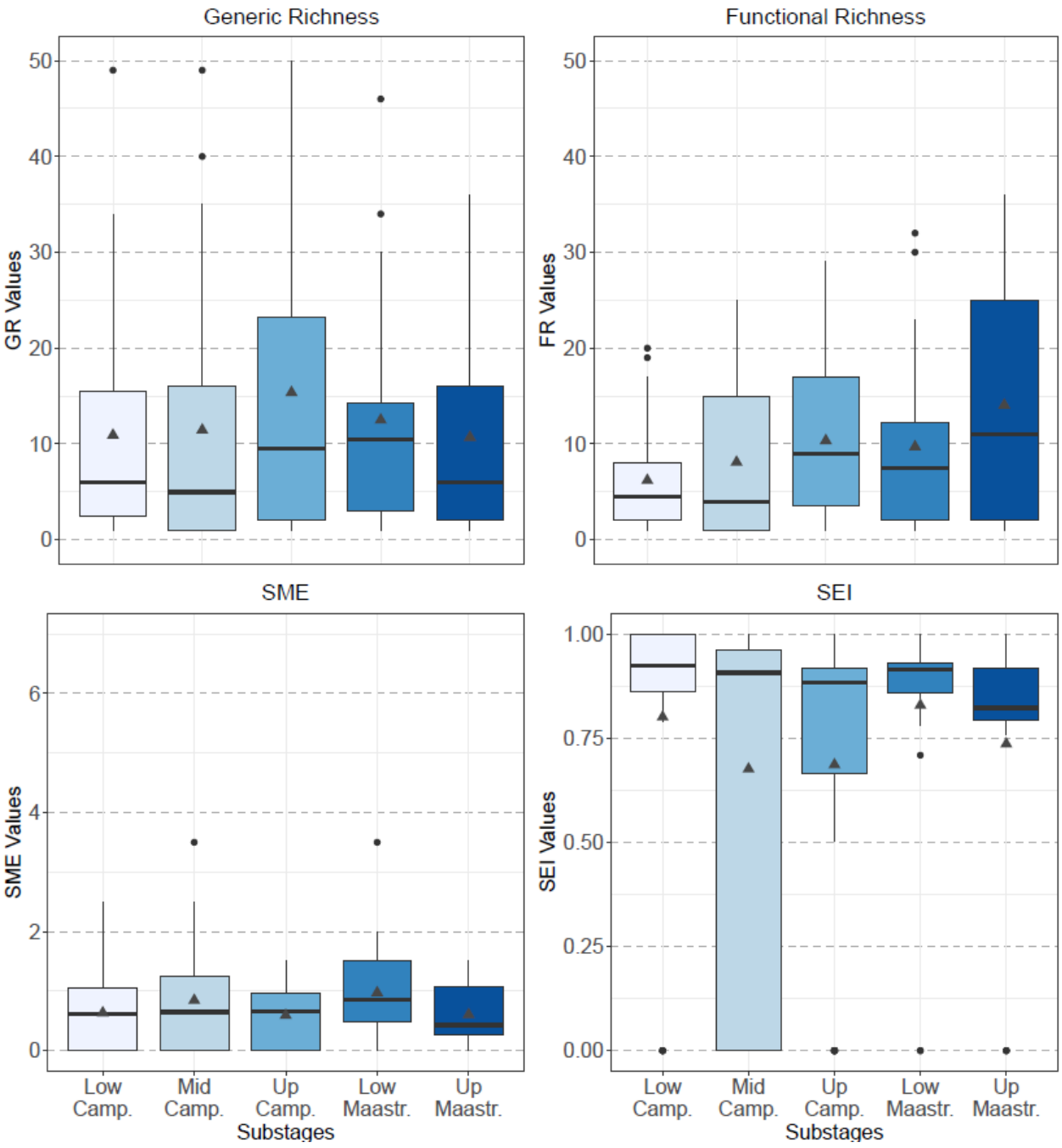



Figure S18. Box and whisker plots of generic richness, FR, SME, and SEI for 360km nodes in each substage. Median values are denoted by a horizontal line and means by triangles.

Table 29. Summary statistics of generic and functional richness, SME and SEI values from 360km nodes.

| | Lower Camp. | Middle Camp. | Upper Camp. | Lower Maastr. | Upper Maastr. |
|---|---|---|---|---|---|
| | **Generic Richness** | | | | |
| Mean | 12.39 | 19.03 | 31.39 | 31.08 | 56.85 |
| Median | 6.00 | 6.00 | 19.00 | 12.50 | 18.50 |
| SD | 14.10 | 26.38 | 36.85 | 49.58 | 73.74 |
| | **Functional Richness** | | | | |
| Mean | 6.21 | 8.10 | 10.35 | 9.71 | 14.05 |
| Median | 4.50 | 4.00 | 9.00 | 7.50 | 11.00 |
| SD | 5.55 | 8.08 | 8.19 | 9.12 | 12.71 |
| | **SME** | | | | |
| Mean | 0.64 | 0.84 | 0.59 | 0.97 | 0.60 |
| Median | 0.61 | 0.65 | 0.65 | 0.85 | 0.43 |
| SD | 0.70 | 0.94 | 0.46 | 0.80 | 0.48 |
| | **SEI** | | | | |
| Mean | 0.80 | 0.68 | 0.69 | 0.83 | 0.74 |
| Median | 0.92 | 0.91 | 0.88 | 0.92 | 0.82 |
| SD | 0.34 | 0.43 | 0.39 | 0.26 | 0.32 |

Table S30. Summary statistics of proportional change in generic and functional richness, SME and SEI values from 360km nodes.

| | | Low Camp. – Mid Camp. | Mid. Camp. – Up. Camp. | Up. Camp – Low Maastr. | Low Maastr. – Up. Maastr. |
|---|---|---|---|---|---|
| **Generic Richness** | Mean | 1.56 | 6.25 | 9.91 | 2.09 |
| | Median | 0.69 | 0.76 | -0.50 | -0.27 |
| | SD | 2.93 | 13.43 | 46.16 | 6.50 |
| **Functional Richness** | Mean | 1.14 | 2.08 | 1.29 | 0.98 |
| | Median | 0.32 | 0.28 | -0.39 | 0.02 |
| | SD | 2.15 | 3.90 | 7.01 | 3.25 |
| **SME** | Mean | -0.05 | -0.21 | 0.52 | 0.05 |
| | Median | -0.24 | -0.27 | 0.15 | -0.01 |
| | SD | 0.36 | 0.34 | 1.08 | 0.61 |
| **SEI** | Mean | 0.01 | -0.05 | 0.05 | 0.00 |
| | Median | 0.00 | -0.04 | 0.02 | 0.01 |
| | SD | 0.05 | 0.09 | 0.16 | 0.12 |

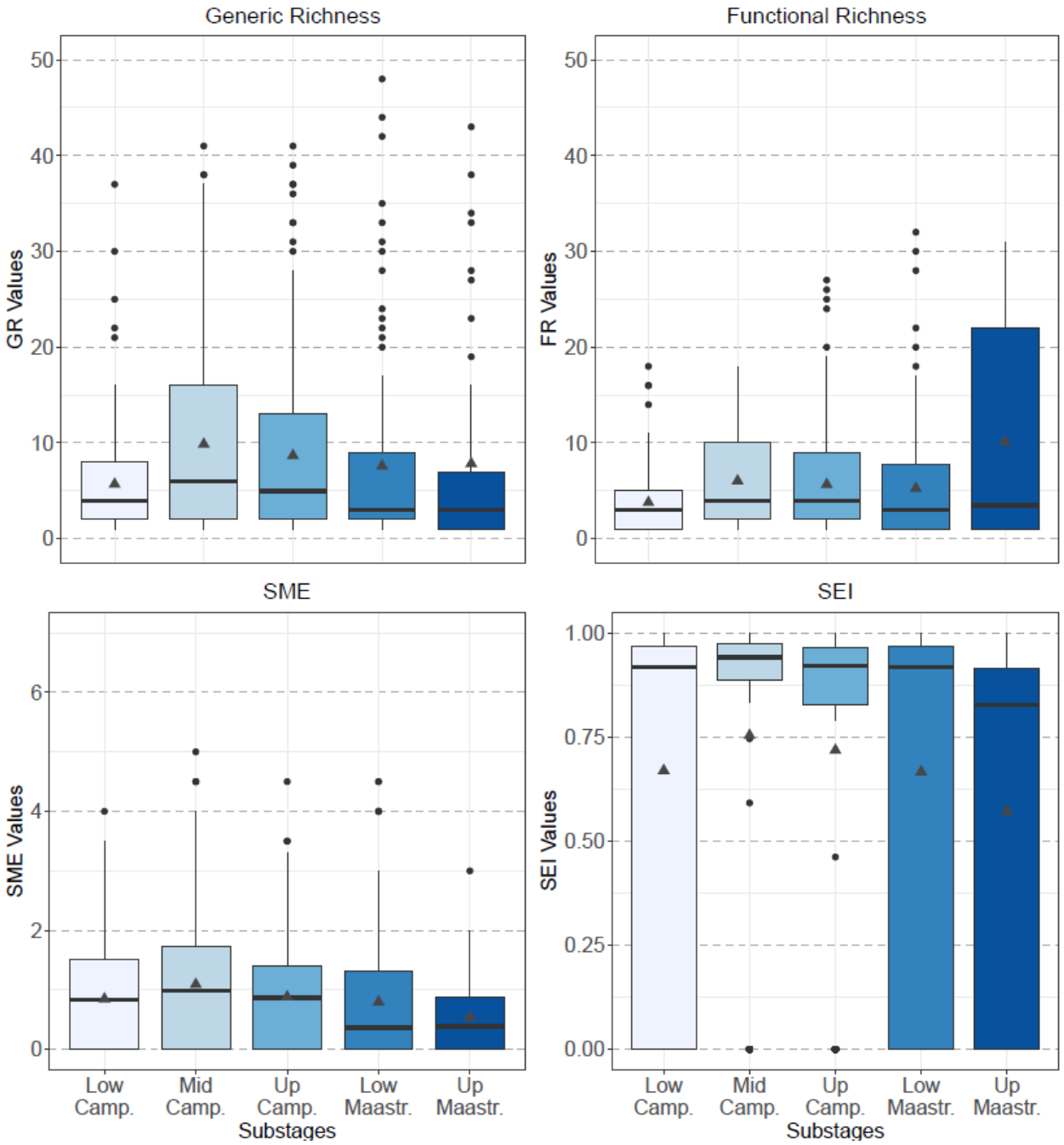



Figure S19. Box and whisker plots of generic richness, FR, SME, and SEI for 60km nodes in each substage. Median values are denoted by a horizontal line and means by triangles.

Table S31. Summary statistics of generic and functional richness, SME and SEI values from 60km nodes.

| | Lower Camp. | Middle Camp. | Upper Camp. | Lower Maastr. | Upper Maastr. |
|---|---|---|---|---|---|
| | **Generic Richness** | | | | |
| Mean | 5.70 | 9.86 | 10.85 | 12.39 | 33.05 |
| Median | 4.00 | 6.00 | 5.00 | 3.00 | 6.00 |
| SD | 5.88 | 10.11 | 15.68 | 26.56 | 44.50 |
| | **Functional Richness** | | | | |
| Mean | 3.79 | 6.04 | 5.65 | 5.26 | 10.11 |
| Median | 3.00 | 4.00 | 4.00 | 3.00 | 3.50 |
| SD | 3.28 | 4.94 | 5.20 | 6.10 | 10.35 |
| | **SME** | | | | |
| Mean | 0.85 | 1.09 | 0.88 | 0.79 | 0.54 |
| Median | 0.83 | 0.99 | 0.86 | 0.37 | 0.38 |
| SD | 0.93 | 1.14 | 0.94 | 1.05 | 0.63 |
| | **SEI** | | | | |
| Mean | 0.67 | 0.75 | 0.72 | 0.67 | 0.57 |
| Median | 0.92 | 0.94 | 0.92 | 0.92 | 0.83 |
| SD | 0.43 | 0.39 | 0.40 | 0.43 | 0.43 |

Table S32. Summary statistics of proportional change in generic and functional richness, SME and SEI values from 60km nodes.

| | | Low Camp. - Mid Camp. | Mid. Camp. - Up. Camp. | Up. Camp. - Low Maastr. | Low Maastr. - Up. Maastr. |
|---|---|---|---|---|---|
| **Generic Richness** | Mean | 3.79 | 1.18 | 1.39 | 8.54 |
| | Median | 1.25 | 0.12 | -0.38 | 0.00 |
| | SD | 6.58 | 2.49 | 10.92 | 28.09 |
| **Functional Richness** | Mean | 2.56 | 0.68 | 0.42 | 2.67 |
| | Median | 0.93 | 0.00 | -0.40 | 0.00 |
| | SD | 4.10 | 1.83 | 2.87 | 7.14 |
| **SME** | Mean | 0.16 | -0.11 | 0.37 | -0.32 |
| | Median | 0.00 | -0.21 | 0.13 | -0.27 |
| | SD | 0.74 | 0.50 | 0.95 | 0.38 |
| **SEI** | Mean | 0.00 | -0.02 | 0.02 | -0.08 |
| | Median | 0.00 | -0.02 | 0.02 | -0.07 |
| | SD | 0.07 | 0.07 | 0.07 | 0.09 |

## Summary of Functional vs Taxonomic Diversity change in 360km2 Nodes across Distance:

Proportional changes in metrics between nodes within a single substage, plotted by the geographic distance between the nodes indicate that change in richness values are not strongly correlated with distance for the Campanian substages (Figure S20-S23). In the Maastrichtian substages, increasing geographic distance is weakly correlated with increasing proportional change in both generic and functional richness (Figure S20-S21).

SME values similarly do not show strong correlations between proportional changes to evenness and geographic distance, though the in all substages except the Lower Campanian, a very week negative trend may exist, indicating that evenness values change less as distances increase (Figure S22). This negative trend is most likely a product of heteroscedasticity in the data, given that low values of proportional change are present across all distances but high values become increasing less common as distances increase. Given that most nodes are fairly closely spaced there will by necessity be fewer data points at very great distances to compare.

SEI values show even less of a correlation between proportional change between nodes and geographic distance (Figure S23). Negative trend lines observed in the figures are most likely a product of the highly heteroscedastic nature of the data and due to the decreasing number of data points at greater distances. The overall distribution of data points forms a fairly uniform cloud which has fewer values at greater distances, suggesting that, when baseline values of -1 are not taken into account, there is no strong correlation between proportional change in SEI and distance.

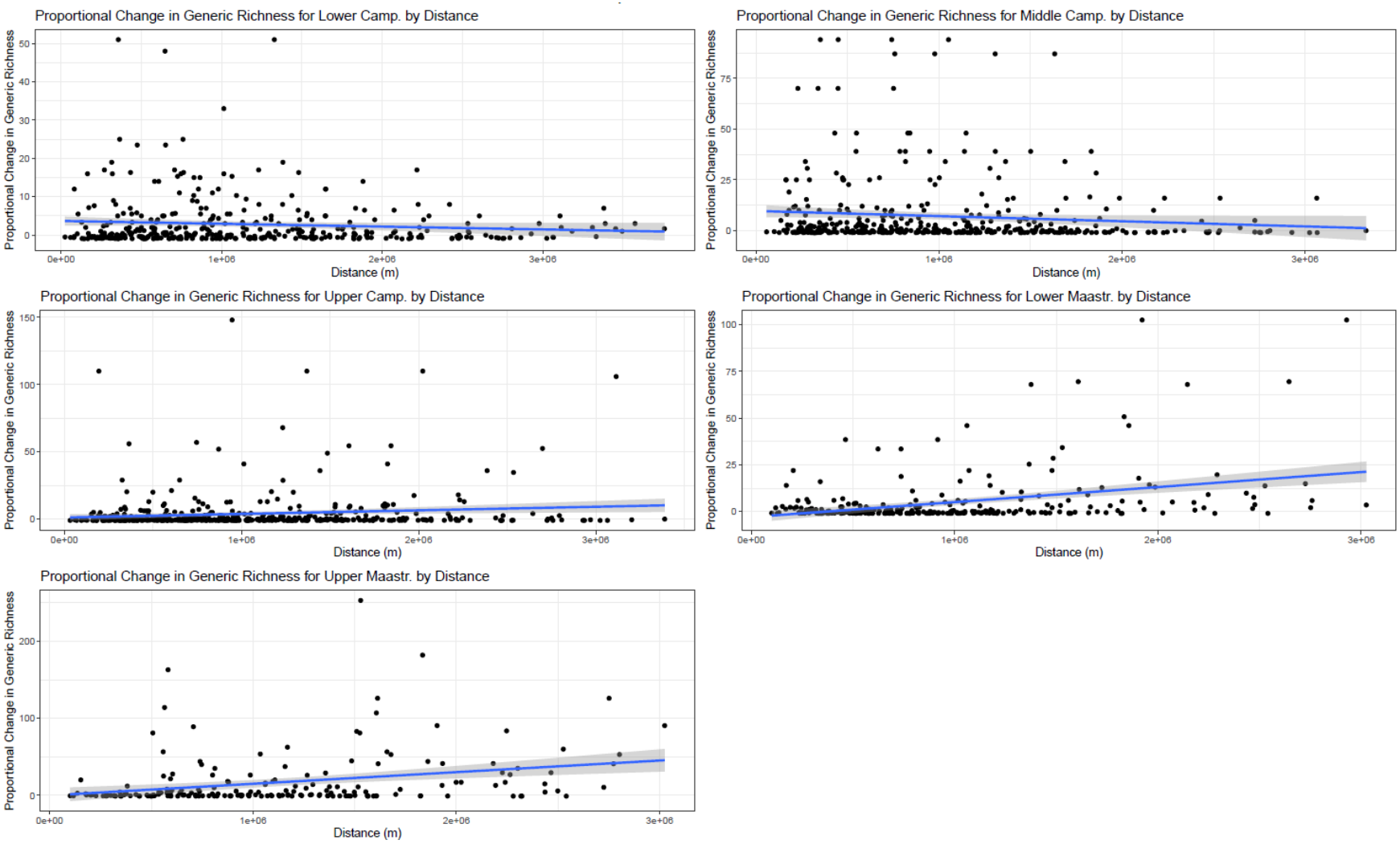



Figure S20. Proportional change in generic richness between all occupied 360km nodes in each substage plotted against distance (m). Blue lines indicate linear regression lines calculated for each.

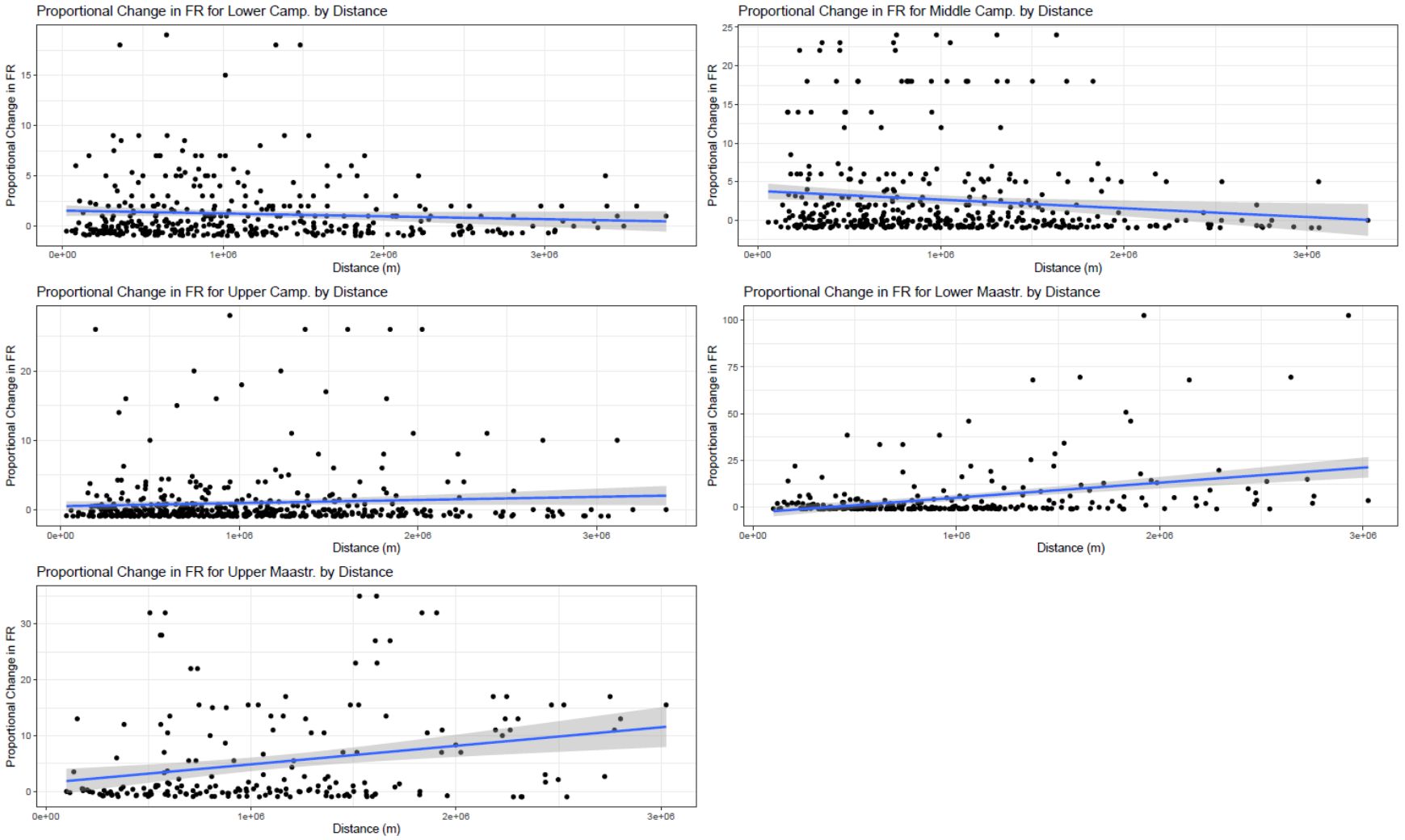



Figure S21. Proportional change in functional richness between all occupied 360km nodes in each substage plotted against distance (m). Blue lines indicate linear regression lines calculated for each.

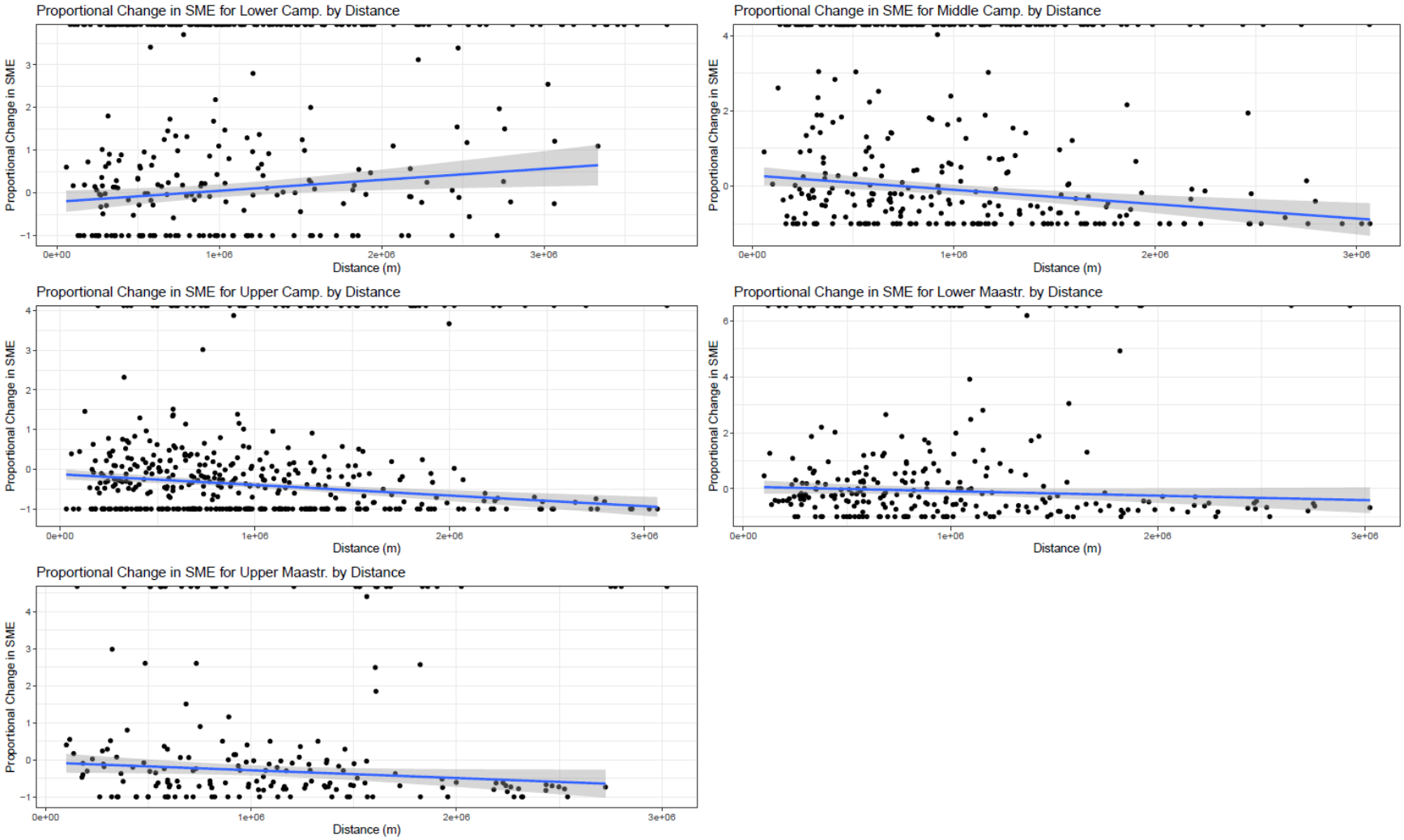



Figure S22. Proportional change in SME between all occupied 360km nodes in each substage plotted against distance (m). Blue lines indicate linear regression lines calculated for each.

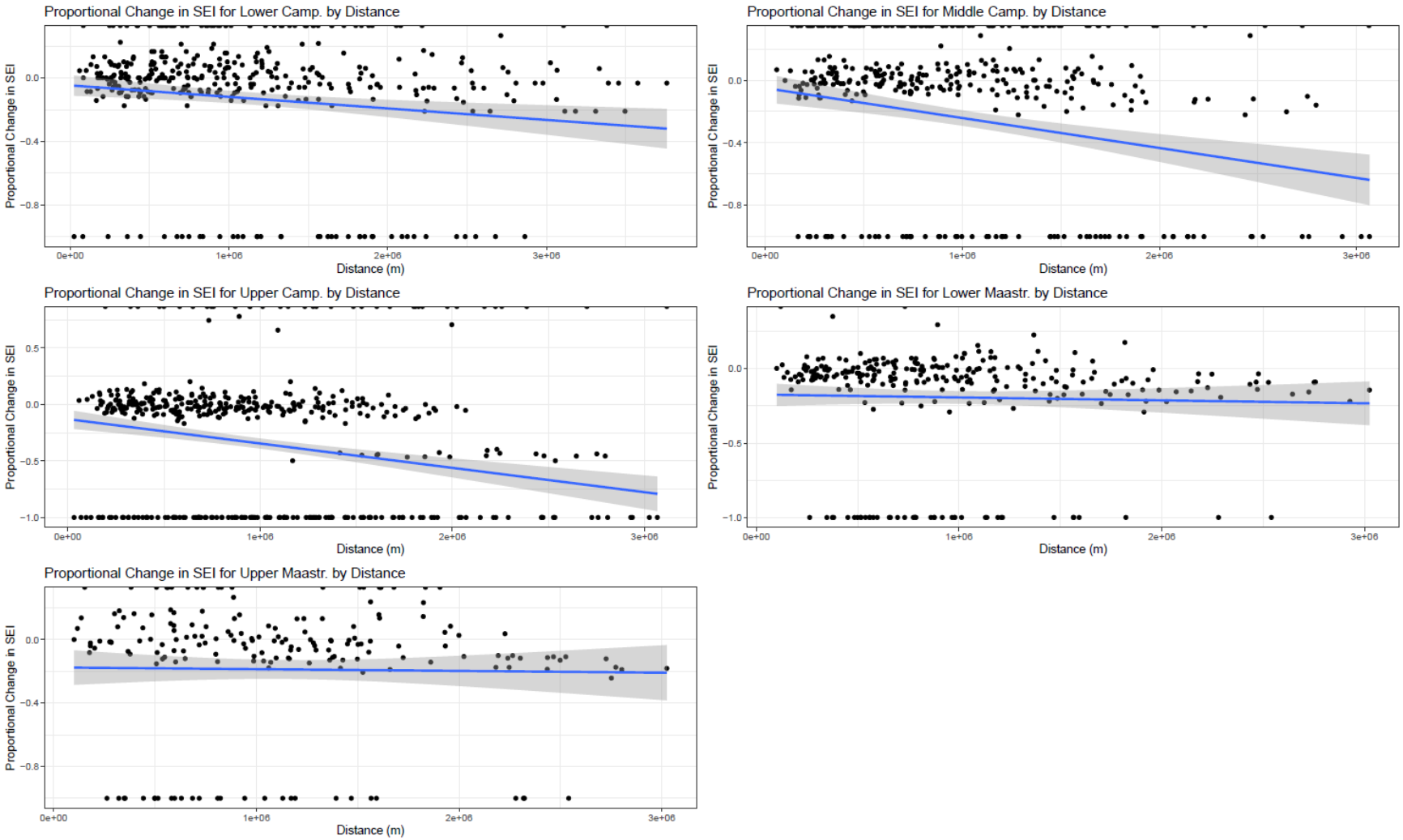



Figure S23. Proportional change in SEI between all occupied 360km nodes in each substage plotted against distance (m). Blue lines indicate linear regression lines calculated for each.

## Summary of Functional vs Taxonomic Diversity change in 60km2 Nodes across Distance:

Proportional changes in metrics between 60km nodes within a single substage, plotted by the geographic distance between the nodes indicate that change in richness values are not strongly correlated with distance for the Campanian substages (Figure S24-27). In the Maastrichtian substages, increasing geographic distance is weakly correlated with increasing proportional change in both generic and functional richness (Figure S24-S25).

Proportional changes in SME values for 60km nodes are very slightly negatively correlated with the geographic between nodes at this scale of analysis (Figure S26). SEI values show no correlation between proportional change across nodes and geographic distance between those nodes when values of -1 are not considered (Figure S27). Negative trend lines observed in the figures are most likely a product of the highly heteroscedastic nature of the data and due to the decreasing number of data points at greater distances. The overall distribution of data points forms a fairly uniform cloud which has fewer values at greater distances, suggesting that, when baseline values of -1 are not taken into account, there is no strong correlation between proportional change in SEI and distance. These patterns are extremely similar to those observed at the 360km node scale, excepting that a negative trend in the proportional change in SME with distance is dampened.

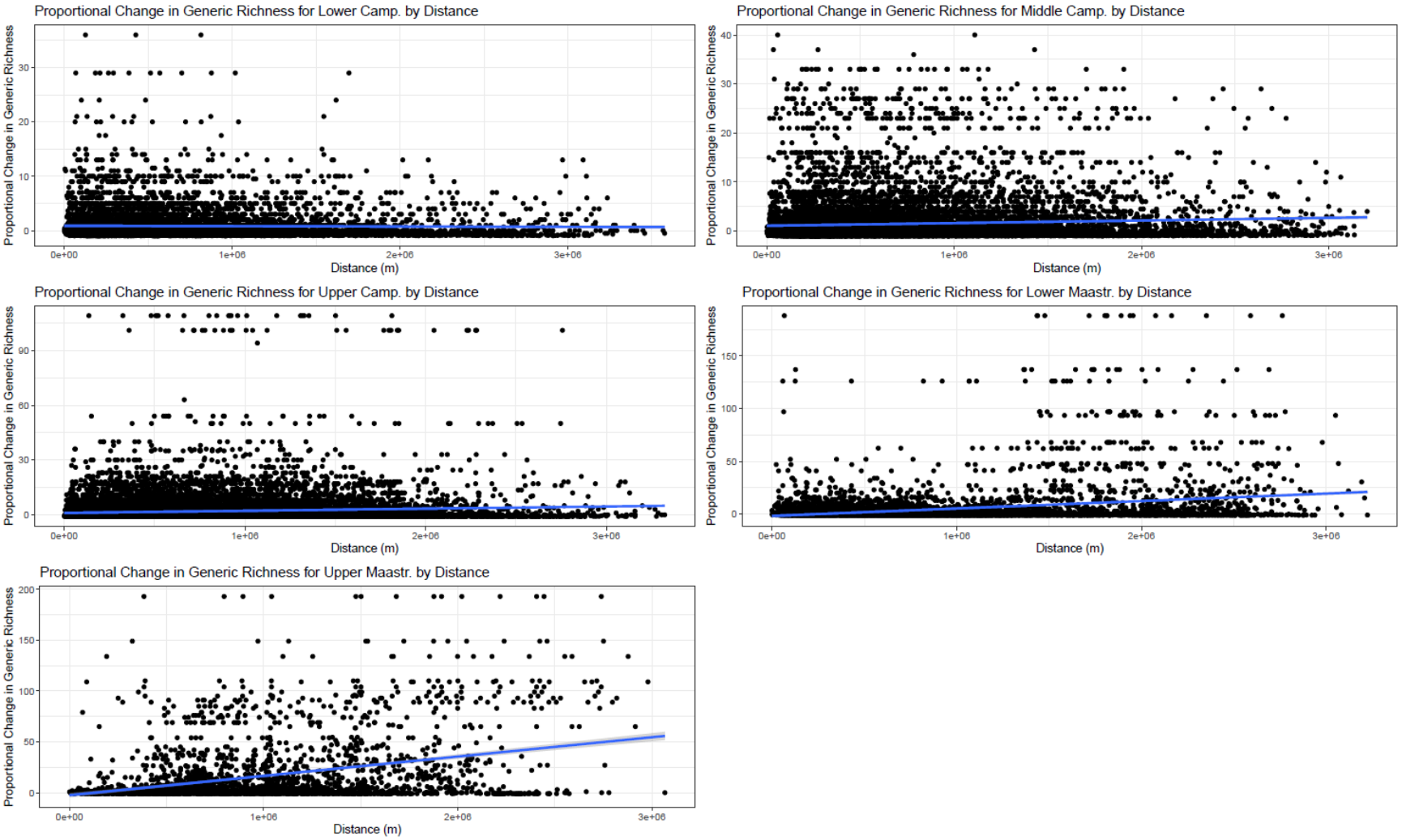



Figure S24. Proportional change in generic richness between all occupied 60km nodes in each substage plotted against distance (m). Blue lines indicate linear regression lines calculated for each.

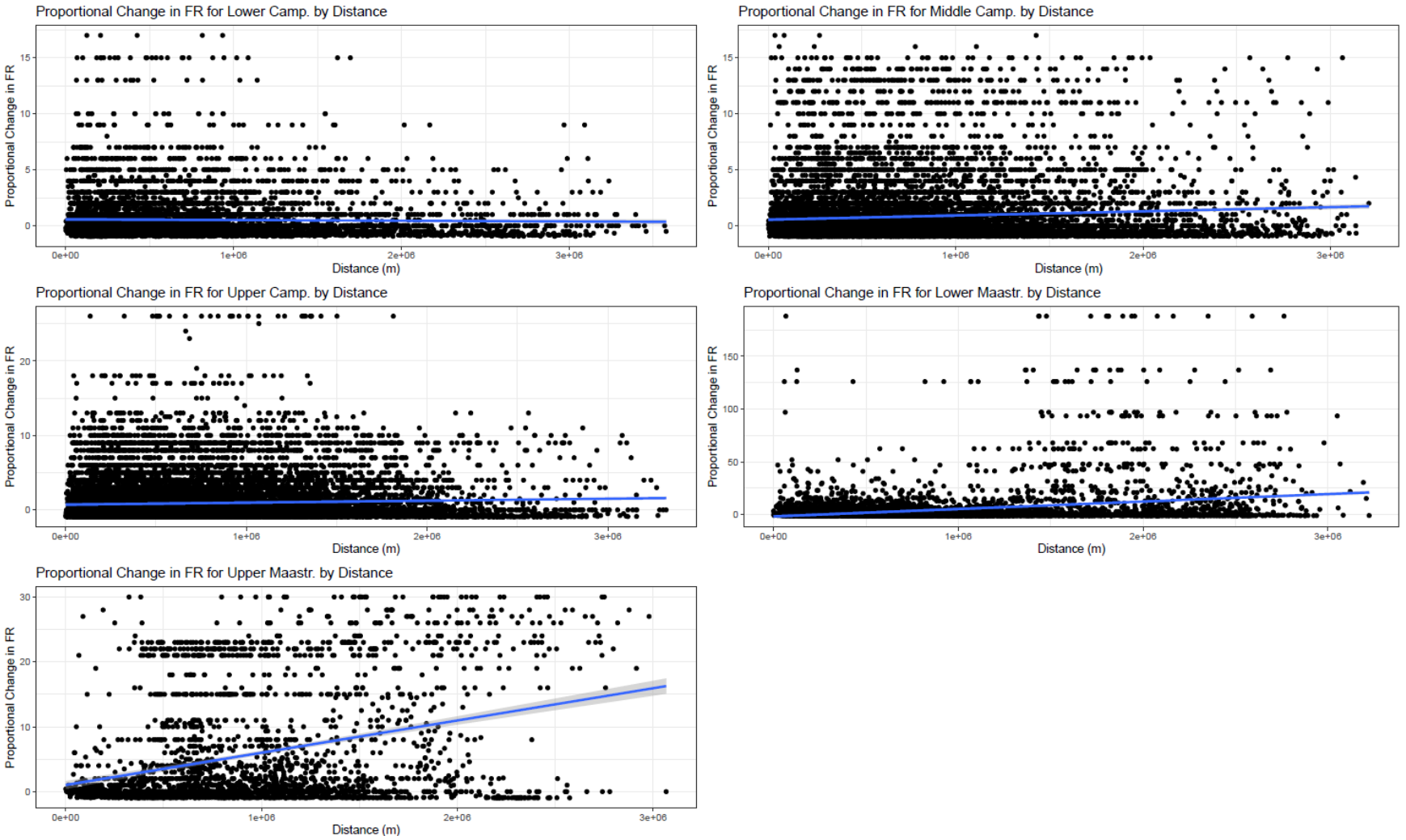



Figure S25. Proportional change in functional richness between all occupied 60km nodes in each substage plotted against distance (m). Blue lines indicate linear regression lines calculated for each.

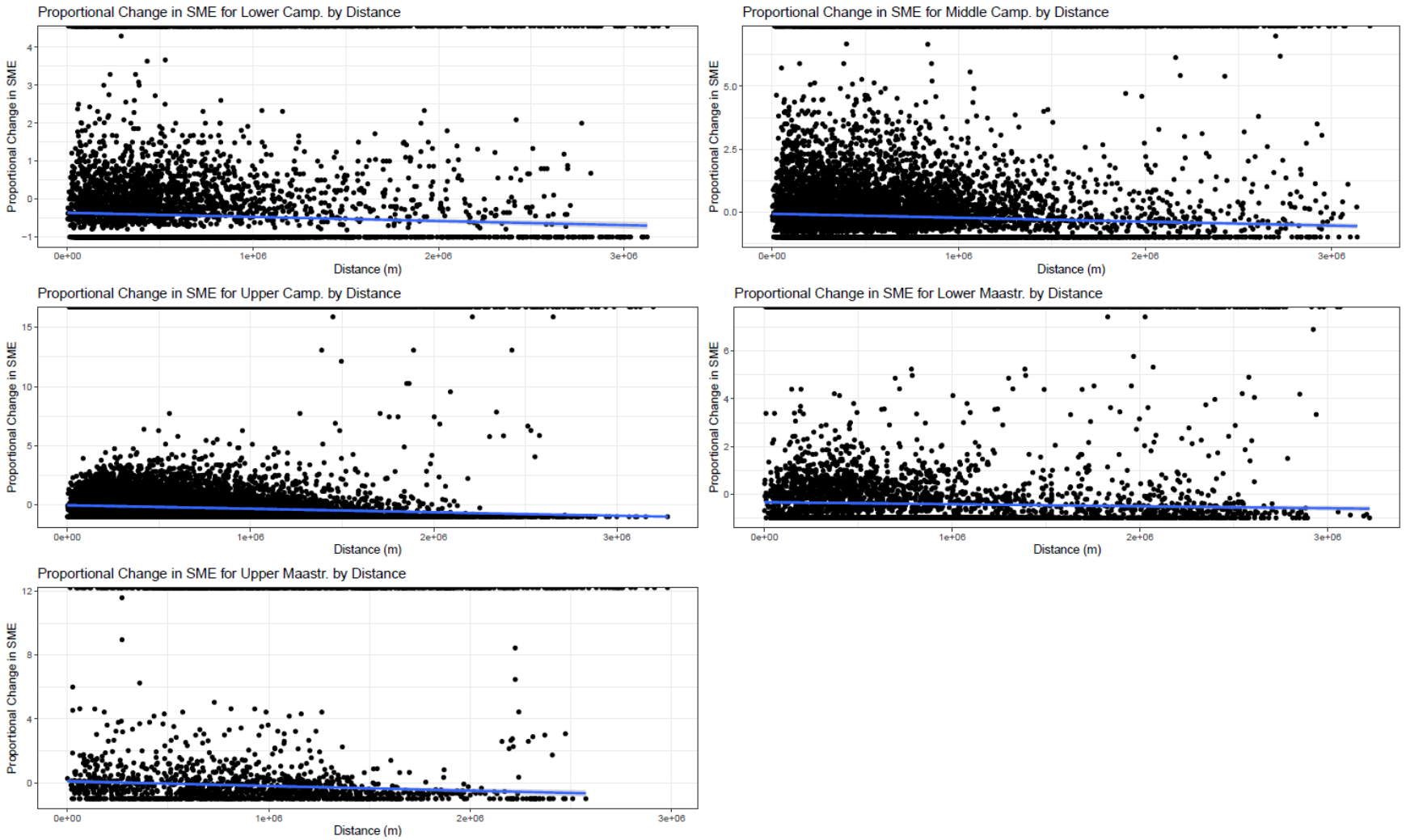



Figure S26. Proportional change in SME between all occupied 60km nodes in each substage plotted against distance (m). Blue lines indicate linear regression lines calculated for each.

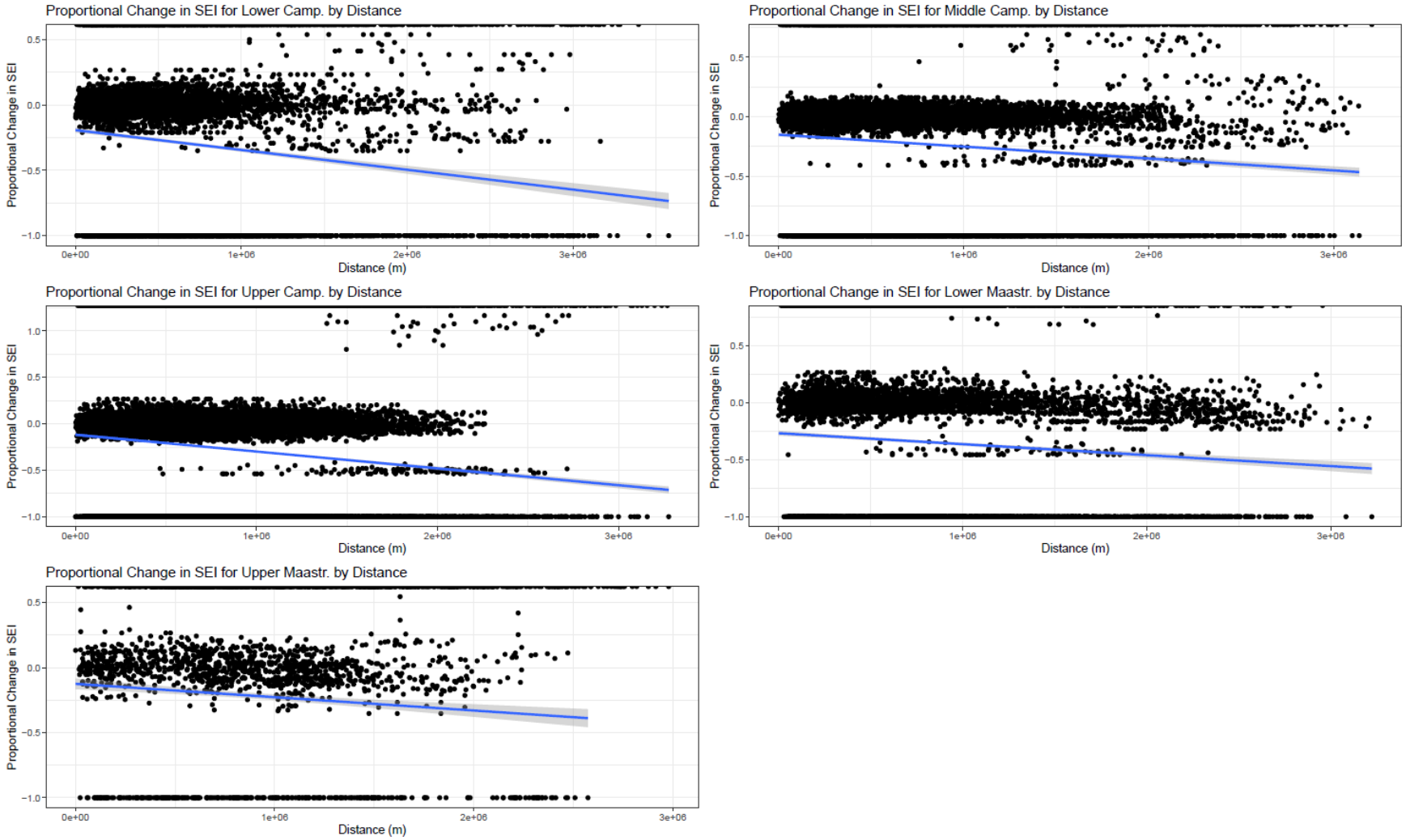



Figure S27. Proportional change in SEI between all occupied 60km nodes in each substage plotted against distance (m). Blue lines indicate linear regression lines calculated for each.

## Summary of 360km2 Functional Entity Networks:

When all substages are modeled in a single network, strong network connections exist across the entire study interval and no clear pattern of provinciality is observed (Figure S30). Two major components are observed in the network at percolation (TH=0.36), but they do not appear to represent any consistent geographic region, instead overlapping similarly to the components observed in the Middle Campanian. Furthermore, the geographically distinct secondary component in the Upper Maastrichtian GCP region shares strong node similarities with both WIS and GCP regions in previous substages, indicating that this is not a geographically consistent set of FE assemblages.

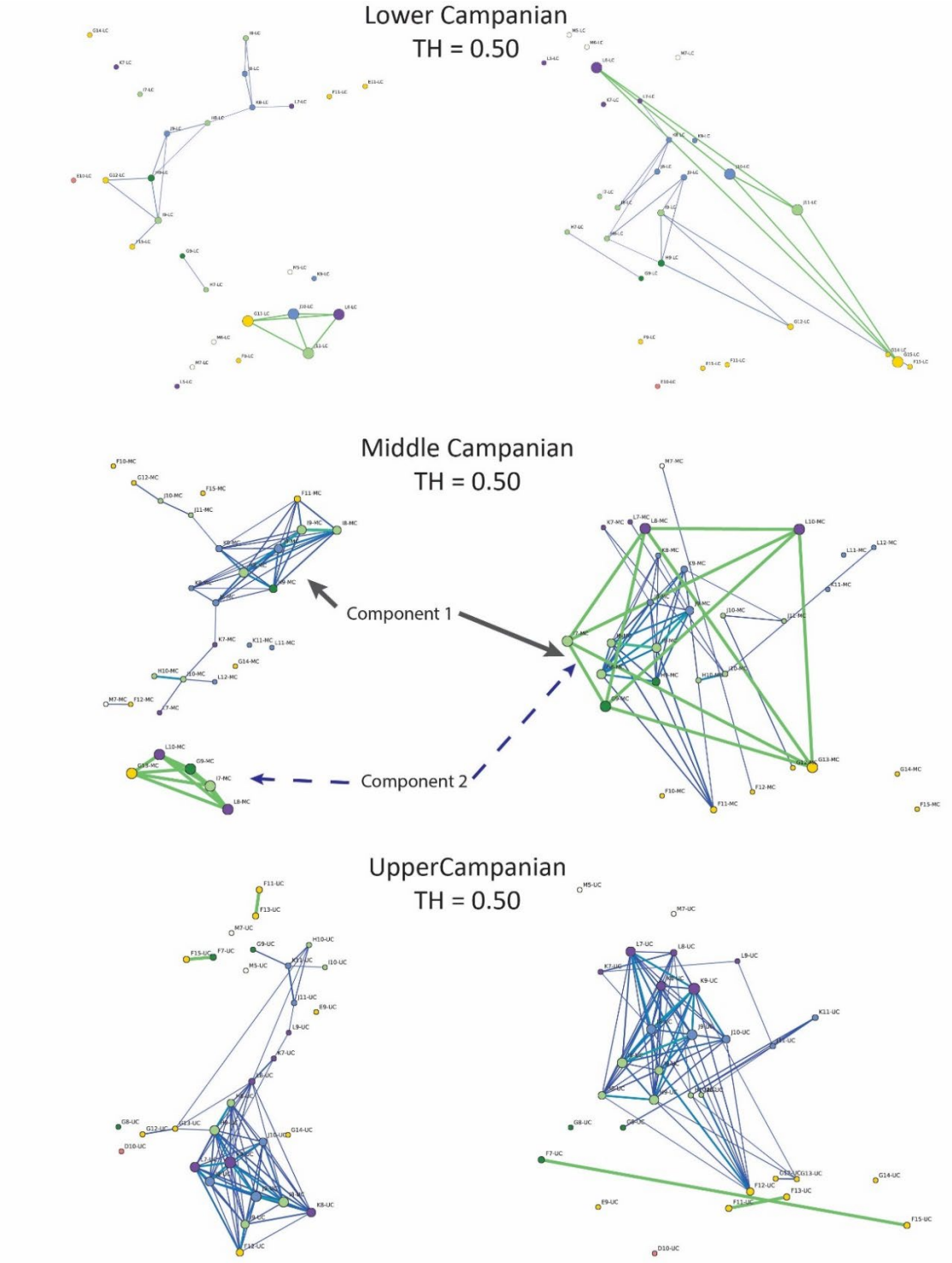



Figure S28. Network models of Campanian substages based on 360km nodes. All networks show connections present at percolation with the threshold level denoted below the substage name. Models on the left are not plotted based on any geographic position, models on the right are plotted based on paleo-latitude coordinates. Node colors indicate paleo-latitudes (warmer = southern paleo-latitudes and cooler = northern paleo-latitudes). Color and size of links between nodes indicates the level of similarity (thicker and lighter = more similar). Middle Campanian components are denoted with arrows.

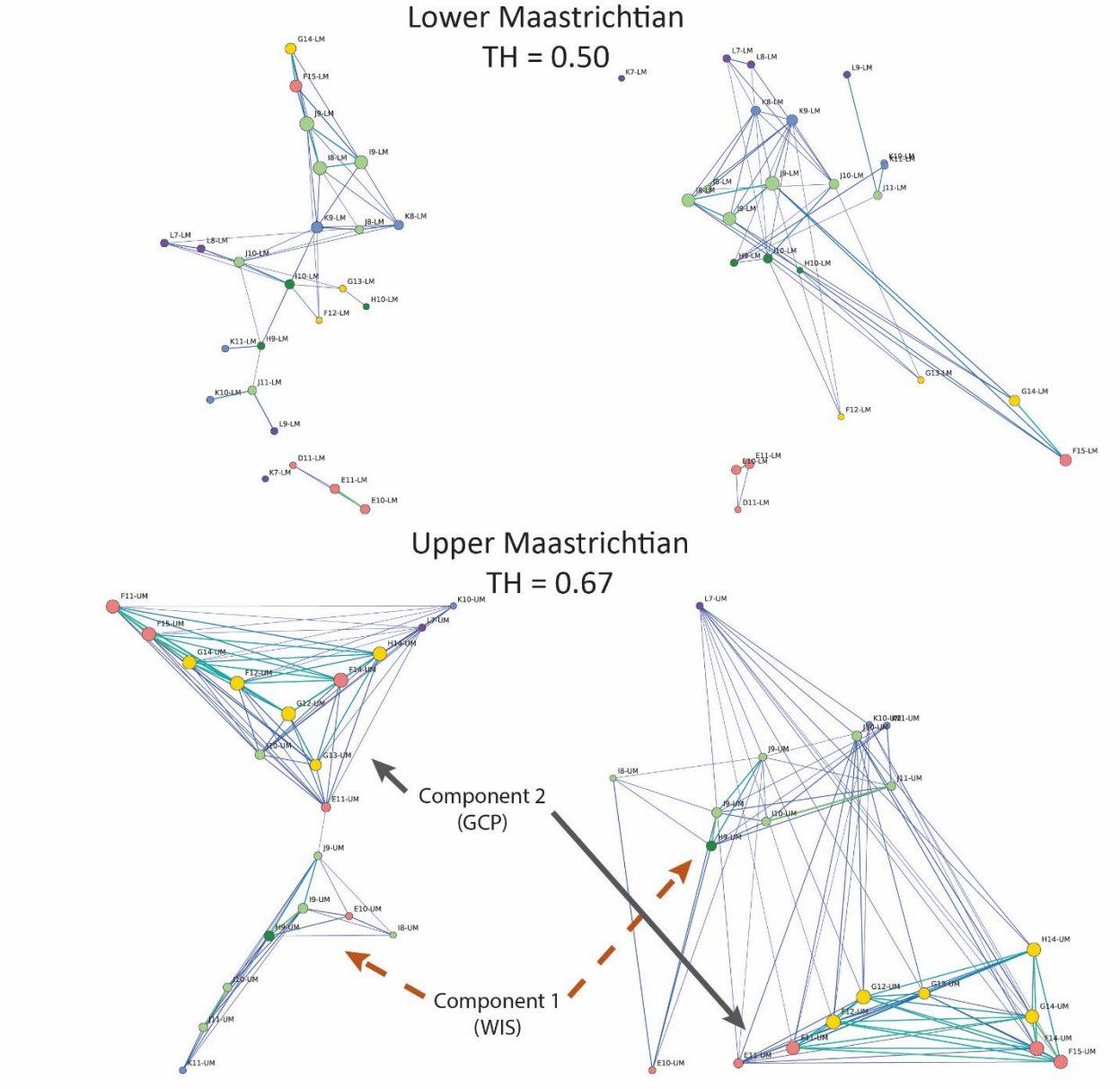



Figure S29. Network models of Maastrichtian substages based on 360km nodes. All networks show connections present at percolation with the threshold level denoted below the substage name. Models on the left are not plotted to show components and models on the right are plotted based on paleo-latitude coordinates. Node colors indicate paleo-latitudes (warmer = southern paleo-latitudes and cooler = northern paleo-latitudes). Color and size of links between nodes indicates the level of similarity (thicker and lighter = more similar). The two components present in the Upper Maastrichtian network are indicated with arrows.

Table S33. List of FE in each major network component for the Middle Campanian and Upper Maastrichtian based on 360km nodes. FEs which are shared by all nodes in the component are denoted with asterisks. FEs which are shared by all nodes in the component are denoted with asterisks. These components were separated based on manually thresholded networks created below percolation: TH=0.52 for the Middle Camp. and TH=0.75 for the Upper Maastr. FEs representative of gastropods are indicated with a gastropod symbol (🜏).

| **Middle Campanian** | | **Upper Maastrichtian** | |
|---|---|---|---|
| Primary component (WIS region) | Secondary Component (WIS region) | Primary component (WIS region) | Secondary Component (GCP region) |
| MU-E-C | MU-NB-SF* | MU-E-O 🜏 | MU-E-C 🜏 |
| MU-NB-C | | MU-N-SF | MU-I-D |
| MU-PI-C | | MU-NB-SF | FU-I-D |
| FU-I-C* | | MU-E-C 🜏 | FU-PI-H 🜏 |
| MU-I-D* | | FU-E-SF 🜏 | IA-B-SF |
| MU-PI-H | | | IU-B-SF |
| FU-I-D | | | FU-I-SF* |
| FU-I-SF* | | | FA-E-SF 🜏 |
| FA-E-SF | | | FU-E-SF 🜏 |
| FU-E-SF | | | IA-E-SF* |
| IA-E-SF* | | | IU-E-SF |
| MU-N-SF | | | MU-E-SF 🜏 |
| MU-NB-SF* | | | MU-N-SF |
| MU-I-SF | | | MU-NB-SF* |
| IA-I-SF | | | MU-I-SF |
| MU-N-C | | | IA-I-SF |
| FU-I-CH | | | MU-N-C |
| MU-I-CH | | | MU-NB-C |
| MU-E-O | | | MU-PI-C 🜏 |
| FA-I-SF | | | FU-I-C |
| IU-E-SF | | | MU-I-C 🜏 |
| MU-E-SF | | | FU-I-CH |
| FA-PI-SF | | | MU-I-CH |
| IA-PI-SF | | | FA-I-D |
| IU-I-SF | | | FU-E-H 🜏 |
| FU-PI-H | | | MU-E-H 🜏 |
| MU-PI-D | | | MU-PI-H 🜏 |
| MU-I-C | | | MU-E-O 🜏 |
| IU-B-SF | | | IU-PI-P |
| MU-E-H | | | MU-PI-D |
| | | | FU-B-SF |
| | | | FA-I-SF |
| | | | FA-PI-SF |
| | | | IA-E-P |
| | | | IA-PI-P |
| | | | IU-I-SF |
| | | | IA-PI-SF |
| | | | MU-E-D |

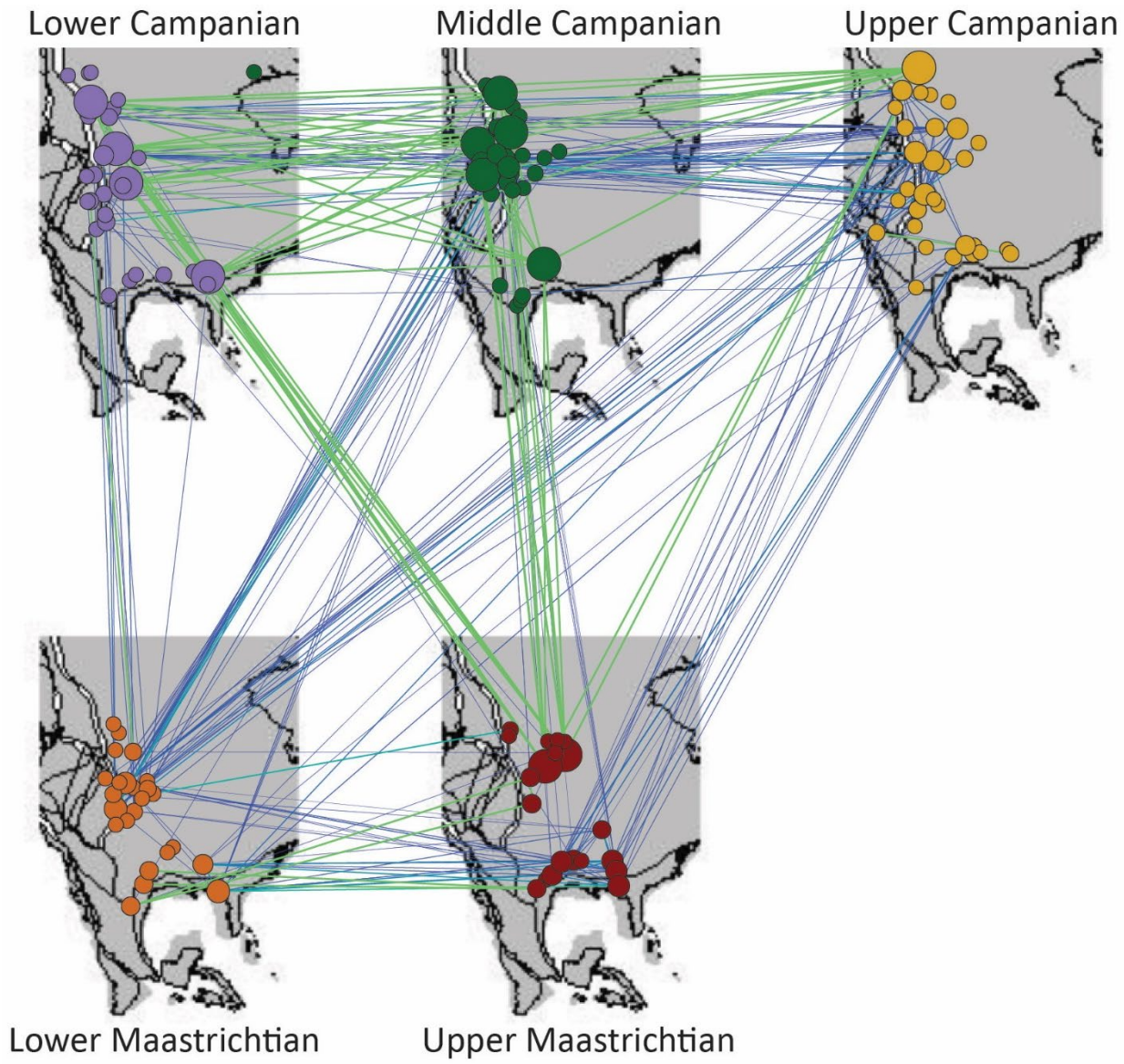



Figure S30. Network model of the entire database using 360km$^2$ nodes. Nodes are colored by substage and plotted on paleo-map reconstructions(Kocsis and Raja 2020) based on relative paleo-location (nodes may be offset to make visualization easier) Node size indicates betweenness centrality and link color and size indicates the degree of functional similarity (lighter and thicker = more similar).

Summary of 60km2 Functional Entity Networks:

In the Lower Maastrichtian a small highly similar group of nodes (which only disconnect from other nodes at thresholds below percolation) geographically overlies the more extensive but less similar primary component, mostly in the WIS region (Figure 8). The larger, less similar component is dominated by functionally diverse nodes with highly variable FEs, none of which are shared by all nodes in the component, while the highly similar component is dominated by functionally depauperate nodes with a single shared FE of mobile, unattached, nekto-benthic suspension feeders (cephalopods; Table S34). The third component in the Lower Maastrichtian that is geographically located in the eastern GCP region is dominated by only mobile, unattached, nektonic, suspension feeders (cephalopods). Therefore, in this Lower Maastrichtian network, high FE assemblage similarity is predominately controlled by a disparity between FE diversity in different nodes, which can most likely be attributed to sampling bias. Since generic richness strongly correlates with both sampling and FR values, it is likely that the separation of components in this substage is a product of sampling disparity between nodes.

In the Upper Maastrichtian, where functional similarity has a clear geographical signal, sampling bias may be similarly influencing network topology. In this substage, WIS region sampling is not as strong as GCP region sampling, which may explain the functionally less diverse WIS component that does not share any FEs between all nodes and the highly diverse GCP component whose nodes all share five total FEs. The Upper Maastrichtian components are characterized by a functional depauperate WIS assemblage is dominated by mobile, unattached, nekto-benthic and nektonic suspension feeders (cephalopods) and facultatively mobile, epifaunal suspension feeders (gastropods/bivalves) and a highly functionally diverse GCP assemblage.

Table S34. List of FE in each major network component for the Lower and Upper Maastrichtian based on 60km nodes. FEs which are shared by all nodes in the component are denoted with asterisks. These components were separated based on manually thresholded networks created below percolation: TH=0.45 for the Lower Maastr and TH=0.50 for the Upper Maastr. In the Upper Maastrichtian, FEs representative of gastropods are indicated with a gastropod symbol (🐌).

| **Lower Maastrichtian** | | | **Upper Maastrichtian** | |
|---|---|---|---|---|
| Component 1 (WIS region) | Component 2 (WIS region) | Component 3 (Eastern GCP region) | Primary component (WIS region) | Secondary Component (GCP region) |
| MU-I-D | MU-NB-SF* | MU-N-SF* | MU-NB-SF | FA-E-SF* 🐌 |
| FU-I-D | | | MU-N-SF | FA-I-D |
| FA-I-SF | | | FU-E-SF 🐌 | FA-I-SF |
| FU-I-SF | | | | FA-PI-SF |
| IU-I-SF | | | | FU-B-SF |
| FU-E-SF | | | | FU-E-H 🐌 |
| IA-E-SF | | | | FU-E-SF* 🐌 |
| MU-NB-SF | | | | FU-I-C |
| MU-E-C | | | | FU-I-CH |
| MU-PI-C | | | | FU-I-D |
| FU-I-CH | | | | FU-I-SF* |
| MU-PI-H | | | | FU-PI-H 🐌 |
| IU-B-SF | | | | IA-B-SF |
| FA-E-SF | | | | IA-E-P |
| FA-PI-SF | | | | IA-E-SF* |
| FU-I-C | | | | IA-I-SF |
| FA-I-D | | | | IA-PI-P |
| MU-E-H | | | | IA-PI-SF |
| FU-PI-H | | | | IU-B-SF |
| MU-PI-D | | | | IU-E-SF |
| IA-B-SF | | | | IU-I-SF |
| MU-E-SF | | | | IU-PI-P |
| MU-N-SF | | | | MU-E-C* 🐌 |
| MU-I-SF | | | | MU-E-D |
| IA-I-SF | | | | MU-E-H 🐌 |
| MU-N-C | | | | MU-E-O 🐌 |
| MU-NB-C | | | | MU-E-SF 🐌 |
| IU-E-SF | | | | MU-I-C 🐌 |
| MU-E-O | | | | MU-I-CH |
| MU-I-CH | | | | MU-I-D |
| MU-I-C | | | | MU-I-SF |
| IA-PI-P | | | | MU-NB-C |
| MU-E-D | | | | MU-NB-SF |
| IA-PI-SF | | | | MU-N-C |
| FU-E-H | | | | MU-N-SF |
| | | | | MU-PI-C 🐌 |
| | | | | MU-PI-D |
| | | | | MU-PI-H 🐌 |

Table S35. List of paleo-environments provided by the PBDB for select occurrences found in network components for the Lower and Upper Maastrichtian (see Table S34). Environmental descriptions which include carbonaceous attributes are in bold. In the Upper Maastrichtian components, a potential difference between siliciclastic dominated and carbonaceous-dominated environments may be present, but more data is necessary to determine this pattern than is possible given the scope of this project.

| **Lower Maastrichtian** | | **Upper Maastrichtian** | |
|---|---|---|---|
| Component 1 (WIS region) | Component 2 (WIS region) | Primary component | Secondary Component |
| Offshore, siliciclastic | Siliciclastic shale | Shale | Argillaceous sandstone |
| Sandstone | Siliciclastic sediments | Sandstone | Siltstone |
| Calcareous carbonaceous SS | **Offshore limestone** | Sandstone | **Sandy chalk** |
| Gray Sandstone | Siltstone | Shale | Silty sandstone |
| Grey Sandstone | | Siliciclastic silty mud/sandstone | **Chalk** |
| Grey Sandstone | | Sandstone/silty shale | **Chalk** |
| Sandstone | | Sandstone | Grey Sandstone |
| Concretionary sideritic calcareous shale | | | |
| Concretionary shale | | | |
| **Calcareous shale and grey limestone (offshore)** | | | |

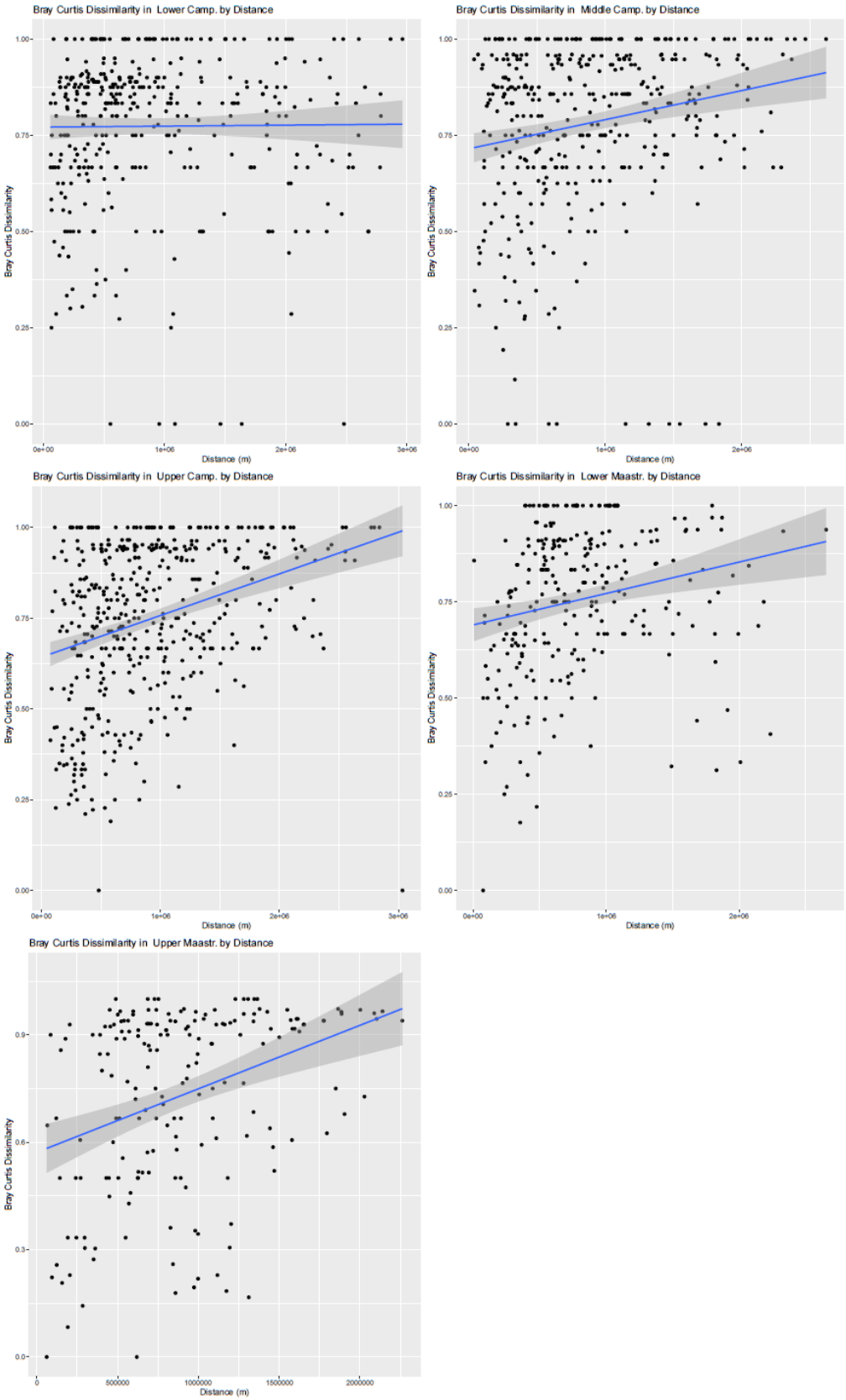



Figure S31. Bray-Curtis dissimilarity values between 360km nodes plotted by distance.

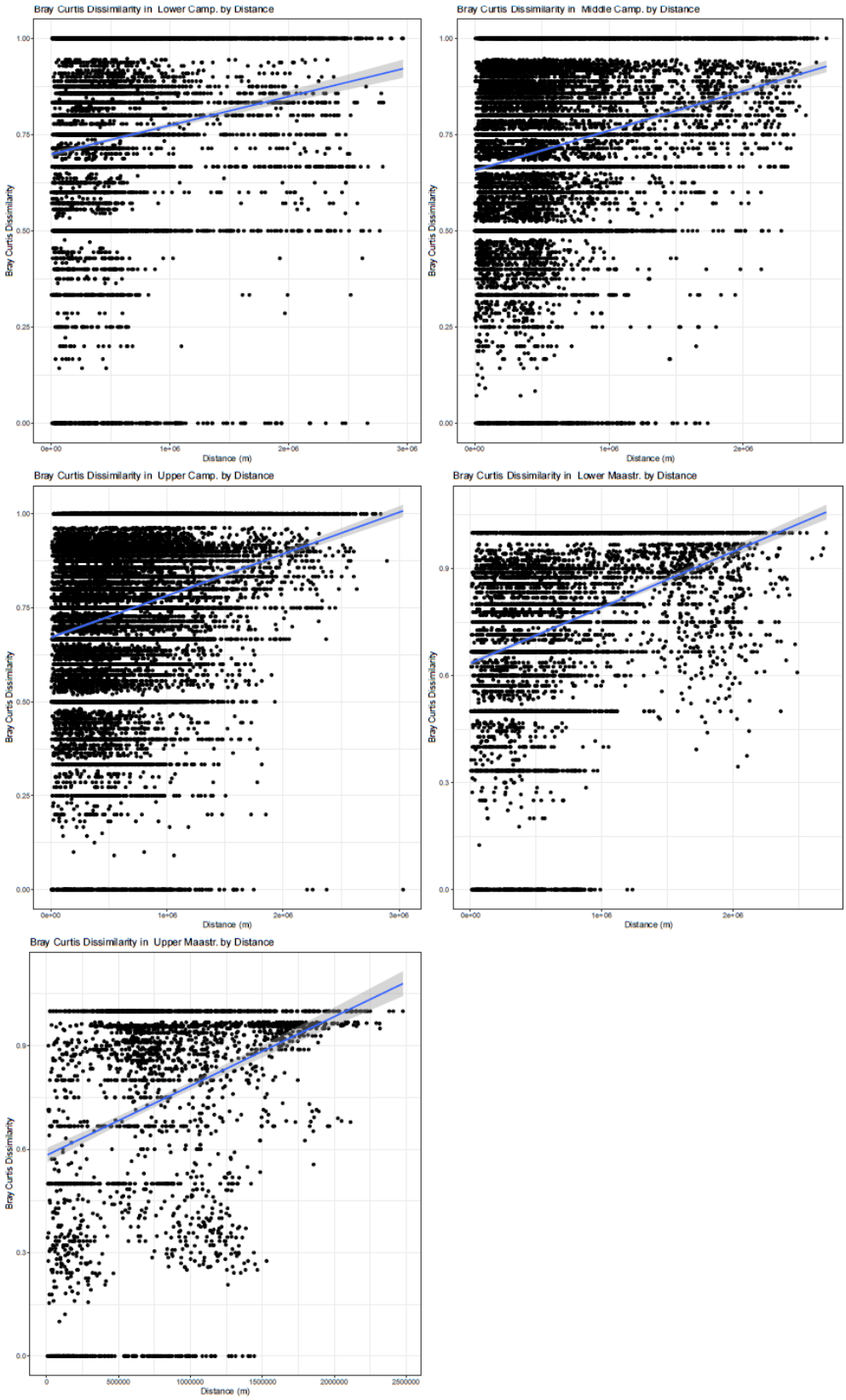



Figure S32. Bray-Curtis dissimilarity values between 60km nodes plotted by distance.

## Summary of 60km2 Class-level Functional Entity Networks:

Class-level analysis of 60km$^2$ networks indicates that bivalves and cephalopods, which make up the majority of FEs in the database, form strong network connections across all regions and lack a geographic or provincial signal (Figure S33-S36). However, it should be noted that Late Maastrichtian bivalves are not common and form poor network connections in general, though not necessarily based on geography (Figure S36). The strongly connected, functionally depauperate components in the Lower Maastrichtian are represented by specific, shared cephalopod FEs (Table S4). Their sampling is not strongly provincial, and therefore most likely produces the lack of provinciality in most substage networks where they form the dominate taxa. Gastropods, on the other hand, are strongly provincial (Figure S37-S38). Indeed, the three of the five FEs which are shared by all nodes in the GCP component of the Upper Maastrichtian are represented by gastropods and eleven of the 38 FEs possible in this region are represented by gastropods. Gastropods, therefore, are considered to have the prevailing influence on provinciality in FE assemblages in the Upper Maastrichtian.

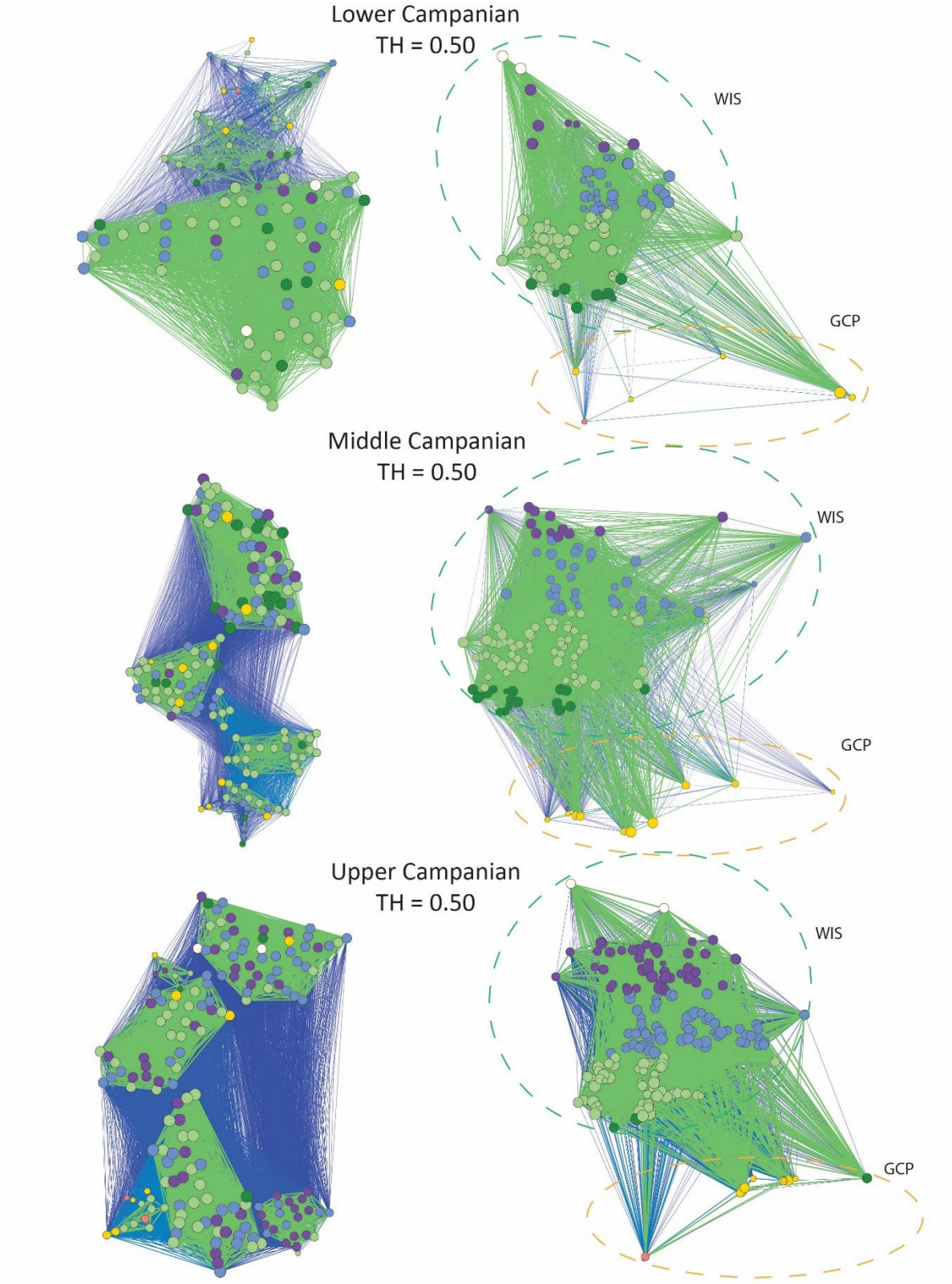



Figure S33. Campanian cephalopod substage networks plotted without any geographic coordinates (left) and based on paleo-coordinates (right). Node colors indicate paleo-latitude where warmer colors indicate southern paleo-latitudes and cooler colors indicate northern paleo-latitude.

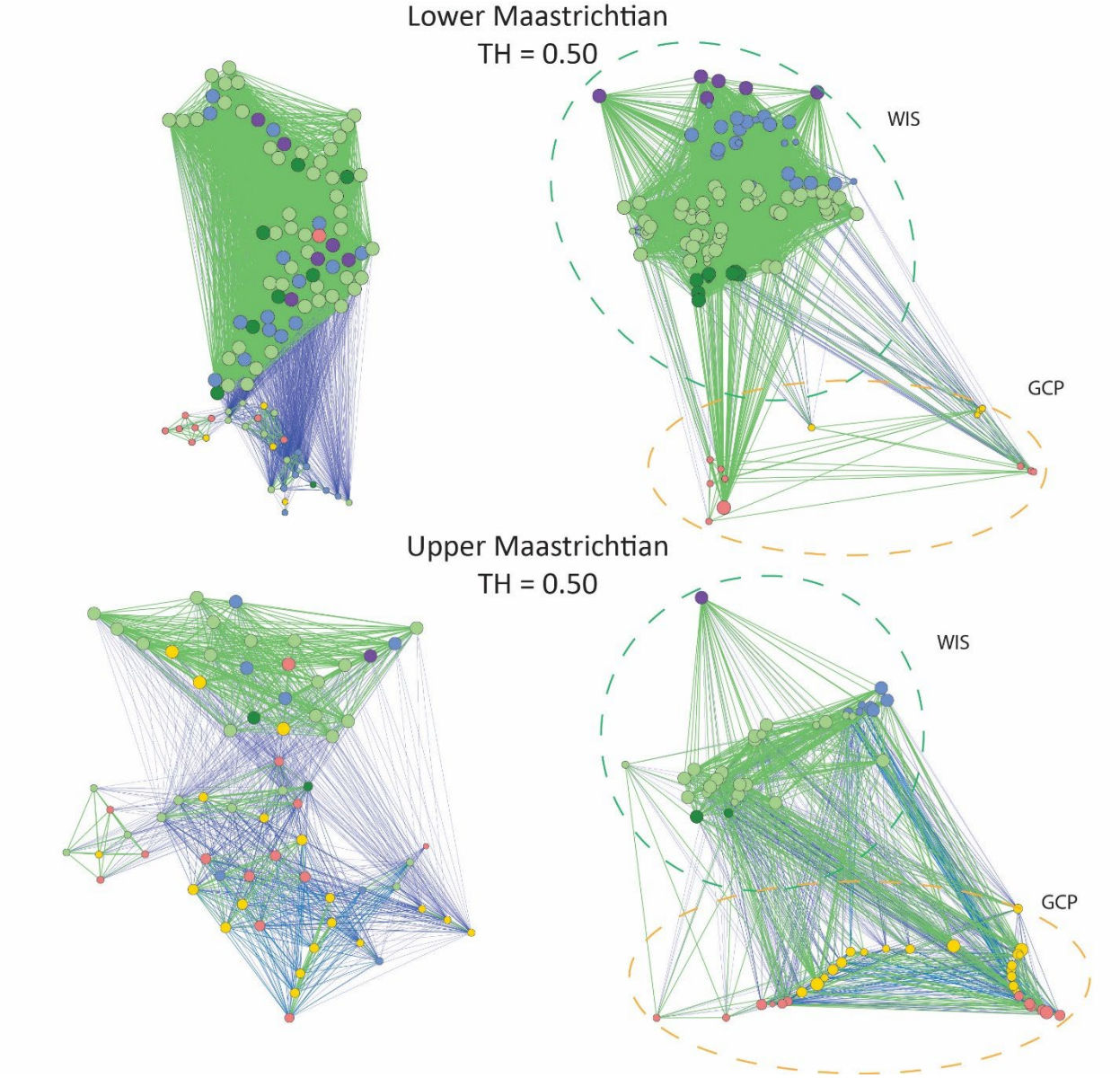



Figure S34. Maastrichtian cephalopods substage networks plotted without any geographic coordinates (left) and based on paleo-coordinates (right). Node colors indicate paleo-latitude where warmer colors indicate southern paleo-latitudes and cooler colors indicate northern paleo-latitude.

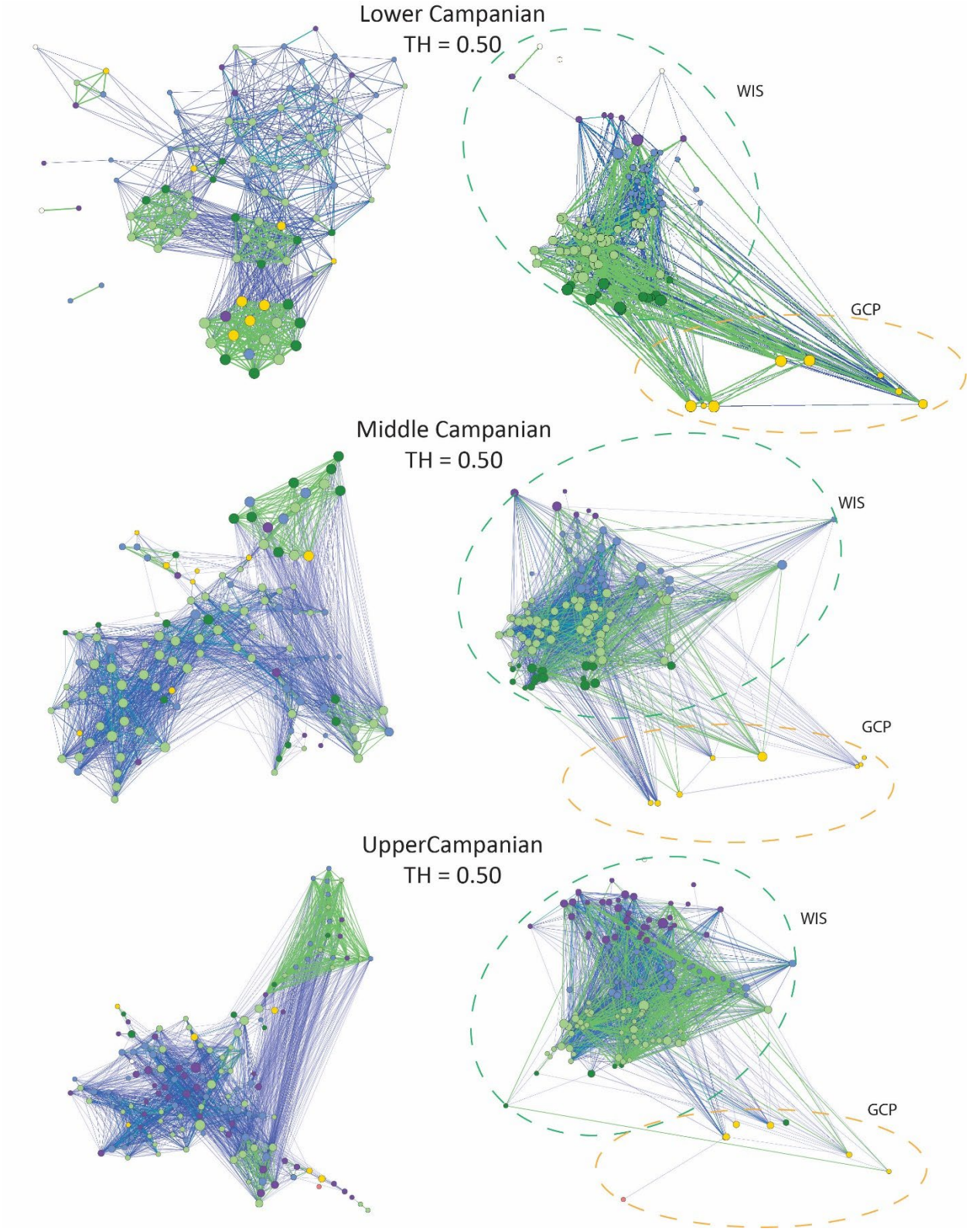



Figure S35. Campanian bivalve substage networks plotted without any geographic coordinates (left) and based on paleo-coordinates (right). Node colors indicate paleo-latitude where warmer colors indicate southern paleo-latitudes and cooler colors indicate northern paleo-latitude.

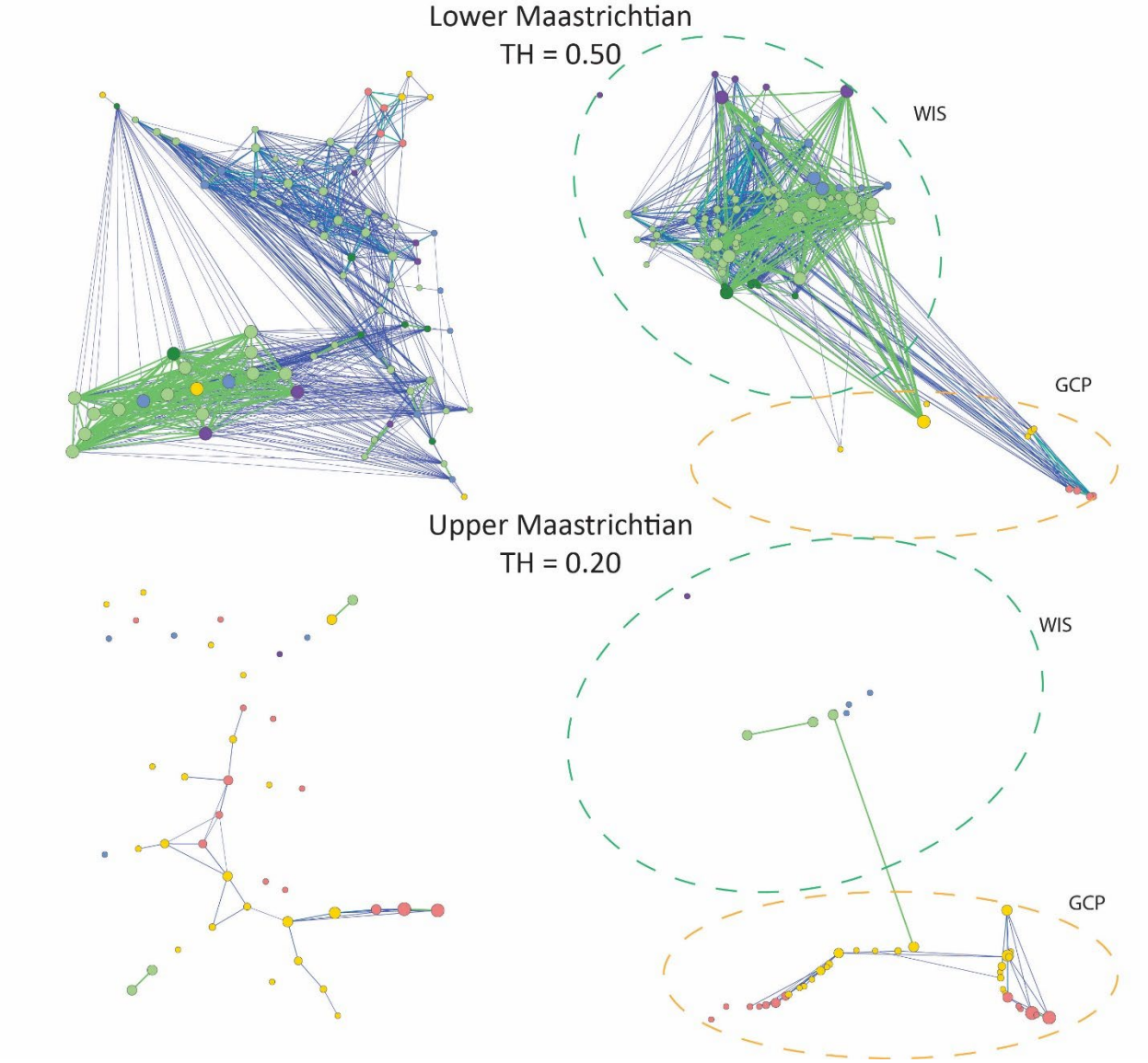



Figure S36. Maastrichtian bivalve substage networks plotted without any geographic coordinates (left) and based on paleo-coordinates (right). Node colors indicate paleo-latitude where warmer colors indicate southern paleo-latitudes and cooler colors indicate northern paleo-latitude.

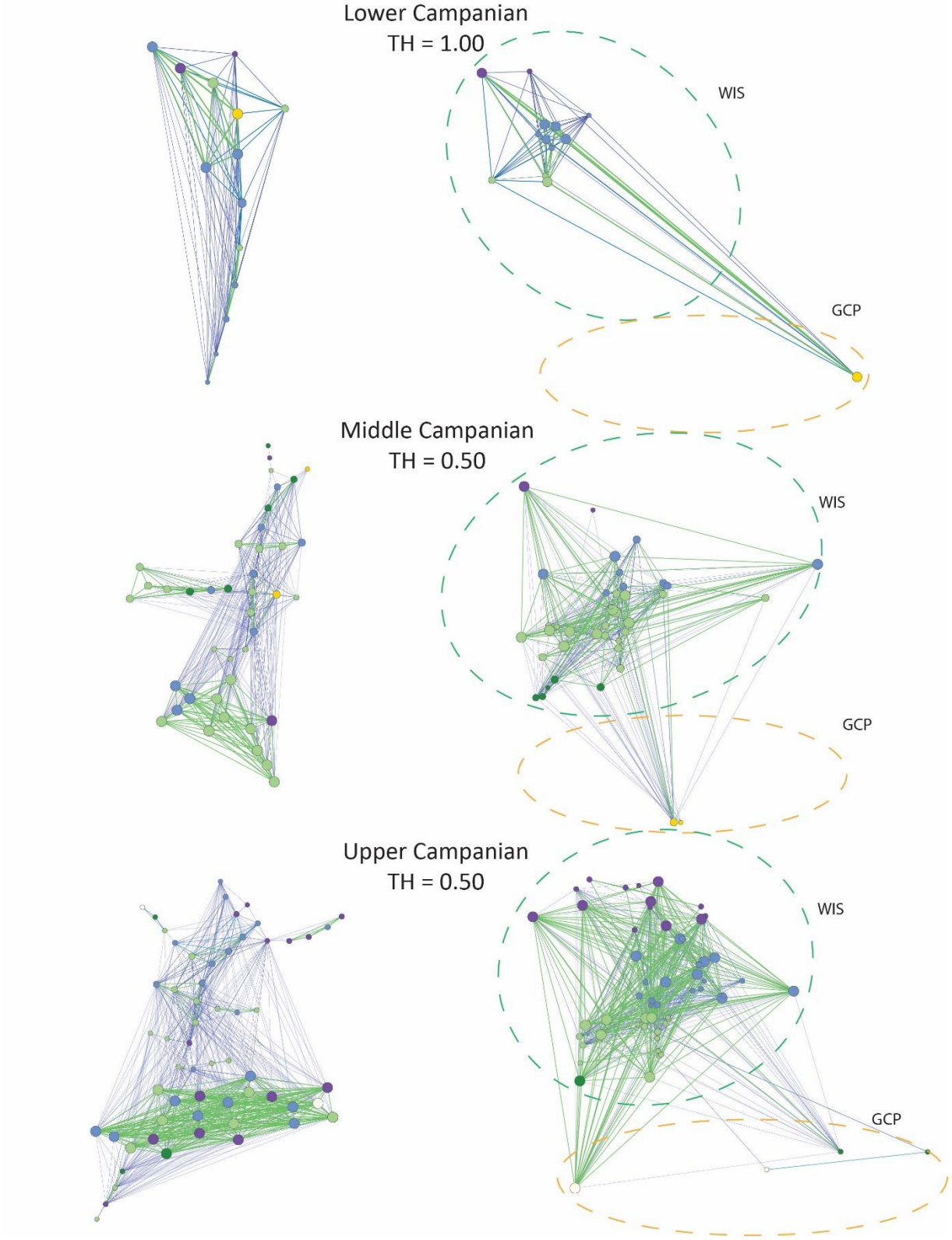



Figure S37. Campanian gastropod substage networks plotted without any geographic coordinates (left) and based on paleo-coordinates (right). Node colors indicate paleo-latitude where warmer colors indicate southern paleo-latitudes and cooler colors indicate northern paleo-latitude.

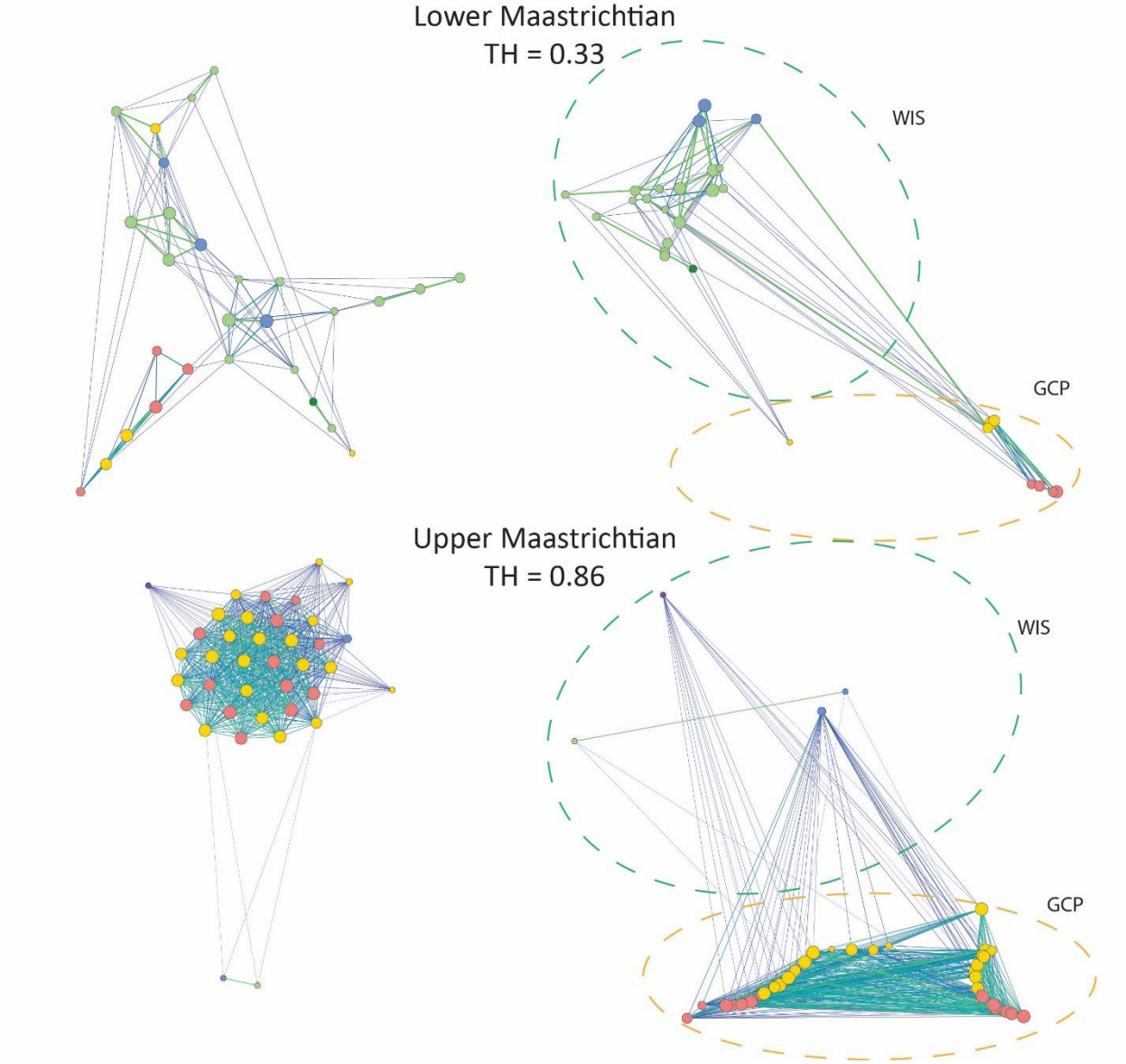



Figure S38. Maastrichtian gastropod substage networks plotted without any geographic coordinates (left) and based on paleo-coordinates (right). Node colors indicate paleo-latitude where warmer colors indicate southern paleo-latitudes and cooler colors indicate northern paleo-latitude.

## References

Alroy, J. (2010). Geographical, environmental and intrinsic biotic controls on Phanerozoic marine diversification. Palaeontology, 53(6), 1211–1235. https://doi.org/10.1111/j.1475-4983.2010.01011.x

Baselga, A., & Orme, C. D. L. (2012). Betapart: An R package for the study of beta diversity. Methods in Ecology and Evolution, 3(5), 808–812. https://doi.org/10.1111/j.2041-210X.2012.00224.x

Kocsis, Á. T., & Raja, N. B. (2020). chronosphere: Earth system history variables.

Magurran, A. E. (2003). Measuring Biological Diversity. Blackwell Science Ltd.

Schumm, M., Edie, S. M., Collins, K. S., Gómez-Bahamón, V., Supriya, K., White, A. E., Price, T. D., & Jablonski, D. (2019). Common latitudinal gradients in functional richness and functional evenness across marine and terrestrial systems. Proceedings of the Royal Society B: Biological Sciences, 286(1908). https://doi.org/10.1098/rspb.2019.0745

Appendix B-2. Functional Ecology References for Table S36

Aberhan, M., 1994, Guild-structure and evolution of Mesozoic benthic shelf communities: Palaios, v. 9, p. 516–545, doi:10.2307/3515126.

Aberhan, M., 1992, Palokologie und zeitliche Verbreitung benthischer Faunengemeinschaften im Unterjura von Chile: Beringeria, v. 5, p. 1–174.

Aberhan, M., Alroy, J., Fursich, F.T., Kiessling, W., Kosnik, M., Madin, J., Patzkowsky, M., and Wagner, P., 2004, Ecological attributes of marine invertebrates: unpublished,.

Aberhan, M., and Kiessling, W., 2015, Persistent ecological shifts in marine molluscan assemblages across the end-Cretaceous mass extinction: Proceedings of the National Academy of Sciences of the United States of America, v. 112, p. 7207–7212, doi:10.1073/pnas.1422248112.

Anderson, L.A., 2014, Relationships of Internal Shell Features to Chemosymbiosis, Life Position, and Geometric Constraints Within the Lucinidae (Bivalvia), in Hembree, D., Platt, B., and Smith, J. eds., Experimental Approaches to Understanding Fossil Organisms. Topics in Geobiology, Dordrecht, Springer, v. 41.

Andrade, C., and Brey, T., 2014, Trophic ecology of limpets among rocky intertidal in Bahia Laredo, Strait of Magellan (Chile): Anales del Instituto de la Patagonia, v. 42, p. 65–70, doi:10.4067/s0718-686x2014000200006.

Ansell, A.D., 1978, On the rate of growth of Nuculana minuta (Muller)(Bivalvia: Nuculanidae): Journal of Molluscan Studies, v. 44, p. 71–82.

Ansell, A.D., and Morton, B., 1987, Alternative predation strategies of a tropical naticid gastropod: Journal of Experimental Marine Biology and Ecology, v. 111, p. 109–119.

Ansell, A.D., and Morton, B., 1985, Aspects of naticid predation in Hong Kong with special reference to the defensive adaptations of Bassina (Callanaitis) calophylla (Bivalvia), in Moron, B. and Dudgeon, D. eds., Proceedings of the Second International Workshop on the Malacofauna of Hong Kong and Southern China, Hong Kong, 1983, Hong Kong, Hong Kong University Press.

Avelar, W., and Cunha, A.D., 2009, The anatomy and functional morphology of Diplodon rhombeus fontainianus (Orbigny, 1835) (Mollusca Bivalvia, Hyriidae).:

Bambach, R.K., Bush, A.M., and Erwin, D.H., 2007, Autecology and the filling of ecospace: Key metazoan radiations: Palaeontology, v. 50, p. 1–22, doi:10.1111/j.1475-4983.2006.00611.x.

Berg, C.J.Jr., 1975, Behavior and ecology of conch (Superfamily Strombacea) on a deep subtidal algal plain: Bulletin of Marine Science, v. 25, p. 307–317.

Berke, S.K., and Woodin, S.A., 2009, Behavioral and Morphological Aspects of Decorating in Oregonia gracilis (Brachyura: Majoidea):, https://about.jstor.org/terms.

Beu, A.G., and Maxwell, P.A., 1987a, A revision of the fossil and living gastropods related to Plesiotriton Fischer, 1884 (Family Cancellariidae, Subfamily Plesiotritoninae n. subfam.). With an appendix:

Genera of Buccinidae Pisaniinae related to Colubraria Schumacher, 1817: New Zealand Geological Survey Paleontological Bulletin, v. 54, p. 1–140.

Bishop, G., Feldmann, R.M., and Vega, F., 1998, The Dakoticancridae (Decapoda, Brachyura) from the Late Cretaceous of North America and Mexico: Contributions to Zoology, v. 67, p. 237–255.

Blake, D.B., 1990, Adaptive zones of the Class Asteroidea (Echinodermata): Bulletin of Marine Science-Miami, v. 46, p. 701–718.

Boyd, D.W., and Newell, N.D., 2002, A unique pterioid bivalve from the Early Triassic of Utah: American Museum Novitates, v. 3375, p. 1–9.

Bryan, J.R., and Jones, D.S., 1989, Fabric of the cretaceous-tertiary marine macrofaunal transition at Braggs, Alabama: Palaeogeography, Palaeoclimatology, Palaeoecology, v. 69, p. 279–301.

Burn, R., and Thompson, T.E., 1998, Order Cephalaspidea, in Beesley, P.L., Ross, G.J.B., and Wells, A. eds., Mollusca: The Southern Synthesis. Fauna of Australia, Melbourne, CSIRO Publishing, v. 5, p. 943–959.

Carter, J.G., 1978, Ecology and evolution of the Gastrochaenacea (Mollusca, Bivalvia) with notes on the evolution of the endolithic habit: Bulletin of the Peabody Museum of Natural History, Yale University, v. 41, p. 1–92.

Carter, R.M., 1968, Functional studies on the Cretaceous oyster Arctostrea: Paleontology, v. 11, p. 458–485.

Cartes, J.E., 1993, Diets of deep-sea brachyuran crabs in the Western Mediterranean Sea: Marine Biology, v. 117, p. 449–457, doi:10.1007/BF00349321.

Casey, R., 1960, A Lower Cretaceous gastropod with fossilized intestines: Paleontology, v. 2, p. 270–276.

Chew, K.K., 1960, Study of food preferences and rate of feeding of the Japanese oyster drill Ocenebra japonica: Special Science Report, US Fish and Widlife Service, v. 365, p. 1–27.

Cleevey, R.J., and Morris, N.J., 1988, Taxonomy and ecology of Cretaceous Cassiopidae (Mesogastropoda): Bulletin of the British Museum (Natural History) Geology, v. 44, p. 233–291.

Clements, J.C., Ellsworth-Power, M., and Rawlings, T.A., 2013, Diet breadth of the northern moonsnail (Lunatia heros) on the Northwestern Atlantic Coast (Naticidae): American Malacological Bulletin, v. 31, p. 331–336, doi:10.4003/006.031.0212.

Crovo, M.E., 1971a, Cypraea cervus and Cypraea zebra in Florida - one species or two ? The Veliger, v. 13, p. 292–295.

Dando, P.R., Southward, A.J., and Southward, E.C., 1986, Chemoautotrophic symbionts in the gills of the bivalve mollusc Lucinoma borealis and the sediment chemistry of its habitat: Proceedings of the Royal Society of London, Series B, Biological Sciences, v. 227, p. 227–247.

Diaz, J.M., and Torres, D.C., 2009, Rediscovery of a Caribbean living fossil: Pholadomya candida G. B. Sowerby I, 1823 (Bivalvia: Anomalodesmata: Pholadomyoidea): Nautilus, v. 123, p. 19–20.

Dominici, S., and Zuschen, M., 2016, Palaeocommunities, diversity and sea-level change from middle Eocene shell beds of the Paris Basin: Journal of the Geological Society, v. 173, p. 889–900.

Dushane, H., 1985, The Family Epitioniidae of Panama Bay: The Festivus, v. 17, p. 68–75.

Dushane, H., 1974, The Panamic-Galapagan Epitoniidae: The Veliger, v. 16, supplement, p. 1–84.

Edie, S.M., Jablonski, D., and Valentine, J.W., 2018, Contrasting responses of functional diversity to major losses in taxonomic diversity: Proceedings of the National Academy of Sciences of the United States of America, v. 115, p. 732–737, doi:10.1073/pnas.1717636115.

Endean, R., 1972, Aspects of molluscan pharmacology, in Florkin, M. and Scheer, B.T. eds., Chemical Zoology, Mollusca, New York, Academic Press, v. 7, p. 421–466.

Fearon, J., and Clapham, M., 2023, Decapod Life Habit: Placeholder reference,.

Fleming, C.A., 1978, The bivalve mollusc genus Limatula: A list of described species and a review of living and fossil species in the Southwest Pacific: Journal of the Royal Society of New Zealand, v. 8, p. 17–91, doi:10.1080/03036758.1978.10419418.

de Forges, R.B., 2006, Découverte en mer du Corail d'une deuxième espèce de glyphéide (Crustacea, Decapoda, Glypheoidea): Zoosystema, v. 28, p. 17–29, www.zoosystema.com.

Fontoura-da-Silva, V., de Souza Dantas, R.J., and Caetano, C.H.S., 2013, Foraging tactics in Mollusca: A review of the feeding behavior of their most obscure classes (Aplacophora, Polyplacophora, Monoplacophora, Scaphopoda and Cephalopoda): Oecologia Australis, v. 17, p. 358–373, doi:10.4257/oeco.2013.1703.04.

Foster, W.J., Garvie, C.L., Weiss, A.M., Muscente, A.D., Aberhan, M., Counts, J.W., and Martindale, R.C., 2020, Resilience of marine invertebrate communities during the early Cenozoic hyperthermals: Scientific Reports, v. 10, doi:10.1038/s41598-020-58986-5.

Fretter, V., 1975, Umbonium vestiarium, a filter-feeding trochid: Journal of Zoology, v. 177, p. 514–552.

Fretter, V., and Graham, A., 1978, The prosobranch molluscs of Britain and Denmark. Part 4 - marine Rissoacea: Journal of Molluscan Studies, Supplement, v. 6, p. 152–241.

Fretter, V., and Graham, A., 1981, The prosobranch molluscs of Britain and Denmark. Part 6 - Cerithiacea, Strombacea, Hipponicacea, Calyptraeacea, Lamellariacea, Cypraeacea, Naticacea, Tonnacea, Heteropoda: Journal of Molluscan Studies, Supplement, 9, p. 285–362.

Fretter, V., Graham, A., and Andrews, E.B., 1986, The prosobranch molluscs of Britain and Denmark. Part 9 - Pyramidellacea: The Journal of Molluscan Studies, Supplement 16, p. 557–649.

Fretter, V., and Manly, R., 1977, Algal associations of Tricolia pullus, Lacuna vincta and Cerithiopsis tubercularis (Gastropoda) with special reference to the settlement of their larvae: Journal of the Marine Biological Association of the United Kingdom, v. 57, p. 1999–1017.

Fuchs, D., Iba, Y., Heyng, A., Iijima, M., Klug, C., Larson, N.L., and Schweigert, G., 2020, The Muensterelloidea: phylogeny and character evolution of Mesozoic stem octopods: Papers in Palaeontology, v. 6, p. 31–92, doi:10.1002/spp2.1254.

Fursich, F.T., and Kirkland, J.I., 1986, Biostratinomy and paleoecology of a Cretaceous brackish lagoon: PALAIOS, v. 1, p. 543–560.

Fursich, F.T., and Pandey, D.K., 1999, Genesis and environmental significance of Upper Cretaceous shell concentrations from the Cauvery Basin, southern India: Palacogeography, Palaeoclimatology, Palaeoecology, v. 145, p. 119–139.

Fushs, D., Iba, Y., Heyng, A., Iijima, M., Klug, C., Larson, N.L., and Schweigert, G., 2019, The Muensterelloidea: phylogeny and character evolution of Mesozoic stem octopods: Papers in Paleontology, v. 6, p. 31–92.

Graham, A., 1966, The R/V Pillsbury deep-sea biological expedition to the Gulf of Guinea, 1964-65. 8. The fore-gut of some marginellid and cancellariid prosobranchs: Studies in Tropical Oceanography, v. 4, p. 134–151.

Grau, G., 1959, Pectinidae of the Eastern Pacific: Allan Hancock Pacific Expeditions, v. 23, p. 1–308.

Griffiths, R.J., 1981, Predation on the bivalves Choromytilus meridionalis (Kr.) by the gastropod Natica (Tectonatica) tecta Anton: Journal of Molluscan Studies, v. 47, p. 112–120.

Guerrero, S., and Reyment, R.A., 1988, Predation and feeding in the naticid gastropod Naticarius intricatoides (Hidalgo): Palaeogeography, Palaeoclimatology, Palaeoecology, v. 68, p. 49–52.

Hadfield, M.G., 1976, Molluscs associated with living tropical corals: Micronesica, v. 12, p. 133–148.

Hansen, T., Farrand, R.B., Montgomery, H.A., Billman, H.G., and Blechschmidt, G., 1987, Sedimentology and extinction patterns across the Cretaceous-Tertiary boundary interval in east Texas: Cretaceous Research, v. 8, p. 229–252.

Hansen, T., and Surlyk, F., 2014, Marine macrofossil communities in the uppermost Maastrichtian chalk of Stevns Klint, Denmark: Palaeogeography, Palaeoclimatology, Palaeoecology, v. 399, p. 323–344, doi:10.1016/j.palaeo.2014.01.025.

Harasewych, M.G., and Petit, R.E., 1982, Notes on the morphology of Cancellaria reticulata (Gastropoda: Cancellariidae): The Nautilus, v. 96, p. 104–113.

Harasewych, M.G., and Petit, R.E., 1984, Notes on the morphology of Olsonella smithii (Gastropoda: Cancellariidae): The Nautilus, v. 98, p. 37–44.

Hardison, L.K., and Kitting, C.L., 1985, Epiphytic algal browsing by Bittium varium (Gastropoda) among Thalassia testudinium turtlegrass: Journal of Phycology, v. 21, p. 1–13.

Haszprunar, G., 1985, Zur Anatomie und systematischen Stellung der Architectonicidae (Mollusca, Allogastropoda): Zoologica Scripta, v. 14, p. 25–43.

Hayami, I., 1984, Natural history and evolution of Cryptopecten (a Cenozoic- Recent pectinid genus): The University Museum, The University of Tokyo, Bulletin, 24, p. 1–149.

Hayami, I., and Noda, M., 1977, Notes of the Morphology of Neithea (Cretaceous Pectinids) with taxonomic revision of Japanese species: Tansactions and Proceedings of the Palaeontological Society of Japan, N.S., v. 105, p. 27–54.

Hayes, T., 1983, The influence of diet on local distribution of Cypraea: Pacific Science, v. 37, p. 27–36.

Healey, J.M., and Lamprell, K., 1998, Superfamily Mactroidea, in Beesley, B.L., Ross, G.J.B., and Wells, A. eds., Mollusca: The Southern Synthesis. Fauna of Australia, Melbourne, CSIRO Publishing, v. 5, p. 336–340.

Heinberg, C., 1999, Lower Danian bivalves, Stevns Klint, Denmark: continuity across the K/T boundary: Palaeogeography, Palaeoclimatology, Palaeoecology, v. 154, p. 87–106.

Hendy, A., Aberhan, M., Alroy, J., Kiessling, W., Lin, A., and LaFlamme, M., 2009, Unpublished ecological data in support of GSA 2009 abstract: A 600 million year record of ecological diversification:

Hickman, C.S., 1998, Superfamily Trochoidea, in Beesley, P.L., Ross, G.J.B., and Wells, A. eds., Mollusca: The Southern Synthesis. Fauna of Australia, Melbourne, CSIRO Publishing, v. 5, p. 671–692.

Hickman, C.S., and Lipps, J.H., 1983, Foraminiferivory; selective ingestion of foraminifera and test alterations produced by the gastropod Olivella: Journal of Foraminiferal Research, v. 13, p. 108–114.

Hickman, C.S., and McLean, J.H., 1990, Systematic revision and suprageneric classification of trochacean gastropods: Natural History Museum of Los Angeles County, Science Series, v. 35, p. 1–169.

Hicks, G.R.F., and Marshall, B.A., 1985, Sex selective predation of deep-sea meiobenthic copepods by pectinacean bivalves and its influence on copepod sex ratios: New Zealand Journal of Marine and Freshwater Research, v. 19, p. 227–231.

Hoffmann, R., and Stevens, K., 2020, The palaeobiology of belemnites – foundation for the interpretation of rostrum geochemistry: Biological Reviews, v. 95, p. 94–123, doi:10.1111/brv.12557.

Houbrick, R., 1981, Growth studies on the genus Cerithium (Gastropoda: Prosobranchia) with notes on ecology and microhabitats: The Nautilus, v. 88, p. 14–27.

Houbrick, R., 1993, Phylogenetic relationships and generic review of the Bittiinae (Prosobranchia: Cerithioidea): Malacologia, v. 35, p. 261–313.

Houbrick, R., 1978, The family Cerithiidae in the Indo-Pacific. Part 1: the genera Rhinoclavis, Pseudovertagus and Clavocerithium: Monographs of Marine Mollusca, v. 1, p. 250–430.

Houbrick, R., and Fretter, R., 1969, Some aspects of the functional anatomy and biology of Cymatium and Bursa: Proceedings of the Malacological Society of London, v. 38, p. 415–429.

Hubber, M., 2010, Compendium of bivalves. A full-color guide to 3,300 of the world's marine bivalves. A status on Bivalvia after 250 years of research: Hackenheim, Conchbooks, 1–901 p.

Hughes, R.N., 1986, Laboratory observations on the feeding behaviour, reproduction and morphology of Galeodea echinophora (Gastropoda, Cassidae): Zoological Journal of the Linnean Society, v. 86, p. 355–365.

Hughes, R.N., and Hughes, H.P.I., 1981, Morphological and behavioural aspects of feeding in the Cassidae (Tonnacea, Mesogastropoda): Malacologia, v. 20, p. 385–402.

Hurst, A., 1965, Studies on the structure and function of the feeding apparatus of Philine aperta with a comparative consideration of some other opisthobranchs: Malacologia, v. 2, p. 221–347.

Jablonski, D., and Bottjer, D.J., 1983, Soft-bottom epifaunal suspension-feeding assemblages in the Late Cretaceous: implications for the evolution of benthic paleocommunities., in Taevesz, M.J.S. and McCall, P.L. eds., Biotic Interactions in Recent and Fossil Benthic Communities, New York, Plenum Press.

Jackson, J.B.C., 1973, The ecology of molluscs of Thalassia communities, Jamaica, West Indies. 1. Distribution, environmental physiology, and ecology of common shallow-water species: Bulletin of Marine Science, v. 23, p. 313–350.

Jell, P.A., and Adrain, J.M., 2003, Available generic names for trilobites: Memoirs of the Queensland Museum, v. 48, p. 331–553.

Jenny, D., Fuchs, D., Arkhipkin, A.I., Hauff, R.B., Fritschi, B., and Klug, C., 2019, Predatory behaviour and taphonomy of a Jurassic belemnoid coleoid (Diplobelida, Cephalopoda): Scientific Reports, v. 9, doi:10.1038/s41598-019-44260-w.

Johnson, C.C., and Kauffman, E.K., 1996, Chapter 9, Maastrichtian extinction patterns of Caribbean Province rudists, in MacLeod, N. and Keller, G. eds., Cretaceous-Tertiary mass extinctions: biotic and environmental change, New York-London, Norton and Company, p. 231–272.

Juinio, M.A.R., and Cobb, J.S., 1992a, Natural diet and feeding habits of the postlarval lobster Homarus americanus: Marine Ecology Progress Series, v. 85, p. 83–91, doi:10.3354/meps085083.

Kay, E.A., 1960, The functional morphology of Cypraea caputserpentis L. and the interpretation of the relationships among the Cypraeacea: Internationale Revue Gesamten Hydrobiologie, v. 45, p. 175–196.

Keupp, H., Hoffmann, R., Stevens, K., and Albersdörfer, R., 2016, Key innovations in Mesozoic ammonoids: the multicuspidate radula and the calcified aptychus: Palaeontology, v. 59, p. 775–791, doi:10.1111/pala.12254.

Kiessling, W., 2004, Ecology opinions:

Kirkland, J.I., 1996, Paleontology of the Greenhorn Cyclothem (Cretaceous: Late Cenomanian to Middle Turonian) at Black Mesa, Northeastern Arizona: New Mexico Museum of Natural History and Science, Bulletin , v. 6, p. 1–131.

Klug, C., Schweigert, G., Fuchs, D., and de Baets, K., 2021, Distraction sinking and fossilized coleoid predatory behaviour from the German Early Jurassic: Swiss Journal of Palaeontology, v. 140, doi:10.1186/s13358-021-00218-y.

Knudsen, J., 1967, The John Murray expedition 1933-34. The deep sea Bivalvia: Scientific Reports. The John Murray Expedition 1933-34, v. 11, p. 237–343.

Kohl, B., and Vokes, H.E., 1994, On the living habits of Acesta bullisi (Vokes) in chemosynthetic bottom communities, Gulf of Mexico: The Nautilus, v. 108, p. 9–14.

Kohn, A.J., 2001, The Conidae of India revisited: Phuket Marine Biological Center Special Publication, v. 25, p. 357–362, doi:10.1080/00222937800770171.

Kohn, A.J., Taylor, J.D., and Wai, J.M., 1997, Diets of predatory gastropods of the families Mitridae and Buccinidae in the Houtman Abolhos Islands, Western Australia, in Wells, F.E. ed., Proceedings of the Seventh International Marine Biological Workshop: The Marine Flora and Fauna of the Houtman Abrolhos Islands, Western Australia, Perth, Western Australian Museum, v. 1, p. 133–139.

Komatsu, T., 2013, Palaeoecology of the mid-Cretaceous siphonate bivalve genus Goshoraia (Mollusca, Veneridae) from Japan: Palaeontology, v. 56, p. 381–397, doi:10.1111/j.1475-4983.2012.01206.x.

Kropp, R.K., 1982, Response of five holothurian species to attack by a predatory gastropod, Tonna perdix: Pacific Science, v. 36, p. 445–452.

Kruta, I., Landman, N.H., Rouget, I., Cecca, F., and Larson, N.L., 2010, The jaw apparatus of the Late Cretaceous ammonite Didymoceras : Journal of Paleontology, v. 84, p. 556–560, doi:10.1666/09-110.1.

Kruta, I., Landman, N., Rouget, I., Cecca, F., and Tafforeau, P., 2011, The role of ammonites in the mesozoic marine food web revealed by jaw preservation: Science, v. 331, p. 70–72, doi:10.1126/science.1198793.

Landman, N.H., Tsujita, C.J., Cobban, W.A., Larson, N.L., Tanabe, K., and Flemming, R.L., 2006, Jaws of late cretaceous placenticeratid ammonites: How preservation affects the interpretation of morphology: American Museum Novitates, p. 1–48, doi:10.1206/0003-0082(2006)500[0001:JOLCPA]2.0.CO;2.

Laxton, J.H., 1971, Feeding in some Australasian Cymatiidae (Gastropoda: Prosobranchia): Zoological Journal of the Linnnean Society, v. 50, p. 1–9.

Leonard-Pingel, J.S., and Jackson, J.B.C., 2013, Drilling intensity varies among neogene tropical american bivalvia in relation to shell form and life habit: Bulletin of Marine Science, v. 89, p. 905–919, doi:10.5343/bms.2012.1058.

Levine, T.D., Hansen, H.B., and Gerald, G.W., 2013, Effects of shell shape, size, and sculpture in burrowing and anchoring abilities in the freshwater mussel Potamilus alatus (Unionidae).:

Levings, S.C., and Garrity, S.D., 1983, Diet and tidal movements of two co-occurring neritid snails: differences in grazing patterns on a tropical rocky shore: Journal of Experimental Marine Biology and Ecology, v. 67, p. 61–287.

Littlewood, D.T.J., 1989a, Predation on cultivated Crassostrea rhizophorae (Guilding) by the gastropod Cymatium pileare (Linnaeus): Journal of Molluscan Studies, v. 55, p. 125–127.

Lopez, A., Montoya, M., and Lopez, J., 1998, A review of the genus Agaronia (Olividae) in the Panamic Province and the description of two new species from Nicaragua: The Veliger, v. 30, p. 295–304.

Maes, V.O., and Raeihle, D., 1976, Systematics and biology of Thala floridana (Gastropoda: Vexillidae): Malacologia, v. 15, p. 43–67.

Malchus, N., Dhondt, A. v, and Troger, K.-A., 1994, Upper Cretaceous bivalves from the Glauconie de Lonzee near Gembloux (SE Belgium): Bulletin de L'Institut Royal des Sciences Naturelles de Belgique: Sciences de la Terre, v. 64, p. 109–149.

Marcus, E., and Marcus, Ev.D.B.-R., 1962, On Leucozonia nassa: Faculdade de Filosofia, Ciencias e Letras da Universidade de Sao Paulo, 261 (Zoologia, 24), p. 11–24.

Marcus, E., and Marcus, Ev.D.B.-R., 1959, Studies on Olividae: Boletim Faculdade de Filosofia, Ciências e Letras da Universidade de São Paulo 232, Zool., v. 22, p. 99–188.

Marshall, B.A., 1978, Cerithiopsidae (Mollusca: Gastropoda) of New Zealand, and a provisional classification of the family: New Zealand Journal of Zoology, v. 5, p. 47–120.

McLay, C.L., 2006, Retroplumidae (Crustacea, Decapoda)from the Indo-Malayan archipelago(Indonesia, Philippine) and the Melanesian arc islands(Solomon Islands, Fiji and New Caledonia),and paleogeographical comments: Tropical Deep-Sea Benthos, p. 375–391.

Menge, J.L., 1974, Prey selection and foraging period of the predacious rocky intertidal snail, Acanthina punctulata: Oecologia, v. 17, p. 293–316.

Metz, G., 1995, Agaronia eats Olivella: The Festivus, v. 27, p. 86–87.

Mikkelsen, P.M., 1996, The evolutionary relationships of Cephalaspidea s.l. (Gastropoda: Opisthobranchia): a phylogenetic analysis: Malacologia, v. 37, p. 375–442.

Mikkelsen, P.M., and Bieter, R., 2003, Systematic revision of the western Atlantic file clams, Lima and Ctenoides (Bivalvia:Limoida:Limidae): Invertebrate Systematics, v. 17, p. 667–710, doi:10.1071/IS03007.

Mikkelsen, P.M., and Bieler, R., 2008, Seashells of Southern Florida: Living marine mollusks of the Florida Keys and regions: 1–503 p.

Miller, B.A., 1970, Feeding mechanisms in the family Terebridae: Reports of the American Malacological Union, Pacific Division, 1970, p. 72–74.

Mitchell, S.F., 2013, Revision of the Antillocaprinidae Mac Gillavry (Hippuritida, Bivalvia) and their position within the Caprinoidea d'Orbigny: Geobios, v. 46, p. 423–446, doi:10.1016/j.geobios.2013.07.003.

Morris, R.H., Abbott, D.P., and Haderlie, E.C., 1980, Intertidal Invertebrates of California: Stanford, Stanford University Press, v. ix, 1–690 p.

Morton, B., 1984, A review of Polymesoda (Geloina) Gray 1842 (Bivalvia: Corbiculacea) from Indo-Pacific mangroves: Asian Marine Biology, v. 1, p. 77–86.

Morton, B., 1979, A comparison of lip structure and function correlated with other aspects of functional morphology of Lima lima, Limaria (Platilimaria) fragilis and Limaria (Platilimaria) hongkongensis sp. nov. (Bivalvia: Limacea): Canadian Journal of Zoology, v. 57, p. 728–742.

Morton, B., 1991, Aspects of predation by Tonna zonatum (Prosobranchia: Tonnoidea) feeding on holothurians in Hong Kong: Journal of Molluscan Studies, v. 57, p. 11–20.

Morton, B., 1983, The biology and functional morphology of Eufistulana mumia (Bivalvia: Gastrochaenacea): Journal of Zoology, v. 200, p. 381–404.

Morton, B., 1985, The reproductive strategy of the mangrove bivalve Polymesoda (Geloina) erosa (Bivalvia: Corbiculoidea) in Hong Kong: Malacological Review, v. 18, p. 83–89.

Morton, B., 1986, The diet and prey capture mechanism of Melo melo (Prosobranchia: Volutidae): Journal of Molluscan Studies, v. 52, p. 156–160.

Morton, B., 2007, The evolution of the watering pot shells (Bivalvia: Anomalodesmata: Clavagellidae and Penicillidae): Records of the Western Australian Museum, v. 24, p. 19, doi:10.18195/issn.0312-3162.24(1).2007.019-064.

Morton, B., Prezant, R.S., and Wilson, B., 1998, Class Bivalvia, in Beesley, P.L., Ross, G.J.B., and Wells, A. eds., Mollusca, The Southern Synthesis. Fauna of Australia, Melbourne, CSIRO Publishing, v. 5, p. 195–234.

Morton, B., and Thurston, M.H., 1989, The functional morphology of Propeamussium lucidum (Bivalvia: Pectinacea), a deep-sea predatory scallop: Journal of Zoology, London, v. 218, p. 471–496.

Munster, H., 1995, Taxonomie und Paliobiologie der Bakevelliidae (Bivalvia): Beringeria, v. 15, p. 1–161.

Nicol, E.A.T., 1932, The Feeding Habits of the Galatheidea: Journal of the Marine Biological Association of the United Kingdom, v. 18, p. 87–106.

Oliver, P.G., 1981, The functional morphology and evolution of recent Limopsidae (Bivalvia, Arcoidea): Malacologia, v. 21, p. 61–93.

Oliver, P.G., and Holmes, A.M., 2006, The Arcoidea (Mollusca: Bivalvia): a review of the current phenetic-based systematics: Zoological Journal of the Linnean Society, v. 148, p. 237–251.

Olsson, A.A., 1970, The cancellariid radula and its interpretation: Paleontographica Americana, v. 7, p. 19–27.

Osorio, C., Jara, F., and Ramirez, M.E., 1993, Diet of Cypraea caputdraconis (Mollusca: Gastropoda) as it relates to food availability in Easter Island: Pacific Science, v. 47, p. 34–42.

Owen, G., 1961, A Note on the Habits and Nutrition of Solemya parkinsoni (Protobranchia: Bivalvia): Journal of Cell Science, v. 102, p. 15–21.

Paine, R.T., 1966, Function of labial spines, composition of diet and size of certain marine gastropods: The Veliger, v. 9, p. 17–24.

Perron, F.E., 1978, Seasonal Burrowing Behavior and Ecology of Aporrhais occidentalis (Gastropoda: Strombacea):, https://about.jstor.org/terms.

Petuch, E.J., 1988, Neogene history of tropical American mollusks. Biogeography & evolutionary patterns of tropical Western Atlantic Mollusca: Charlottesville, CERF, 1–217 p.

Pohlo, R., 1982, Evolution of Tellinacea (Bivalvia) : Journal of Molluscan Studies, v. 48, p. 245–256.

Pojeta, J., and Sohl, N.F., 1987, Ascaulocardium armatum (Morton, 1833), New Genus (Late Cretaceous): The Ultimate Variation on the Bivalve Paradigm: Memoir (The Paleontological Society, https://about.jstor.org/terms.

Ponder, W.F., 1985a, A review of the genera of the Rissoidae (Mollusca: Mesogastropoda: Rissoacea): Records of the Australian Museum, Supplement 4, p. 1–221.

Ponder, W.F., 1968, Anatomical notes on two species of the Colubrariidae (Mollusca, Prosobranchia): Transactions of the Royal Society of New Zealand, v. 10, p. 217–223.

Ponder, W.F., 1985b, The anatomy and relationships of Elachisina Dall (Gastropoda: Rissoacea): Journal of Molluscan Studies, v. 51, p. 23–34.

Ponder, W.F., 1972, The morphology of some mitriform gastropods with special reference to their alimentary and reproductive systems (Neogastropoda): Malacologia, v. 11, p. 295–342.

Ponder, W.F., and Keyzer, R.G., 1998, Superfamily Rissoidea, in Beesley, P.L., Ross, G.J.B., and Wells, A. eds., Mollusca: The Southern Synthesis. Fauna of Australia, Melbourne, CSIRO Publishing, v. 3, p. 745–766.

Quinn, J.F., 1980, A new species and subspecies of Oocorythidae (Gastropoda: Tonnacea) from the western Atlantic: The Nautilus, v. 94, p. 149–158.

Radwin, G.E., and D'Attilio, A., 1976, Murex shells of the world. An illustrated guide to the Muricidae: Stanford, Stanford University Press, v. x, 1–284 p.

Radwin, G.E., and Wells, H.W., 1968, Comparative radular morphology and feeding habits of muricid gastropods from the Gulf of Mexico: Bulletin of Marine Science, v. 1, p. 72–85.

Reidel, F., 1995, An outline of cassoidean phylogeny (Mollusca, Gastropoda): Contributions to Tertiary and Quaternary Geology, v. 32, p. 97–132.

Reidel, F., 1994, Recognition of the superfamily Ficoidea Meek 1864 and definition of the Thalassocynidae fam. nov: Zoologische Jahrbücher Abteilung für Systematik, v. 121, p. 457–474.

Rhoads, D., Speden, I., and Waage, K., 1972, Trophic Group Analysis of Upper Cretaceous (Maestrichtian) Bivalve Assemblages from South Dakota: AAPG Bulletin, v. 56, p. 1100–1113.

Robertson, R., 1996, Fargoa bartschi (Winkley, 1909): a little-known Atlantic and Gulf coast American odostomian (Pyramidellidae) and its generic relationships: American Malacological Bulletin, v. 13, p. 11–21.

Robertson, R., 1967, Heliacus (Gastropoda: Architectonicidae) symbiotic with Zoanthiniaria (Coelenterata): Science, v. 156, p. 246–248.

Robertson, R., 1983, Observations on the life history of the wentletrap Epitonium albidum in the West Indies: American Malacological Bulletin, v. 1, p. 1–12.

Robertson, R., 1970, Review of the predators and parasites of stony corals, with special reference to symbiotic prosobranch gastropods: Pacific Science, v. 24, p. 43–54.

Robertson, R., 1963, Wentletraps (Epitoniidae) feeding on sea anemones and corals: Proceedings of the Malacological Society of London, v. 35, p. 51–63.

Robertson, R., and Mau-Lastovicka, T., 1979, The ectoparasitism of Boonea and Fargoa (Gastropoda: Pyramidellacea): Biological Bulletin, v. 157, p. 320–333.

Roig, M.S., 1926, Los Equiodermos fosiles de Cuba: Contribucion a la paleontologia Cubana, p. 1–179.

Ros-Franch, S., Marquez-Aliaga, A., and Damborencea, S.E., 2015, Comprehensive database on Induan (Lower Triassic) to Sinemurian (Lower Jurassic) marine bivalve genera and their paleobiogeographic record: Paleontological Contributions, p. 3–219, doi:10.17161/pc.1808.13433.

Rowden, A.A., and Jones, M.B., 1995, The burrow structure of the mud shrimp callianassa subterranea (Decapoda: Thalassinidea) from the north sea: Journal of Natural History, v. 29, p. 1155–1165, doi:10.1080/00222939500770491.

Rudwick, M.J.S., 1970, Living and Fossil Brachiopods: London, Hutchinson & Co, 1–199 p.

Sahlmann, C., Chan, T.Y., and Chan, B.K.K., 2011a, Feeding modes of deep-sea lobsters (Crustacea: Decapoda: Nephropidae and Palinuridae) in Northwest Pacific waters: Functional morphology of mouthparts, feeding behaviour and gut content analysis: Zoologischer Anzeiger, v. 250, p. 55–66, doi:10.1016/j.jcz.2010.11.003.

Saul, L.R., 1973, Evidence for the origin of the Mactridae (Bivalvia) in the Cretaceous: University of California Press, p. 1–59.

Savazzi, E., 1991, Burrowing in the inarticulate brachiopod Lingula anatina: Palaeogeography, Palaeoclimatology, Palaeoecology, v. 85, p. 101–106.

Schembri, P.J., 1982, Feeding behaviour of fifteen species of hermit crabs (Crustacea: Decapoda: Anomura) from the otago region, southeastern new zealand: Journal of Natural History, v. 16, p. 859–878, doi:10.1080/00222938200770691.

Schweitzer, C.E., Lacovara, K.J., Smith, J.B., Lamanna, M.C., Lyon, M.A., and Attia, Y., 2003, Mangrove-dwelling crabs (Decapoda: Brachyura: Necrocarcinidae) associated with dinosaurs from the Upper Cretaceous (Cenomanian) of Egypt: Journal of Paleontology, v. 77, p. 888–894, doi:10.1666/0022-3360(2003)077<0888:mcdbna>2.0.co;2.

Scott, B.J., and Kenny, R., 1998, Superfamily Neritoidea, in Beesley, P.L., Ross, G.J.B., and Wells, A. eds., Mollusca: The Southern Synthesis. Fauna of Australia, Melbourne, CSIRO Publishing, v. 5, p. 565–1115.

Seamon, N., and Seamon, E., 1967, Collecting in the Netherlands Antilles (notes on the feeding habits of captive Murex brevifrons and Voluta musica): New York Shell Club Notes, v. 134, p. 3–6.

Sessa, J.A., Bralower, T.J., Patzkowsky, M.E., Handley, J.C., and Ivany, L.C., 2012, Environmental and biological controls on the diversity and ecology of Late Cretaceous through early Paleogene marine ecosystems in the U.S. Gulf Coastal Plain: Paleobiology, v. 38, p. 218–239, doi:10.1666/10042.1.

Skinner, D.G., and Hill, B.J., 1987, Feeding and reproductive behaviour and their effect on catchability of the spanner crab Ranina ranina: Marine Biology, v. 94, p. 211–218.

Slack-Smith, S., 1998, Superfamily Glossoidea, in Beesley, P.L., Ross, G.J.B., and Wells, A. eds., Mollusca: The Southern Synthesis. Fauna of Australia, Melbourne, CSIRO Publishing xv, v. 5.

Smith, J.T., 1991, Cenozoic giant pectinids from California and the Tertiary Caribbean Province: Lyropecten, “Macrochlamis”, Vertipecten, and Nodipecten species: United States Geological Survey Professional Paper 1391, p. 1–137.

Smith, A.B., 2009, Classification of the Echinoidea: Online database,.

Smith, A.M., Key, M.M., and Gordon, D.P., 2006, Skeletal mineralogy of bryozoans: Taxonomic and temporal patterns: Earth-Science Reviews, v. 78, p. 287–306, doi:10.1016/j.earscirev.2006.06.001.

Smith, C.P.A., Landman, N.H., Bardin, J., and Kruta, I., 2021, New evidence from exceptionally “well-preserved” specimens sheds light on the structure of the ammonite brachial crown: Scientific Reports, v. 11, doi:10.1038/s41598-021-89998-4.

Speden, I.G., 1970, The Type Fox Hills Formation, Cretaceous (Maestrichtian), South Dakota. Part 2. Systematics of the Bivalvia: Peabody Museum of Natural History, Bulletin, v. 33, p. 1–222.

Stanley, S.M., 1972, Functional morphology and evolution of bysally attached bivalve molluscs: Journal of Paleontology, v. 46, p. 165–296.

Stanley, S.M., 1970a, Relation of shell form to life habits in the Bivalvia (Mollusca), in Stanley, S.M. ed., Relation of shell form to life habits in the Bivalvia (Mollusca),.

Stasek, C.R., 1961, The ciliation and function of the labial palps of Acila castrensis (Protobranchia, Nuculidae) with an evaluation of the role of the protobranch organs of feeding in the evolution of the Bivalvia: Proceedings of the Zoological Society of London, v. 137, p. 511–538.

Statnlmis, E.J., Reede-Dekker, T., van Etten, Y., de Wiljes, J.J., and Videler, J.J., 1996, Behaviour and time allocation of the burrowing shrimp Callianassa subterranea (Decapoda, Thalassinidea).:

Steneck, R.S., and Watling, L., 1982, Feeding capabilities and limitation of herbivorous molluscs: a functional group approach: Marine Biology, v. 68, p. 299–319.

Stenzel, H.B., 1971, Oysters, in Moore, R.C. ed., Treatise on Invertebrate Paleontology. Part N. Mollusca 6, Bivalvia, Vol. 3, Boulder, Colorado & Lawrence, Kansas, Geological Society of America & Univeristy of Kansas Press, p. N953–N1197.

Stephenson, L.W., 1941, The larger invertebrate fossils of the Navarro Group of Texas: University of Texas Publication number 4101, p. 1–641.

Stephenson, Ll.W., 1923, Volume V: The Cretaceous Formations of North Carolina: Part 1. Invertebrate fossils of the Upper Cretaceous formations.:

Stevcic, Z., 1973, Society of Systematic Biologists The Systematic Position of the Family Raninidae: Systematic Zoology, v. 22, p. 625–632.

Stupakoff, I., 1986, Observations on the feeding behavior of the gastropod Pleuroploca princeps (Fasciolariidae) in the Galapagos Islands: The Nautilus, v. 100, p. 92–95.

Sundberg, F.A., 1980, Late Cretaceous Paleoecology of the Holz Shale, Orange County, California:, https://www.jstor.org/stable/1304314.

Takeda, Y., Tanabe, K., Sasaki, T., Uesugi, K., and Hoshino, M., 2016, Non-destructive analysis of in situ ammonoid jaws by synchrotron radiation X-ray micro-computed tomography: Palaeontologica Electronica, v. 19.3, p. 1–13.

Tanabe, K., Aiba, D., and Abe, J., 2021, The jaw apparatus of the Late Cretaceous heteromorph ammonoid Turrilites costatus from central Hokkaido, Japan: Bulletin of the Mikassa City Museum, v. 24, p. 1–8.

Tanabe, K., and Fukuda, Y., 1987, Mouth part histology and morphology, in Saunders, W.B. and Landman, N.H. eds., Nautilus. The Biology and Paleobiology of a Living Fossil., New York, Plenum Press, p. 312–322.

Tanabe, K., Landman, N.H., and Kruta, I., 2012, Microstructure and mineralogy of the outer calcareous layer in the lower jaws of Cretaceous Tetragonitoidea and Desmoceratoidea (Ammonoidea): Lethaia, v. 45, p. 191–199, doi:10.1111/j.1502-3931.2011.00272.x.

Tanabe, K., Misaki, A., Landman, N.H., and Kato, T., 2013a, The jaw apparatuses of Cretaceous Phylloceratina (Ammonoidea): Lethaia, v. 46, p. 399–408.

Tanabe, K., and Shigeta, Y., 2019, Lower jaws of two species of Menuites (Pachydiscidae, Ammonoidea) from the middle Campanian (Upper Cretaceous) in the Soya area, northern Hokkaido, Japan: Bulletin of the National Museum of Nature and Science, Series C, v. 45, p. 19–27, https://www.researchgate.net/publication/338178645.

Tanabe, K., Tsujino, Y., Okuhira, K., and Misaki, A., 2015, The jaw apparatus of the Late Cretaceous heteromorph ammonoid Pravitoceras: Journal of Paleontology, v. 89, p. 611–616, doi:10.1017/jpa.2015.27.

Taylor, J.D., 1984, A partial food web involving predatory gastropods on a fringing reef: Journal of Experimental Marine Biology and Ecology, v. 74, p. 273–290.

Taylor, J.D., 1968a, Coral reef and associated invertebrate communities (mainly molluscan) around Mahé, Seychelles: Philosophical Transactions of the Royal Society of London, Series B, v. 254, p. 129–206.

Taylor, J.D., 1982, Diets of sublittoral predatory gastropods of Hong Kong, in Morton, B.S. and Tseng, C.K. eds., Proceedings of the First International Marine Biology Workshop, Hong Kong, Hong Kong University Press.

Taylor, J.D., 1989, The diet of coral-reef Mitridae (Gastropoda) from Guam: with a review of other species of the family: Journal of Natural History, v. 23, p. 261–278.

Taylor, J.D., and Glover, E.A., 2000, Diet of olives: Oliva tigridella Duclos, 1835 in Queensland: Molluscan Research, v. 20, p. 19–24.

Taylor, J.D., and Reid, D.G., 1984, The abundance and trophic classification of molluscs upon coral reefs in the Sudanese Red Sea: Journal of Natural History, v. 18, p. 175–209.

Thomas, R.D.K., 1978, Shell Form and the Ecological Range of Living and Extinct Arcoida:, https://about.jstor.org/terms.

Todd, J.A., 2001, Bivalve Life Habits: unpublished,.

Tong, I.K.Y., 1986, The feeding ecology of Thais clavigera and Morula musiva (Gastropoda: Muricidae) in Hong Kong: Asian Marine Biology, v. 3, p. 163–178.

Turner, R.D., 1969, Superfamily Pholadacea Lamarck, 1908, in Cox, L.R. and et al. eds., Treatise on Invertebrate Zoology. Part N. Volume 2 (of 3). Mollusca, Bivalvia, Lawrence, The Geological Society of America, Boulder and the University of Kansas.

Turnsek, D., 1997, Mesozoic corals of Slovenia: Ljubljana, ZRC SAZU, 1–512 p.

Waller, T.R., 1969, The evolution of the Argopecten gibbus stock (Mollusca: Bivalvia), with emphasis on the Tertiary and Quaternary species of eastern North America: Journal of Paleontology, The Paleontographical Society Memoir, v. 3, p. 1–125.

Warmke, G.L., and Almodovar, L.R., 1963, Some associations of marine mollusks and algae in Puerto Rico: Malacologia, v. 1, p. 163–167.

Warren, A., and Bouchet, P., 1988, A new species of Vanikoridae from the western Mediterranean, with remarks on the northeast Atlantic species of the family: Bolletino Malacologico, v. 24, p. 73–100.

Wassenberg, T., and Hill, B., 1989, Diets of four decapod crustaceans (Linuparus trigonus, Metanephrops andamanicus, M. australiensis and M. boschmai) from the continental shelf around Australia: Marine Biology, v. 103, p. 161–167.

Wells, H.W., 1958, Predation of Pelecypods and Gastropods by Fasciolaria hunteria (Perry): Bulletin of Marine Science of the Gulf and Caribbean, v. 8, p. 152–166.

Whittington, H.B. et al., 1997, Trilobita. Introduction, Order Agnostina, Order Redlichiida. Volume 1: v. 1, 1–503 p.

Willan, R.C., 1998, Superfamily Tellinoidea, in Beesely, P.L., Ross, G.J.B., and Wells, A. eds., Mollusca: The Southern Synthesis. Fauna of Australia, Melbourne, CSIRO Publishing, v. 5, p. 342–348.

Wippich, M.G.E., and Lehmann, J., 2004, Allocrioceras from the Cenomanian (mid-cretaceous) of the Lebanon and its bearing on the palaeobiology interpretation of heteromorphic ammonites: Palaeontology, v. 47, p. 1093–1107, doi:10.1111/j.0031-0239.2004.00408.x.

Wise, J.B., 1996, Morphology and phylogenetic relationships of certain pyramidellid taxa (Heterobranchia): Malacologia, v. 37, p. 443–511.

Yonge, C.M., 1977, Form and Evolution in the Anomiacea (Mollusca: Bivalvia)--Pododesmus, Anomia, Patro, Enigmonia (Anomiidae): Placunanomia, Placuna (Placunidae Fam. Nov.): Philosophical Transactions of the Royal Society B: Biological Sciences, v. 276, p. 453–523.

Yonge, C.M., 1971, On the functional morphology and adaptive radiation in the bivalve Super- family Saxicavacea (Hiatella (=Saxicava), Saxicavella, Panomya, Panope, Cyr- todaria: Malacologi, v. 11, p. 1–44.

Yonge, C.M., 1939, The protobranchiate Mollusca: a functional interpretation of their structure and evolution: Philosophical Transactions of the Royal Society, Series B, Biological Sciences, v. 230, p. 79–147.

Yonge, C.M., 1955, Adaptation to rock-boring in Botula and Lithophaga (Lamellibranchia, Mytilidae) with a discussion on the evolution of the habit: Quarterly Journal of Microscopical Science, v. 96, p. 383–410.

Yonge, C.M., and Thompson, T.E., 1976, Living Marine Molluscs: London, Collins, 1–288 p.

Zanzerl, H., and Dufour, S.C., 2017, The burrowing behaviour of symbiotic and asymbiotic Thyasirid bivalves: Article in Journal of Conchology, v. 42, p. 299–308, https://www.researchgate.net/publication/312586311.

Appendix B-3. R Code

```
###### Code for mapping FD data and analyzing FD based on mFD package (Magneville et al. 2021) #####


  #### Load necessary packages ####

library(ggplot2) # for making plots

library(ggmap) # for plotting on maps

library(tidyverse)

library(sf) # shapefiles

library(mapview)

library(raster)

library(rgdal)

library(dismo)

library(XML)

library(maps)

library(mFD) # primary package for running FE analysis

library(reshape2) # manipulate dfs

library(vegan)

library(dplyr) # df stuff

library(tidyr)

library(gridExtra) # for plotting muplitple plots on one image

library(goodpractice)

library(tidyr)



  #### Download data and set directory ####

# Set working directory to get the file

setwd("C:/Users/Ceara/Documents/Province Project/FE_Analysis_Code-and-Files")
```

```
# Read in database file
data <- read.csv("FE_data_compress_fe_3-7-23_basins.csv") # DATABASE HAS BASIN ID'S

# Option to subset out different basin's data
# data <- subset(data, Basin == "WIS") # SUBSET OUT WIS BASIN OCC
# data <- subset(data, Basin == "GCP") # SUBSET OUT GCP BASIN OCC
nrow(data)
colnames(data) # check file

 # The data here has all columns from the "raw" file, including:
 # "Key", "CollectionSource", "Lat", "Long", "LocalityDescrip", "Country",
 # "State", "County", "Town", "Phylum", "Class", "Order", "Family",
 # "Updated_Genus", "Updated_Sp", "Updated_subtax", "Early_Int.Substag",
 # "Late_Int.Substag", "Corr_Stage", "HigherTax_Class", "Formation", "Group",
 # "Member", "Zone_Comments", "Substage_from_Zone.Mbr",
 # "Inform_Zone...Merewethere.et.al.2015.", "X", "PaperCited", "collection_no",
 # "Museum", "EcolRef", "motility", "life_habitat", "feeding"

#### Function for SUBSAMPLING WITH SQS ####

# function to calc sqs value, based on
https://strata.uga.edu/8370/rtips/shareholderQuorumSubsampling.html
# this looks like it was put together by Steve Holland.... not sure how to credit
run_sqs <-function(abundance, quota=0.8, trials=100, ignore.singletons=FALSE,
exclude.dominant=FALSE) {
 # abundance is a vector of integers representing the abundance of every species

 if ((quota <= 0 || quota >= 1)) {
   stop('The SQS quota must be greater than 0.0 and less than 1.0')
```

```
}

# compute basic statistics
specimens <- sum(abundance)
numTaxa <- length(abundance)
singletons <- sum(abundance==1)
doubletons <- sum(abundance==2)
highest <- max(abundance)
mostFrequent <- which(abundance==highest)[1]

if (exclude.dominant == FALSE) {
  highest <- 0
  mostFrequent <- 0
}

# compute Good's u
u <- 0
if (exclude.dominant == TRUE) {
  u <- 1 - singletons / (specimens - highest)
} else {
  u <- 1 - singletons / specimens
}

if (u == 0) {
  stop('Coverage is zero because all taxa are singletons')
}

# re-compute taxon frequencies for SQS
frequencyInitial <- abundance - (singletons + doubletons / 2) / numTaxa
```

```
 frequency <- frequencyInitial / (specimens - highest)

 # return if the quorum target is higher than estimated coverage
 if ((quota > sum(frequency)) || (quota >= sum(abundance))) {
  stop('SQS quota is too large, relative to the estimated coverage')
 }

 # create a vector, length equal to total number of specimens,
 # each value is the index of that species in the abundance array
 ids <- unlist(mapply(rep, 1:numTaxa, abundance))

 # subsampling trial loop
 richness <- rep(0, trials) # subsampled taxon richness
 for (trial in 1:trials) {
  pool <- ids # pool from which specimens will be sampled
  specimensRemaining <- length(pool) # number of specimens remaining to be sampled
  seen <- rep(0, numTaxa) # keeps track of whether taxa have been sampled
  subsampledFrequency <- rep(0, numTaxa) # subsampled frequencies of the taxa
  coverage <- 0

  while (coverage < quota) {
   # draw a specimen
   drawnSpecimen <- sample(1:specimensRemaining, size=1)
   drawnTaxon <- pool[drawnSpecimen]

   # increment frequency for this taxon
   subsampledFrequency[drawnTaxon] <- subsampledFrequency[drawnTaxon] + 1

   # if taxon has not yet been found, increment the coverage
```

```
    if (seen[drawnTaxon] == 0) {
      if (drawnTaxon != mostFrequent && (ignore.singletons == 0 || abundance[drawnTaxon] > 1)) {
        coverage <- coverage + frequency[drawnTaxon]
      }
      seen[drawnTaxon] <- 1

      # increment the richness if the quota hasn't been exceeded,
      # and randomly throw back some draws that put the coverage over quota
      if (coverage < quota || runif(1) <= frequency[drawnTaxon]) {
        richness[trial] <- richness[trial] + 1
      } else {
        subsampledFrequency[drawnTaxon] <- subsampledFrequency[drawnTaxon] - 1
      }
    }

    # decrease pool of specimens not yet drawn
    pool[drawnSpecimen] <- pool[specimensRemaining]
    specimensRemaining <- specimensRemaining - 1
  }
 }

 # compute subsampled richness
 s2 <- richness[richness>0]
 subsampledRichness <- exp(mean(log(s2))) * length(s2)/length(richness)
 return(round(subsampledRichness, 1))
}

# get the full db list of taxa and then get their frequencies
genus_vector <- data[,c("Updated_Genus")]
```

```
genus_freq <-table(genus_vector) # all data frequencies of genera

# run sqs on entire db abundance vector
data_sqs_val <- run_sqs(genus_freq)
data_sqs_val

# function to get abundance list for each substage
sub_func_for_sqs <- function(data,substage){
  dat <- as.data.frame(subset(data, Substage_from_Zone.Mbr ==substage))
  dat <- dat[,c("Updated_Genus")]
  dat2 <- table(dat)
  return(dat2)
}

maa_up_for_sqs <- sub_func_for_sqs(data,"MAA (up)")
maa_low_for_sqs <- sub_func_for_sqs(data,"MAA (low)")
cam_up_for_sqs <- sub_func_for_sqs(data,"CAM (up)")
cam_mid_for_sqs <- sub_func_for_sqs(data,"CAM (mid)")
cam_low_for_sqs <- sub_func_for_sqs(data,"CAM (low)")

# run sqs on the substages
maa_up_sqs_val <- run_sqs(maa_up_for_sqs, quota=0.9)
maa_low_sqs_val <- run_sqs(maa_low_for_sqs, quota=0.9)
cam_up_sqs_val <- run_sqs(cam_up_for_sqs, quota=0.9)
cam_mid_sqs_val <- run_sqs(cam_mid_for_sqs, quota=0.9)
cam_low_sqs_val <- run_sqs(cam_low_for_sqs, quota=0.9)


maa_up_sqs_val
```

```
maa_low_sqs_val

cam_up_sqs_val

cam_mid_sqs_val

cam_low_sqs_val


#### CONFIGURE DATA #####

  #### Overlay Grid and Create Smaller Data Frame (remove unnecessary columns) ####


# Set new WD to read in grid made in ArcGIS pro

setwd("C:/Users/ceara/Documents/Province Project/FE_Analysis_Code-and-
Files/OccurrenceData_copiedfromProvinceProject/VettedCombinedDB")


sixtykmgrid <- st_read("SpatialGrids/grid_60km.shp")

# sixtykmgrid <- st_read("SpatialGrids/grid_360km.shp") # option to use 360km grid instead


# Transform database into shapefile

sf_data <- st_as_sf(data, coords = c("Long", "Lat"),  crs = 4326)


# Function to spatially join DB with grids

spatial_join <- function(shapefile_data,grid){

  data1 <- st_join(shapefile_data,left=TRUE,grid["PageName"])

  return(data1)

}


# Spatially join Grid cell Page Names to data table

join60_data <- spatial_join(sf_data,sixtykmgrid)

join60_data <- as.data.frame(join60_data)

nrow(join60_data) # number of occurrences
```

```
# number of unque genera in the database
length(unique(join60_data$Updated_Genus))

#### Option to SUBSET OUT JUST SPECIFIC CLASS OF ORGANISMS FOR COMPARISON ####

 ## Subseting out just specific Class types, using only the most abundant
  # classes of Bivalvia, Gastropoda, and Cephalopoda

# Look at the frequencies of different taxonomic classes
table(join60_data$Class)

# Make subsetted df for each of the three most common
bivalve_db <- subset(join60_data, Class == "Bivalvia")
gastropod_db <- subset(join60_data, Class == "Gastropoda")
cephalopod_db <- subset(join60_data, Class == "Cephalopoda")

# Apply the specific taxonomic classes as the base data to use in the rest of the code
# join60_data <- bivalve_db
# join60_data <- gastropod_db
# join60_data <- cephalopod_db

#### Option to SUBSET OUT GRID CELLS with greater than 3 unique genera from DB ####

# Function to remove all rows that represent grid cells with fewer than 3 unique genera
get_greater_3_database <- function(data){
 # Function to get vector of grid cell names with greater than 3 unique genera
 get_great_3_grids <- function(data,Age){
  # Get just the columns of Genus, Age, and Grid cell name
  gen_age_grid <- data[,c("Updated_Genus","Substage_from_Zone.Mbr","PageName")]
```

```
  # Remove duplicates from the list
  unique_gen_age_grid <- unique(gen_age_grid)

  # Get just the substage information in question
  substage_unique_gen_age_grid <- unique_gen_age_grid %>% filter(Substage_from_Zone.Mbr == Age)

  # Get the number of unique genera in each grid cell in the substage
  substage_grid_freq <- data.frame(table(substage_unique_gen_age_grid$PageName))

  # Extract just the Grid cell names that have 3 or more unique genera
  substage_grid_greater_3_gen <- substage_grid_freq %>% filter(Freq >2)
  colnames(substage_grid_greater_3_gen) <- c("PageName","freq_gen")

  return(substage_grid_greater_3_gen)
}

# use above function to create list of Grid cells with >3 occ
maa_up_grid_greater_3_gen <- get_great_3_grids(data,"MAA (up)")
maa_low_grid_greater_3_gen <- get_great_3_grids(data,"MAA (low)")
cam_up_grid_greater_3_gen <- get_great_3_grids(data,"CAM (up)")
cam_mid_grid_greater_3_gen <- get_great_3_grids(data,"CAM (mid)")
cam_low_grid_greater_3_gen <- get_great_3_grids(data,"CAM (low)")

# get subset of the original db for each substage
maa_up_data1 <- as.data.frame(subset(data, Substage_from_Zone.Mbr =="MAA (up)"))
maa_low_data1 <- as.data.frame(subset(data, Substage_from_Zone.Mbr =="MAA (low)"))
cam_up_data1 <- as.data.frame(subset(data, Substage_from_Zone.Mbr =="CAM (up)"))
cam_mid_data1 <- as.data.frame(subset(data, Substage_from_Zone.Mbr =="CAM (mid)"))
```

```
 cam_low_data1 <- as.data.frame(subset(data, Substage_from_Zone.Mbr =="CAM (low)"))


 # merge the two dfs together to get extra column that has NA if less than 3 genera in the grid cell in
that SS

 maa_up_data1_merge <- merge(maa_up_data1,maa_up_grid_greater_3_gen,by="PageName",all.x =
TRUE)

 maa_low_data1_merge <- merge(maa_low_data1,maa_low_grid_greater_3_gen,by="PageName",all.x
= TRUE)

 cam_up_data1_merge <- merge(cam_up_data1,cam_up_grid_greater_3_gen,by="PageName",all.x =
TRUE)

 cam_mid_data1_merge <- merge(cam_mid_data1,cam_mid_grid_greater_3_gen,by="PageName",all.x
= TRUE)

 cam_low_data1_merge <- merge(cam_low_data1,cam_low_grid_greater_3_gen,by="PageName",all.x
= TRUE)


 # remove rows with NA in the freq_gen column

 maa_up_data1_data2 <- as.data.frame(subset(maa_up_data1_merge, !is.na(freq_gen)))

 maa_low_data1_data2 <- as.data.frame(subset(maa_low_data1_merge, !is.na(freq_gen)))

 cam_up_data1_data2 <- as.data.frame(subset(cam_up_data1_merge, !is.na(freq_gen)))

 cam_mid_data1_data2 <- as.data.frame(subset(cam_mid_data1_merge, !is.na(freq_gen)))

 cam_low_data1_data2 <- as.data.frame(subset(cam_low_data1_merge, !is.na(freq_gen)))


 # bind the SS dfs back together

 all_substage_greater_3_data <- rbind(maa_up_data1_data2,maa_low_data1_data2,

                                      cam_up_data1_data2,cam_mid_data1_data2,

                                      cam_low_data1_data2)


 # remove the column of freq_gen

 all_substage_greater_3_data <- all_substage_greater_3_data[,c(1:ncol(all_substage_greater_3_data)-
1)]
```

```
  return(all_substage_greater_3_data)
}

# Use above function to remove all rows from the DB that represent grid cells with
# fewer than 3 unique genera
join60_great_3_data <- get_greater_3_database(join60_data)

  #### IF GOING TO SUBSAMPLE BY WAY OF USING ONLY GRID CELLS WITH >3 UNIQUE GENERA,
    ## USE THE ABOVE CODE AND THEN THE NEXT LINE TO REPLACE THE DB USED IN THE REST OF
    ## THE CODE WITH ONE WITH >3 UNIQUE GENERA GRID CELLS

# Apply the truncated df (with only >3 occ) as the base data to use in the rest of the code
# join60_data <- join60_great_3_data

#### Option to SUBSET OUT GRID CELLS with greater than SQS limit of unique genera from DB ####

# Function to remove all rows that represent grid cells with fewer than 3 unique genera
get_greater_sqs_database <- function(data){
  # Function to get vector of grid cell names with greater than CAM for SQS limit of 16 unique genera
  get_great_cam_sqs_grids <- function(data,Age){
    # Get just the columns of Genus, Age, and Grid cell name
    gen_age_grid <- data[,c("Updated_Genus","Substage_from_Zone.Mbr","PageName")]

    # Remove duplicates from the list
    unique_gen_age_grid <- unique(gen_age_grid)

    # Get just the substage information in question
    substage_unique_gen_age_grid <- unique_gen_age_grid %>% filter(Substage_from_Zone.Mbr == Age)
```

```
  # Get the number of unique genera in each grid cell in the substage
  substage_grid_freq <- data.frame(table(substage_unique_gen_age_grid$PageName))

  # Extract just the Grid cell names that have 3 or more unique genera
  substage_grid_greater_cam_sqs_gen <- substage_grid_freq %>% filter(Freq >15)
  colnames(substage_grid_greater_cam_sqs_gen) <- c("PageName","freq_gen")

  return(substage_grid_greater_cam_sqs_gen)
}

# Function to get vector of grid cell names with greater than MAA for SQS limit of 18 unique genera
get_great_maa_sqs_grids <- function(data,Age){
  # Get just the columns of Genus, Age, and Grid cell name
  gen_age_grid <- data[,c("Updated_Genus","Substage_from_Zone.Mbr","PageName")]

  # Remove duplicates from the list
  unique_gen_age_grid <- unique(gen_age_grid)

  # Get just the substage information in question
  substage_unique_gen_age_grid <- unique_gen_age_grid %>% filter(Substage_from_Zone.Mbr == Age)

  # Get the number of unique genera in each grid cell in the substage
  substage_grid_freq <- data.frame(table(substage_unique_gen_age_grid$PageName))

  # Extract just the Grid cell names that have 3 or more unique genera
  substage_grid_greater_maa_sqs_gen <- substage_grid_freq %>% filter(Freq >17)
  colnames(substage_grid_greater_maa_sqs_gen) <- c("PageName","freq_gen")

  return(substage_grid_greater_maa_sqs_gen)
```

```
}

# use above function to create list of Grid cells with >3 occ
maa_up_grid_greater_maa_sqs_gen <- get_great_maa_sqs_grids(data,"MAA (up)")
maa_low_grid_greater_maa_sqs_gen <- get_great_maa_sqs_grids(data,"MAA (low)")
cam_up_grid_greater_cam_sqs_gen <- get_great_cam_sqs_grids(data,"CAM (up)")
cam_mid_grid_greater_cam_sqs_gen <- get_great_cam_sqs_grids(data,"CAM (mid)")
cam_low_grid_greater_cam_sqs_gen <- get_great_cam_sqs_grids(data,"CAM (low)")

# get subset of the original db for each substage
maa_up_data1 <- as.data.frame(subset(data, Substage_from_Zone.Mbr =="MAA (up)"))
maa_low_data1 <- as.data.frame(subset(data, Substage_from_Zone.Mbr =="MAA (low)"))
cam_up_data1 <- as.data.frame(subset(data, Substage_from_Zone.Mbr =="CAM (up)"))
cam_mid_data1 <- as.data.frame(subset(data, Substage_from_Zone.Mbr =="CAM (mid)"))
cam_low_data1 <- as.data.frame(subset(data, Substage_from_Zone.Mbr =="CAM (low)"))

# merge the two dfs together to get extra column that has NA if less than 3 genera in the grid cell in
that SS
maa_up_data1_merge <-
merge(maa_up_data1,maa_up_grid_greater_maa_sqs_gen,by="PageName",all.x = TRUE)
maa_low_data1_merge <-
merge(maa_low_data1,maa_low_grid_greater_maa_sqs_gen,by="PageName",all.x = TRUE)
cam_up_data1_merge <-
merge(cam_up_data1,cam_up_grid_greater_cam_sqs_gen,by="PageName",all.x = TRUE)
cam_mid_data1_merge <-
merge(cam_mid_data1,cam_mid_grid_greater_cam_sqs_gen,by="PageName",all.x = TRUE)
cam_low_data1_merge <-
merge(cam_low_data1,cam_low_grid_greater_cam_sqs_gen,by="PageName",all.x = TRUE)

# remove rows with NA in the freq_gen column
maa_up_data1_data2 <- as.data.frame(subset(maa_up_data1_merge, !is.na(freq_gen)))
```

```
 maa_low_data1_data2 <- as.data.frame(subset(maa_low_data1_merge, !is.na(freq_gen)))
 cam_up_data1_data2 <- as.data.frame(subset(cam_up_data1_merge, !is.na(freq_gen)))
 cam_mid_data1_data2 <- as.data.frame(subset(cam_mid_data1_merge, !is.na(freq_gen)))
 cam_low_data1_data2 <- as.data.frame(subset(cam_low_data1_merge, !is.na(freq_gen)))

 # bind the SS dfs back together
 all_substage_greater_sqs_data <- rbind(maa_up_data1_data2,maa_low_data1_data2,
                          cam_up_data1_data2,cam_mid_data1_data2,
                          cam_low_data1_data2)

 # remove the column of freq_gen
 all_substage_greater_sqs_data <-
all_substage_greater_sqs_data[,c(1:ncol(all_substage_greater_sqs_data)-1)]

 return(all_substage_greater_sqs_data)
}

# Use above function to remove all rows from the DB that represent grid cells with
# fewer than sqs number of unique genera
join60_great_sqs_data <- get_greater_sqs_database(join60_data)

#### IF GOING TO SUBSAMPLE BY WAY OF USING ONLY GRID CELLS WITH >SQS estimate of UNIQUE
GENERA,
## USE THE ABOVE CODE AND THEN THE NEXT LINE TO REPLACE THE DB USED IN THE REST OF
## THE CODE WITH ONE WITH >3 UNIQUE GENERA GRID CELLS

# join60_data <- join60_great_sqs_data
```

```
## Get the number of nodes removed from each substage.
x = subset(join60_data,Substage_from_Zone.Mbr=="MAA (up)")
a = length(unique(x$PageName))
y = subset(join60_great_sqs_data,Substage_from_Zone.Mbr=="MAA (up)")
b = length(unique(y$PageName))


x1 = subset(join60_data,Substage_from_Zone.Mbr=="MAA (low)")
a1 = length(unique(x1$PageName))
y1 = subset(join60_great_sqs_data,Substage_from_Zone.Mbr=="MAA (low)")
b1 = length(unique(y1$PageName))

x2 = subset(join60_data,Substage_from_Zone.Mbr=="CAM (up)")
a2 = length(unique(x2$PageName))
y2 = subset(join60_great_sqs_data,Substage_from_Zone.Mbr=="CAM (up)")
b2 = length(unique(y2$PageName))


x3 = subset(join60_data,Substage_from_Zone.Mbr=="CAM (mid)")
a3 = length(unique(x3$PageName))
y3 = subset(join60_great_sqs_data,Substage_from_Zone.Mbr=="CAM (mid)")
b3 = length(unique(y3$PageName))


x4 = subset(join60_data,Substage_from_Zone.Mbr=="CAM (low)")
a4 = length(unique(x4$PageName))
y4 = subset(join60_great_sqs_data,Substage_from_Zone.Mbr=="CAM (low)")
b4 = length(unique(y4$PageName))
```

```
# Number removed

(a-b)+(a1-b1)+(a2-b2)+(a3-b3)+(a4-b4)


# Origional number of nodes

a+a1+a2+a3+a4


#### CHANGE DIRECTORY FOR RESULTS TO BE SAVED IN! (This should be changed based on subsampling/subsetting of the original df) ####

# setwd("C:/Users/Ceara/Documents/Province Project/FE_Analysis_Code-and-Files/FE_base_R_Analysis_OutputFiles/base_data_all_outputs")

# setwd("C:/Users/Ceara/Documents/Province Project/FE_Analysis_Code-and-Files/FE_base_R_Analysis_OutputFiles/greater_3_data_all_outputs")

setwd("C:/Users/Ceara/Documents/Province Project/FE_Analysis_Code-and-Files/FE_base_R_Analysis_OutputFiles/compress_data_all_outputs")

# setwd("C:/Users/Ceara/Documents/Province Project/FE_Analysis_Code-and-Files/FE_base_R_Analysis_OutputFiles/compress_data_360_outputs")

# setwd("C:/Users/Ceara/Documents/Province Project/FE_Analysis_Code-and-Files/FE_base_R_Analysis_OutputFiles/compress_greater_3_data_all_outputs")

# setwd("C:/Users/Ceara/Documents/Province Project/FE_Analysis_Code-and-Files/FE_base_R_Analysis_OutputFiles/compress_wis_data_all_outputs")

# setwd("C:/Users/Ceara/Documents/Province Project/FE_Analysis_Code-and-Files/FE_base_R_Analysis_OutputFiles/compress_gcp_data_all_outputs")

# setwd("C:/Users/Ceara/Documents/Province Project/FE_Analysis_Code-and-Files/FE_base_R_Analysis_OutputFiles/compress_bivalve_data_outputs")

# setwd("C:/Users/Ceara/Documents/Province Project/FE_Analysis_Code-and-Files/FE_base_R_Analysis_OutputFiles/compress_gastropod_data_outputs")

# setwd("C:/Users/Ceara/Documents/Province Project/FE_Analysis_Code-and-Files/FE_base_R_Analysis_OutputFiles/compress_cephalopod_data_outputs")

# setwd("C:/Users/Ceara/Documents/Province Project/FE_Analysis_Code-and-Files/FE_base_R_Analysis_OutputFiles/compress_wis_360_data_outputs")

# setwd("C:/Users/Ceara/Documents/Province Project/FE_Analysis_Code-and-Files/FE_base_R_Analysis_OutputFiles/compress_sqs_all_data_outputs")

# setwd("C:/Users/Ceara/Documents/Province Project/FE_Analysis_Code-and-Files/FE_base_R_Analysis_OutputFiles/compress_sqs_360_data_outputs")
```

```
#### Create DF of trait information (motility, life_habitat, and feeding) and simplify DB columns ####

# Create data frame of functional traits for imputing into mFD analysis
trait_df <- data.frame(c("motility", "life_habitat", "feeding"),
                       c("N", "N", "N"))
names(trait_df) <- c("trait_name", "trait_type")
head(trait_df)

# Convert data to simple dfs that can be used in mFD analyses (subset columns)
simplify_data <- function(data){
  data1 <- data.frame(data) # make into a df
  data2 <- subset(data1, select = c("Updated_Genus", "Substage_from_Zone.Mbr",
                                    "motility", "life_habitat", "feeding",
                                    "PageName","geometry"))
  return(data2)
}

simple60_data <- simplify_data(join60_data)
ncol(simple60_data)
colnames(simple60_data)



#### Separate Data into Substages ####

# Function for splitting into substages, remove duplicate Genus-FE rows, make
# genus name the index
substg_split <- function(data,x){
  data1 <- as.data.frame(subset(data, Substage_from_Zone.Mbr ==x))
```

```
  data2 <-distinct(data1,Updated_Genus,PageName,.keep_all=TRUE) # delete duplicate genus-Grid Cell values
  return(data2)
}



# Create individual substage trait data frames for different genera using substg_split function
maa_up_60_sbstg<- substg_split(simple60_data,"MAA (up)")
maa_low_60_sbstg <- substg_split(simple60_data,"MAA (low)")
cam_up_60_sbstg <- substg_split(simple60_data,"CAM (up)")
cam_mid_60_sbstg <- substg_split(simple60_data,"CAM (mid)")
cam_low_60_sbstg <- substg_split(simple60_data,"CAM (low)")



# Write CSV for each subsampled substage dataset with grid cell values
write.csv(maa_up_60_sbstg,file="MaaUP60_substage_FE.csv")
write.csv(maa_low_60_sbstg,file="MaaLOW60_substage_FE.csv")
write.csv(cam_up_60_sbstg,file="CamUP60_substage_FE.csv")
write.csv(cam_mid_60_sbstg,file="CamMID60_substage_FE.csv")
write.csv(cam_low_60_sbstg,file="CamLOW60_substage_FE.csv")

#### Create df and pres-abs matrix of Substages to use in mFD analysis ####
# Function to remove extra columns from the substage data frames (keep only FE
# traits and grid loc "PageName" with genus as index still)
substg_trait_df <- function(data){
  data1 <- as.data.frame(subset(data, select = c("Updated_Genus","motility", "life_habitat", "feeding", "PageName")))
  data2 <- data1 %>% distinct(Updated_Genus, .keep_all = TRUE)
  data2 <- data.frame(data2[,-1], row.names = data2[,1]) # make the first column with genus names the index
```

```
  return(data2)
}

# Remove unnecessary columns from the data frames to get just the trait
# information for each gen
maa_up_60_trait_data<- substg_trait_df(maa_up_60_sbstg)
maa_low_60_trait_data <- substg_trait_df(maa_low_60_sbstg)
cam_up_60_trait_data <- substg_trait_df(cam_up_60_sbstg)
cam_mid_60_trait_data <- substg_trait_df(cam_mid_60_sbstg)
cam_low_60_trait_data <- substg_trait_df(cam_low_60_sbstg)

# Check number of rows (# genera, also generic richness because no duplicates)
nrow(maa_up_60_trait_data)
nrow(maa_low_60_trait_data)
nrow(cam_up_60_trait_data)
nrow(cam_mid_60_trait_data)
nrow(cam_low_60_trait_data)


# Function to create a simple df of trait values with all variables as factors
# need this later in the code, when calculating sp.to.fe,
df_factor_simple <- function(data){
  data2 <- as.data.frame(unclass(data[c(1:4)]), stringsAsFactors = TRUE) # make the columns into factors
  row.names(data2) <- row.names(data) # make the first column with genus names the index
  data3 <- data2[,c("motility","life_habitat","feeding")]
  return(data3)
}

# create simple factor variable df for all substages
```

```
maa_up_60_trait_data_factor <- df_factor_simple(maa_up_60_trait_data)
maa_low_60_trait_data_factor <- df_factor_simple(maa_low_60_trait_data)
cam_up_60_trait_data_factor <- df_factor_simple(cam_up_60_trait_data)
cam_mid_60_trait_data_factor <- df_factor_simple(cam_mid_60_trait_data)
cam_low_60_trait_data_factor <- df_factor_simple(cam_low_60_trait_data)

class(cam_low_60_trait_data_factor$motility) # check class of column



# Get a presence-absence matrix for the substages based on grid cells
## NOTE: I redo this later by creating Latitudinal Bins that will act as "collections", instead of grid cells
# see below section on latitudinal bins.

## Function for data substages removing all columns but Genus
# names and Grid Cell names, and for creating presence-absence matrixes for
# each grid cell as collection locations
substg_genus_grid <- function(data){
  data1 <- as.data.frame(subset(data)[,c("Updated_Genus","PageName")]) # subset out just the genus name and site
  data2 <- dcast(data1, PageName~Updated_Genus, length) # transform into a pres-abs matrix
  data3 <- data.frame(data2[,-1], row.names = data2[,1]) # make the site names the row names
  return(data3)
}

# Create data frames for pres-abs for each substage based on grid cells
# use the substage df that still has Grid Names and geometry
maa_up_60_pres_ab <- substg_genus_grid(maa_up_60_sbstg)
maa_low_60_pres_ab <- substg_genus_grid(maa_low_60_sbstg)
```

```
cam_up_60_pres_ab <- substg_genus_grid(cam_up_60_sbstg)
cam_mid_60_pres_ab <- substg_genus_grid(cam_mid_60_sbstg)
cam_low_60_pres_ab <- substg_genus_grid(cam_low_60_sbstg)

# check the number of columns (# genera) against the number of rows in original
# df (again, # genera)
ncol(maa_up_60_pres_ab)
ncol(maa_low_60_pres_ab)
ncol(cam_up_60_pres_ab)
ncol(cam_mid_60_pres_ab)
ncol(cam_low_60_pres_ab)



#### RUN mFD ANALYSIS FOR DATABASE ####
 #### Summarize Substage data using sp.tr.summary function ####

# Summarize data using function from mFD pakage:
maa_up_60_traits_summ <- mFD::sp.tr.summary(tr_cat = trait_df,
                                sp_tr = maa_up_60_trait_data_factor,
                                stop_if_NA = TRUE)

maa_low_60_traits_summ <- mFD::sp.tr.summary(tr_cat = trait_df,
                                 sp_tr = maa_low_60_trait_data_factor,
                                 stop_if_NA = TRUE)

cam_up_60_traits_summ <- mFD::sp.tr.summary(tr_cat = trait_df,
                                 sp_tr = cam_up_60_trait_data_factor,
                                 stop_if_NA = TRUE)
```

```
cam_mid_60_traits_summ <- mFD::sp.tr.summary(tr_cat = trait_df,
                                sp_tr = cam_mid_60_trait_data_factor,
                                stop_if_NA = TRUE)

cam_low_60_traits_summ <- mFD::sp.tr.summary(tr_cat = trait_df,
                                sp_tr = cam_low_60_trait_data_factor,
                                stop_if_NA = TRUE)

# Check trait types
maa_up_60_traits_summ$"tr_types"

# Check trait type details on levels
maa_up_60_traits_summ$"mod_list"

# Create df of Trait summary list for each substage:
maa_up_60_traits_list <- maa_up_60_traits_summ$"tr_summary_list"
maa_low_60_traits_list <- maa_low_60_traits_summ$"tr_summary_list"
cam_up_60_traits_list <- cam_up_60_traits_summ$"tr_summary_list"
cam_mid_60_traits_list <- cam_mid_60_traits_summ$"tr_summary_list"
cam_low_60_traits_list <- cam_low_60_traits_summ$"tr_summary_list"



# write the summary lists of traits to csv files:
write.csv(maa_up_60_traits_list,file="substages_outputs/maa_up_60_traits_list.csv")
write.csv(maa_low_60_traits_list,file="substages_outputs/maa_low_60_traits_list.csv")
write.csv(cam_up_60_traits_list,file="substages_outputs/cam_up_60_traits_list.csv")
write.csv(cam_mid_60_traits_list,file="substages_outputs/cam_mid_60_traits_list.csv")
write.csv(cam_low_60_traits_list,file="substages_outputs/cam_low_60_traits_list.csv")
```

```
 # NOTE: This just gives summary information of the trait data, not actual

  # counts of FE, etc. Can run on more SS if necessary using above lines.

 #### Compute sp.to.fe mFD function to get functional entity information for each Substage ####


## Function for summarizing genera into functional entities

  # This function will take the genera from each substage, classify them into FE, then give information about

  # the distribution of genera in FE, list of FE, etc.

sum_fe <- function(data){

 data1 <- mFD::sp.to.fe(

  sp_tr      = data, # df of trait values for each genus, with genera names as row index IDs

  tr_cat     = trait_df, # df of traits, includes just "motility","feeding", and "life_habitat"

  fe_nm_type  = "fe_rank", # tell to name FE by ranked # genera in each (can't use trait names, too many)

  check_input = TRUE) # tell function to give error description if issues in dfs

 return(data1)

}


  ### Possible outputs from above function: ###

    # $fe_nm > gives a vector of the FE names (will be ranked list based on # genera in each, not based on traits)

    # $sp_fe > gives a vector of genera names and their FE in the substage; order based on decreasing # genera

    # $fe_tr > gives a df of the FE and their traits (motility, life habitat, feeding) in columns with FE as rows

    # $fe_nb_sp > gives a vector with species number per FE. If all FE have only one species, a warning message

      # is returned. FE are ordered according to the decreasing number of species they gather.

    # $details_fe > gives a list containing:
```

```
        # $fe_codes > a vector containing character referring to traits values (like a barcode) with names as
                    # in fe_nm_type and sorted according to fe_nb_sp
        # tr_uval > a list containing for each trait a vector of its unique values or a data frame for fuzzy-coded traits
        # fuzzy_E > a list with for each fuzzy-coded trait a data frame with names of entities (E) and names of species (sp)
        # tr_nb_uval > a vector with number of unique values per trait (or combinations for fuzzy-coded traits)
        # max_nb_fe > the maximum number of FE possible given number of unique values per trait.



# Classify each substage into FE using above function
maa_up_60_traits <- sum_fe(maa_up_60_trait_data_factor)
maa_low_60_traits <- sum_fe(maa_low_60_trait_data_factor)
cam_up_60_traits <- sum_fe(cam_up_60_trait_data_factor)
cam_mid_60_traits <- sum_fe(cam_mid_60_trait_data_factor)
cam_low_60_traits <- sum_fe(cam_low_60_trait_data_factor)

    ## NOTE: the above trait outputs will be used later to calc FR and evenness for each SS overall using the $fe_nb_sp
      # option; see later section.



# Look at the FE names
maa_up_60_traits$"fe_nm"
maa_low_60_traits$"fe_nm"
cam_up_60_traits$"fe_nm"
cam_mid_60_traits$"fe_nm"
cam_low_60_traits$"fe_nm"  # different numbers of FE found for each substage, but roughly the same amount
```

```
# Look at which FE each genus belongs to
sp_fe_maa_up_60_traits <- maa_up_60_traits$"sp_fe"
sp_fe_maa_low_60_traits <- maa_low_60_traits$"sp_fe"
sp_fe_cam_up_60_traits <- cam_up_60_traits$"sp_fe"
sp_fe_cam_mid_60_traits <- cam_mid_60_traits$"sp_fe"
sp_fe_cam_low_60_traits <- cam_low_60_traits$"sp_fe"

# Look at number of genera per FE
fe_nb_sp_maa_up_60_traits <- maa_up_60_traits$"fe_nb_sp" # yes, genus names kept!! hurray!
fe_nb_sp_maa_up_60_traits

#### Use SQS to check subsampling size for adequate representation of FEs ####

## This is intended to give an idea about how low sampling can go before FE are
  # not well represented in each SS. Important for nodes, where there are very
  # few total genera, and even fewer unique ones...
## Use the above table lists of FE for each genera.

# Create Frequency lists for each FE in each SS:
maa_up_60_fe_freq <-table(sp_fe_maa_up_60_traits) # all data frequencies of genera
maa_low_60_fe_freq <-table(sp_fe_maa_low_60_traits) # all data frequencies of genera
cam_up_60_fe_freq <-table(sp_fe_cam_up_60_traits) # all data frequencies of genera
cam_mid_60_fe_freq <-table(sp_fe_cam_up_60_traits) # all data frequencies of genera
cam_low_60_fe_freq <-table(sp_fe_cam_up_60_traits) # all data frequencies of genera

# Use SQS subsampling function (see above) to get estimate for sampling level:
maa_up_sqs_sample_for_fe <- run_sqs(maa_up_60_fe_freq)
maa_low_sqs_sample_for_fe <- run_sqs(maa_low_60_fe_freq)
```

```
cam_up_sqs_sample_for_fe <- run_sqs(cam_up_60_fe_freq)
cam_mid_sqs_sample_for_fe <- run_sqs(cam_mid_60_fe_freq)
cam_low_sqs_sample_for_fe <- run_sqs(cam_low_60_fe_freq)

# CHANGE IF USING WIS ONLY: Bind into a df and give row names for each SS:
sqs_sample_size_fe <- data.frame(c(maa_up_sqs_sample_for_fe, maa_low_sqs_sample_for_fe,
                cam_up_sqs_sample_for_fe, cam_mid_sqs_sample_for_fe,
                cam_low_sqs_sample_for_fe))

# sqs_sample_size_fe <- data.frame(c(maa_low_sqs_sample_for_fe,
                        # cam_up_sqs_sample_for_fe, cam_mid_sqs_sample_for_fe,
                        # cam_low_sqs_sample_for_fe))

rownames(sqs_sample_size_fe) <- c("Maa UP","Maa LOW","Cam UP","Cam MID","Cam LOW")

# rownames(sqs_sample_size_fe) <- c("Maa LOW","Cam UP","Cam MID","Cam LOW")

colnames(sqs_sample_size_fe) <- "SQS Size Estimate"

# Write to csv
write.csv(sqs_sample_size_fe,file="grid_collections_outputs/sqs_sample_size_fe.csv")



#### ANALYSIS FOR GRID CELLS AS COLLECTIONS ####
 #### Compute alpha.fd.fe mFD function to get alpha and beta diversity information for each Substage
####

## Function to calc functional trait values
calc_fe <- function(trait_data,pres_ab_data){
```

```
  alpha_fd_fe_MAAup <- mFD::alpha.fd.fe(
    asb_sp_occ      = pres_ab_data,
    sp_to_fe        = trait_data,
    ind_nm          = c("fred", "fored", "fvuln"),
    check_input     = TRUE,
    details_returned = TRUE)
}

# Calculate functional richness, redundancy, over redundancy, and vulnerability for each substage in grid cells
maa_up_60_fe <- calc_fe(maa_up_60_traits, maa_up_60_pres_ab)
maa_low_60_fe<- calc_fe(maa_low_60_traits, maa_low_60_pres_ab)
cam_up_60_fe <- calc_fe(cam_up_60_traits, cam_up_60_pres_ab)
cam_mid_60_fe <- calc_fe(cam_mid_60_traits, cam_mid_60_pres_ab)
cam_low_60_fe <- calc_fe(cam_low_60_traits, cam_low_60_pres_ab)

# Get a matrix of fe values for each grid cell in a time bin
maa_up_60_fdfe <- (maa_up_60_fe$"asb_fdfe")
maa_low_60_fdfe <- (maa_low_60_fe$"asb_fdfe")
cam_up_60_fdfe <- (cam_up_60_fe$"asb_fdfe")
cam_mid_60_fdfe <- (cam_mid_60_fe$"asb_fdfe")
cam_low_60_fdfe <- (cam_low_60_fe$"asb_fdfe")

# Look at the df of the alpha-beta diversity outputs
maa_up_60_fdfe
maa_low_60_fdfe
cam_up_60_fdfe
cam_mid_60_fdfe
cam_low_60_fdfe
```

```
# Export results tables summarizing the number of spp, fe, and functional
  # richness and evenness values for each substage
write.csv(maa_up_60_fdfe,file="grid_collections_outputs/MaaUP60_fe_alpha_beta_results.csv")
write.csv(maa_low_60_fdfe,file="grid_collections_outputs/MaaLOW60_fe_alpha_beta_results.csv")
write.csv(cam_up_60_fdfe,file="grid_collections_outputs/CamUP60_fe_alpha_beta_results.csv")
write.csv(cam_mid_60_fdfe,file="grid_collections_outputs/CamMID60_fe_alpha_beta_results.csv")
write.csv(cam_low_60_fdfe,file="grid_collections_outputs/CamLOW60_fe_alpha_beta_results.csv")

# Function to check for diff between two data frames genus name lists, only use
  # if error in cal FR above
check_diff_sp <- function(trait_data,pres_ab_data){
  df <- as.data.frame(trait_data$"sp_fe")
  df <- cbind(newColName = rownames(df), df)
  rownames(df) <- NULL
  df <- df[,1]

  dd <- as.data.frame(colnames(pres_ab_data))
  dd <- dd[,1]
  head(dd)

  df <- sort(df)
  dd <- sort(dd)
  diff <-setdiff(df,dd)
  return(diff)
}

#### Look at relationship between Generic and Functional Richness ####
```

```
# Use iNext to create rarifaction curve of the number genera per number of FEs
library(iNEXT)

data(spider)
str(spider)

out <- iNEXT(spider, q=0, datatype="abundance", endpoint=500)

aa <- fe_data1[,2:3]

bb <- iNEXT(aa, q=0, datatype="abundance", size=NULL, endpoint=NULL, knots=40, se=TRUE, conf=0.95,
    nboot=50)



out <- iNEXT(spider, q=c(0, 1, 2), datatype="abundance", endpoint=500)
# Sample-size-based R/E curves, separating plots by "site"
ggiNEXT(out, type=1, facet.var="site")

# function to get the Age data as first column
df_func <- function(data){
 ###  Get the functional ecology information for the substage (this is a dataframe) and reconfig: ###
 data1 <- cbind(Bin = rownames(data), data) # make the grid name the 1st column again
 rownames(data1) <- 1:nrow(data1) # make a new index for the rownames
 data1 <- as.data.frame(data1)
 i <- c(2:6)
 data1[ , i] <- apply(data1[ , i], 2,          # Specify own function within apply to make numeric
                  function(x) as.numeric(as.character(x)))
 return(as.data.frame(data1))
}
```

```
fe_data1 <- df_func(maa_up_60_fdfe)
fe_data2 <- df_func(maa_low_60_fdfe)
fe_data3 <- df_func(cam_up_60_fdfe)
fe_data4 <- df_func(cam_mid_60_fdfe)
fe_data5 <- df_func(cam_low_60_fdfe)

data_fe_all <- rbind(fe_data1,fe_data2,fe_data3,fe_data4,fe_data5)



# perform Shapiro test for normality on the GR and FR values from grid cells
shapiro.test(data_fe_all[,"nb_sp"]) # very much not normal
shapiro.test(data_fe_all[,"nb_fe"]) # very much not normal



# make plot of all the GR vs FR values for each grid cell in all substages
p <- ggplot(fe_data1, aes(x = nb_sp, y = nb_fe)) +
  geom_point(color = "darkred") +
  geom_point(data = fe_data2, aes(x = nb_sp, y = nb_fe), color = "#D55E00") +
  geom_point(data = fe_data3, aes(x = nb_sp, y = nb_fe), color = "goldenrod3") +
  geom_point(data = fe_data4, aes(x = nb_sp, y = nb_fe), color = "darkolivegreen4") +
  geom_point(data = fe_data5, aes(x = nb_sp, y = nb_fe), color = "steelblue4")

# add different models to the data to see which fits best
p + stat_smooth(method = "lm", formula = y ~ x, size = 1, se = FALSE, colour = "black") +
  stat_smooth(method = "lm", formula = y ~ x + I(x^2), size = 1, se = FALSE, colour = "blue") +
  stat_smooth(method = "loess", formula = y ~ x, size = 1, se = FALSE, colour = "red") +
  stat_smooth(method = "gam", formula = y ~ s(x), size = 1, se = FALSE, colour = "green") +
  stat_smooth(method = "gam", formula = y ~ s(x, k = 3), size = 1, se = FALSE, colour = "violet") +
```

```
  stat_smooth(method = "lm", formula = y ~ sqrt(x), size = 1, se = FALSE, colour = "goldenrod") +
  stat_smooth(method = "lm", formula = log(y) ~ log(x), size = 1, se = FALSE, colour = "brown4")

# look at most likely models of those above
p + stat_smooth(method = "lm", formula = y ~ x + I(x^2), size = 1, se = FALSE, colour = "blue") +
  stat_smooth(method = "loess", formula = y ~ x, size = 1, se = FALSE, colour = "red") +
  stat_smooth(method = "gam", formula = y ~ s(x, k = 3), size = 1, se = FALSE, colour = "violet") +
  stat_smooth(method = "lm", formula = y ~ sqrt(x), size = 1, se = FALSE, colour = "goldenrod")



# look at model that appears to fit best of those above
p + stat_smooth(method = "lm", formula = y ~ sqrt(x), size = 1, colour = "goldenrod") +

  p + stat_smooth(method = "lm", formula = y ~ x + I(x^2), size = 1, colour = "blue")

# It appears, based on the above plots, that other than the loess and gam models,
# the square root and squared models fit the data best, so tested below:

# create square root model
model_sr<-lm(fe_data1$nb_sp ~ I(sqrt(fe_data1$nb_fe)), data = fe_data1)
summary(model_sr)

# create squared model (this one appears to have a higher R2 so probably slightly better fit)
model_sq<-lm(fe_data1$nb_sp ~ I((fe_data1$nb_fe)^2), data = fe_data1)
summary(model_sq)

names(summary(model_sq))

summary(model_sq)$r.squared # extremely high R2 value, indicates good fit
```

```
summary(model_sq)$coefficients

summary(model_sq)$adj.r.squared


# plot the information related to the regression model

par(mfrow = c(2, 2))

plot(model_sq)


# export plots as pdf

pdf("grid_collections_outputs/squared_regression_model_GRvsFR_plots.pdf", width = 11, height = 11) # Open a new pdf file

par(mfrow = c(2, 2))

plot(model_sq)

dev.off() # Close the file


# export plot of GR vs FR within grid cells for each SS and regression line:

pdf("grid_collections_outputs/GRvsFR_regression_plots.pdf", width = 6, height = 6) # Open a new pdf file

p + stat_smooth(method = "lm", formula = y ~ x + I(x^2), size = 1, colour = "blue")

dev.off() # Close the file



# Based on the above plots, only one point counts as an outlier which may influence

# the regression (>1 cook's distance), but the residuals indicate a linear

# relationship in the data (Residual vs Fitted). The Scale-location plot suggests

# that the data have fairly homogeneous variance values, except at low fitted

# values where the variance is quite low. The Q-Q plots show non-normality,

# which is expected given the Shapiro-Wilk tests performed above.
```

```
# Results show that there is a high correlation between GR and FR, which appears to potentially
# plateau at the very highest levels of GR (>150), and potentially follow a squared
# model function (i.e., peaking at ~150 genera before the correlation becomes negative).
# Either way, this should be considered when examining the results of the analysis.
# When using lower sampling (correlated strongly with GR), GR is low, and therefore
# FR will probably be low as well. Might be the case that you can't really examine
# FD at local scales in the fossil record, given the limitations on the data. This issue
# may also become even more difficult when using subsampling like sqs, since it will
# necessarily decrease sampling overall and impacting sampling especially in localized
# regions.

 #### Use alpha.fd.fe.plot to Plot the results for each grid cell in each substage ####
 # (par(mfrow) wont work for ggplots so this is my work around)

   # Note: can't keep title of each subplot when putting them together. very annoying. Not sure
    # how to fix this since it seems like it's just part of the base code for plotting these.
    # Will have to manually add titles if necessary.



# Run a for loop function to create the plots and add them to the empty list
plot_alpha_beta_fe <- function(fe_data,sbstg_data){
 fe_plots = list() # create an empty list to store the grid plots in
 for(i in 1:length(unique(sbstg_data$PageName))){
  grid <- unique(sbstg_data$PageName)
  fe_plots[[i]] = mFD::alpha.fd.fe.plot(
    alpha_fd_fe      = fe_data,
    plot_asb_nm      = c(grid[i]),
    plot_ind_nm      = c("fred", "fored", "fvuln"),
```

```
    name_file       = NULL,
    color_fill_fored  = "darkolivegreen2",
    color_line_fred   = "darkolivegreen4",
    color_fill_bar    = "grey80",
    color_fill_fvuln  = "lightcoral",
    color_arrow_fvuln = "indianred4",
    size_line_fred    = 0.5,
    size_arrow_fvuln  = 0.5,
    check_input     = TRUE)
 }
 return(fe_plots)
}

maa_up_60_fe_plots <- plot_alpha_beta_fe(maa_up_60_fe,maa_up_60_sbstg)
maa_low_60_fe_plots <- plot_alpha_beta_fe(maa_low_60_fe,maa_low_60_sbstg)
cam_up_60_fe_plots <- plot_alpha_beta_fe(cam_up_60_fe,cam_up_60_sbstg)
cam_mid_60_fe_plots <- plot_alpha_beta_fe(cam_mid_60_fe,cam_mid_60_sbstg)
cam_low_60_fe_plots <- plot_alpha_beta_fe(cam_low_60_fe,cam_low_60_sbstg)

# export plots as single pdf
pdf("grid_collections_outputs/alpha_beta_allgrids_MAA_up_plots.pdf", width = 15, height = 25) # Open a new pdf file
grid.arrange(grobs = maa_up_60_fe_plots,
        top = "Upper Maastr. Alpha/Beta Diversity per Grid Cell") # Write the grid.arrange in the file
dev.off() # Close the file


# export plots as single pdf
```

```
pdf("grid_collections_outputs/alpha_beta_allgrids_MAA_low_plots.pdf", width = 15, height = 25) # Open a new pdf file

grid.arrange(grobs = maa_low_60_fe_plots,

        top = "Lower Maastr. Alpha/Beta Diversity per Grid Cell") # Write the grid.arrange in the file

dev.off() # Close the file


# export plots as single pdf

pdf("grid_collections_outputs/alpha_beta_allgrids_CAM_up_plots.pdf", width = 15, height = 25) # Open a new pdf file

grid.arrange(grobs = cam_up_60_fe_plots,

        top = "Upper Camp. Alpha/Beta Diversity per Grid Cell") # Write the grid.arrange in the file

dev.off() # Close the file


# export plots as single pdf

pdf("grid_collections_outputs/alpha_beta_allgrids_CAM_mid_plots.pdf", width = 15, height = 25) # Open a new pdf file

grid.arrange(grobs = cam_mid_60_fe_plots,

        top = "Middle Camp. Alpha/Beta Diversity per Grid Cell") # Write the grid.arrange in the file

dev.off() # Close the file


# export plots as single pdf

pdf("grid_collections_outputs/alpha_beta_allgrids_CAM_low_plots.pdf", width = 15, height = 25) # Open a new pdf file

grid.arrange(grobs = cam_low_60_fe_plots,

        top = "Lower Camp. Alpha/Beta Diversity per Grid Cell") # Write the grid.arrange in the file

dev.off() # Close the file


# Clearly, many grid cells contain very few functional entities AND genera

# think about removing cells with fewer than 3 genera...
```

```
#### Create df of FR for each Grid Cell through time####

  # This section is firstly trying to make matrices which will show richness values in each
   # grid cell across the five substages. Many grid cells do not have consistent occupation,
   # however. Next, will be adding the general location of the grid cells (lat/lon) to the FR df
   # so that these tables can be exported and plotted in on maps.

# Function to configure to columns with site names as 1st and FR values as second for FUNCTIONAL RICHNESS
config_for_loc_agg <- function(data, Age){
 # Create df that just take the functional richness values in each grid cell
 data_rich_summ <- (data$"asb_fdfe")[,2]

 data_rich_summ <- data.frame(data_rich_summ)
 data_rich_summ <- cbind(newColName = rownames(data_rich_summ), data_rich_summ) # make grid cell first column
 rownames(data_rich_summ) <- 1:nrow(data_rich_summ)
 colnames(data_rich_summ) <- c("Site","Numb_FE")

 data_rich_summ$Age <- Age
 return(data_rich_summ)
}

# Use config function (above) to alter then added new column with Age
maa_up_60_func_rich_summ <- config_for_loc_agg(maa_up_60_fe, "MAA up")
maa_low_60_func_rich_summ <- config_for_loc_agg(maa_low_60_fe, "MAA low")
cam_up_60_func_rich_summ <- config_for_loc_agg(cam_up_60_fe, "CAM up")
cam_mid_60_func_rich_summ <- config_for_loc_agg(cam_mid_60_fe, "CAM mid")
cam_low_60_func_rich_summ <- config_for_loc_agg(cam_low_60_fe, "CAM low")
```

```
# Bind the data into a single df with 2 columns (stacked the different dfs)
func_rich_stack <- rbind(maa_up_60_func_rich_summ,maa_low_60_func_rich_summ,
                         cam_up_60_func_rich_summ,cam_mid_60_func_rich_summ,
                         cam_low_60_func_rich_summ)

# Used dcast to reconfigure the data so that Age and Site would be identifiers with unique values
func_rich_table_all <- dcast(func_rich_stack, Site ~ Age, value.var = "Numb_FE")



# This is a table showing the number of FE in each grid cell for each substage
head(func_rich_table_all)

write.csv(func_rich_table_all, file="grid_collections_outputs/Grid_FR_through_time.csv")

  #### Create df of GR for each Grid Cell through time####

# Function to configure to columns with site names as 1st and FR values as second for GENERIC
RICHNESS
config_for_loc_agg_gen_rich <- function(data, Age){
  # Create df that just take the functional richness values in each grid cell
  data_rich_summ <- (data$"asb_fdfe")[,1]

  data_rich_summ <- data.frame(data_rich_summ)
  data_rich_summ <- cbind(newColName = rownames(data_rich_summ), data_rich_summ) # make grid
cell first column
  rownames(data_rich_summ) <- 1:nrow(data_rich_summ)
  colnames(data_rich_summ) <- c("Site","Numb_G")
```

```
  data_rich_summ$Age <- Age
  return(data_rich_summ)
}

# Use config function (above) to alter then added new column with Age
maa_up_60_generic_rich_summ <- config_for_loc_agg_gen_rich(maa_up_60_fe, "MAA up")
maa_low_60_generic_rich_summ <- config_for_loc_agg_gen_rich(maa_low_60_fe, "MAA low")
cam_up_60_generic_rich_summ <- config_for_loc_agg_gen_rich(cam_up_60_fe, "CAM up")
cam_mid_60_generic_rich_summ <- config_for_loc_agg_gen_rich(cam_mid_60_fe, "CAM mid")
cam_low_60_generic_rich_summ <- config_for_loc_agg_gen_rich(cam_low_60_fe, "CAM low")

# Bind the data into a single df with 2 columns (stacked the different dfs)
generic_rich_stack <- rbind(maa_up_60_generic_rich_summ,maa_low_60_generic_rich_summ,
                cam_up_60_generic_rich_summ,cam_mid_60_generic_rich_summ,
                cam_low_60_generic_rich_summ)

# Used dcast to reconfigure the data so that Age and Site would be identifiers with unique values
generic_rich_table_all <- dcast(generic_rich_stack, Site ~ Age, value.var = "Numb_G")


# This is a table showing the number of FE in each grid cell for each substage
head(generic_rich_table_all)

write.csv(generic_rich_table_all, file="grid_collections_outputs/Grid_GR_through_time.csv")

  #### Get the average location for each grid cell per substage (use to group by lat and compare across
lat) ####
```

```
# Function to separate geometry back out into lat-long information:
sep_lat_substage <- function(data,substage){

  # separate the geometry values back into lat/lon
  data <- data %>%
    mutate(geometry = gsub('[POINT ()]', '', geometry)) %>%
    separate(col = geometry, into = c('Longitude', 'Latitude'), sep = '\\,')

  # remove the first character in lat string, don't know why but the above code adds a "c"
  data$Longitude <- sub('.', '', data$Longitude)

  # make the lat/lon columns into numeric values
  data$Longitude <- as.numeric(data$Longitude)
  data$Latitude <- as.numeric(data$Latitude)

  # separate the df by substages so can take average locations for each grid cell at each substage
  substg_data <- as.data.frame(subset(data, Substage_from_Zone.Mbr ==substage))
  substg_data %>% distinct(Updated_Genus,PageName,.keep_all=TRUE) # delete duplicate value

  return(substg_data)
}

# Create new df of all data with the lat/lon separated back out
maa_up_60_loc_data <-sep_lat_substage(join60_data, "MAA (up)")
maa_low_60_loc_data <-sep_lat_substage(join60_data, "MAA (low)")
cam_up_60_loc_data <-sep_lat_substage(join60_data, "CAM (up)")
cam_mid_60_loc_data <-sep_lat_substage(join60_data, "CAM (mid)")
cam_low_60_loc_data <-sep_lat_substage(join60_data, "CAM (low)")
```

```
# Function to separate geometry back out into lat-long information for entire db:
sep_lat <- function(data){

  # separate the geometry values back into lat/lon
  data <- data %>%
    mutate(geometry = gsub('[POINT ()]', '', geometry)) %>%
    separate(col = geometry, into = c('Longitude', 'Latitude'), sep = '\\,')

  # remove the first character in lat string, don't know why but the above code adds a "c"
  data$Longitude <- sub('.', '', data$Longitude)

  # make the lat/lon columns into numeric values
  data$Longitude <- as.numeric(data$Longitude)
  data$Latitude <- as.numeric(data$Latitude)

  return(data)
}

join60_data_lat_sep <-sep_lat(join60_data)
# this above format not used in separating by lat bins (below section), instead
# use the db table made by combining substages that have been given paleo coords,
# see next few lines.

# load the chronosphere package to run this code. This should allow for reconstructions of plat and plong:
# https://www.evolv-ed.net/post/chronosphere-paleomap/chronosphere-paleomap/
library(chronosphere)
```

```
# Function to get the plat and plon inforamtion for each occ:
estimate_ploc <- function(data,age){
  # get just the lat long locations from the table
  data_loc_2col <- data[,c("Longitude","Latitude")]

  # Reconstruct paleocoord using age of rough age of substage (approximate, will use the same age for
M Cam because limited options...)
  data_ploc <- as.data.frame(reconstruct(data_loc_2col, age=c(age)))

  colnames(data_ploc) <- c("plong","plat")

  ## Add the paleocoord back into the original location df
  data$plong <- data_ploc$plong
  data$plat <- data_ploc$plat
  return(data)
}

# Use above function to get the plat and plon information for each substage site
maa_up_60_paleo_loc_data <-estimate_ploc(maa_up_60_loc_data, 80)
maa_low_60_paleo_loc_data <-estimate_ploc(maa_low_60_loc_data, 80)
cam_up_60_paleo_loc_data <-estimate_ploc(cam_up_60_loc_data, 75)
cam_mid_60_paleo_loc_data <-estimate_ploc(cam_mid_60_loc_data, 70)
cam_low_60_paleo_loc_data <-estimate_ploc(cam_low_60_loc_data, 70)

# Write paleo location information to a new file set (MIGHT WANT THESE LATER IF CHRONOSPHERE
STOPS WORKING AGAIN...)
```

```
write.csv(maa_up_60_paleo_loc_data, file="maa_up_60_paleo_loc_data.csv")
write.csv(maa_low_60_paleo_loc_data, file="maa_low_60_paleo_loc_data.csv")
write.csv(cam_up_60_paleo_loc_data, file="cam_up_60_paleo_loc_data.csv")
write.csv(cam_mid_60_paleo_loc_data, file="cam_mid_60_paleo_loc_data.csv")
write.csv(cam_low_60_paleo_loc_data, file="cam_low_60_paleo_loc_data.csv")

# read in the csv data that was produced before for paleocoor information, to save comp time
maa_up_60_paleo_loc_data <- read.csv("maa_up_60_paleo_loc_data.csv")
maa_low_60_paleo_loc_data <- read.csv("maa_low_60_paleo_loc_data.csv")
cam_up_60_paleo_loc_data <- read.csv("cam_up_60_paleo_loc_data.csv")
cam_mid_60_paleo_loc_data <- read.csv("cam_mid_60_paleo_loc_data.csv")
cam_low_60_paleo_loc_data <- read.csv("cam_low_60_paleo_loc_data.csv")

# join the above dfs together to create a complete df of all occurrence information
join60_data_paleo_loc <- rbind(maa_up_60_paleo_loc_data,
    maa_low_60_paleo_loc_data,
    cam_up_60_paleo_loc_data,
    cam_mid_60_paleo_loc_data,
    cam_low_60_paleo_loc_data)

# write the entire db as a new file with paleo locations:
write.csv(join60_data_paleo_loc, file="join60_data_60_paleo_loc_data.csv")

  ## NOTE: the above df used later when binning by latitude, used to create "bins" that can be used
instead of
    # grid cells when calculating alpha-beta functional diversity


# Create function to aggregate location information based on grid cell and merge with grid IDs
```

```
agg_paleo_loc <- function(loc_data,rich_data){
 agg_lat <- aggregate(x = loc_data$plat,              # Specify data column
        by = list(loc_data$PageName),            # Specify group indicator
        FUN = mean)
 agg_lon <- aggregate(x = loc_data$plong,             # Specify data column
                by = list(loc_data$PageName),           # Specify group indicator
                FUN = mean)
 grid_loc <- merge(agg_lat,agg_lon,"Group.1") #merge into single df
 colnames(grid_loc) <- c("Site","pLat","pLon")
 rich_latlong <- merge(rich_data,grid_loc,by = "Site", all.x=TRUE)
 return(rich_latlong)
 }

# Aggregate and merge location information for each substage (above function)
   # these locations will be used to plot the locations of values on maps based on average grid cell locations
maa_up_60_grid_paleo_local <- agg_paleo_loc(maa_up_60_paleo_loc_data, maa_up_60_func_rich_summ)
maa_low_60_grid_paleo_local <- agg_paleo_loc(maa_low_60_paleo_loc_data, maa_low_60_func_rich_summ)
cam_up_60_grid_paleo_local <- agg_paleo_loc(cam_up_60_paleo_loc_data, cam_up_60_func_rich_summ)
cam_mid_60_grid_paleo_local <- agg_paleo_loc(cam_mid_60_paleo_loc_data, cam_mid_60_func_rich_summ)
cam_low_60_grid_paleo_local <- agg_paleo_loc(cam_low_60_paleo_loc_data, cam_low_60_func_rich_summ)

#Aggregate and merge location information for entire database (above function)
 # these locations will be used to group lat bins in plots based on average grid
 # cell locations
grid_local <- agg_paleo_loc(join60_data_paleo_loc, func_rich_table_all)
```

#### Create histograms of grid metric values for each substage ####

```
# function to create histograms of all metrics calculated by mFD for alpha and beta diversity
hist_plots <- function(data){
  sp_hist <- ggplot() +
    geom_histogram(as.data.frame(data), mapping = aes(x = nb_sp),
                   binwidth=5, fill="skyblue4", color="#e9ecef", alpha=0.9) +
    labs(title="Grid Cell Generic Richness (GR)",x="# Genera", y = "Frequency of Grid Cells") +
    theme(plot.title = element_text(hjust = 0.5))
  fe_hist <- ggplot() +
    geom_histogram(as.data.frame(data), mapping = aes(x = nb_fe),
                   binwidth=3, fill="skyblue4", color="#e9ecef", alpha=0.9) +
    labs(title="Grid Cell Functional Richness (FR)",x="# Functional Entities", y = "Frequency of Grid Cells")
+
    theme(plot.title = element_text(hjust = 0.5))
  fred_hist <- ggplot() +
    geom_histogram(as.data.frame(data), mapping = aes(x = fred),
                   binwidth=0.3, fill="skyblue4", color="#e9ecef", alpha=0.9) +
    labs(title="Grid Cell Functional Redundancy (FRed)",x="GR/FR", y = "Frequency of Grid Cells") +
    theme(plot.title = element_text(hjust = 0.5))
  fored_hist <- ggplot() +
    geom_histogram(as.data.frame(data), mapping = aes(x = fored),
                   binwidth=0.1, fill="skyblue4", color="#e9ecef", alpha=0.9) +
    labs(title="Grid Cell Functional Over-Redundancy (FORed)",x="FRed Values", y = "Frequency of Grid
Cells") +
    theme(plot.title = element_text(hjust = 0.5))
```

```
  fvuln_hist <- ggplot() +
    geom_histogram(as.data.frame(data), mapping = aes(x = fvuln),
                   binwidth=0.15, fill="skyblue4", color="#e9ecef", alpha=0.9) +
    labs(title="Grid Cell Functional Vulnerability (FVuln)",x="FVuln Values", y = "Frequency of Grid Cells") +
    theme(plot.title = element_text(hjust = 0.5))
  dist_list <- list(sp_hist,fe_hist,fred_hist,fored_hist,fvuln_hist)
}

# Create series of histograms for metric values in each substage
maa_up_60_grid_hist <- hist_plots(maa_up_60_fdfe)
maa_low_60_grid_hist <- hist_plots(maa_low_60_fdfe)
cam_up_60_grid_hist <- hist_plots(cam_up_60_fdfe)
cam_mid_60_grid_hist <- hist_plots(cam_mid_60_fdfe)
cam_low_60_grid_hist <- hist_plots(cam_low_60_fdfe)



pdf("grid_collections_outputs/hist_MAA_up_allgrids.pdf", width = 15, height = 25) # Open a new pdf file
grid.arrange(grobs = maa_up_60_grid_hist,
             top = "Upper Maastr. Histograms of Diversity Metrics per Grid Cell") # Write the grid.arrange in the file
dev.off() #close the file


pdf("grid_collections_outputs/hist_MAA_low_allgrids.pdf", width = 15, height = 25) # Open a new pdf file
grid.arrange(grobs = maa_low_60_grid_hist,
             top = "Lower Maastr. Histograms of Diversity Metrics per Grid Cell") # Write the grid.arrange in the file
dev.off() #close the file


pdf("grid_collections_outputs/hist_CAM_up_allgrids.pdf", width = 15, height = 25) # Open a new pdf file
```

```
grid.arrange(grobs = cam_up_60_grid_hist,
        top = "Upper Camp. Histograms of Diversity Metrics per Grid Cell") # Write the grid.arrange in the file
dev.off() #close the file

pdf("grid_collections_outputs/hist_CAM_mid_allgrids.pdf", width = 15, height = 25) # Open a new pdf file
grid.arrange(grobs = cam_mid_60_grid_hist,
        top = "Middle Camp. Histograms of Diversity Metrics per Grid Cell") # Write the grid.arrange in the file
dev.off() #close the file

pdf("grid_collections_outputs/hist_CAM_low_allgrids.pdf", width = 15, height = 25) # Open a new pdf file
grid.arrange(grobs = cam_low_60_grid_hist,
        top = "Lower Camp. Histograms of Diversity Metrics per Grid Cell") # Write the grid.arrange in the file
dev.off() #close the file

 #### Create bar plots of FE metrics for each grid cell in each Substage ####
# Function to create series of bar plots for each substage's grid cells
bar_plot <- function(fe_data, loc_data, substage){

 ###  Get the functional ecology information for the substage (this is a dataframe) and reconfig: ###
 data1 <- cbind(newColName = rownames(fe_data), fe_data) # make the genus name the 1st column again
 rownames(data1) <- 1:nrow(data1) # make a new index for the rownames
 data1 <- as.data.frame(data1)
 i <- c(2:6)
 data1[ , i] <- apply(data1[ , i], 2,          # Specify own function within apply to make numeric
                 function(x) as.numeric(as.character(x)))
```

```
colnames(data1) <- c("Grid", "nb_sp", "nb_fe", "fred", "fored", "fvuln") # set colnames



### Extract Location Information: ###
loc_data1 <- loc_data[c("Site",substage,"pLat","pLon")] # Extract columns for just one substage
colnames(loc_data1) <- c("Grid",substage,"pLat","pLon") # Rename columns
loc_data1$Lat_1 <- loc_data1$pLat # Add another column to use for binning lat
# Next, use mutate and cut to create binned categories of lat
loc_data2 <- loc_data1 %>% mutate(Lat_1 = cut(loc_data1$Lat_1,
                                  breaks=c(-Inf, seq(round(min(loc_data1$pLat)/5)*5,
                                                     round(max(loc_data1$pLat)/5)*5, by = 5), Inf)))



### Merge the new bins with the fe df: ###
data3 <- merge(data1,loc_data2, by= "Grid", all.x = TRUE)

# get the unique lat bins available for this substage and change colname
bins <- data.frame(unique(data3$Lat_1))
colnames(bins) <- "Lat_1"

# Create column with the sorts of labels actually want to use based on ifelse
bins$lat_bins <-with(bins, ifelse(Lat_1 == "(-Inf,30]", '23-30°N',
                          ifelse(Lat_1 == "(30,35]", '30-35°N',
                              ifelse(Lat_1 == "(35,40]", '35-40°N',
                                  ifelse(Lat_1 == "(40,45]", '40-45°N',
                                      ifelse(Lat_1 == "(45,50]", '45-50°N',
                                          ifelse(Lat_1 == "(50,55]", '50-55°N',
                                              ifelse(Lat_1 == "(55,60]", '55-60°N',
                                                  '60-65°N' ))))))))
```

```
# sort the vector so they're in order from lowest to highest
bin_labels <- sort(bins$lat_bins)

### Arrange the df based on lat bins and then create a object with the correct order: ###
data4 <- arrange(data3,Lat_1)
grids <- data4$Grid

### Plot the data using geom_col: ###
sp_bar <- ggplot(data3, aes(x=Grid, y=nb_sp, fill=Lat_1)) +
 geom_col(color="#e9ecef") +
 geom_text(aes(label=nb_sp), position=position_dodge(width=0.9), vjust=-0.25, size = 2.5) +
 labs(title="Generic Richness",x="Grid Cell", y = "Number of Genera", fill="Paleo-Latitude") +
 theme(plot.title = element_text(hjust = 0.5),
      axis.text.x = element_text(angle = 90, vjust = 0.5, hjust=1)) +
 scale_fill_manual(values=c("red4",
                            "tomato3",
                            "darkorange3",
                            "goldenrod2",
                            "darkolivegreen4",
                            "aquamarine4",
                            "steelblue"),
                            labels = bin_labels)+
 scale_x_discrete(limits=grids) +
 ylim(0, 1 + max(data3$nb_sp))


fe_bar <- ggplot(data3, aes(x=Grid, y=nb_fe, fill=Lat_1)) +
 geom_col(color="#e9ecef") +
```

```
  geom_text(aes(label=signif(nb_fe,2)), position=position_dodge(width=0.9), vjust=-0.25, size = 2.5) +
  labs(title="Functional Richness",x="Grid Cell", y = "Number of Functional Entities", fill="Paleo-
Latitude") +
  theme(plot.title = element_text(hjust = 0.5),
      axis.text.x = element_text(angle = 90, vjust = 0.5, hjust=1)) +
  scale_fill_manual(values=c("red4",
                             "tomato3",
                             "darkorange3",
                             "goldenrod2",
                             "darkolivegreen4",
                             "aquamarine4",
                             "steelblue"),
                             labels = bin_labels)+
  scale_x_discrete(limits=grids) +
  ylim(0, 1 + max(data3$nb_fe))


 fred_bar <- ggplot(data3, aes(x=Grid, y=fred, fill=Lat_1)) +
  geom_col(color="#e9ecef") +
  geom_text(aes(label=signif(fred,2)), position=position_dodge(width=0.9), vjust=-0.25, size = 2.5) +
  labs(title="Functional Redundancy (FRed)",x="Grid Cell", y = "FRed Value", fill="Paleo-Latitude") +
  theme(plot.title = element_text(hjust = 0.5),
      axis.text.x = element_text(angle = 90, vjust = 0.5, hjust=1)) +
  scale_fill_manual(values=c("red4",
                             "tomato3",
                             "darkorange3",
                             "goldenrod2",
                             "darkolivegreen4",
                             "aquamarine4",
                             "steelblue"),
```

```
                        labels = bin_labels)+
  scale_x_discrete(limits=grids) +
  ylim(0, 1 + max(data3$fred))

 fored_bar <- ggplot(data3, aes(x=Grid, y=fored, fill=Lat_1)) +
  geom_col(color="#e9ecef") +
  geom_text(aes(label=signif(fored,2)), position=position_dodge(width=0.9), vjust=-0.25, size = 2.5) +
  labs(title="Functional Over-Redundancy",x="Grid Cell", y = "FORed Value", fill="Paleo-Latitude") +
  theme(plot.title = element_text(hjust = 0.5),
      axis.text.x = element_text(angle = 90, vjust = 0.5, hjust=1)) +
  scale_fill_manual(values=c("red4",
                        "tomato3",
                        "darkorange3",
                        "goldenrod2",
                        "darkolivegreen4",
                        "aquamarine4",
                        "steelblue"),
                        labels = bin_labels)+
  scale_x_discrete(limits=grids) +
  ylim(0, 1 + max(data3$fored))

 fvuln_bar <- ggplot(data3, aes(x=Grid, y=fvuln, fill=Lat_1)) +
  geom_col(color="#e9ecef") +
  geom_text(aes(label=signif(fvuln,2)), position=position_dodge(width=0.9), vjust=-0.25, size = 2.5) +
  labs(title="Functional Vulnerability (FVuln)",x="Grid Cell", y = "FVuln Value", fill="Paleo-Latitude") +
  theme(plot.title = element_text(hjust = 0.5),
      axis.text.x = element_text(angle = 90, vjust = 0.5, hjust=1)) +
  scale_fill_manual(values=c("red4",
                        "tomato3",
```

```
                    "darkorange3",
                    "goldenrod2",
                    "darkolivegreen4",
                    "aquamarine4",
                    "steelblue"),
                    labels = bin_labels)+
  scale_x_discrete(limits=grids) +
  ylim(0, 1 + max(data3$fvuln))

 bar_list <- list(sp_bar,fe_bar,fred_bar,fored_bar,fvuln_bar)
}

# Create series of bar plots for metric values in each substage
maa_up_60_grid_bar <- bar_plot(maa_up_60_fdfe,grid_local,"MAA up")
maa_low_60_grid_bar <- bar_plot(maa_low_60_fdfe,grid_local,"MAA low")
cam_up_60_grid_bar <- bar_plot(cam_up_60_fdfe,grid_local,"CAM up")
cam_mid_60_grid_bar <- bar_plot(cam_mid_60_fdfe,grid_local,"CAM mid")
cam_low_60_grid_bar <- bar_plot(cam_low_60_fdfe,grid_local,"CAM low")


# Extract
pdf("grid_collections_outputs/bar_MAA_up_allgrids.pdf", width = 50, height = 25) # Open a new pdf file
grid.arrange(grobs = maa_up_60_grid_bar,
          top = "Upper Maastr. Bar Plots of Diversity Metrics per Grid Cell") # Write the grid.arrange in the
file
dev.off() #close the file

pdf("grid_collections_outputs/bar_MAA_low_allgrids.pdf", width = 50, height = 25) # Open a new pdf
file
```

```
grid.arrange(grobs = maa_low_60_grid_bar,
        top = "Lower Maastr. Bar Plots of Diversity Metrics per Grid Cell") # Write the grid.arrange in the
file
dev.off() #close the file

pdf("grid_collections_outputs/bar_CAM_up_allgrids.pdf", width = 50, height = 25) # Open a new pdf file
grid.arrange(grobs = cam_up_60_grid_bar,
        top = "Upper Camp. Bar Plots of Diversity Metrics per Grid Cell") # Write the grid.arrange in the
file
dev.off() #close the file

pdf("grid_collections_outputs/bar_CAM_mid_allgrids.pdf", width = 50, height = 25) # Open a new pdf
file
grid.arrange(grobs = cam_mid_60_grid_bar,
        top = "Middle Camp. Bar Plots of Diversity Metrics per Grid Cell") # Write the grid.arrange in the
file
dev.off() #close the file

pdf("grid_collections_outputs/bar_CAM_low_allgrids.pdf", width = 50, height = 25) # Open a new pdf
file
grid.arrange(grobs = cam_low_60_grid_bar,
        top = "Lower Camp. Bar Plots of Diversity Metrics per Grid Cell") # Write the grid.arrange in the
file
dev.off() #close the file

 #### Simpsons Measure of Evenness per Grid Cell ####

## function for calculating Simpson's measure of evenness that can be used with the apply funct
# n = total # genera in a FE in a site/time; N = total # genera in all FE at a site/time;
# S = total # FE at the site/time; D = sum(n*(n-1))/N(N-1); inversD = inverse of D value;
#final= inverseD/S
```

```
simpsons_calc <- function(data){
  data1 <- data.frame(data$"details_fdfe")

  # simpsons measure of evenness calculation:
  simp_calc <- function(data1){
    data1 <- data.frame(data1)
    n <- data1[data1!=0] # get vector of only values greater than 0 (no empty FE)
    if (length(n)>1 && max(n)>1){ # use if statement to only run the calc on nodes with >1 FE and more
than 1 genera in at least one FE
      N <- sum(n)
      S <- length(n)
      D = sum(n*(n-1))/(N*(N-1))
      inversD <- 1/D
      sD <- inversD/S
    } else {
      0
    }
  }

  final <- apply(data1,1,simp_calc) # 1 indicates that it applies function to rows

  data3 <- data.frame(final)
  data4 <- cbind(newColName = rownames(data3), data3)
  rownames(data4) <- 1:nrow(data4)
  colnames(data4) <- c("Site","Simp_Even_Index")

  return(data4)
}
```

```
# Calculate Simpson's measure of evenness for each grid cell in substages using above function
maa_up_60_simp_even <- simpsons_calc(maa_up_60_fe)
maa_low_60_simp_even <- simpsons_calc(maa_low_60_fe)
cam_up_60_simp_even <- simpsons_calc(cam_up_60_fe)
cam_mid_60_simp_even <- simpsons_calc(cam_mid_60_fe)
cam_low_60_simp_even <- simpsons_calc(cam_low_60_fe)

 #### Create df of Simpsons for each Grid Cell through time ####

# Function to configure to columns with site names as 1st and evenness values as second, age as 3rd
config_for_loc_even_agg <- function(data, Age){
 # Create df that just take the functional richness values in each grid cell
 data_even_summ <- data

 data_even_summ <- data.frame(data_even_summ)
 colnames(data_even_summ) <- c("Site","Evenness")

 data_even_summ$Age <- Age
 return(data_even_summ)
}

# Use config function (above) to alter then added new column with Age
maa_up_60_simp_summ <- config_for_loc_even_agg(maa_up_60_simp_even, "MAA up")
maa_low_60_simp_summ <- config_for_loc_even_agg(maa_low_60_simp_even, "MAA low")
cam_up_60_simp_summ <- config_for_loc_even_agg(cam_up_60_simp_even, "CAM up")
cam_mid_60_simp_summ <- config_for_loc_even_agg(cam_mid_60_simp_even, "CAM mid")
cam_low_60_simp_summ <- config_for_loc_even_agg(cam_low_60_simp_even, "CAM low")
```

```
# Bind the data into a single df with 2 columns (stacked the different dfs)

simp_stack <-
rbind(maa_up_60_simp_summ,maa_low_60_simp_summ,cam_up_60_simp_summ,cam_mid_60_simp
_summ,cam_low_60_simp_summ)


# Used dcast to reconfigure the data so that Age and Site would be identifiers with unique values

simp_table_all <- dcast(simp_stack, Site ~ Age, value.var = "Evenness")


# This is a table showing the number of FE in each grid cell for each substage

head(simp_table_all)


# Export the df as a csv file

write.csv(simp_table_all, file="grid_collections_outputs/Grid_Simpsons_through_time.csv")


 #### Create Bar plots of Simpsons Measure of Evenness ####


# Function to create plot bar graphs of simpson's evenness for each grid cell per substage

bar_plot_even <- function(even_data,loc_data,substage_column, substage_name){


 colnames(even_data) <- c("Grid","Simp_Even_Index")


 ### Extract Location Information: ###

 loc_data1 <- loc_data[c("Site",substage_column,"pLat","pLon")] # Extract columns for just one
substage

 colnames(loc_data1) <- c("Grid",substage_column,"pLat","pLon") # Rename columns

 loc_data1$Lat_1 <- loc_data1$pLat # Add another column to use for binning lat

 # Next, use mutate and cut to create binned categories of lat

 loc_data2 <- loc_data1 %>% mutate(Lat_1 = cut(loc_data1$Lat_1,
                                    breaks=c(-Inf, seq(round(min(loc_data1$pLat)/5)*5,
                                        round(max(loc_data1$pLat)/5)*5, by = 5), Inf)))
```

```
### Merge the new bins with the fe df: ###
data3 <- merge(even_data,loc_data2, by= "Grid", all.x = TRUE)

# get the unique lat bins available for this substage and change colname
bins <- data.frame(unique(data3$Lat_1))
colnames(bins) <- "Lat_1"

# Create column with the sorts of labels actually want to use based on ifelse
bins$lat_bins <-with(bins, ifelse(Lat_1 == "(-Inf,30]", '23-30°N',
                          ifelse(Lat_1 == "(30,35]", '30-35°N',
                              ifelse(Lat_1 == "(35,40]", '35-40°N',
                                  ifelse(Lat_1 == "(40,45]", '40-45°N',
                                      ifelse(Lat_1 == "(45,50]", '45-50°N',
                                          ifelse(Lat_1 == "(50,55]", '50-55°N',
                                              ifelse(Lat_1 == "(55,60]", '55-60°N',
                                                  '60-65°N' ))))))))

# sort the vector so they're in order from lowest to highest
bin_labels <- sort(bins$lat_bins)

### Arrange the df based on lat bins and then create a object with the correct order: ###
data4 <- arrange(data3,Lat_1)

grids <- data4$Grid

### Plot the data using geom_col: ###
bar <- ggplot(data3, aes(x=Grid, y=Simp_Even_Index, fill=Lat_1)) +
  geom_col(color="#e9ecef") +
```

```
  geom_text(aes(label=signif(Simp_Even_Index,2)), position=position_dodge(width=0.9), vjust=-0.25,
size = 2.5) +
  labs(title=paste(substage_name,"Simpson's Diversity Index"),x="Grid Cell", y = "Evenness") +
  theme(plot.title = element_text(hjust = 0.5),
      axis.text.x = element_text(angle = 90, vjust = 0.5, hjust=1)) +
  labs(fill="Paleo-Latitude") +
  scale_fill_manual(values=c("red4",
                             "tomato3",
                             "darkorange3",
                             "goldenrod2",
                             "darkolivegreen4",
                             "aquamarine4",
                             "steelblue"),
                             labels = bin_labels)+
  scale_x_discrete(limits=grids) +
  ylim(0, 0.25 + max(data3$Simp_Even_Index))

 return(bar)
}

# Use above function to create bar plots of Simpson's Evenness for each SS
maa_up_simp_even_barplots <- bar_plot_even(maa_up_60_simp_even,grid_local,"MAA up", "Upper
Maastr.")
maa_low_simp_even_barplots <- bar_plot_even(maa_low_60_simp_even,grid_local,"MAA up", "Lower
Maastr.")
cam_up_simp_even_barplots <- bar_plot_even(cam_up_60_simp_even,grid_local,"MAA up", "Upper
Camp.")
cam_mid_simp_even_barplots <- bar_plot_even(cam_mid_60_simp_even,grid_local,"MAA up", "Middle
Camp.")
cam_low_simp_even_barplots <- bar_plot_even(cam_low_60_simp_even,grid_local,"MAA up", "Lower
Camp.")
```

```
# Make list of the barplots of simpsons evenness
simp_even_barplots <- list(maa_up_simp_even_barplots, maa_low_simp_even_barplots,
            cam_up_simp_even_barplots, cam_mid_simp_even_barplots,
            cam_low_simp_even_barplots)



# Export bar plots of the Simpson's diversity index per grid cell for each substage
pdf("grid_collections_outputs/bar_plots_SimpEven_grids.pdf", width = 50, height = 25) # Open a new pdf file
grid.arrange(grobs = simp_even_barplots,
         top = "Bar Plots of Simpson's Evenness per Grid Cell") # Write the grid.arrange in the file
dev.off() #close file


 #### Shannon Equability Index for each grid cell#####


# function to get overall shannons evenness for each substage
shannons_calc <- function(data){


 # equation for shannons calc
 shannon_equit <- function(data){
  data1 <- data.frame(data)
  data1 <- data1[data1!=0] # get vector of only values greater than 0 (no empty FE)


  if (length(data1)>1){ # use if statement to only run the calc on nodes with >1 FE and more than 1 genera in at least one FE


  S <- length(which(data1!=0)) # get the numb of FE
  Si <- sum(data1) # find out how many genera total present in all FE
```

```
  Pvals <- data1/Si # get proportions of each FE based on dividing # genera in each by total # genera
  nlPvals <- log(Pvals) # take natural log of the proportions
  H <- -sum(nlPvals*Pvals) # calculate Shannon diversity index (H) by multiplying nl of p by p and summing and multiply by neg 1
  LNS <- log(S) # get natural log of number of FE present
  Hi <- H/LNS # calculate Shannon Equitability Index (Hi) by dividing H by nat log of S
  } else {
   0
  }
 }

 data1 <- data.frame(data$"details_fdfe")

 final <- apply(data1,1,shannon_equit) # 1 indicates that it applies function to rows

 data3 <- data.frame(final)
 data4 <- cbind(newColName = rownames(data3), data3) # make row name (i.e., grid cell) first column again
 rownames(data4) <- 1:nrow(data4)
 colnames(data4) <- c("Site","Shan_Equit_Index")

 return(data4)
}

# Calculate Shannons Equability Index for each grid cell in substages using above function
maa_up_60_shan_even <- shannons_calc(maa_up_60_fe)
maa_low_60_shan_even <- shannons_calc(maa_low_60_fe)
cam_up_60_shan_even <- shannons_calc(cam_up_60_fe)
```

```
cam_mid_60_shan_even <- shannons_calc(cam_mid_60_fe)
cam_low_60_shan_even <- shannons_calc(cam_low_60_fe)



 #### Create df of Shannon for each Grid Cell through time ####

# Function to configure to columns with site names as 1st and evenness values as second, age as 3rd
config_for_loc_even_agg <- function(data, Age){
 # Create df that just take the functional richness values in each grid cell
 data_even_summ <- data

 data_even_summ <- data.frame(data_even_summ)
 colnames(data_even_summ) <- c("Site","Evenness")

 data_even_summ$Age <- Age
 return(data_even_summ)
}

# Use config function (above) to alter then added new column with Age
maa_up_60_shan_summ <- config_for_loc_even_agg(maa_up_60_shan_even, "MAA up")
maa_low_60_shan_summ <- config_for_loc_even_agg(maa_low_60_shan_even, "MAA low")
cam_up_60_shan_summ <- config_for_loc_even_agg(cam_up_60_shan_even, "CAM up")
cam_mid_60_shan_summ <- config_for_loc_even_agg(cam_mid_60_shan_even, "CAM mid")
cam_low_60_shan_summ <- config_for_loc_even_agg(cam_low_60_shan_even, "CAM low")

# Bind the data into a single df with 2 columns (stacked the different dfs)
shan_stack <-
rbind(maa_up_60_shan_summ,maa_low_60_shan_summ,cam_up_60_shan_summ,cam_mid_60_shan_
summ,cam_low_60_shan_summ)
```

```
# Used dcast to reconfigure the data so that Age and Site would be identifiers with unique values
shan_table_all <- dcast(shan_stack, Site ~ Age, value.var = "Evenness")

# This is a table showing the number of FE in each grid cell for each substage
head(shan_table_all)

# Export the df as a csv file
write.csv(shan_table_all, file="grid_collections_outputs/Grid_Shannon_through_time.csv")

 #### Create Bar plots of Shannon Equability Index ####

# Plot bar graphs of simpson's evenness for each grid cell per substage
bar_plot_shan_even <- function(even_data, loc_data, substage_column, substage_name){

 colnames(even_data) <- c("Grid","Shan_Even_Index")

 ### Extract Location Information: ###
 loc_data1 <- loc_data[c("Site",substage_column,"pLat","pLon")] # Extract columns for just one substage
 colnames(loc_data1) <- c("Grid",substage_column,"pLat","pLon") # Rename columns
 loc_data1$Lat_1 <- loc_data1$pLat # Add another column to use for binning lat
 # Next, use mutate and cut to create binned categories of lat
 loc_data2 <- loc_data1 %>% mutate(Lat_1 = cut(loc_data1$Lat_1,
                                   breaks=c(-Inf, seq(round(min(loc_data1$Lat_1)/5)*5,
                                                      round(max(loc_data1$Lat_1)/5)*5, by = 5), Inf)))

 ### Merge the new bins with the fe df: ###
```

```
data3 <- merge(even_data,loc_data2, by= "Grid", all.x = TRUE)



# get the unique lat bins available for this substage and change colname
bins <- data.frame(unique(data3$Lat_1))
colnames(bins) <- "Lat_1"

# Create column with the sorts of labels actually want to use based on ifelse
bins$lat_bins <-with(bins, ifelse(Lat_1 == "(-Inf,30]", '23-30°N',
                          ifelse(Lat_1 == "(30,35]", '30-35°N',
                              ifelse(Lat_1 == "(35,40]", '35-40°N',
                                  ifelse(Lat_1 == "(40,45]", '40-45°N',
                                      ifelse(Lat_1 == "(45,50]", '45-50°N',
                                          ifelse(Lat_1 == "(50,55]", '50-55°N',
                                              ifelse(Lat_1 == "(55,60]", '55-60°N',
                                                  '60-65°N' ))))))))

# sort the vector so they're in order from lowest to highest
bin_labels <- sort(bins$lat_bins)

### Arrange the df based on lat bins and then create a object with the correct order: ###
data4 <- arrange(data3,Lat_1)

grids <- data4$Grid



### Plot the data using geom_col: ###
bar <- ggplot(data3, aes(x=Grid, y=Shan_Even_Index, fill=Lat_1)) +
  geom_col(color="#e9ecef") +
```

```
  geom_text(aes(label=signif(Shan_Even_Index,2)), position=position_dodge(width=0.9), vjust=-0.25,
size = 2.5) +
  labs(title=paste(substage_name, "Shannon Equitability Index"),x="Grid Cell", y = "Evenness", fill =
"Paleo-Latitude") +
  theme(plot.title = element_text(hjust = 0.5),
      axis.text.x = element_text(angle = 90, vjust = 0.5, hjust=1)) +
  scale_fill_manual(values=c("red4",
                             "tomato3",
                             "darkorange3",
                             "goldenrod2",
                             "darkolivegreen4",
                             "aquamarine4",
                             "steelblue"),
                             labels = bin_labels)+
  scale_x_discrete(limits=grids) +
  ylim(0, 0.1 + max(data3$Shan_Even_Index))

 return(bar)

}

# Plot bar plots for each substage
maa_up_shan_even_barplots <- bar_plot_shan_even(maa_up_60_shan_even,grid_local,"MAA
up","Upper Maastr.")
maa_low_shan_even_barplots <- bar_plot_shan_even(maa_low_60_shan_even,grid_local,"MAA
low","Lower Maastr.")
cam_up_shan_even_barplots <- bar_plot_shan_even(cam_up_60_shan_even,grid_local,"CAM
up","Upper Camp.")
cam_mid_shan_even_barplots <- bar_plot_shan_even(cam_mid_60_shan_even,grid_local,"CAM
mid","Middle Camp.")
```

```
cam_low_shan_even_barplots <- bar_plot_shan_even(cam_low_60_shan_even,grid_local,"CAM
low","Lower Camp.")


# Create list of shannon bar plots

bar_list_shan <- list(maa_up_shan_even_barplots, maa_low_shan_even_barplots,

                      cam_up_shan_even_barplots, cam_mid_shan_even_barplots,

                      cam_low_shan_even_barplots)


# plot bar plots of the Shannon Equitability index per grid cell for each substage

pdf("grid_collections_outputs/bar_plots_ShanEven_grid.pdf", width = 50, height = 25) # Open a new pdf
file

grid.arrange(grobs = bar_list_shan,

             top = "Bar Plots of Shannon Equitability Index per Grid Cell") # Write the grid.arrange in the file

dev.off() #close file


##### ANALYSIS LATITUDE BINS AS COLLECTIONS #####


  #### Create pres-absence matrices to use in mFD analysis ####


# Function for taking the substage df that have latitudinal bins (see above section on location
information)

  # and creating a presence-absence matrix based on bins as "collection" locations

substg_gen_latbins_pres_abs <- function(data){

  # select out the relevant columns from the df

  data1 <-
subset(data,select=c("Updated_Genus","Substage_from_Zone.Mbr","motility","life_habitat","feeding",

                                    "plong","plat"))

  # create new column with lat information to mess with

  data1$Lat_1 <- data1$plat
```

```
 # create breaks in the df based on latitude
 data2 <- data1 %>% mutate(Lat_1 = cut(data1$Lat_1,
                                       breaks=c(-Inf, seq(round(min(data1$plat)/5)*5,
                                                          round(max(data1$plat)/5)*5, by = 5), Inf)))

 # subset out just the genus name and lat bin
 data3 <- as.data.frame(subset(data2,select=c("Updated_Genus","Lat_1")))

 # remove duplicates
 data3 %>% distinct(Updated_Genus,Lat_1, .keep_all = TRUE)

 # transform into a pres-abs matrix
 data4 <- dcast(data3, Lat_1~Updated_Genus, length)

 # make the lat bin names the row names
 data5 <- data.frame(data4[,-1], row.names = data4[,1])

 # make all values over 1 into 1 and keep all else 0 (presence-absence)
 data6 <- ifelse(data5 > 0, 1, 0)

 return(data6)
}



#Create data frames for pres-abs for each substage based on latitudinal bins (use paleo-lat dfs)
maa_up_60_lat_pres_ab <- substg_gen_latbins_pres_abs(maa_up_60_paleo_loc_data)
maa_low_60_lat_pres_ab <- substg_gen_latbins_pres_abs(maa_low_60_paleo_loc_data)
cam_up_60_lat_pres_ab <- substg_gen_latbins_pres_abs(cam_up_60_paleo_loc_data)
cam_mid_60_lat_pres_ab <- substg_gen_latbins_pres_abs(cam_mid_60_paleo_loc_data)
```

```
cam_low_60_lat_pres_ab <- substg_gen_latbins_pres_abs(cam_low_60_paleo_loc_data)



## Function to calc functional trait values
calc_fe <- function(trait_data,pres_ab_data){
  alpha_fd_fe_MAAup <- mFD::alpha.fd.fe(
    asb_sp_occ      = pres_ab_data,
    sp_to_fe        = trait_data,
    ind_nm          = c("fred", "fored", "fvuln"),
    check_input     = TRUE,
    details_returned = TRUE)
}

  #### Compute alpha.fd.fe from mFD for Latitudinal Bins ####

## NOTE: use the same df of trait values for entire SS overall calculated using sp.to.fe"

# Calculate functional richness, redundancy, over redundancy, and vulnerability for each substage in lat bins
  # see above section for the function used here (same function used for getting alpha-beta diversity for grid cells)
maa_up_60_fe_lat <- calc_fe(maa_up_60_traits, maa_up_60_lat_pres_ab)
maa_low_60_fe_lat<- calc_fe(maa_low_60_traits, maa_low_60_lat_pres_ab)
cam_up_60_fe_lat <- calc_fe(cam_up_60_traits, cam_up_60_lat_pres_ab)
cam_mid_60_fe_lat <- calc_fe(cam_mid_60_traits, cam_mid_60_lat_pres_ab)
cam_low_60_fe_lat <- calc_fe(cam_low_60_traits, cam_low_60_lat_pres_ab)



#get a matrix of fe values for each latitudinal bin in a substage
```

```
maa_up_60_fdfe_lat <- (maa_up_60_fe_lat$"asb_fdfe")
maa_low_60_fdfe_lat <- (maa_low_60_fe_lat$"asb_fdfe")
cam_up_60_fdfe_lat <- (cam_up_60_fe_lat$"asb_fdfe")
cam_mid_60_fdfe_lat <- (cam_mid_60_fe_lat$"asb_fdfe")
cam_low_60_fdfe_lat <- (cam_low_60_fe_lat$"asb_fdfe")

maa_up_60_fdfe_lat
maa_low_60_fdfe_lat
cam_up_60_fdfe_lat
cam_mid_60_fdfe_lat
cam_low_60_fdfe_lat



# Function to fix row names (so all the same)
make_lat_names <- function(data){
  data <- data.frame(data)
  data$bins <- rownames(data)
  data$bins <-with(data, ifelse(bins == "(-Inf,30]", '(-Inf,30]',
                          ifelse(bins == "(-Inf,35]", '(30,35]',
                              ifelse(bins == "(30,35]", '(30,35]',
                                  ifelse(bins == "(35,40]", '(35,40]',
                                      ifelse(bins == "(-Inf,45]", '(40,45]',
                                      ifelse(bins == "(40,45]", '(40,45]',
                                          ifelse(bins == "(45,50]", '(45,50]',
                                              ifelse(bins == "(50,55]", '(50,55]',
                                                  ifelse(bins == "(55,60]", '(55,60]',
                                                      ifelse(bins == "(55, Inf]", '(55,60]',
                                                          ifelse(bins == '(60,65]', '(60, Inf]',
                                                              '(60, Inf]' ))))))))))))
```

```
  rownames(data) <- data$bins
  return(data[,c("nb_sp","nb_fe","fred","fored","fvuln")])
}

# Use above function to make sure lat bin names are all the same when they mean the same thing
maa_up_60_fdfe_lat <- make_lat_names(maa_up_60_fdfe_lat)
maa_low_60_fdfe_lat <- make_lat_names(maa_low_60_fdfe_lat)
cam_up_60_fdfe_lat <- make_lat_names(cam_up_60_fdfe_lat)
cam_mid_60_fdfe_lat <- make_lat_names(cam_mid_60_fdfe_lat)
cam_low_60_fdfe_lat <- make_lat_names(cam_low_60_fdfe_lat)



# export results tables summarizing the number of spp, fe, and functional richness and evenness values for each substage
write.csv(maa_up_60_fdfe_lat,file="lat_bins_outputs/LatBins_MaaUP60_fe_alpha_beta_results.csv")
write.csv(maa_low_60_fdfe_lat,file="lat_bins_outputs/LatBins_MaaLOW60_fe_alpha_beta_results.csv"
)
write.csv(cam_up_60_fdfe_lat,file="lat_bins_outputs/LatBins_CamUP60_fe_alpha_beta_results.csv")
write.csv(cam_mid_60_fdfe_lat,file="lat_bins_outputs/LatBins_CamMID60_fe_alpha_beta_results.csv")
write.csv(cam_low_60_fdfe_lat,file="lat_bins_outputs/LatBins_CamLOW60_fe_alpha_beta_results.csv")



  #### Create Bar Plots of Metrics across lat bins for all substages: ####


# Function to create series of bar plots for each substage's latitudinal bins
bar_plot_lat <- function(fe_data){
```

```
###  Get the functional ecology information for the substage (this is a dataframe) and reconfig: ###
data1 <- cbind(Bin = rownames(fe_data), fe_data) # make the genus name the 1st column again
rownames(data1) <- 1:nrow(data1) # make a new index for the rownames
data1 <- as.data.frame(data1)
i <- c(2:6)
data1[ , i] <- apply(data1[ , i], 2,           # Specify own function within apply to make numeric
                 function(x) as.numeric(as.character(x)))

# get the unique lat bins available for this substage and change colname
bins <- data.frame(unique(data1$Bin))
colnames(bins) <- "Lat_1"

# Create column with the sorts of labels actually want to use based on ifelse
bins$lat_bins <-with(bins, ifelse(Lat_1 == "(-Inf,30]", '23-30°N',
                           ifelse(Lat_1 == "(-Inf,35]", '30-35°N',
                                ifelse(Lat_1 == "(30,35]", '30-35°N',
                                     ifelse(Lat_1 == "(35,40]", '35-40°N',
                                          ifelse(Lat_1 == "(40,45]", '40-45°N',
                                               ifelse(Lat_1 == "(45,50]", '45-50°N',
                                                    ifelse(Lat_1 == "(50,55]", '50-55°N',
                                                         ifelse(Lat_1 == "(55,60]", '55-60°N',
                                                              ifelse(Lat_1 == "(55, Inf]", '55-60°N',
                                                                   '60-65°N' ))))))))))

# sort the vector so they're in order from lowest to highest
bin_labels <- sort(bins$lat_bins)

### Plot the data using geom_col: ###
sp_bar <- ggplot(data1, aes(x=Bin, y=nb_sp, fill=Bin)) +
```

```
  geom_col(color="#e9ecef") +
  labs(title="Generic Richness",x="Paleo-Latitude Bin", y = "Number of Genera", fill="Paleo-Latitude") +
  theme(plot.title = element_text(hjust = 0.5),
      axis.text.x = element_text(angle = 0, vjust = 0.5, hjust=0.5)) +
  guides(fill = "none") +
  scale_fill_manual(values=c("red4",
                             "tomato3",
                             "darkorange3",
                             "goldenrod2",
                             "darkolivegreen4",
                             "aquamarine4",
                             "steelblue"),
                             labels = bin_labels)+
  scale_x_discrete(labels= bin_labels) +
  ylim(0, 5 + max(data1$nb_sp))+
  geom_text(aes(label=signif(nb_sp,3)), position=position_dodge(width=0.9), vjust=-0.25)


 fe_bar <- ggplot(data1, aes(x=Bin, y=nb_fe, fill=Bin)) +
  geom_col(color="#e9ecef") +
  labs(title="Functional Richness",x="Paleo-Latitude Bin", y = "Number of Functional Entities",
fill="Paleo-Latitude") +
  theme(plot.title = element_text(hjust = 0.5),
      axis.text.x = element_text(angle = 0, vjust = 0.5, hjust=0.5)) +
  guides(fill = "none") +
  scale_fill_manual(values=c("red4", # set scale of colors manually, need 5 colors since 5 lat bins
                             "tomato3",
                             "darkorange3",
                             "goldenrod2",
                             "darkolivegreen4",
```

```
                    "aquamarine4",
                    "steelblue"),
                    labels = bin_labels)+
  scale_x_discrete(labels= bin_labels) +
  ylim(0, 1 + max(data1$nb_fe)) +
  geom_text(aes(label=signif(nb_fe,3)), position=position_dodge(width=0.9), vjust=-0.25)

 fred_bar <- ggplot(data1, aes(x=Bin, y=fred, fill=Bin)) +
  geom_col(color="#e9ecef") +
  labs(title="Functional Redundancy (FRed)",x="Paleo-Latitude Bin", y = "FRed Value", fill="Paleo-
Latitude") +
  theme(plot.title = element_text(hjust = 0.5),
      axis.text.x = element_text(angle = 0, vjust = 0.5, hjust=0.5)) +
  guides(fill = "none") +
  scale_fill_manual(values=c("red4",
                    "tomato3",
                    "darkorange3",
                    "goldenrod2",
                    "darkolivegreen4",
                    "aquamarine4",
                    "steelblue"),
                    labels = bin_labels)+
  scale_x_discrete(labels= bin_labels) +
  ylim(0, 0.25 + max(data1$fred))+
  geom_text(aes(label=signif(fred,3)), position=position_dodge(width=0.9), vjust=-0.25)

 fored_bar <- ggplot(data1, aes(x=Bin, y=fored, fill=Bin)) +
  geom_col(color="#e9ecef") +
```

```
  labs(title="Functional Over-Redundancy",x="Paleo-Latitude Bin", y = "FORed Value", fill="Paleo-Latitude") +
  theme(plot.title = element_text(hjust = 0.5),
      axis.text.x = element_text(angle = 0, vjust = 0.5, hjust=0.5)) +
  guides(fill = "none") +
  scale_fill_manual(values=c("red4",
                             "tomato3",
                             "darkorange3",
                             "goldenrod2",
                             "darkolivegreen4",
                             "aquamarine4",
                             "steelblue"),
                             labels = bin_labels)+
  scale_x_discrete(labels= bin_labels) +
  ylim(0, 0.25 + max(data1$fored)) +
  geom_text(aes(label=signif(fored,3)), position=position_dodge(width=0.9), vjust=-0.25)


 fvuln_bar <- ggplot(data1, aes(x=Bin, y=fvuln, fill=Bin)) +
  geom_col(color="#e9ecef") +
  labs(title="Functional Vulnerability (FVuln)","Paleo-Latitude Bin", y = "FVuln Value", fill="Paleo-Latitude") +
  theme(plot.title = element_text(hjust = 0.5),
      axis.text.x = element_text(angle = 0, vjust = 0.5, hjust=0.5)) +
  guides(fill = "none") +
  scale_fill_manual(values=c("red4",
                             "tomato3",
                             "darkorange3",
                             "goldenrod2",
                             "darkolivegreen4",
```

```
                    "aquamarine4",
                    "steelblue"),
                    labels = bin_labels)+
  scale_x_discrete(labels= bin_labels) +
  ylim(0, 0.25 + max(data1$fvuln))+
  geom_text(aes(label=signif(fvuln,3)), position=position_dodge(width=0.9), vjust=-0.25)

 bar_list <- list(sp_bar,fe_bar,fred_bar,fored_bar,fvuln_bar)
}

# Create series of bar plots for metric values in each substage using above function
maa_up_60_lat_bar <- bar_plot_lat(maa_up_60_fdfe_lat)
maa_low_60_lat_bar <- bar_plot_lat(maa_low_60_fdfe_lat)
cam_up_60_lat_bar <- bar_plot_lat(cam_up_60_fdfe_lat)
cam_mid_60_lat_bar <- bar_plot_lat(cam_mid_60_fdfe_lat)
cam_low_60_lat_bar <- bar_plot_lat(cam_low_60_fdfe_lat)


pdf("lat_bins_outputs/LatBins_MAA_up_bars.pdf", width = 15, height = 25) # Open a new pdf file
grid.arrange(grobs = maa_up_60_lat_bar,
        top = "Upper Maastr. Bar Plots of Diversity Metrics per Latitudinal Bin") # Write the grid.arrange in the file
dev.off() #close the file

pdf("lat_bins_outputs/LatBins_MAA_low_bars.pdf", width = 15, height = 25) # Open a new pdf file
grid.arrange(grobs = maa_low_60_lat_bar,
        top = "Lower Maastr. Bar Plots of Diversity Metrics per Latitudinal Bin") # Write the grid.arrange in the file
dev.off() #close the file
```

```
pdf("lat_bins_outputs/LatBins_CAM_up_bars.pdf", width = 15, height = 25) # Open a new pdf file
grid.arrange(grobs = cam_up_60_lat_bar,
        top = "Upper Camp. Bar Plots of Diversity Metrics per Latitudinal Bin") # Write the grid.arrange in the file
dev.off() #close the file

pdf("lat_bins_outputs/LatBins_CAM_mid_bars.pdf", width = 15, height = 25) # Open a new pdf file
grid.arrange(grobs = cam_mid_60_lat_bar,
        top = "Middle Camp. Bar Plots of Diversity Metrics per Latitudinal Bin") # Write the grid.arrange in the file
dev.off() #close the file

pdf("lat_bins_outputs/LatBins_CAM_low_bars.pdf", width = 15, height = 25) # Open a new pdf file
grid.arrange(grobs = cam_low_60_lat_bar,
        top = "Lower Camp. Bar Plots of Diversity Metrics per Latitudinal Bin") # Write the grid.arrange in the file
dev.off() #close the file



 #### Create Line Plots of richness for each Latitudinal Bin in each Substage ####

# Function to create line plots of GR vs FR across Lat bins
plot_line_lat <- function(fe_data1,fe_data2,fe_data3,fe_data4,fe_data5,yaxis,legendlabel){
 # function to get the Age data as first column
 df_func <- function(data){
  ###  Get the functional ecology information for the substage (this is a dataframe) and reconfig: ###
  data1 <- cbind(Bin = rownames(data), data) # make the genus name the 1st column again
  rownames(data1) <- 1:nrow(data1) # make a new index for the rownames
```

```
  data1 <- as.data.frame(data1)
  i <- c(2:6)
  data1[ , i] <- apply(data1[ , i], 2,          # Specify own function within apply to make numeric
                       function(x) as.numeric(as.character(x)))
  return(as.data.frame(data1))
}

fe_data1 <- df_func(fe_data1)
fe_data2 <- df_func(fe_data2)
fe_data3 <- df_func(fe_data3)
fe_data4 <- df_func(fe_data4)
fe_data5 <- df_func(fe_data5)


# function to get the bin names (latitude bins) for each SS
get_bin_names <- function(data){
  # get the unique lat bins available for this substage and change colname
  bins <- data.frame(unique(data$Bin))
  colnames(bins) <- "Lat_1"

  # Create column with the sorts of labels actually want to use based on ifelse
  bins$lat_bins <-with(bins, ifelse(Lat_1 == "(-Inf,30]", '23-30°N',
                                    ifelse(Lat_1 == "(-Inf,35]", '30-35°N',
                                           ifelse(Lat_1 == "(30,35]", '30-35°N',
                                                  ifelse(Lat_1 == "(35,40]", '35-40°N',
                                                         ifelse(Lat_1 == "(40,45]", '40-45°N',
                                                                ifelse(Lat_1 == "(45,50]", '45-50°N',
                                                                       ifelse(Lat_1 == "(50,55]", '50-55°N',
                                                                              ifelse(Lat_1 == "(55,60]", '55-60°N',
```

```
                                                    ifelse(Lat_1 == "(55, Inf]", '55-60°N',
                                                           '60-65°N' ))))))))))

  # sort the vector so they're in order from lowest to highest
  bin_labels <- sort(bins$lat_bins)
  return(bin_labels)
}

bin_labels1 <- get_bin_names(fe_data1)
bin_labels2 <- get_bin_names(fe_data2)
bin_labels3 <- get_bin_names(fe_data3)
bin_labels4 <- get_bin_names(fe_data4)
bin_labels5 <- get_bin_names(fe_data5)



bin1_line <- ggplot(fe_data1, aes(x=Bin)) +
  geom_line(aes(y = nb_sp, color = "#D55E00",group=1)) +
  geom_line(aes(y = 5*nb_fe, color="steelblue",group=2),linetype = "dashed") +
  geom_point(aes(y = nb_sp), color = "#D55E00") +
  geom_point(aes(y = 5*nb_fe), color="steelblue") +
  scale_y_continuous(
    # Features of the first axis
    name = "# Genera",
    # Add a second axis and specify its features
    sec.axis = sec_axis(~ . * 0.2,name=yaxis)
  ) +
  scale_color_discrete(labels=c("Generic Richness", legendlabel)) +
  labs(title=paste("Upper Maastr.")) +
  xlab(label="Latitude") +
```

```
  scale_x_discrete(labels= bin_labels1) +
  theme(legend.title=element_blank(),plot.title = element_text(hjust = 0.5))

bin2_line <- ggplot(fe_data2, aes(x=Bin)) +
  geom_line(aes(y = nb_sp, color = "#D55E00",group=1)) +
  geom_line(aes(y = 5*nb_fe, color="steelblue",group=2),linetype = "dashed") +
  geom_point(aes(y = nb_sp), color = "#D55E00") +
  geom_point(aes(y = 5*nb_fe), color="steelblue") +
  scale_y_continuous(
    # Features of the first axis
    name = "# Genera",
    # Add a second axis and specify its features
    sec.axis = sec_axis(~ . * 0.2,name=yaxis)
  ) +
  scale_color_discrete(labels=c("Generic Richness", legendlabel)) +
  labs(title=paste("Lower Maastr.")) +
  xlab(label="Latitude") +
  scale_x_discrete(labels= bin_labels2) +
  theme(legend.title=element_blank(),plot.title = element_text(hjust = 0.5))

bin3_line <- ggplot(fe_data3, aes(x=Bin)) +
  geom_line(aes(y = nb_sp, color = "#D55E00",group=1)) +
  geom_line(aes(y = 5*nb_fe, color="steelblue",group=2),linetype = "dashed") +
  geom_point(aes(y = nb_sp), color = "#D55E00") +
  geom_point(aes(y = 5*nb_fe), color="steelblue") +
  scale_y_continuous(
    # Features of the first axis
    name = "# Genera",
    # Add a second axis and specify its features
```

```
    sec.axis = sec_axis(~ . * 0.2,name=yaxis)
  ) +
  scale_color_discrete(labels=c("Generic Richness", legendlabel)) +
  labs(title=paste("Upper Camp.")) +
  xlab(label="Latitude") +
  scale_x_discrete(labels= bin_labels3) +
  theme(legend.title=element_blank(),plot.title = element_text(hjust = 0.5))

bin4_line <- ggplot(fe_data4, aes(x=Bin)) +
  geom_line(aes(y = nb_sp, color = "#D55E00",group=1)) +
  geom_line(aes(y = 5*nb_fe, color="steelblue",group=2),linetype = "dashed") +
  geom_point(aes(y = nb_sp), color = "#D55E00") +
  geom_point(aes(y = 5*nb_fe), color="steelblue") +
  scale_y_continuous(
    # Features of the first axis
    name = "# Genera",
    # Add a second axis and specify its features
    sec.axis = sec_axis(~ . * 0.2,name=yaxis)
  ) +
  scale_color_discrete(labels=c("Generic Richness", legendlabel)) +
  labs(title=paste("Middle Camp.")) +
  xlab(label="Latitude") +
  scale_x_discrete(labels= bin_labels4) +
  theme(legend.title=element_blank(),plot.title = element_text(hjust = 0.5))

bin5_line <- ggplot(fe_data5, aes(x=Bin)) +
  geom_line(aes(y = nb_sp, color = "#D55E00",group=1)) +
  geom_line(aes(y = 5*nb_fe, color="steelblue",group=2),linetype = "dashed") +
  geom_point(aes(y = nb_sp), color = "#D55E00") +
```

```
   geom_point(aes(y = 5*nb_fe), color="steelblue") +
   scale_y_continuous(
    # Features of the first axis
    name = "# Genera",
    # Add a second axis and specify its features
    sec.axis = sec_axis(~ . * 0.2,name=yaxis)
   ) +
   scale_color_discrete(labels=c("Generic Richness", legendlabel)) +
   labs(title=paste("Lower Camp.")) +
   xlab(label="Latitude") +
   scale_x_discrete(labels= bin_labels5) +
   theme(legend.title=element_blank(),plot.title = element_text(hjust = 0.5))

 bin_list <- list(bin1_line,bin2_line,bin3_line,bin4_line,bin5_line)
 return(bin_list)
}

# create list of plots for changes in fr across lat in diff substages
plot_bins_gr_fr <- plot_line_lat(maa_up_60_fdfe_lat,maa_low_60_fdfe_lat,cam_up_60_fdfe_lat,
                        cam_mid_60_fdfe_lat,cam_low_60_fdfe_lat,
                        "# Functional Entities","Functional Richness")


pdf("lat_bins_outputs/avg_GR_FR_lat_bin_allsubstages.pdf", width = 20, height = 25) # Open a new pdf
file
grid.arrange(grobs = plot_bins_gr_fr,
          top = "Average Generic and Functional Richness across Latitude") # Write the grid.arrange in the
file
dev.off() #close the file
```

```
  #### Export table of FR and GR through time across Latitude Bins ####

# Function to configure to columns with site names as 1st and evenness values as second, age as 3rd
config_for_loc_rich_agg_latbins <- function(data, Age){
  # Create df that just take the functional richness values in each grid cell
  data_rich_summ <- data.frame(data[,c("nb_sp","nb_fe")])
  data_rich_summ$bins <- row.names(data_rich_summ)
  row.names(data_rich_summ) <- 1:nrow(data_rich_summ)
  data_rich_summ <- data_rich_summ[,c("bins","nb_sp","nb_fe")]
  data_rich_summ$Age <- Age
  return(data_rich_summ)
}

# Use config function (above) to alter then added new column with Age
maa_up_60_rich_summ_lat_bins <- config_for_loc_rich_agg_latbins(maa_up_60_fdfe_lat, "MAA up")
maa_low_60_rich_summ_lat_bins <- config_for_loc_rich_agg_latbins(maa_low_60_fdfe_lat, "MAA
low")
cam_up_60_rich_summ_lat_bins <- config_for_loc_rich_agg_latbins(cam_up_60_fdfe_lat, "CAM up")
cam_mid_60_rich_summ_lat_bins <- config_for_loc_rich_agg_latbins(cam_mid_60_fdfe_lat, "CAM
mid")
cam_low_60_rich_summ_lat_bins <- config_for_loc_rich_agg_latbins(cam_low_60_fdfe_lat, "CAM
low")

# Bind the GR data into a single df with 2 columns (stacked the different dfs)
GR_stack <-
rbind(maa_up_60_rich_summ_lat_bins[,c("bins","nb_sp","Age")],maa_low_60_rich_summ_lat_bins[,c("
bins","nb_sp","Age")],
```

```
cam_up_60_rich_summ_lat_bins[,c("bins","nb_sp","Age")],cam_mid_60_rich_summ_lat_bins[,c("bins",
"nb_sp","Age")],
        cam_low_60_rich_summ_lat_bins[,c("bins","nb_sp","Age")])

# Bind the FR data into a single df with 2 columns (stacked the different dfs)
FR_stack <-
rbind(maa_up_60_rich_summ_lat_bins[,c("bins","nb_fe","Age")],maa_low_60_rich_summ_lat_bins[,c("
bins","nb_fe","Age")],

cam_up_60_rich_summ_lat_bins[,c("bins","nb_fe","Age")],cam_mid_60_rich_summ_lat_bins[,c("bins","
nb_fe","Age")],
        cam_low_60_rich_summ_lat_bins[,c("bins","nb_fe","Age")])


# Used dcast to reconfigure the data so that Age and Site would be identifiers with unique values
GR_lat_bins_table_all <- dcast(GR_stack, bins ~ Age, value.var = "nb_sp")

# Used dcast to reconfigure the data so that Age and Site would be identifiers with unique values
FR_lat_bins_table_all <- dcast(FR_stack, bins ~ Age, value.var = "nb_fe")

# This are tables showing the number of genera and FE in each grid cell for each substage
head(GR_lat_bins_table_all)
head(FR_lat_bins_table_all)

# Export the dfS as a csv file
write.csv(GR_lat_bins_table_all, file="lat_bins_outputs/Lat_bins_GR_through_time.csv")
write.csv(FR_lat_bins_table_all, file="lat_bins_outputs/Lat_bins_FR_through_time.csv")

 #### Calculate and create Bar Plot of Simpsons Evenness for Latitudinal Bins ####
```

```
# simpsons measure of evenness calculation function

simpsons_calc <- function(data){

 data1 <- data.frame(data$"details_fdfe")


 # simpsons measure of evenness calculation:

 simp_calc <- function(data1){

  data1 <- data.frame(data1)

  n <- data1[data1!=0] # get vector of only values greater than 0 (no empty FE)

  if (length(n)>1 && max(n)>1){ # use if statement to only run the calc on nodes with >1 FE and more
than 1 genera in at least one FE

    N <- sum(n)

    S <- length(n)

    D = sum(n*(n-1))/(N*(N-1))

    inversD <- 1/D

    sD <- inversD/S

  } else {

    0

  }

 }


 final <- apply(data1,1,simp_calc) # 1 indicates that it applies function to rows


 data3 <- data.frame(final)

 data4 <- cbind(newColName = rownames(data3), data3)

 rownames(data4) <- 1:nrow(data4)

 colnames(data4) <- c("Site","Simp_Even_Index")


 return(data4)

}
```

```
# Calculate Simpson's measure of evenness for each lat bin in substages using above function
maa_up_60_simp_even_lat <- simpsons_calc(maa_up_60_fe_lat)
maa_low_60_simp_even_lat <- simpsons_calc(maa_low_60_fe_lat)
cam_up_60_simp_even_lat <- simpsons_calc(cam_up_60_fe_lat)
cam_mid_60_simp_even_lat <- simpsons_calc(cam_mid_60_fe_lat)
cam_low_60_simp_even_lat <- simpsons_calc(cam_low_60_fe_lat)


# plot bar graphs of simpson's evenness for each grid cell per substage
bar_plot_even_lat <- function(maa_up_even_data, maa_low_even_data,
                      cam_up_even_data,cam_mid_even_data,cam_low_even_data){

 colnames(maa_up_even_data) <- c("Bin","Simp_Even_Index")
 colnames(maa_low_even_data) <- c("Bin","Simp_Even_Index")
 colnames(cam_up_even_data) <- c("Bin","Simp_Even_Index")
 colnames(cam_mid_even_data) <- c("Bin","Simp_Even_Index")
 colnames(cam_low_even_data) <- c("Bin","Simp_Even_Index")


 # function to get the bin names (latitude bins) for each SS
 get_bin_names <- function(data){
  # get the unique lat bins available for this substage and change colname
  bins <- data.frame(unique(data$Bin))
  colnames(bins) <- "Lat_1"

  # Create column with the sorts of labels actually want to use based on ifelse
  bins$lat_bins <-with(bins, ifelse(Lat_1 == "(-Inf,30]", '23-30°N',
                            ifelse(Lat_1 == "(-Inf,35]", '30-35°N',
```

```
                                ifelse(Lat_1 == "(30,35]", '30-35°N',
                                    ifelse(Lat_1 == "(35,40]", '35-40°N',
                                        ifelse(Lat_1 == "(40,45]", '40-45°N',
                                            ifelse(Lat_1 == "(45,50]", '45-50°N',
                                                ifelse(Lat_1 == "(50,55]", '50-55°N',
                                                    ifelse(Lat_1 == "(55,60]", '55-60°N',
                                                        ifelse(Lat_1 == "(55, Inf]", '55-60°N',
                                                            '60-65°N' ))))))))))

  # sort the vector so they're in order from lowest to highest
  bin_labels <- sort(bins$lat_bins)
  return(bin_labels)
}

bin_labels1 <- get_bin_names(maa_up_even_data)
bin_labels2 <- get_bin_names(maa_low_even_data)
bin_labels3 <- get_bin_names(cam_up_even_data)
bin_labels4 <- get_bin_names(cam_mid_even_data)
bin_labels5 <- get_bin_names(cam_low_even_data)


### Plot the data using geom_col: ###
maa_up_bar <- ggplot(maa_up_even_data, aes(x=Bin, y=Simp_Even_Index, fill=Bin)) +
  geom_col(color="#e9ecef") +
  labs(title="Upper Maastr. Simpson's Diversity Index",x="Paleo-Latitude Bin", y = "SME") +
  theme(plot.title = element_text(hjust = 0.5),
        axis.text.x = element_text(angle = 0, vjust = 0.5, hjust=0.5)) +
  guides(fill= "none") +
  scale_fill_manual(values=c("red4",
```

```
                    "tomato3",
                    "darkorange3",
                    "goldenrod2",
                    "darkolivegreen4",
                    "aquamarine4",
                    "steelblue"))+
  scale_x_discrete(labels= bin_labels1) +
  ylim(0, 0.25 + max(maa_up_even_data$Simp_Even_Index)) +
  geom_text(aes(label=signif(Simp_Even_Index,3)), position=position_dodge(width=0.9), vjust=-0.25)

 maa_low_bar <- ggplot(maa_low_even_data, aes(x=Bin, y=Simp_Even_Index, fill=Bin)) +
  geom_col(color="#e9ecef") +
  labs(title="Lower Maastr. Simpson's Diversity Index",x="Paleo-Latitude Bin", y = "SME") +
  theme(plot.title = element_text(hjust = 0.5),
      axis.text.x = element_text(angle = 0, vjust = 0.5, hjust=0.5)) +
  guides(fill= "none") +
  scale_fill_manual(values=c("red4",
                    "tomato3",
                    "darkorange3",
                    "goldenrod2",
                    "darkolivegreen4",
                    "aquamarine4",
                    "steelblue"))+
  scale_x_discrete(labels= bin_labels2) +
  ylim(0, 0.25 + max(maa_low_even_data$Simp_Even_Index)) +
  geom_text(aes(label=signif(Simp_Even_Index,3)), position=position_dodge(width=0.9), vjust=-0.25)

 cam_up_bar <- ggplot(cam_up_even_data, aes(x=Bin, y=Simp_Even_Index, fill=Bin)) +
  geom_col(color="#e9ecef") +
```

```
  labs(title="Upper Camp. Simpson's Diversity Index",x="Paleo-Latitude Bin", y = "SME") +
  theme(plot.title = element_text(hjust = 0.5),
      axis.text.x = element_text(angle = 0, vjust = 0.5, hjust=0.5)) +
  guides(fill= "none") +
  scale_fill_manual(values=c("red4",
                             "tomato3",
                             "darkorange3",
                             "goldenrod2",
                             "darkolivegreen4",
                             "aquamarine4",
                             "steelblue"))+
  scale_x_discrete(labels= bin_labels3) +
  ylim(0, 0.25 + max(cam_up_even_data$Simp_Even_Index)) +
  geom_text(aes(label=signif(Simp_Even_Index,3)), position=position_dodge(width=0.9), vjust=-0.25)


cam_mid_bar <- ggplot(cam_mid_even_data, aes(x=Bin, y=Simp_Even_Index, fill=Bin)) +
  geom_col(color="#e9ecef") +
  labs(title="Middle Camp. Simpson's Diversity Index",x="Paleo-Latitude Bin", y = "SME") +
  theme(plot.title = element_text(hjust = 0.5),
      axis.text.x = element_text(angle = 0, vjust = 0.5, hjust=0.5)) +
  guides(fill= "none") +
  scale_fill_manual(values=c("red4",
                             "tomato3",
                             "darkorange3",
                             "goldenrod2",
                             "darkolivegreen4",
                             "aquamarine4",
                             "steelblue"))+
  scale_x_discrete(labels= bin_labels4) +
```

```
  ylim(0, 0.25 + max(cam_mid_even_data$Simp_Even_Index)) +
  geom_text(aes(label=signif(Simp_Even_Index,3)), position=position_dodge(width=0.9), vjust=-0.25)

 cam_low_bar <- ggplot(cam_low_even_data, aes(x=Bin, y=Simp_Even_Index, fill=Bin)) +
  geom_col(color="#e9ecef") +
  labs(title="Lower Camp. Simpson's Diversity Index",x="Paleo-Latitude Bin", y = "SME") +
  theme(plot.title = element_text(hjust = 0.5),
      axis.text.x = element_text(angle = 0, vjust = 0.5, hjust=0.5)) +
  guides(fill= "none") +
  scale_fill_manual(values=c("red4",
                             "tomato3",
                             "darkorange3",
                             "goldenrod2",
                             "darkolivegreen4",
                             "aquamarine4",
                             "steelblue"))+
  scale_x_discrete(labels= bin_labels5) +
  ylim(0, 0.25 + max(cam_low_even_data$Simp_Even_Index)) +
  geom_text(aes(label=signif(Simp_Even_Index,3)), position=position_dodge(width=0.9), vjust=-0.25)

 bar_list <- list(maa_up_bar,maa_low_bar,cam_up_bar,cam_mid_bar,cam_low_bar)
}

simp_even_barplots_lat <-
bar_plot_even_lat(maa_up_60_simp_even_lat,maa_low_60_simp_even_lat,cam_up_60_simp_even_lat
,
                              cam_mid_60_simp_even_lat,cam_low_60_simp_even_lat)

# plot bar plots of the Simpson's diversity index per grid cell for each substage
```

```
pdf("lat_bins_outputs/bar_plots_SimpEven_lat_bins.pdf", width = 15, height = 25) # Open a new pdf file
grid.arrange(grobs = simp_even_barplots_lat,
          top = "Bar Plots of Simpson's Evenness per Latitudinal Bin") # Write the grid.arrange in the file
dev.off() #close file


 #### Create line plots of Simpsons for Lat Bins ####


# function to compare GR and simpsons evenness in line plots for each substage
plot_line_lat_simp <- function(fe_data1,fe_data2,fe_data3,fe_data4,fe_data5,
                       data1,data2,data3,data4,data5,yaxis,legendlabel){


 # function to get the Age data as first column and make relevant columns numeric
 df_func <- function(data){
  ###  Get the functional ecology information for the substage (this is a dataframe) and reconfig: ###
  data1 <- cbind(Bin = rownames(data), data) # make the lat bin name the 1st column again
  rownames(data1) <- 1:nrow(data1) # make a new index for the rownames
  data1 <- as.data.frame(data1)
  i <- c(2:6)
  data1[ , i] <- apply(data1[ , i], 2,           # Specify own function within apply to make numeric
                function(x) as.numeric(as.character(x)))
  return(as.data.frame(data1))
 }


 # use above function on the fe df output to get lat bins as first column and make nb_fe and nb_sp
numeric
 fe_data1 <- df_func(fe_data1$asb_fdfe)
 fe_data2 <- df_func(fe_data2$asb_fdfe)
 fe_data3 <- df_func(fe_data3$asb_fdfe)
 fe_data4 <- df_func(fe_data4$asb_fdfe)
```

```
fe_data5 <- df_func(fe_data5$asb_fdfe)



# bind the lat bin and gr (nb_sp column) with the fr for each
bind_data1 <- cbind(fe_data1[,c(1,2)],data1[,2])
bind_data2 <- cbind(fe_data2[,c(1,2)],data2[,2])
bind_data3 <- cbind(fe_data3[,c(1,2)],data3[,2])
bind_data4 <- cbind(fe_data4[,c(1,2)],data4[,2])
bind_data5 <- cbind(fe_data5[,c(1,2)],data5[,2])

# change the column names
colnames(bind_data1) <- c("Bin","GR","other")
colnames(bind_data2) <- c("Bin","GR","other")
colnames(bind_data3) <- c("Bin","GR","other")
colnames(bind_data4) <- c("Bin","GR","other")
colnames(bind_data5) <- c("Bin","GR","other")



# function to get the bin names (latitude bins) for each SS
get_bin_names <- function(data){
  # get the unique lat bins available for this substage and change colname
  bins <- data.frame(unique(data$Bin))
  colnames(bins) <- "Lat_1"

  # Create column with the sorts of labels actually want to use based on ifelse
  bins$lat_bins <-with(bins, ifelse(Lat_1 == "(-Inf,30]", '23-30°N',
                                    ifelse(Lat_1 == "(-Inf,35]", '30-35°N',
                                           ifelse(Lat_1 == "(30,35]", '30-35°N',
                                                  ifelse(Lat_1 == "(35,40]", '35-40°N',
```

```
                                    ifelse(Lat_1 == "(40,45]", '40-45°N',
                                        ifelse(Lat_1 == "(45,50]", '45-50°N',
                                            ifelse(Lat_1 == "(50,55]", '50-55°N',
                                                ifelse(Lat_1 == "(55,60]", '55-60°N',
                                                    ifelse(Lat_1 == "(55, Inf]", '55-60°N',
                                                        '60-65°N' ))))))))))

  # sort the vector so they're in order from lowest to highest
  bin_labels <- sort(bins$lat_bins)
  return(bin_labels)
}

bin_labels1 <- get_bin_names(bind_data1)
bin_labels2 <- get_bin_names(bind_data2)
bin_labels3 <- get_bin_names(bind_data3)
bin_labels4 <- get_bin_names(bind_data4)
bin_labels5 <- get_bin_names(bind_data5)


bin1_line <- ggplot(bind_data1, aes(x=Bin)) +
  geom_line(aes(y = GR, color = "#D55E00",group=1)) +
  geom_line(aes(y = 50*other, color="steelblue",group=2),linetype = "dashed") + # have multiplied the
other variable by 50
  geom_point(aes(y = GR), color = "#D55E00") +
  geom_point(aes(y = 50*other), color="steelblue") +
  scale_y_continuous(
    # Features of the first axis
    name = "# Genera",
    # Add a second axis and specify its features
```

```
    sec.axis = sec_axis(~ . * 0.02,name=yaxis) # scale the other y axis to match multiplication of other
variable
  ) +
  scale_color_discrete(labels=c("Generic Richness", legendlabel)) +
  labs(title=paste("Upper Maastr.")) +
  xlab(label="Latitude") +
  theme(legend.title=element_blank(),plot.title = element_text(hjust = 0.5)) +
  scale_x_discrete(labels= bin_labels1)

 bin2_line <- ggplot(bind_data2, aes(x=Bin)) +
  geom_line(aes(y = GR, color = "#D55E00",group=1)) +
  geom_line(aes(y = 50*other, color="steelblue",group=2),linetype = "dashed") + # have multiplied the
other variable by 50
  geom_point(aes(y = GR), color = "#D55E00") +
  geom_point(aes(y = 50*other), color="steelblue") +
  scale_y_continuous(
   # Features of the first axis
   name = "# Genera",
   # Add a second axis and specify its features
   sec.axis = sec_axis(~ . * 0.02,name=yaxis)# scale the other y axis to match multiplication of other
variable
  )  +
  scale_color_discrete(labels=c("Generic Richness", legendlabel)) +
  labs(title=paste("Lower Maastr.")) +
  xlab(label="Latitude") +
  theme(legend.title=element_blank(),plot.title = element_text(hjust = 0.5)) +
  scale_x_discrete(labels= bin_labels2)

 bin3_line <- ggplot(bind_data3, aes(x=Bin)) +
  geom_line(aes(y = GR, color = "#D55E00",group=1)) +
```

```
  geom_line(aes(y = 50*other, color="steelblue",group=2),linetype = "dashed") + # have multiplied the other variable by 50
  geom_point(aes(y = GR), color = "#D55E00") +
  geom_point(aes(y = 50*other), color="steelblue") +
  scale_y_continuous(
    # Features of the first axis
    name = "# Genera",
    # Add a second axis and specify its features
    sec.axis = sec_axis(~ . * 0.02,name=yaxis)# scale the other y axis to match multiplication of other variable
  ) +
  scale_color_discrete(labels=c("Generic Richness", legendlabel)) +
  labs(title=paste("Upper Camp.")) +
  xlab(label="Latitude") +
  theme(legend.title=element_blank(),plot.title = element_text(hjust = 0.5)) +
  scale_x_discrete(labels= bin_labels3)


 bin4_line <- ggplot(bind_data4, aes(x=Bin)) +
  geom_line(aes(y = GR, color = "#D55E00",group=1)) +
  geom_line(aes(y = 50*other, color="steelblue",group=2),linetype = "dashed") + # have multiplied the other variable by 50
  geom_point(aes(y = GR), color = "#D55E00") +
  geom_point(aes(y = 50*other), color="steelblue") +
  scale_y_continuous(
    # Features of the first axis
    name = "# Genera",
    # Add a second axis and specify its features
    sec.axis = sec_axis(~ . * 0.02,name=yaxis) # scale the other y axis to match multiplication of other variable
  ) +
```

```
    scale_color_discrete(labels=c("Generic Richness", legendlabel)) +
    labs(title=paste("Middle Camp.")) +
    xlab(label="Latitude") +
    theme(legend.title=element_blank(),plot.title = element_text(hjust = 0.5)) +
    scale_x_discrete(labels= bin_labels4)

  bin5_line <- ggplot(bind_data5, aes(x=Bin)) +
    geom_line(aes(y = GR, color = "#D55E00",group=1)) +
    geom_line(aes(y = 50*other, color="steelblue",group=2),linetype = "dashed") + # have multiplied the
other variable by 50
    geom_point(aes(y = GR), color = "#D55E00") +
    geom_point(aes(y = 50*other), color="steelblue") +
    scale_y_continuous(
      # Features of the first axis
      name = "# Genera",
      # Add a second axis and specify its features
      sec.axis = sec_axis(~ . * 0.02,name=yaxis) # scale the other y axis to match multiplication of other
variable
    )  +
    scale_color_discrete(labels=c("Generic Richness", legendlabel)) +
    labs(title=paste("Lower Camp.")) +
    xlab(label="Latitude") +
    theme(legend.title=element_blank(),plot.title = element_text(hjust = 0.5)) +
    scale_x_discrete(labels= bin_labels5)

  bin_list <- list(bin1_line,bin2_line,bin3_line,bin4_line,bin5_line)
  return(bin_list)
}
```

```
# create list of plots for changes in simpsons and gr across lat in diff substages

plot_bins_gr_simp <-
plot_line_lat_simp(maa_up_60_fe_lat,maa_low_60_fe_lat,cam_up_60_fe_lat,cam_mid_60_fe_lat,cam_
low_60_fe_lat,


maa_up_60_simp_even_lat,maa_low_60_simp_even_lat,cam_up_60_simp_even_lat,cam_mid_60_sim
p_even_lat,

                        cam_low_60_simp_even_lat, "D/S","Simpson's Measure of Evenness")



#### NOTE: I have rescaled the evenness data by multiplying by 100, ggplot not able to plot 2 scales

# but I also rescaled and the second y axis to match (by 0.01).


# Save plots of changes in simpsons and gr across latitude

pdf("lat_bins_outputs/lat_bin_GR_Simp_allsubstages.pdf", width = 20, height = 25) # Open a new pdf
file

grid.arrange(grobs = plot_bins_gr_simp,

        top = "Average Generic Richness and Simpson's Measure of Evenness across Latitude") # Write
the grid.arrange in the file

dev.off() #close the file



 #### Export table of Simpsons through time across Latitude Bins ####


# Function to configure to columns with site names as 1st and evenness values as second, age as 3rd

config_for_loc_even_agg_latbins <- function(data, Age){

 # Create df that just take the functional evenness values in each lat bin

 data_even_summ <- data.frame(data)

 colnames(data_even_summ) <- c("bins","even")
```

```
  data_even_summ$Age <- Age

  return(data_even_summ)

}


# Use config function (above) to alter then added new column with Age

maa_up_60_simp_summ_lat_bins <- config_for_loc_even_agg_latbins(maa_up_60_simp_even_lat,
"MAA up")

maa_low_60_simp_summ_lat_bins <- config_for_loc_even_agg_latbins(maa_low_60_simp_even_lat,
"MAA low")

cam_up_60_simp_summ_lat_bins <- config_for_loc_even_agg_latbins(cam_up_60_simp_even_lat,
"CAM up")

cam_mid_60_simp_summ_lat_bins <- config_for_loc_even_agg_latbins(cam_mid_60_simp_even_lat,
"CAM mid")

cam_low_60_simp_summ_lat_bins <- config_for_loc_even_agg_latbins(cam_low_60_simp_even_lat,
"CAM low")



# Function to fix row names (so all the same)

make_lat_names_even <- function(data){

  data <- data.frame(data)

  data$bins <-with(data, ifelse(bins == "(-Inf,30]", '(-Inf,30]',

                          ifelse(bins == "(-Inf,35]", '(30,35]',

                              ifelse(bins == "(30,35]", '(30,35]',

                                  ifelse(bins == "(35,40]", '(35,40]',

                                      ifelse(bins == "(-Inf,45]", '(40,45]',

                                      ifelse(bins == "(40,45]", '(40,45]',

                                          ifelse(bins == "(45,50]", '(45,50]',

                                              ifelse(bins == "(50,55]", '(50,55]',

                                                  ifelse(bins == "(55,60]", '(55,60]',

                                                      ifelse(bins == "(55, Inf]", '(55,60]',
```

```
                                                ifelse(bins == '(60,65]', '(60, Inf]',
                                                    '(60, Inf]' ))))))))))))
  return(data)
}

# Use above function to make sure lat bin names are all the same when they mean the same thing
maa_up_60_simp_summ_lat_bins <- make_lat_names_even(maa_up_60_simp_summ_lat_bins)
maa_low_60_simp_summ_lat_bins <- make_lat_names_even(maa_low_60_simp_summ_lat_bins)
cam_up_60_simp_summ_lat_bins <- make_lat_names_even(cam_up_60_simp_summ_lat_bins)
cam_mid_60_simp_summ_lat_bins <- make_lat_names_even(cam_mid_60_simp_summ_lat_bins)
cam_low_60_simp_summ_lat_bins <- make_lat_names_even(cam_low_60_simp_summ_lat_bins)

# Bind the GR data into a single df with 2 columns (stacked the different dfs)
simp_stack_lat_bins <- rbind(maa_up_60_simp_summ_lat_bins,maa_low_60_simp_summ_lat_bins,
                     cam_up_60_simp_summ_lat_bins,cam_mid_60_simp_summ_lat_bins,
                     cam_low_60_simp_summ_lat_bins)

# Used dcast to reconfigure the data so that Age and Site would be identifiers with unique values
simp_stack_lat_bins_table_all <- dcast(simp_stack_lat_bins, bins ~ Age, value.var = "even")

# This are tables showing the Simpsons values in each lat bin for each substage
head(simp_stack_lat_bins_table_all)

# Export the df as a csv file
write.csv(simp_stack_lat_bins_table_all, file="lat_bins_outputs/Lat_bins_Simp_through_time.csv")

  #### Calculate and create Bar Plot of Shannon Equability for Latitudinal Bins ####

# Shannon equitability calculation function
```

```
shannons_calc <- function(data){

  # equation for shannons calc
  shannon_equit <- function(data){
    data1 <- data.frame(data)
    data1 <- data1[data1!=0] # get vector of only values greater than 0 (no empty FE)

    if (length(data1)>1){ # use if statement to only run the calc on nodes with >1 FE and more than 1 genera in at least one FE

      S <- length(which(data1!=0)) # get the numb of FE
      Si <- sum(data1) # find out how many genera total present in all FE
      Pvals <- data1/Si # get proportions of each FE based on dividing # genera in each by total # genera
      nlPvals <- log(Pvals) # take natural log of the proportions
      H <- -sum(nlPvals*Pvals) # calculate Shannon diversity index (H) by multiplying nl of p by p and summing and multiply by neg 1
      LNS <- log(S) # get natural log of number of FE present
      Hi <- H/LNS # calculate Shannon Equitability Index (Hi) by dividing H by nat log of S
    } else {
      0
    }
  }

  data1 <- data.frame(data$"details_fdfe")

  final <- apply(data1,1,shannon_equit) # 1 indicates that it applies function to rows

  data3 <- data.frame(final)
  data4 <- cbind(newColName = rownames(data3), data3) # make row name (i.e., grid cell) first column again
```

```
  rownames(data4) <- 1:nrow(data4)
  colnames(data4) <- c("Site","Shan_Equit_Index")

  return(data4)
}

# Calculate Shannon Equitability Index for each lat bin in substages using above function
maa_up_60_shan_even_lat <- shannons_calc(maa_up_60_fe_lat)
maa_low_60_shan_even_lat <- shannons_calc(maa_low_60_fe_lat)
cam_up_60_shan_even_lat <- shannons_calc(cam_up_60_fe_lat)
cam_mid_60_shan_even_lat <- shannons_calc(cam_mid_60_fe_lat)
cam_low_60_shan_even_lat <- shannons_calc(cam_low_60_fe_lat)



# plot bar graphs of Shannon Equitability Index for each grid cell per substage
bar_plot_equit_lat <- function(maa_up_even_data, maa_low_even_data,
                       cam_up_even_data,cam_mid_even_data,cam_low_even_data){

  colnames(maa_up_even_data) <- c("Bin","Shan_Equit_Index")
  colnames(maa_low_even_data) <- c("Bin","Shan_Equit_Index")
  colnames(cam_up_even_data) <- c("Bin","Shan_Equit_Index")
  colnames(cam_mid_even_data) <- c("Bin","Shan_Equit_Index")
  colnames(cam_low_even_data) <- c("Bin","Shan_Equit_Index")

  # function to get the bin names (latitude bins) for each SS
  get_bin_names <- function(data){
    # get the unique lat bins available for this substage and change colname
    bins <- data.frame(unique(data$Bin))
    colnames(bins) <- "Lat_1"
```

```
  # Create column with the sorts of labels actually want to use based on ifelse
  bins$lat_bins <-with(bins, ifelse(Lat_1 == "(-Inf,30]", '23-30°N',
                                    ifelse(Lat_1 == "(-Inf,35]", '30-35°N',
                                           ifelse(Lat_1 == "(30,35]", '30-35°N',
                                                  ifelse(Lat_1 == "(35,40]", '35-40°N',
                                                         ifelse(Lat_1 == "(40,45]", '40-45°N',
                                                                ifelse(Lat_1 == "(45,50]", '45-50°N',
                                                                       ifelse(Lat_1 == "(50,55]", '50-55°N',
                                                                              ifelse(Lat_1 == "(55,60]", '55-60°N',
                                                                                     ifelse(Lat_1 == "(55, Inf]", '55-60°N',
                                                                                            '60-65°N' ))))))))))

  # sort the vector so they're in order from lowest to highest
  bin_labels <- sort(bins$lat_bins)
  return(bin_labels)
}

bin_labels1 <- get_bin_names(maa_up_even_data)
bin_labels2 <- get_bin_names(maa_low_even_data)
bin_labels3 <- get_bin_names(cam_up_even_data)
bin_labels4 <- get_bin_names(cam_mid_even_data)
bin_labels5 <- get_bin_names(cam_low_even_data)



### Plot the data using geom_col: ###
maa_up_bar <- ggplot(maa_up_even_data, aes(x=Bin, y=Shan_Equit_Index, fill=Bin)) +
  geom_col(color="#e9ecef") +
```

```
  labs(title="Upper Maastr. Shannon Equitability Index",x="Paleo-Latitude Bin", y = "SEI") +
  theme(plot.title = element_text(hjust = 0.5),
      axis.text.x = element_text(angle = 0, vjust = 0.5, hjust=0.5)) +
  guides(fill="none") +
  scale_fill_manual(values=c("red4",
                             "tomato3",
                             "darkorange3",
                             "goldenrod2",
                             "darkolivegreen4",
                             "aquamarine4",
                             "steelblue"))+
  scale_x_discrete(labels= bin_labels1) +
  ylim(0, 0.25 + max(maa_up_even_data$Shan_Equit_Index)) +
  geom_text(aes(label=signif(Shan_Equit_Index,3)), position=position_dodge(width=0.9), vjust=-0.25)

maa_low_bar <- ggplot(maa_low_even_data, aes(x=Bin, y=Shan_Equit_Index, fill=Bin)) +
  geom_col(color="#e9ecef") +
  labs(title="Lower Maastr. Shannon Equitability Index",x="Paleo-Latitude Bin", y = "SEI") +
  theme(plot.title = element_text(hjust = 0.5),
      axis.text.x = element_text(angle = 0, vjust = 0.5, hjust=0.5)) +
  guides(fill="none") +
  scale_fill_manual(values=c("red4",
                             "tomato3",
                             "darkorange3",
                             "goldenrod2",
                             "darkolivegreen4",
                             "aquamarine4",
                             "steelblue"))+
  scale_x_discrete(labels= bin_labels2) +
```

```
  ylim(0, 0.25 + max(maa_low_even_data$Shan_Equit_Index)) +
  geom_text(aes(label=signif(Shan_Equit_Index,3)), position=position_dodge(width=0.9), vjust=-0.25)


 cam_up_bar <- ggplot(cam_up_even_data, aes(x=Bin, y=Shan_Equit_Index, fill=Bin)) +
  geom_col(color="#e9ecef") +
  labs(title="Upper Camp. Shannon Equitability Index",x="Paleo-Latitude Bin", y = "SEI") +
  theme(plot.title = element_text(hjust = 0.5),
      axis.text.x = element_text(angle = 0, vjust = 0.5, hjust=0.5)) +
  guides(fill="none") +
  scale_fill_manual(values=c("red4",
                             "tomato3",
                             "darkorange3",
                             "goldenrod2",
                             "darkolivegreen4",
                             "aquamarine4",
                             "steelblue"))+
  scale_x_discrete(labels= bin_labels3) +
  ylim(0, 0.25 + max(cam_up_even_data$Shan_Equit_Index)) +
  geom_text(aes(label=signif(Shan_Equit_Index,3)), position=position_dodge(width=0.9), vjust=-0.25)


 cam_mid_bar <- ggplot(cam_mid_even_data, aes(x=Bin, y=Shan_Equit_Index, fill=Bin)) +
  geom_col(color="#e9ecef") +
  labs(title="Middle Camp. Shannon Equitability Index",x="Paleo-Latitude Bin", y = "SEI") +
  theme(plot.title = element_text(hjust = 0.5),
      axis.text.x = element_text(angle = 0, vjust = 0.5, hjust=0.5)) +
  guides(fill="none") +
  scale_fill_manual(values=c("red4",
                             "tomato3",
                             "darkorange3",
```

```
                    "goldenrod2",
                    "darkolivegreen4",
                    "aquamarine4",
                    "steelblue"))+
  scale_x_discrete(labels= bin_labels4) +
  ylim(0, 0.25 + max(cam_mid_even_data$Shan_Equit_Index)) +
  geom_text(aes(label=signif(Shan_Equit_Index,3)), position=position_dodge(width=0.9), vjust=-0.25)

 cam_low_bar <- ggplot(cam_low_even_data, aes(x=Bin, y=Shan_Equit_Index, fill=Bin)) +
  geom_col(color="#e9ecef") +
  labs(title="Lower Camp. Shannon Equitability Index",x="Paleo-Latitude Bin", y = "SEI") +
  theme(plot.title = element_text(hjust = 0.5),
      axis.text.x = element_text(angle = 0, vjust = 0.5, hjust=0.5)) +
  guides(fill="none") +
  scale_fill_manual(values=c("red4",
                    "tomato3",
                    "darkorange3",
                    "goldenrod2",
                    "darkolivegreen4",
                    "aquamarine4",
                    "steelblue"))+
  scale_x_discrete(labels= bin_labels5) +
  ylim(0, 0.25 + max(cam_low_even_data$Shan_Equit_Index)) +
  geom_text(aes(label=signif(Shan_Equit_Index,3)), position=position_dodge(width=0.9), vjust=-0.25)

 bar_list <- list(maa_up_bar,maa_low_bar,cam_up_bar,cam_mid_bar,cam_low_bar)
}
```

```
shan_even_barplots_lat <- bar_plot_equit_lat(maa_up_60_shan_even_lat,
                                             maa_low_60_shan_even_lat,
                                             cam_up_60_shan_even_lat,
                                             cam_mid_60_shan_even_lat,
                                             cam_low_60_shan_even_lat)

# plot bar plots of the Simpson's diversity index per grid cell for each substage
pdf("lat_bins_outputs/bar_plots_ShanEquit_lat_bins.pdf", width = 15, height = 25) # Open a new pdf file
grid.arrange(grobs = shan_even_barplots_lat,
          top = "Bar Plots of Shannon Equitability per Latitudinal Bin") # Write the grid.arrange in the file
dev.off() #close file




 #### Create line plots of Shannon for Lat Bins ####

# function to compare GR and shannon equitability index in line plots for each substage
plot_line_lat_shan <- function(fe_data1,fe_data2,fe_data3,fe_data4,fe_data5,
                               data1,data2,data3,data4,data5,yaxis,legendlabel){

 # function to get the Age data as first column and make relevant columns numeric
 df_func <- function(data){
  ###  Get the functional ecology information for the substage (this is a dataframe) and reconfig: ###
  data1 <- cbind(Bin = rownames(data), data) # make the lat bin name the 1st column again
  rownames(data1) <- 1:nrow(data1) # make a new index for the rownames
  data1 <- as.data.frame(data1)
  i <- c(2:6)
  data1[ , i] <- apply(data1[ , i], 2,           # Specify own function within apply to make numeric
```

```
                function(x) as.numeric(as.character(x)))
  return(as.data.frame(data1))
 }

 # use above function on the fe df output to get lat bins as first column and make nb_fe and nb_sp
numeric
 fe_data1 <- df_func(fe_data1$asb_fdfe)
 fe_data2 <- df_func(fe_data2$asb_fdfe)
 fe_data3 <- df_func(fe_data3$asb_fdfe)
 fe_data4 <- df_func(fe_data4$asb_fdfe)
 fe_data5 <- df_func(fe_data5$asb_fdfe)


 # bind the lat bin and gr (nb_sp column) with the fr for each
 bind_data1 <- cbind(fe_data1[,c(1,2)],data1[,2])
 bind_data2 <- cbind(fe_data2[,c(1,2)],data2[,2])
 bind_data3 <- cbind(fe_data3[,c(1,2)],data3[,2])
 bind_data4 <- cbind(fe_data4[,c(1,2)],data4[,2])
 bind_data5 <- cbind(fe_data5[,c(1,2)],data5[,2])

 # change the column names
 colnames(bind_data1) <- c("Bin","GR","other")
 colnames(bind_data2) <- c("Bin","GR","other")
 colnames(bind_data3) <- c("Bin","GR","other")
 colnames(bind_data4) <- c("Bin","GR","other")
 colnames(bind_data5) <- c("Bin","GR","other")
```

```
# function to get the bin names (latitude bins) for each SS
get_bin_names <- function(data){
 # get the unique lat bins available for this substage and change colname
 bins <- data.frame(unique(data$Bin))
 colnames(bins) <- "Lat_1"

 # Create column with the sorts of labels actually want to use based on ifelse
 bins$lat_bins <-with(bins, ifelse(Lat_1 == "(-Inf,30]", '23-30°N',
                          ifelse(Lat_1 == "(-Inf,35]", '30-35°N',
                              ifelse(Lat_1 == "(30,35]", '30-35°N',
                                  ifelse(Lat_1 == "(35,40]", '35-40°N',
                                      ifelse(Lat_1 == "(40,45]", '40-45°N',
                                          ifelse(Lat_1 == "(45,50]", '45-50°N',
                                              ifelse(Lat_1 == "(50,55]", '50-55°N',
                                                  ifelse(Lat_1 == "(55,60]", '55-60°N',
                                                      ifelse(Lat_1 == "(55, Inf]", '55-60°N',
                                                          '60-65°N' ))))))))))

 # sort the vector so they're in order from lowest to highest
 bin_labels <- sort(bins$lat_bins)
 return(bin_labels)
}

bin_labels1 <- get_bin_names(bind_data1)
bin_labels2 <- get_bin_names(bind_data2)
bin_labels3 <- get_bin_names(bind_data3)
bin_labels4 <- get_bin_names(bind_data4)
bin_labels5 <- get_bin_names(bind_data5)
```

```
 bin1_line <- ggplot(bind_data1, aes(x=Bin)) +
   geom_line(aes(y = GR, color = "#D55E00",group=1)) +
   geom_line(aes(y = 50*other, color="steelblue"),group=2,linetype = "dashed") + # have multiplied the other variable by 50
   geom_point(aes(y = GR), color = "#D55E00") +
   geom_point(aes(y = 50*other), color="steelblue") +
   scale_y_continuous(
     # Features of the first axis
     name = "# Genera",
     # Add a second axis and specify its features
     sec.axis = sec_axis(~ . * 0.02,name=yaxis) # scale the other y axis to match multiplication of other variable
   ) +
   scale_color_discrete(labels=c("Generic Richness", legendlabel)) +
   labs(title=paste("Upper Maastr.")) +
   xlab(label="Latitude") +
   theme(legend.title=element_blank(),plot.title = element_text(hjust = 0.5)) +
   scale_x_discrete(labels= bin_labels1)

 bin2_line <- ggplot(bind_data2, aes(x=Bin)) +
   geom_line(aes(y = GR, color = "#D55E00",group=1)) +
   geom_line(aes(y = 50*other, color="steelblue"),group=2,linetype = "dashed") + # have multiplied the other variable by 50
   geom_point(aes(y = GR), color = "#D55E00") +
   geom_point(aes(y = 50*other), color="steelblue") +
   scale_y_continuous(
     # Features of the first axis
     name = "# Genera",
     # Add a second axis and specify its features
```

```
    sec.axis = sec_axis(~ . * 0.02,name=yaxis)# scale the other y axis to match multiplication of other
variable
  )  +
  scale_color_discrete(labels=c("Generic Richness", legendlabel)) +
  labs(title=paste("Lower Maastr.")) +
  xlab(label="Latitude") +
  theme(legend.title=element_blank(),plot.title = element_text(hjust = 0.5)) +
  scale_x_discrete(labels= bin_labels2)

 bin3_line <- ggplot(bind_data3, aes(x=Bin)) +
  geom_line(aes(y = GR, color = "#D55E00",group=1)) +
  geom_line(aes(y = 50*other, color="steelblue"),group=2,linetype = "dashed") + # have multiplied the
other variable by 50
  geom_point(aes(y = GR), color = "#D55E00") +
  geom_point(aes(y = 50*other), color="steelblue") +
  scale_y_continuous(
    # Features of the first axis
    name = "# Genera",
    # Add a second axis and specify its features
    sec.axis = sec_axis(~ . * 0.02,name=yaxis)# scale the other y axis to match multiplication of other
variable
  )  +
  scale_color_discrete(labels=c("Generic Richness", legendlabel)) +
  labs(title=paste("Upper Camp.")) +
  xlab(label="Latitude") +
  theme(legend.title=element_blank(),plot.title = element_text(hjust = 0.5)) +
  scale_x_discrete(labels= bin_labels3)

 bin4_line <- ggplot(bind_data4, aes(x=Bin)) +
  geom_line(aes(y = GR, color = "#D55E00",group=1)) +
```

```
  geom_line(aes(y = 50*other, color="steelblue"),group=2,linetype = "dashed") + # have multiplied the other variable by 50
  geom_point(aes(y = GR), color = "#D55E00") +
  geom_point(aes(y = 50*other), color="steelblue") +
  scale_y_continuous(
    # Features of the first axis
    name = "# Genera",
    # Add a second axis and specify its features
    sec.axis = sec_axis(~ . * 0.02,name=yaxis) # scale the other y axis to match multiplication of other variable
  ) +
  scale_color_discrete(labels=c("Generic Richness", legendlabel)) +
  labs(title=paste("Middle Camp.")) +
  xlab(label="Latitude") +
  theme(legend.title=element_blank(),plot.title = element_text(hjust = 0.5)) +
  scale_x_discrete(labels= bin_labels4)


 bin5_line <- ggplot(bind_data5, aes(x=Bin)) +
  geom_line(aes(y = GR, color = "#D55E00",group=1)) +
  geom_line(aes(y = 50*other, color="steelblue"),group=2,linetype = "dashed") + # have multiplied the other variable by 50
  geom_point(aes(y = GR), color = "#D55E00") +
  geom_point(aes(y = 50*other), color="steelblue") +
  scale_y_continuous(
    # Features of the first axis
    name = "# Genera",
    # Add a second axis and specify its features
    sec.axis = sec_axis(~ . * 0.02,name=yaxis) # scale the other y axis to match multiplication of other variable
  ) +
```

```
    scale_color_discrete(labels=c("Generic Richness", legendlabel)) +
    labs(title=paste("Lower Camp.")) +
    xlab(label="Latitude") +
    theme(legend.title=element_blank(),plot.title = element_text(hjust = 0.5)) +
    scale_x_discrete(labels= bin_labels5)

  bin_list <- list(bin1_line,bin2_line,bin3_line,bin4_line,bin5_line)
  return(bin_list)
}



# create list of plots for changes in simpsons and gr across lat in diff substages
plot_bins_gr_shan <- plot_line_lat_shan(maa_up_60_fe_lat,maa_low_60_fe_lat,
                                        cam_up_60_fe_lat,cam_mid_60_fe_lat,
                                        cam_low_60_fe_lat,
                                        maa_up_60_shan_even_lat,
                                        maa_low_60_shan_even_lat,
                                        cam_up_60_shan_even_lat,
                                        cam_mid_60_shan_even_lat,
                                        cam_low_60_shan_even_lat, "Hi","Shannon Equitability Index")



#### NOTE: I have rescaled the evenness data by multiplying by 50, ggplot not able to plot 2 scales
# but I also resealed and the second y axis to match (by 0.02).

# Save plots of changes in simpsons and gr across latitude
pdf("lat_bins_outputs/lat_bin_GR_Shan_allsubstages.pdf", width = 20, height = 25) # Open a new pdf
file
grid.arrange(grobs = plot_bins_gr_shan,
```

```
        top = "Average Generic Richness and Simpson's Measure of Evenness across Latitude") # Write the grid.arrange in the file

dev.off() #close the file




 #### Export table of Shannon through time across Latitude Bins ####


# Function to configure to columns with site names as 1st and evenness values as second, age as 3rd

config_for_loc_even_agg_latbins <- function(data, Age){

 # Create df that just take the functional evenness values in each lat bin

 data_even_summ <- data.frame(data)

 colnames(data_even_summ) <- c("bins","even")

 data_even_summ$Age <- Age

 return(data_even_summ)

}


# Use config function (above) to alter then added new column with Age

maa_up_60_shan_summ_lat_bins <- config_for_loc_even_agg_latbins(maa_up_60_shan_even_lat, "MAA up")

maa_low_60_shan_summ_lat_bins <- config_for_loc_even_agg_latbins(maa_low_60_shan_even_lat, "MAA low")

cam_up_60_shan_summ_lat_bins <- config_for_loc_even_agg_latbins(cam_up_60_shan_even_lat, "CAM up")

cam_mid_60_shan_summ_lat_bins <- config_for_loc_even_agg_latbins(cam_mid_60_shan_even_lat, "CAM mid")

cam_low_60_shan_summ_lat_bins <- config_for_loc_even_agg_latbins(cam_low_60_shan_even_lat, "CAM low")
```

```
# Function to fix row names (so all the same)
make_lat_names_even <- function(data){
  data <- data.frame(data)
  data$bins <-with(data, ifelse(bins == "(-Inf,30]", '(-Inf,30]',
                          ifelse(bins == "(-Inf,35]", '(30,35]',
                                 ifelse(bins == "(30,35]", '(30,35]',
                                        ifelse(bins == "(35,40]", '(35,40]',
                                               ifelse(bins == "(-Inf,45]", '(40,45]',
                                               ifelse(bins == "(40,45]", '(40,45]',
                                                      ifelse(bins == "(45,50]", '(45,50]',
                                                             ifelse(bins == "(50,55]", '(50,55]',
                                                                    ifelse(bins == "(55,60]", '(55,60]',
                                                                           ifelse(bins == "(55, Inf]", '(55,60]',
                                                                                  ifelse(bins == '(60,65]', '(60, Inf]',
                                                                                         '(60, Inf]' ))))))))))))
  return(data)
}

# Use above function to make sure lat bin names are all the same when they mean the same thing
maa_up_60_shan_summ_lat_bins <- make_lat_names_even(maa_up_60_shan_summ_lat_bins)
maa_low_60_shan_summ_lat_bins <- make_lat_names_even(maa_low_60_shan_summ_lat_bins)
cam_up_60_shan_summ_lat_bins <- make_lat_names_even(cam_up_60_shan_summ_lat_bins)
cam_mid_60_shan_summ_lat_bins <- make_lat_names_even(cam_mid_60_shan_summ_lat_bins)
cam_low_60_shan_summ_lat_bins <- make_lat_names_even(cam_low_60_shan_summ_lat_bins)

# Bind the GR data into a single df with 2 columns (stacked the different dfs)
shan_stack_lat_bins <- rbind(maa_up_60_shan_summ_lat_bins,maa_low_60_shan_summ_lat_bins,
                             cam_up_60_shan_summ_lat_bins,cam_mid_60_shan_summ_lat_bins,
                             cam_low_60_shan_summ_lat_bins)
```

```
# Used dcast to reconfigure the data so that Age and Site would be identifiers with unique values
shan_stack_lat_bins_table_all <- dcast(shan_stack_lat_bins, bins ~ Age, value.var = "even")

# This are tables showing the Simpsons values in each lat bin for each substage
head(shan_stack_lat_bins_table_all)

# Export the df as a csv file
write.csv(shan_stack_lat_bins_table_all, file="lat_bins_outputs/Lat_bins_Shan_through_time.csv")

##### ANALYSIS USING ALL DATA TO GET VALUES FOR EACH SUBSTAGE OVERALL (NO GRIDS OR LAT BINS) #####

#Function for removing duplicate Genus-FE-Age rows, make genera row index, and randomly subsample the data (maybe)
substg_clean <- function(data){
  data1 <- as.data.frame(data)
  data2 <-distinct(data1,Updated_Genus,motility,life_habitat,feeding,.keep_all=TRUE) # delete duplicate value
  data3 <- data.frame(data2[,-1], row.names = data2[,1]) # make the first column with genus names the index
  #data4 <- data3[sample(nrow(data3), 84), ] # randomly subsample to get 84 occ
  return(data3)
}

# Use above function to simplify the df to just the age, traits, grid name, and geometry
all_sbstgs<- substg_clean(simple60_data)
```

```
## function for grouping data by substage and removing all columns but Genus names and Grid Cell names as a pres-abs matrix
substg_gen_grid_all <- function(data){
  data1 <- cbind(newColName = rownames(data), data) # make the genus name the 1st column again
  rownames(data1) <- 1:nrow(data1) # make a new index for the rownames
  data2 <- as.data.frame(subset(data1)[,c(1,2)]) # subset out just the genus name and age
  colnames(data2) <- c("Updated_Genus","Age") # Give the columns names
  data3 <- dcast(data2, Age~Updated_Genus, length) # transform into a pres-abs matrix
  data4 <- data.frame(data3[,-1], row.names = data3[,1]) # make the site names the row names
  return(data4)
}

# create df for pres-abs for all data
substg_pres_ab_all <- substg_gen_grid_all(all_sbstgs)



# function to create a simple df of trait values with all variables as factors
# need this later in the code, when calculating sp.to.fe otherwise will lose genera names
df_factor_simple_all <- function(data){
  data1 <- cbind(newColName = rownames(data), data) # make the genus name the 1st column again
  data2 <- as.data.frame(unclass(data1[c(1,3:5)]), stringsAsFactors = TRUE)
  data3 <- data.frame(data2[,-1], row.names = data2[,1])
  return(data3)
}

# create simple factor variable df for all substages
all_data_simp <- df_factor_simple_all(all_sbstgs)

# summarize data using function from mFD pakage:
```

```
all_traits_summ <- mFD::sp.tr.summary(tr_cat = trait_df,
                                      sp_tr = all_data_simp,
                                      stop_if_NA = TRUE)

# check trait types
all_traits_summ$"tr_types"

# check trait type details on levels
all_traits_summ$"mod_list"



  ## This section uses the trait data created for each substage prior to alpha-beta FE analysis
    # (which looks at "collections" of grid cells). It therefore is based on FE summary information
    # for the entire SS as a whole, not divided by location.
  #### Find Overall values for each SS without identifying collection locations ####
# Function to get GR, FR, Simpsons, and Shannons values for each SS overall:
rich_even_summ <- function(data_traits, Age){
  # extract vector of the number of genera in each FE for each SS overall:
  data_fe_nb_sp <- data_traits$fe_nb_sp

  # get the total # of FE in each SS, this is the Functional Richness (FR) for the SS overall
  data_nb_fe <- length(data_fe_nb_sp)

  # Extract vector of unique genera and their FE for each SS:
  data_sp_fe <- data_traits$sp_fe

  # get the total # of genera in each SS, this is the Generic Richness (GR) for the SS overall
  data_nb_g <- length(data_sp_fe)
```

```
 # simpsons measure of evenness calculation:
 simp_calc <- function(data1){
  data1 <- data.frame(data1)
  n <- data1[data1!=0] # get vector of only values greater than 0 (no empty FE)
  if (length(n)>1 && max(n)>1){ # use if statement to only run the calc on nodes with >1 FE and more
than 1 genera in at least one FE
    N <- sum(n)
    S <- length(n)
    D = sum(n*(n-1))/(N*(N-1))
    inversD <- 1/D
    sD <- inversD/S
  } else {
    0
  }
 }
 # get the Simpson's Measure of Evenness Value:
 data_simp <- simp_calc(data_fe_nb_sp)

 # equation for shannons calc
 shannon_equit <- function(data){
  data1 <- data.frame(data)
  data1 <- data1[data1!=0] # get vector of only values greater than 0 (no empty FE)

  if (length(data1)>1){ # use if statement to only run the calc on nodes with >1 FE and more than 1
genera in at least one FE

    S <- length(which(data1!=0)) # get the numb of FE
    Si <- sum(data1) # find out how many genera total present in all FE
```

```
    Pvals <- data1/Si # get proportions of each FE based on dividing # genera in each by total # genera
    nlPvals <- log(Pvals) # take natural log of the proportions
    H <- -sum(nlPvals*Pvals) # calculate Shannon diversity index (H) by multiplying nl of p by p and summing and multiply by neg 1
    LNS <- log(S) # get natural log of number of FE present
    Hi <- H/LNS # calculate Shannon Equitability Index (Hi) by dividing H by nat log of S
   } else {
    0
   }
 }

 data_shan <- shannon_equit(data_fe_nb_sp)

 final_df <- data.frame(firstcolumn = data_nb_g, secondcolumn= data_nb_fe, thirdcolumn= data_simp, fourthcolumn= data_shan)
 rownames(final_df) <- Age
 colnames(final_df) <- c("GR","FR","SME","SEI")

 return(final_df)

}

# Use above function to get the overall values of GR, FR, Simpsons, and Shannons for each SS
maa_up_60_rich_even <- rich_even_summ(maa_up_60_traits,"Maa UP")
maa_low_60_rich_even <- rich_even_summ(maa_low_60_traits,"Maa LOW")
cam_up_60_rich_even <- rich_even_summ(cam_up_60_traits,"Cam UP")
cam_mid_60_rich_even <- rich_even_summ(cam_mid_60_traits,"Cam MID")
cam_low_60_rich_even <- rich_even_summ(cam_low_60_traits,"Cam LOW")
```

```
# bind the df of each SS richness and evenness from above into single DF

all_rich_even <-
rbind(cam_low_60_rich_even,cam_mid_60_rich_even,cam_up_60_rich_even,maa_low_60_rich_even,
maa_up_60_rich_even)


write.csv(all_rich_even, file="substages_outputs/all_richness_evenness_through_time.csv")


 #### Plot of functional and generic richness through 5 substages ####


all_rich_even_age <-   cbind(Age = rownames(all_rich_even), all_rich_even) # make the genus name the
1st column again

rownames(all_rich_even_age) <- 1:nrow(all_rich_even_age) # make a new index for the rownames


 # plot the richness values (GR and FR) for each substage to compare

gr_fr_plot <- ggplot(all_rich_even_age, aes(x=Age)) +

  geom_line(aes(y = GR, color = "#D55E00",group=1)) +

  geom_line(aes(y = 10*FR, color="steelblue",group=2),linetype="dashed") + # multiply FR by 10 so more
visible next to GR

  geom_point(aes(y = GR), color = "#D55E00") +

  geom_point(aes(y = 10*FR), color="steelblue") + # multiply FR by 10 so more visible next to GR

  scale_y_continuous(

    # Features of the first axis

    name = "# Genera",

    # Add a second axis and specify its features

    sec.axis = sec_axis(~ . * 0.1,name="# FE") # make sure scale divides y axis proportional to above
mutiplication

  ) +

  scale_color_discrete(labels=c("Generic Richness", "Functional Richness")) +

  labs(title=paste("Generic and Functional Richness through Time")) +

  theme(text = element_text(size = 16),legend.position="none",plot.title = element_text(hjust = 0.5)) +
```

```
  scale_x_discrete(labels=c("Lower Camp.", "Middle Camp.","Upper Camp.","Lower Maastr.","Upper Maastr."))


# plot the GR vs Simp for each substage to compare
gr_simp_plot <- ggplot(all_rich_even_age, aes(x=Age)) +
  geom_line(aes(y = GR, color = "#D55E00",group=1)) +
  geom_line(aes(y = 500*SME, color="steelblue",group=2),linetype="dashed") + # multiply SME by 500 so more visible next to GR
  geom_point(aes(y = GR), color = "#D55E00") +
  geom_point(aes(y = 500*SME), color="steelblue") + # multiply SME by 500 so more visible next to GR
  scale_y_continuous(
    # Features of the first axis
    name = "# Genera",
    # Add a second axis and specify its features
    sec.axis = sec_axis(~ . * 0.02,name="SME") # make sure scale divides y axis proportional to above mutiplication
  ) +
  scale_color_discrete(labels=c("Generic Richness", "Simpson's Measure of Evenness")) +
  labs(title=paste("Generic Richness and Simpson's Measure of Evenness through Time")) +
  theme(text = element_text(size = 16),legend.position="none",plot.title = element_text(hjust = 0.5)) +
  scale_x_discrete(labels=c("Lower Camp.", "Middle Camp.","Upper Camp.","Lower Maastr.","Upper Maastr."))


# plot the GR vs Shan for each substage to compare
gr_shan_plot <- ggplot(all_rich_even_age, aes(x=Age)) +
  geom_line(aes(y = GR, color = "#D55E00",group=1)) +
  geom_line(aes(y = 500*SEI, color="steelblue",group=2),linetype="dashed") + # multiply SEI by 500 so more visible next to GR
  geom_point(aes(y = GR), color = "#D55E00") +
  geom_point(aes(y = 500*SEI), color="steelblue") + # multiply SEI by 500 so more visible next to GR
```

```
  scale_y_continuous(

    # Features of the first axis

    name = "# Genera",

    # Add a second axis and specify its features

    sec.axis = sec_axis(~ . * 0.02,name="SEI") # make sure scale divides y axis proportional to above
mutiplication

  ) +

  scale_color_discrete(labels=c("Generic Richness", "Shannon Equitability Index")) +

  labs(title=paste("Generic Richness and Shannon Equitability Index through Time")) +

  theme(text = element_text(size = 16),legend.position="none",plot.title = element_text(hjust = 0.5)) +

  scale_x_discrete(labels=c("Lower Camp.", "Middle Camp.","Upper Camp.","Lower Maastr.","Upper
Maastr."))



rich_even_plots <- list(gr_fr_plot,gr_simp_plot,gr_shan_plot)



grid.arrange(grobs = rich_even_plots,

          top = "Change in Richness and Evenness Values Through Time") # Write the grid.arrange in the
file


# Save plots of changes in metrics across substages

pdf("substages_outputs/Substage_metrics_change_lineplots.pdf", width = 13, height = 16) # Open a
new pdf file

grid.arrange(grobs = rich_even_plots,

          top = "Change in Richness and Evenness Values Through Time") # Write the grid.arrange in the
file

dev.off() #close the file
```

```
  #### Make "bubble plot" of functional entity change in # genera (as relative abundance) per substage ####

# Function to get # genera present in each FE for a substage
fe_counts_names <- function(data){
  counts <- as.data.frame(data$"fe_nb_sp")
  names <- as.data.frame(data$details_fe$fe_codes)

  counts1 <- cbind(newColName = rownames(counts), counts) # make grid cell first column
  rownames(counts1) <- 1:nrow(counts1)
  colnames(counts1) <- c("FE","genera")

  names1 <- cbind(newColName = rownames(names), names) # make grid cell first column
  rownames(names1) <- 1:nrow(names1)
  colnames(names1) <- c("FE","names")

  count_names <- merge(counts1, names1, by ="FE", x.all=TRUE) # merge the FE genera counts with names
  count_names <- count_names[,c(3,2)] # rearrange columns

}

# Create df of # genera per FE in each substage using above function
maa_up_60_fe_counts <- fe_counts_names(maa_up_60_traits)
maa_low_60_fe_counts <- fe_counts_names(maa_low_60_traits)
cam_up_60_fe_counts <- fe_counts_names(cam_up_60_traits)
cam_mid_60_fe_counts <- fe_counts_names(cam_mid_60_traits)
cam_low_60_fe_counts <- fe_counts_names(cam_low_60_traits)
```

```
# Function to get relative abundance of genera for each FE
find_rel_abund <- function(data){
  data1 <- data.frame(data[,1])
  for(i in 1:length(data[,2])){
    data1$rel_ab[i] <- data[i,2]/sum(data[,2])
  }

  colnames(data1) <- c("names","rel_ab")
  return(data1)
}

# Get relative abundance of genera in each FE in each substage using above function
maa_up_60_fe_rel_abd <- find_rel_abund(maa_up_60_fe_counts)
maa_low_60_fe_rel_abd <- find_rel_abund(maa_low_60_fe_counts)
cam_up_60_fe_rel_abd <- find_rel_abund(cam_up_60_fe_counts)
cam_mid_60_fe_rel_abd <- find_rel_abund(cam_mid_60_fe_counts)
cam_low_60_fe_rel_abd <- find_rel_abund(cam_low_60_fe_counts)



# Compile above df into list to merge
rel_abd_list <- list(cam_low_60_fe_rel_abd,cam_mid_60_fe_rel_abd,
                     cam_up_60_fe_rel_abd,maa_low_60_fe_rel_abd,
                     maa_up_60_fe_rel_abd)

# merge all the df into one (this makes sure no FE are omitted in any SS)
rel_abd_list_complete <- rel_abd_list %>% reduce(full_join, by='names')
```

```
colnames(rel_abd_list_complete) <- c("names","CamLOW","CamMID","CamUP","MaaLOW","MaaUP")



# Function to create trait names for the plot
make_IDs <- function(data){
  # split the names column by "-"
  names_split <- str_split_fixed(data$names, "_", 4)

  # split the subsequent columns by str values to get just trait IDs
  motility_str <- str_split_fixed(names_split[ ,1], "Y", 2)
  habitat_str <- str_split_fixed(names_split[ ,3], "AT", 2)
  feeding_str <- str_split_fixed(names_split[ ,4], "G", 2)

  # concatenate and capitalize the ID values into a single vector
  trait_names <- paste(toupper(motility_str[ ,2]), toupper(habitat_str[ ,2]), toupper(feeding_str[ ,2]),
sep="-")
  return(trait_names)
}

# Get the trait name IDs using the above function
trait_names <- make_IDs(rel_abd_list_complete)

# Add trait names IDs to the abundance df
rel_abd_list_complete$names <- trait_names

write.csv(rel_abd_list_complete,file="substages_outputs/trait_names_rel_abundance.csv")



# Function to create a bubble plot like Foster and Twitchet 2014 Fig 3
```

```
bubble_gen_per_fe <- function(data){
 # make copy of the compiled list of # genera per FE in each substage to manipulate
 data2 <- data

 # Change all values to either 1-5 based on substage (exclude NAs)
 data2$MaaUP[data2$MaaUP >0 ] <- 5
 data2$MaaLOW[data2$MaaLOW >0 ] <- 4
 data2$CamUP[data2$CamUP >0 ] <- 3
 data2$CamMID[data2$CamMID >0 ] <- 2
 data2$CamLOW[data2$CamLOW >0 ] <- 1

 # merge the df with # genera in each substage with the df with 1-5 code numbers for all non NA cells
 data3 <- merge(data,data2,by="names",x.all=TRUE)

 # create new rows of the diff in relative abundance between two successive intervals
 data3$cl_cm <- data3$CamLOW.x-data3$CamMID.x
 data3$cm_cu <- data3$CamMID.x-data3$CamUP.x
 data3$cu_ml <- data3$CamUP.x-data3$MaaLOW.x
 data3$ml_mu <- data3$MaaLOW.x-data3$MaaUP.x

 # create threshold breaks in the columns of diff in relative abundances
 data3$TH_cl_cm<-cut(data3$cl_cm,breaks=c(-Inf,-0.05,0.05,Inf),labels=c("<=-0.05","no",">0.05"))
 data3$TH_cm_cu<-cut(data3$cm_cu,breaks=c(-Inf,-0.05,0.05,Inf),labels=c("<=-0.05","no",">0.05"))
 data3$TH_cu_ml<-cut(data3$cu_ml,breaks=c(-Inf,-0.05,0.05,Inf),labels=c("<=-0.05","no",">0.05"))
 data3$TH_ml_mu<-cut(data3$ml_mu,breaks=c(-Inf,-0.05,0.05,Inf),labels=c("<=-0.05","no",">0.05"))

 # create bubble plot of the df to show # genera (based on relative size) in each FE across substages
 r <- ggplot(data3, aes(x=names)) +
```

```
  geom_point(aes(y = MaaUP.y, size=MaaUP.x, fill = TH_ml_mu),shape=21) + # add points with select color, size by # genera

  geom_point(aes(y = MaaLOW.y, size=MaaLOW.x, fill = TH_cu_ml),shape=21) +

  geom_point(aes(y = CamUP.y, size=CamUP.x, fill = TH_cm_cu),shape=21) +

  geom_point(aes(y = CamMID.y, size=CamMID.x, fill = TH_cl_cm),shape=21) +

  geom_point(aes(y = CamLOW.y, size=CamLOW.x), fill = "darkgrey",shape=21) + # put this color atrib outside aes because it will set absolutely and not follow any variable

  labs(title="Change in # Genera in Functional Entities through Time") +

  labs(size="Relative Abundances") + # add a legend title, but only for size (will remove color from legend)

  guides(fill="none") +

  xlab("Functional Entity") + ylab("Age") + # add the x and y axis labels

  theme(text = element_text(size = 16),plot.title = element_text(hjust = 0.5), # make the title centered

      axis.text.x = element_text(angle = 90, vjust = 0.5, hjust=1), # change x axis ticks to be vert

      axis.text.y = element_text(angle = 90, vjust = 0.5, hjust=0.5),

      legend.position = "top",

      legend.direction="horizontal") +

  scale_y_discrete(limits=c("Lower Camp.", "Middle Camp.","Upper Camp.","Lower Maastr.","Upper Maastr.")) + # change the name of y axis ticks

  scale_fill_manual(values = c("<=-0.05" = "#D55E00",

                      "no"="#E69F00",

                      ">0.05"="#0072B2")) +

  scale_size_continuous(range = c(1, 12), breaks = c(-Inf,0.001,0.02,0.25,0.5,1))


 return(r)

}

# Use above function to create bubble plot of the # genera per FE in each substage

gen_fe_bubble_plot <- bubble_gen_per_fe(rel_abd_list_complete)


gen_fe_bubble_plot
```

```
# Export the bubble plot
pdf("substages_outputs/FE_Bubble_Plot_raw.pdf", width = 14, height = 9) # Open a new pdf file
gen_fe_bubble_plot
dev.off() #close the file
```

# APPENDIX C. Supplementary Materials for Chapter 4

The following outlines detailed methods and results, including environmental data collection, aggregation, and interpolation, as well as niche analysis methods using the ecospat package. Detailed results tables and figures are included here.

## Appendix C-1. Detailed Methods

### Sedimentary data aggregation

All environmental proxy data was collected using best practices outlined in Myers et al (2015) and as applied in Purcell et al. (2023). Lithologic information, including grain size and type, stratigraphic structures, bedding style and thickness, bioturbation level, and other notable features were recorded for each locality included. Grain size and type were converted to percentages while stratigraphic structures, thickness, and all other variables were converted into categorical code values for each unit described (Table S1). Units were aggregated based on three levels of temporal binning: (1) within ammonite biozone intervals (hereafter referred to as biozone intervals), (2) substages, and (3) stage levels (see Table S2 for explanation of variable calculations when aggregated). Not all sedimentary variables recorded and aggregated in the dataset were included in the final analysis, either because the variable is not considered to be well represented in the dataset (e.g., limestone bedding style) or if the variable was recorded to examine potential bias (e.g., bioturbation confidence). Confidence values indicate that very few sedimentary locations were based on sites with literature descriptions that could be interpreted with high confidence (Table S2), however the ratio of covered to uncovered portions of an aggregated section indicated that almost all sedimentary locations were described almost completely and did not include more than 25% covered units, which would create “gaps in the data aggregation (Table S2). Ammonite zones were selected based on the number of stratigraphic localities described, with a minimum of 15 total locations necessary to represent a stratigraphic unit. Ammonite biozone time bins are based primarily on the biostratigraphic level described in the source material itself.

Table S1. Explanation of code values used in sedimentary variables.

| **Bedding Thickness** | | **Limestone Bedding Style** | |
|---|---|---|---|
| **Thickness** | **Avg Size Value (m)** | **Style** | **Code** |
| Laminated (<1cm) | 0.005 | Planar | 1 |
| Very Thin (1-3cm) | 0.02 | Wavy | 2 |
| Thin (3-10cm) | 0.065 | Irregular/Rubbly | 3 |
| Medium (10-30cm) | 0.2 | Nodular | 4 |
| Thick (30-100cm) | 0.65 | | |
| Very Thick (>100cm) | 1.25 | | |
| **Siliciclastic Sedimentary Structures** | | **Bioturbation** | |
| **Structures** | **Code** | **Bioturbation** | **Code** |
| Planar/indistinct/unstated | 1 | No mentioned explicitly | 0.5 |
| Cross Lam/Trough Crossbedding | 2 | At least 1 burrow type mentioned | 1 |
| Ripples/Wavy/Graded | 3 | Multiple burrows/Burrows common | 2 |
| Hummocky/Nodular/Scored/Chnl structures | 4 | Burrows abundant | 3 |
| | | Bioturbated and above | 4 |
| | | Basin | 6 |

Table S2. Explanation of each variable calculated in the final data table of sedimentary information collected from literature sources.

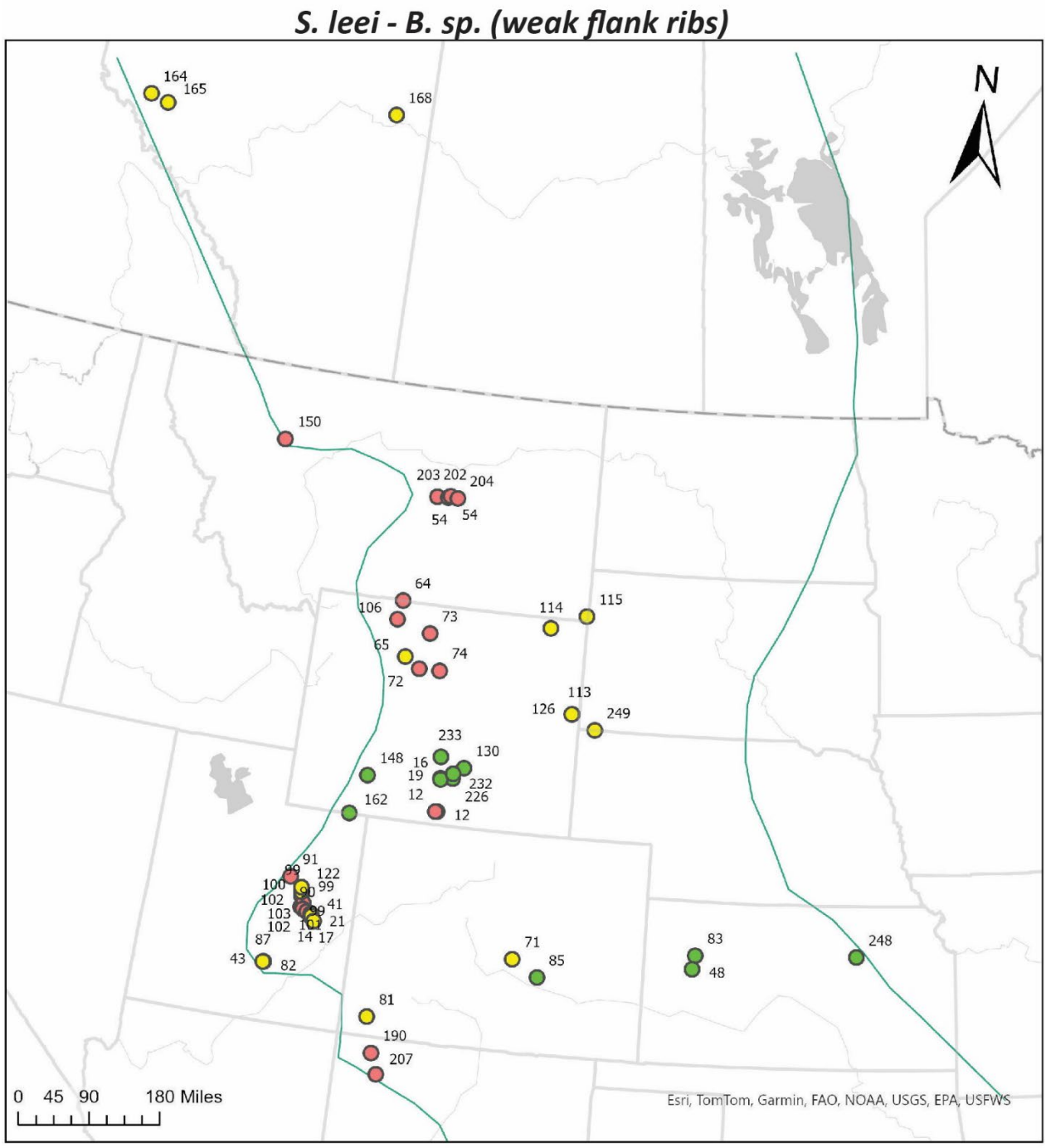
| Variable | Explanation of Table Computations/Rational |
|---|---|
| Sedimentary Grain Percentages | Sum of lengths (m) of grain type/size within all units divided by total aggregation length (m). Interpretation of sediment type amounts based on stratigraphic descriptions and columns can be seen in Tables 1 and 2 in Myers et al. (2015). |
| Average Bed Thickness | Thickness separated by siliciclastic and limestone beds. Sum of all percentages of silic or limestone beds within a unit times the average bed thickness recorded. This value was then multiplied by the unit percent within the aggregated section. Finally, normalized average thickness values were summed for the total aggregated section. If no thickness value recorded in stratigraphic descriptions or displayed in stratigraphic columns from the source materials, assumed average bed thickness of 0.15m. |
| Siliciclastic Bed Style | Siliciclastic bed styles were given a value from 1 to 4 based on assumed energy level during deposition (Table S1). These code values were then multiplied by the unit percentage within the total aggregated section and summed. If siliciclastic beds were present but style was not described, the unit was given a siliciclastic bed style of 1 (planar). |
| Average bioturbation | Bioturbation within a unit was given a value from 1 to 4 based on degree of mixing, number and types of burrows, and burrow distinction (Table S1). These code values were then multiplied by the unit percentage within the total aggregated section and summed. If bioturbation was not mentioned, a value of 0.5 was given, and only if a unit was described as lacking bioturbation explicitly was the unit give a value of 0 (no bioturbation). |
| Limestone Bedding Style | Limestone bed styles were given a value from 1 to 4 based on assumed energy level during deposition (Table S1). These code values were then multiplied by the unit percentage within the total aggregated section and summed. If limestone beds were present but style was not described, the unit was given a limestone bed style of 1 (planar). |
| Ratio of Covered vs. Uncovered Length | To measure some aspect of uncertainty in unit descriptions, the ratio of aggregated unit lengths that were undescribed (usually covered units) to described unit lengths was calculated. |
| Limestone Bedding style Confidence | The relative confidence of limestone bed style measurements was quantified by applying a binary value of either confidence (0) or no confidence (1) to limestone bed style codes. These values were then multiplied by the percentage that unit represents within the aggregated section, and the resulting values were summed. Lower values closer to zero are therefore considered to have greater confidence than higher values. Confidence was only applied to units if a limestone bed style was described in the original source material directly. |
| Siliciclastic Bedding Style Confidence | The relative confidence of siliciclastic bed style measurements was quantified by applying a binary value of either confidence (0) or no confidence (1) to siliciclastic bed style codes. These values were then multiplied by the percentage that unit represents within the aggregated section, and the resulting values were summed. Lower values closer to zero are therefore considered to have greater confidence than higher values. Confidence was only applied to units if a siliciclastic bed style was described in the original source material directly. |
| Bioturbation Confidence | The relative confidence of bioturbation level measurements was quantified by applying a binary value of either confidence (0) or no confidence (1) to bioturbation level codes. These values were then multiplied by the percentage that unit represents within the aggregated section, and the resulting values were summed. Lower values closer to zero are therefore considered to have greater confidence than higher values. Confidence was only applied to units if a bioturbation level was described in the original source material directly. |
| Thickness Confidence | The relative confidence of bed thickness measurements was quantified by applying a binary value of either confidence (0) or no confidence (1) to both siliciclastic and limestone bed thicknesses. These values were then averaged and multiplied by the percentage that unit represents within the aggregated section, and the resulting values were summed. Lower values closer to zero are therefore considered to have greater confidence than higher values. Confidence was only applied to units if bed thickness was described in the original source material directly. |

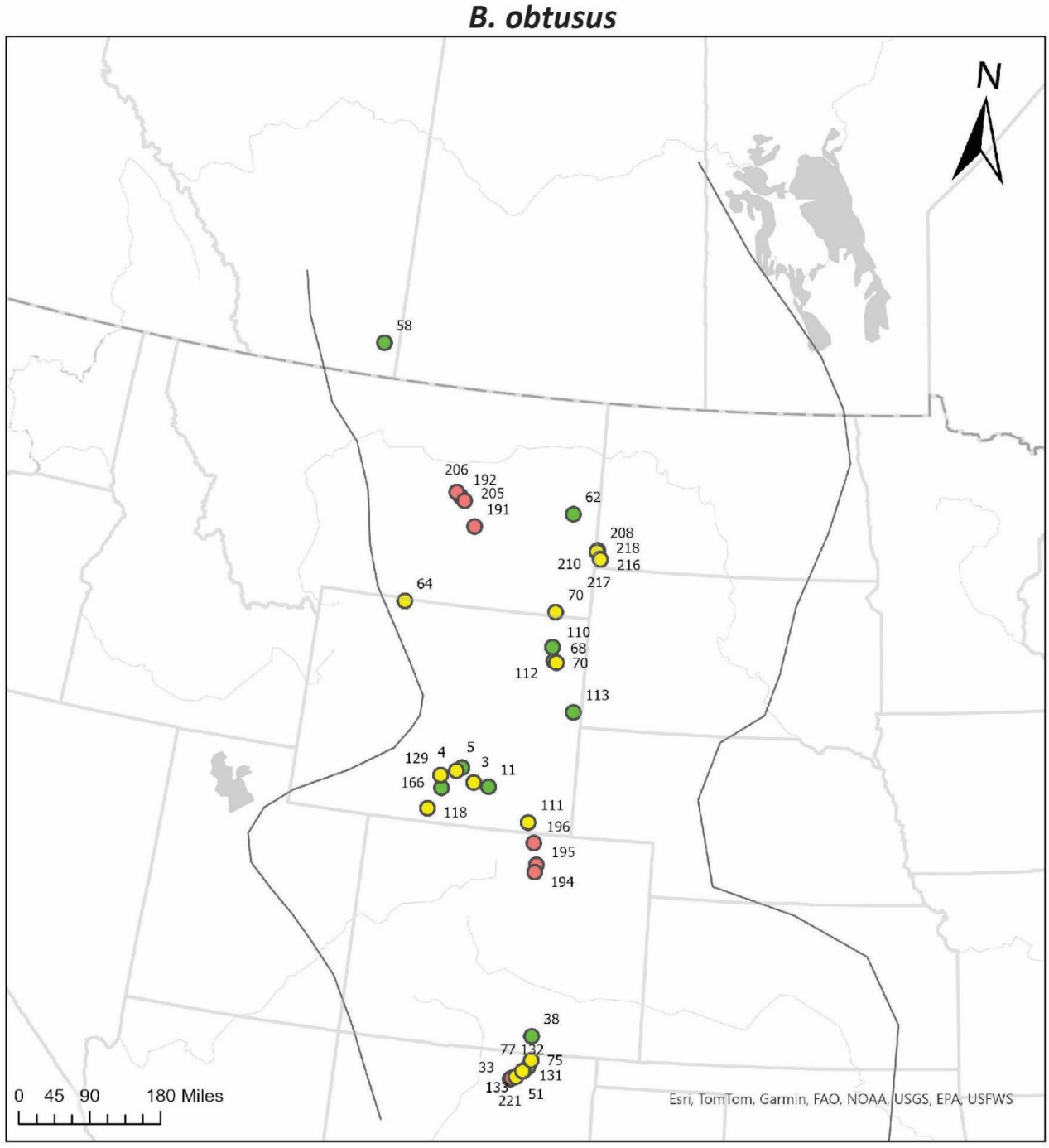
If biostratigraphic information was not included in the source material, a secondary source was used. All stratigraphic information was classified as having either high, medium, or low stratigraphic confidence. High confidence indicates that the biostratigraphic information was well-constrained (i.e., the upper and lower bounds of the zones were observed by the original authors and/or was well supported based on nearby stratigraphy); medium confidence indicates that the biostratigraphic information was partially well-constrained (either part of the section was well-constrained or the biostratigraphic age of the units was well-described in a nearby location by a secondary source); low confidence indicates that the biostratigraphic information was only roughly inferred by a secondary source (Table S3).

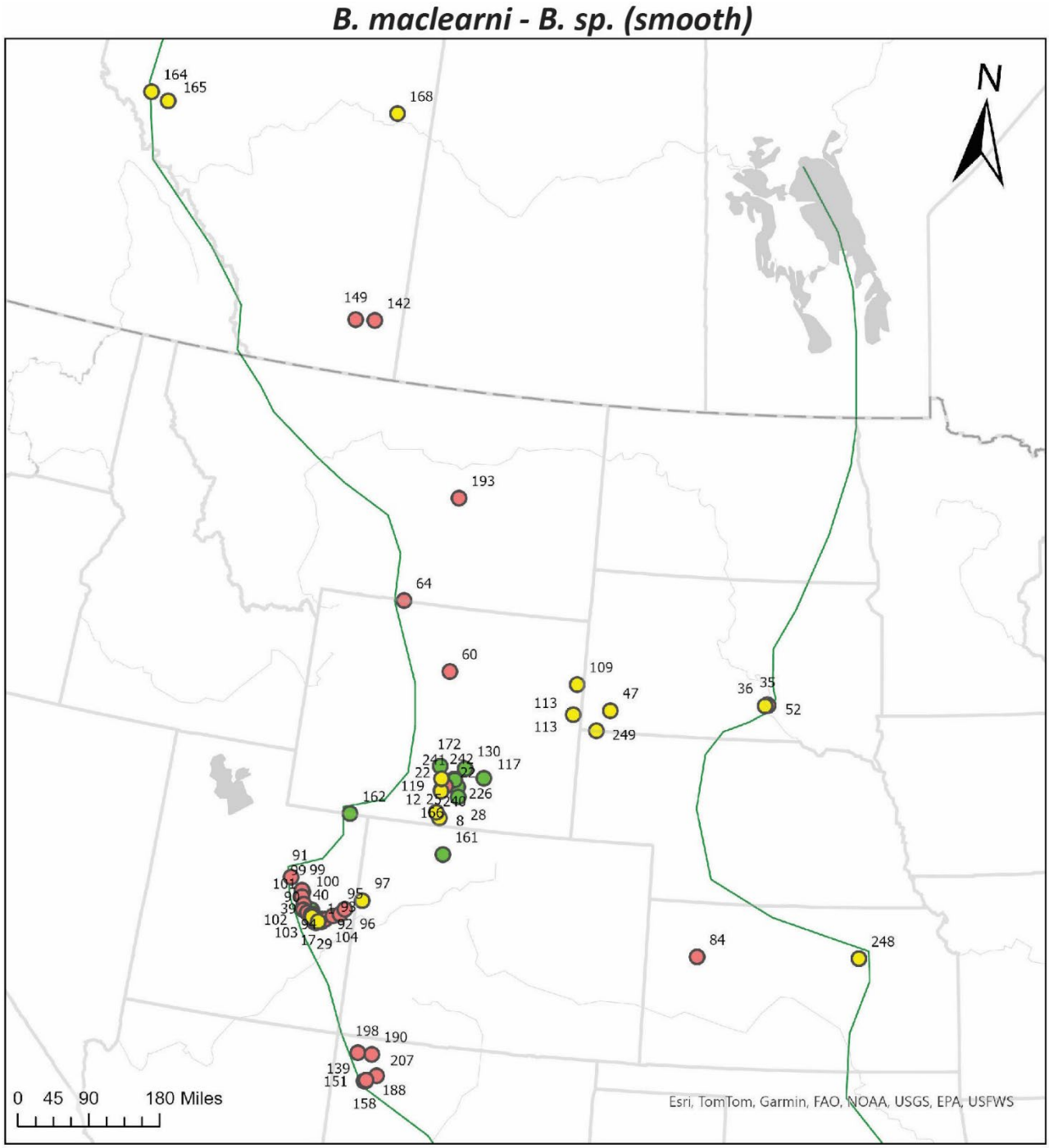
Table S3. Number of stratigraphic localities and their assigned stratigraphic confidence level for each ammonite biozone interval.

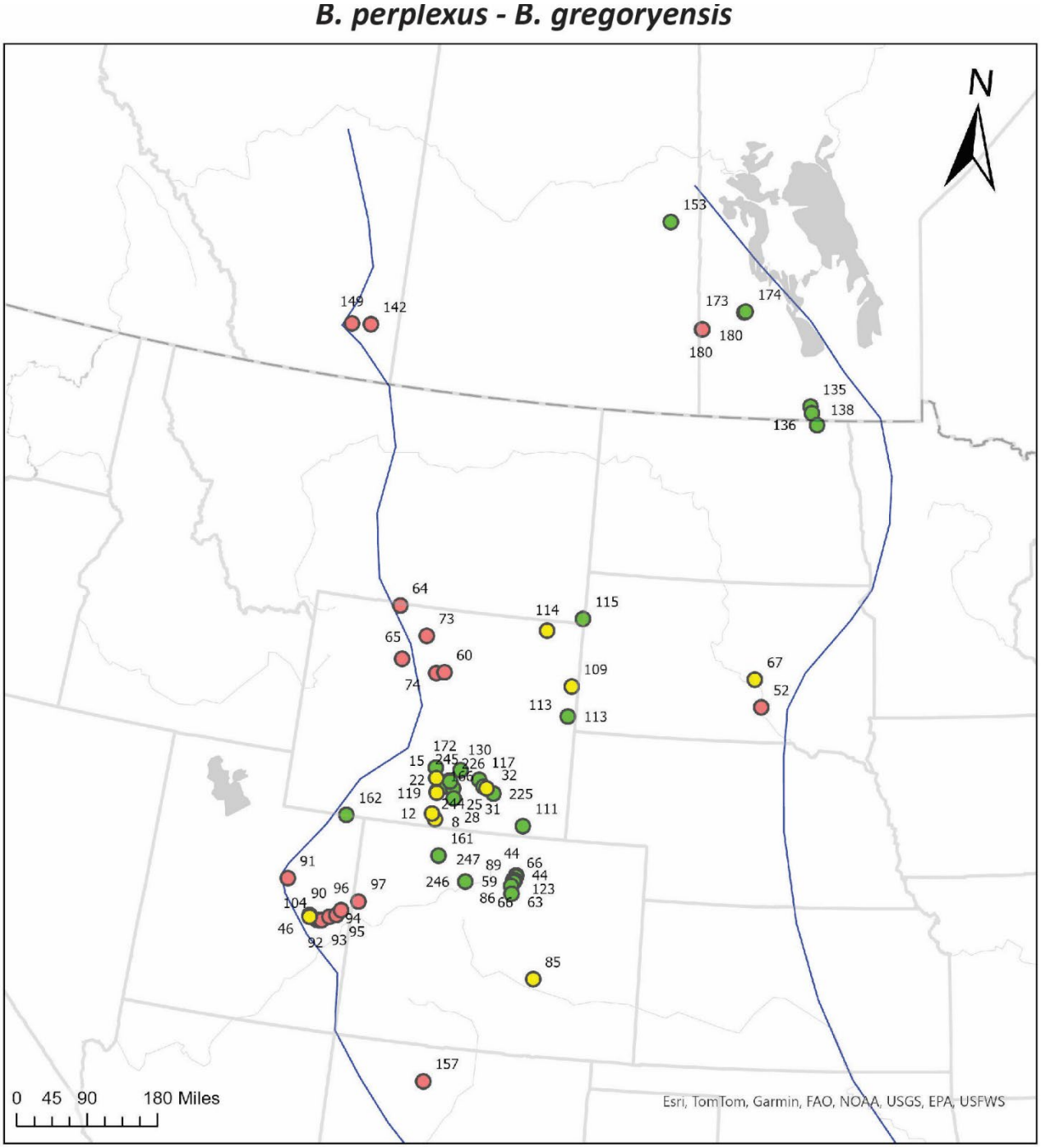
| Interval | High Confidence | Medium Confidence | Low Confidence | Total Localities |
|---|---|---|---|---|
| *Hoploscaphites birkelundae- Hoploscaphites nebrascensis* | 5 | 10 | 5 | **20** |
| *Baculites clinolobatus* | 5 | 19 | 15 | **39** |
| *Baculites baculus- Baculites grandis* | 10 | 18 | 10 | **38** |
| *Baculites reesidei- Baculites eliasi* | 28 | 6 | 3 | **37** |
| *Baculites compressus- Baculites cuneatus* | 13 | 7 | 8 | **28** |
| *Didymoceras cheyennense* | 3 | 9 | 5 | **17** |
| *Didymoceras nebrascense- Exiteloceras jenneyi* | 12 | 21 | 3 | **36** |
| *Baculites reduncus- Baculites scotti* | 13 | 7 | 11 | **31** |
| *Baculites perplexus- Baculites gregoryensis* | 38 | 10 | 19 | **67** |
| *Baculites maclearni- Baculites sp. (smooth)* | 24 | 17 | 30 | **71** |
| *Baculites obtusus* | 25 | 9 | 29 | **63** |
| *Scaphites leei- Baculites sp. (weak flank ribs)* | 19 | 19 | 21 | **59** |

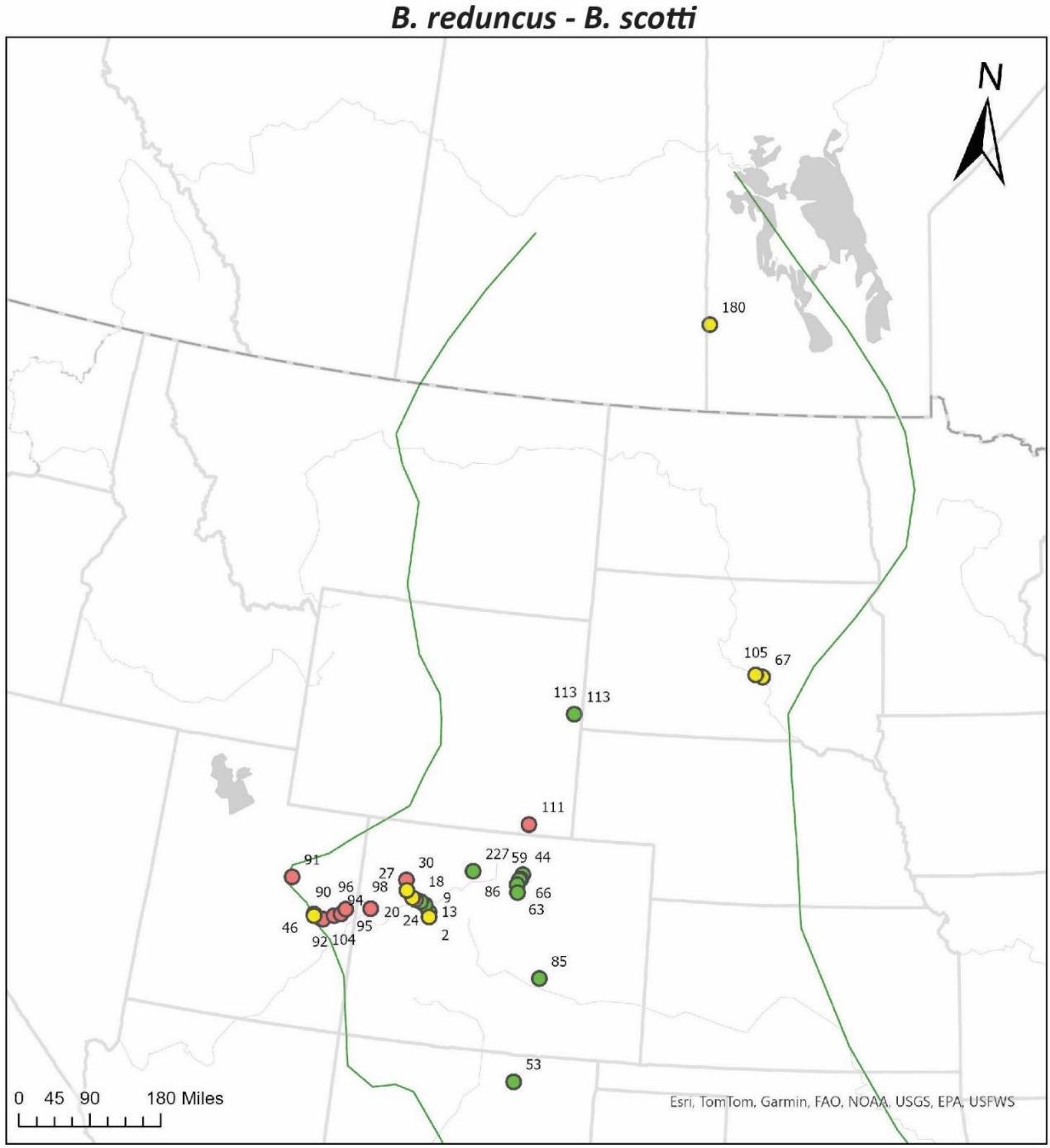
Figure S1. Maps of sedimentary data locations. Points are colored based on biostratigraphic confidence (red = low, yellow = medium, green = high) and labeled with numeric values. These values can be looked up in Table S4 and in the supplementary excel file (Table S27). See Table S3 for full biozone names.

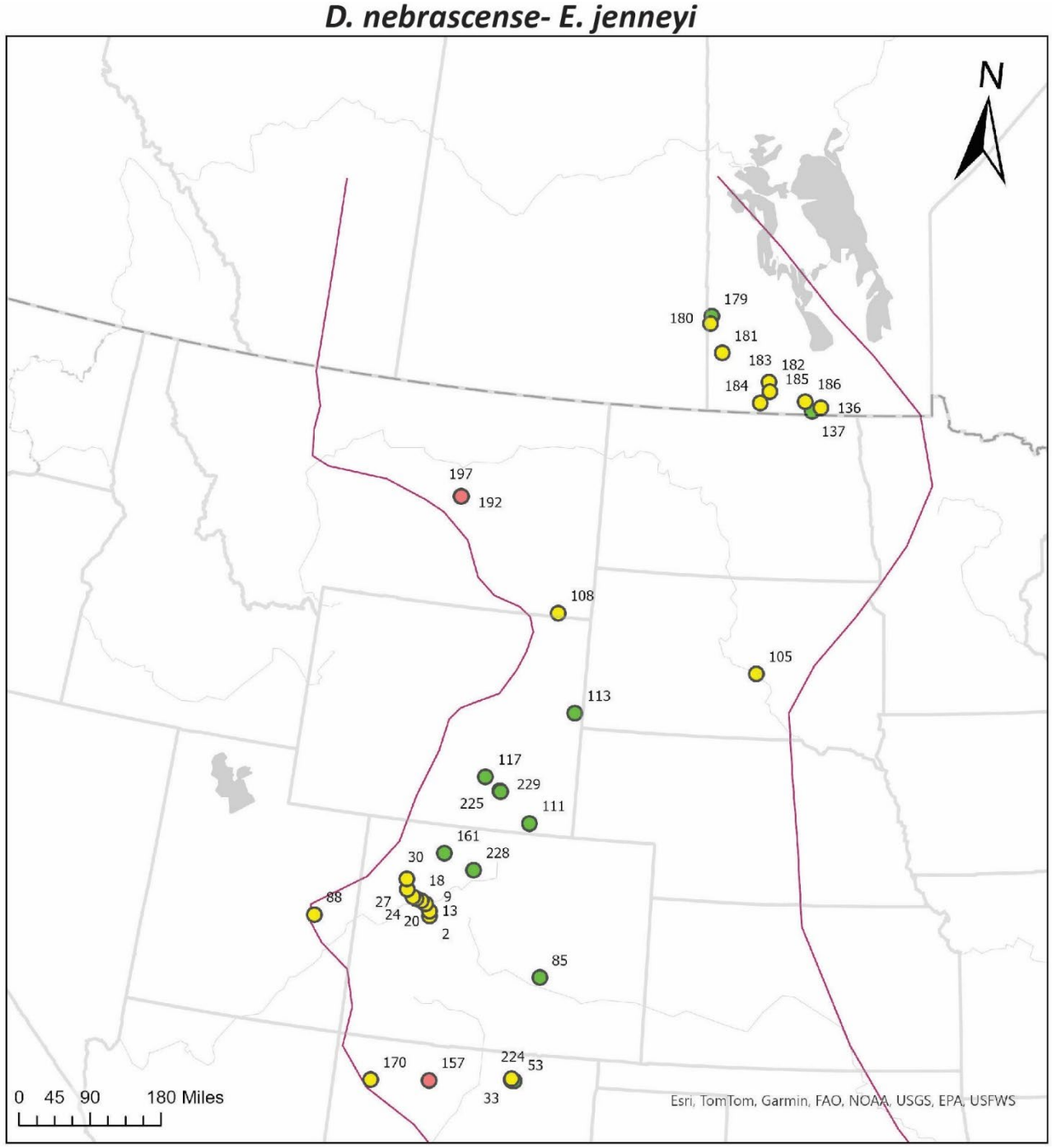


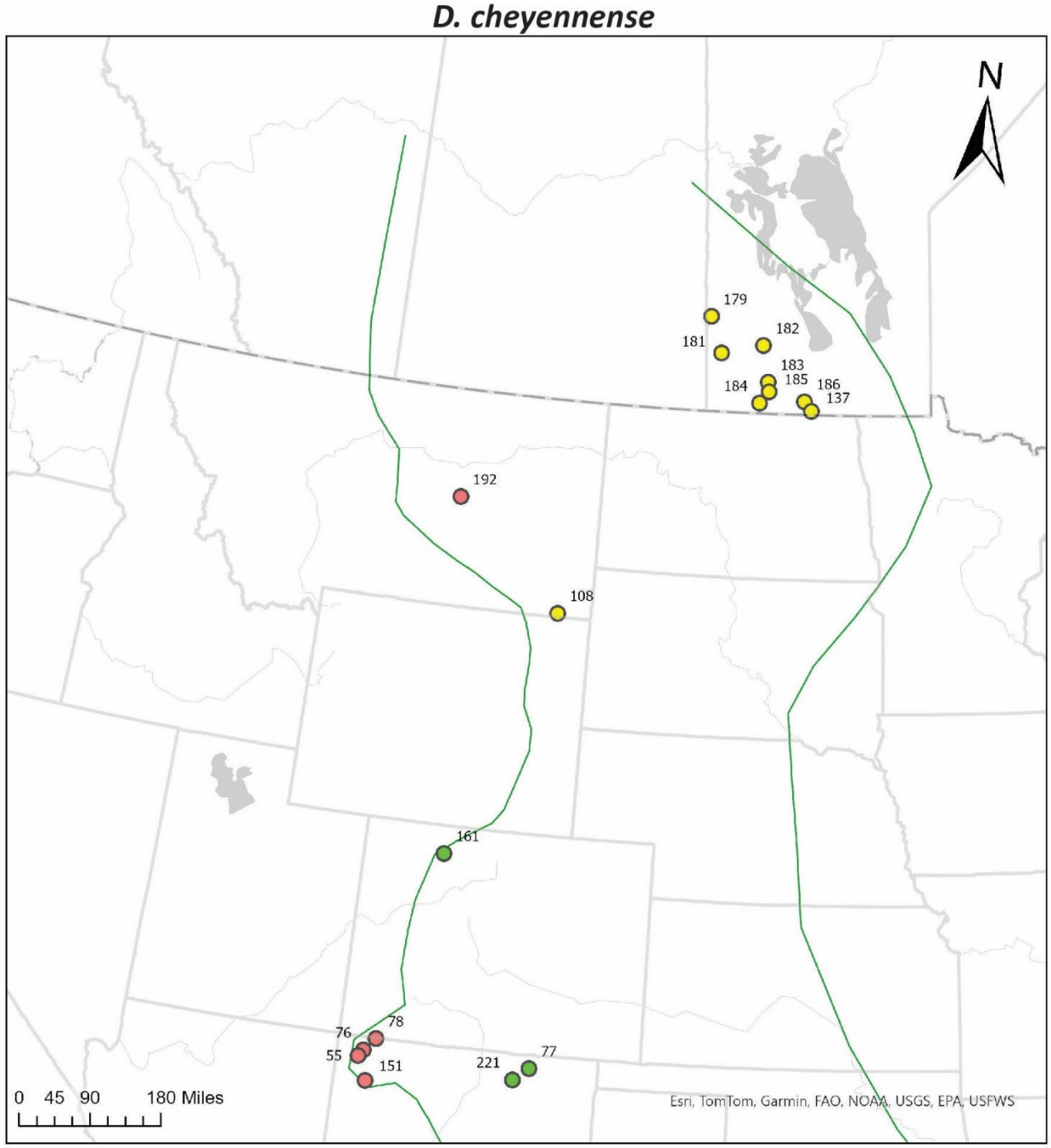

B. obtusus
N
58
206
192
205
191
62
208
218
210
216
217
64
70
110
68
70
112
113
129
4
5
3
11
166
118
111
196
195
194
38
77
132
75
33
131
133
51
221
0 45 90 180 Miles
Esri, TomTom, Garmin, FAO, NOAA, USGS, EPA, USFWS

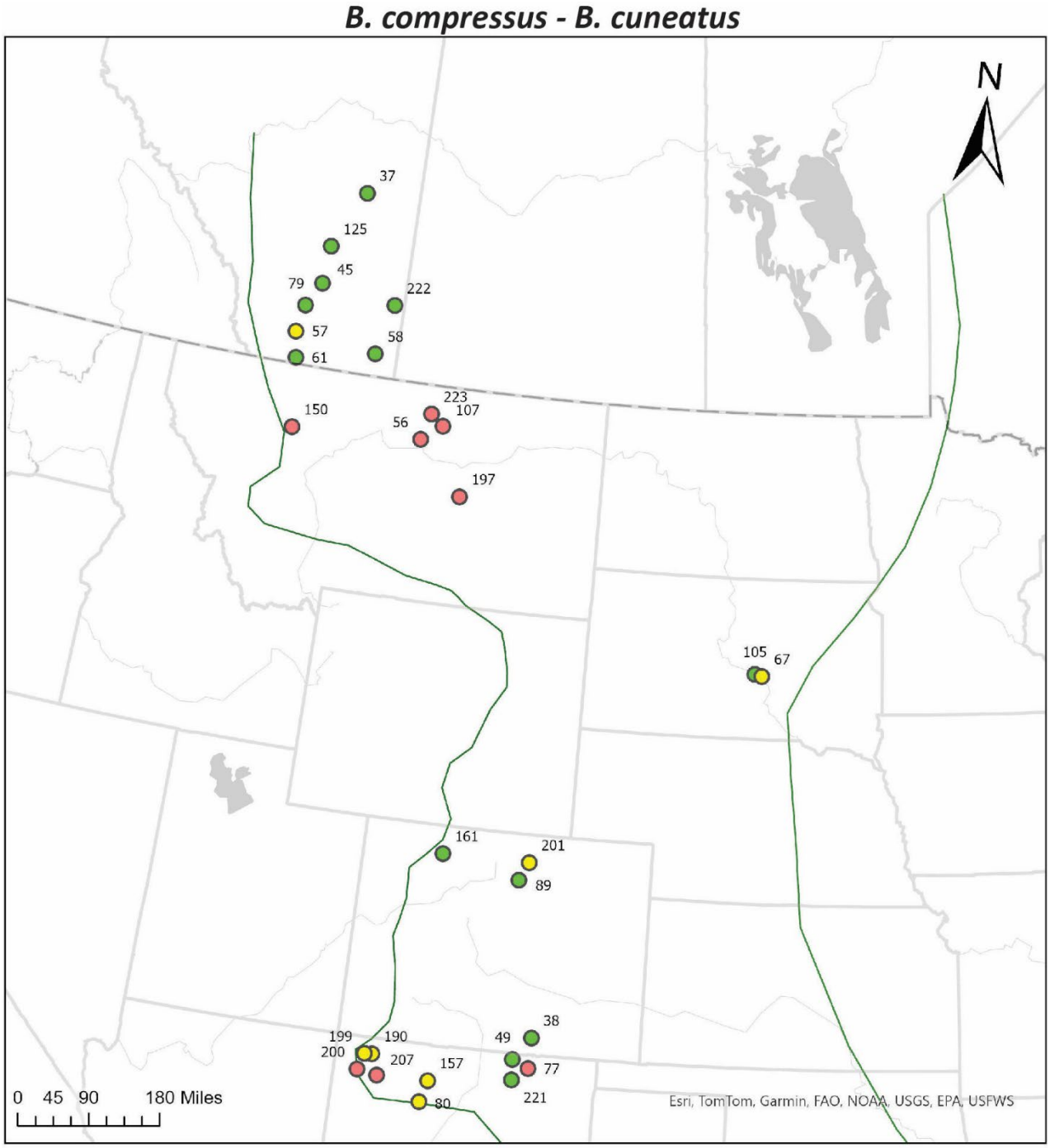

B. maclearni - B. sp. (smooth)
N
164
165
168
149
142
193
64
60
109
47
113
113
249
36
35
52
172
241
242
130
117
22
119
12
25
240
226
166
8
28
162
161
91
99
99
100
97
40
95
102
103
17
29
104
96
84
248
198
190
207
139
151
188
158
0 45 90 180 Miles
Esri, TomTom, Garmin, FAO, NOAA, USGS, EPA, USFWS

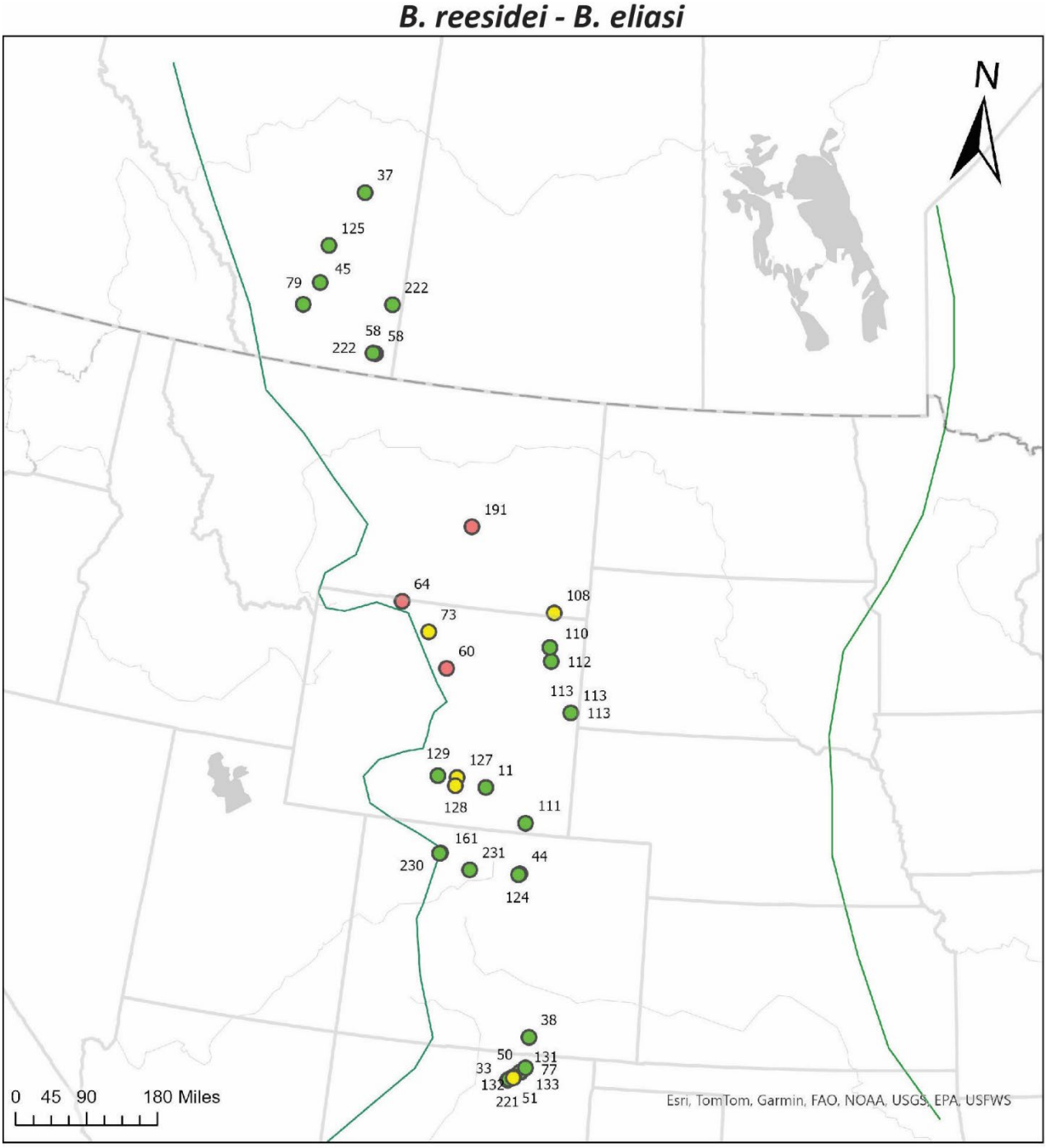

B. perplexus - B. gregoryensis
N
153
149
142
173
174
180
180
135
138
136
64
73
65
60
74
114
115
109
113
113
67
52
172
130
15
245
226
117
22
32
119
225
162
12
8
28
31
111
161
91
247
89
44
66
246
59
44
123
86
68
63
104
90
96
97
94
46
92
93
95
85
157
0 45 90 180 Miles
Esri, TomTom, Garmin, FAO, NOAA, USGS, EPA, USFWS

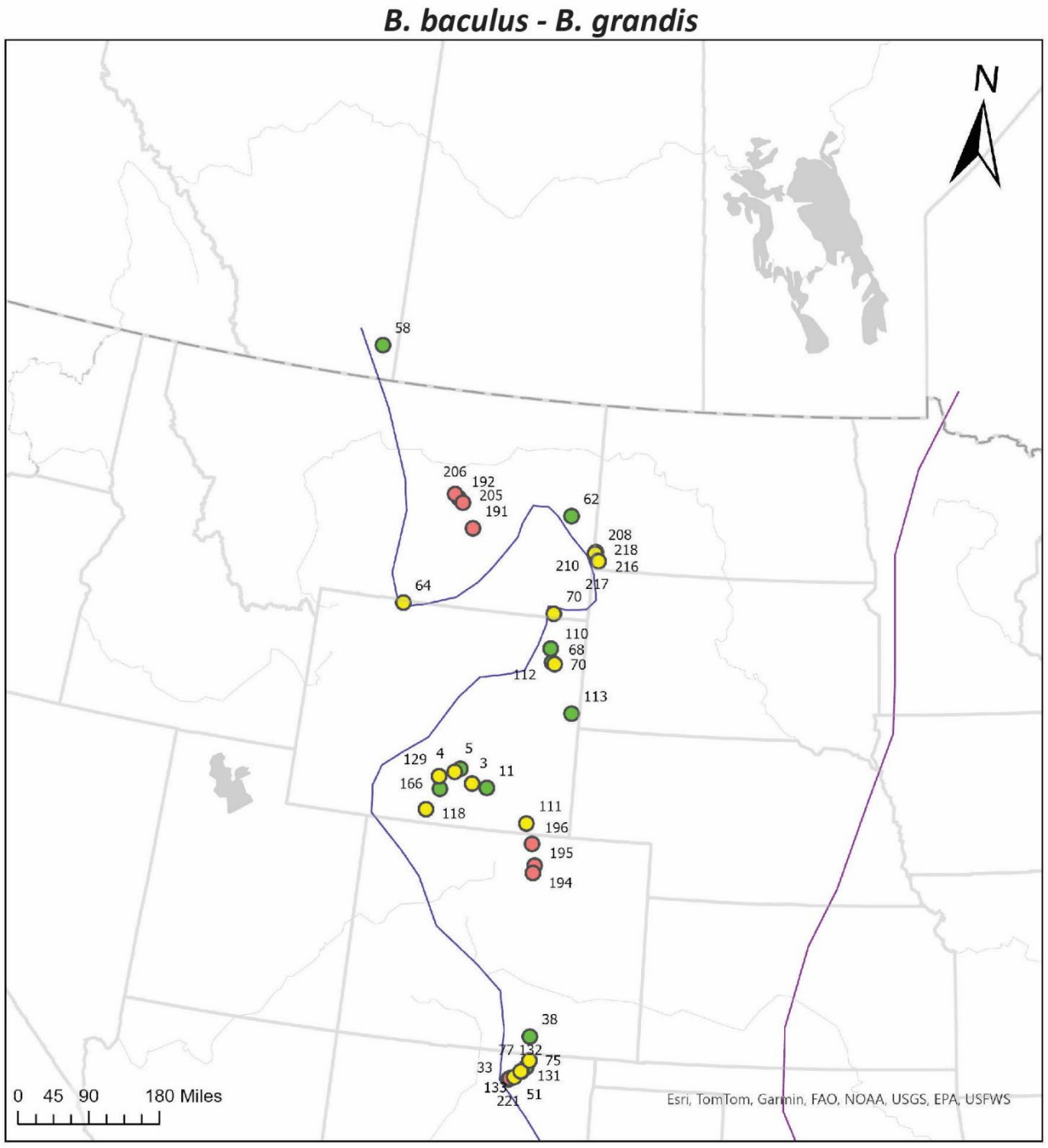

B. reduncus - B. scotti
N
180
105
67
113
113
111
91
227
59
44
30
27
18
98
90
96
94
9
86
66
63
20
24
13
46
95
92
104
2
85
53
0 45 90 180 Miles
Esri, TomTom, Garmin, FAO, NOAA, USGS, EPA, USFWS

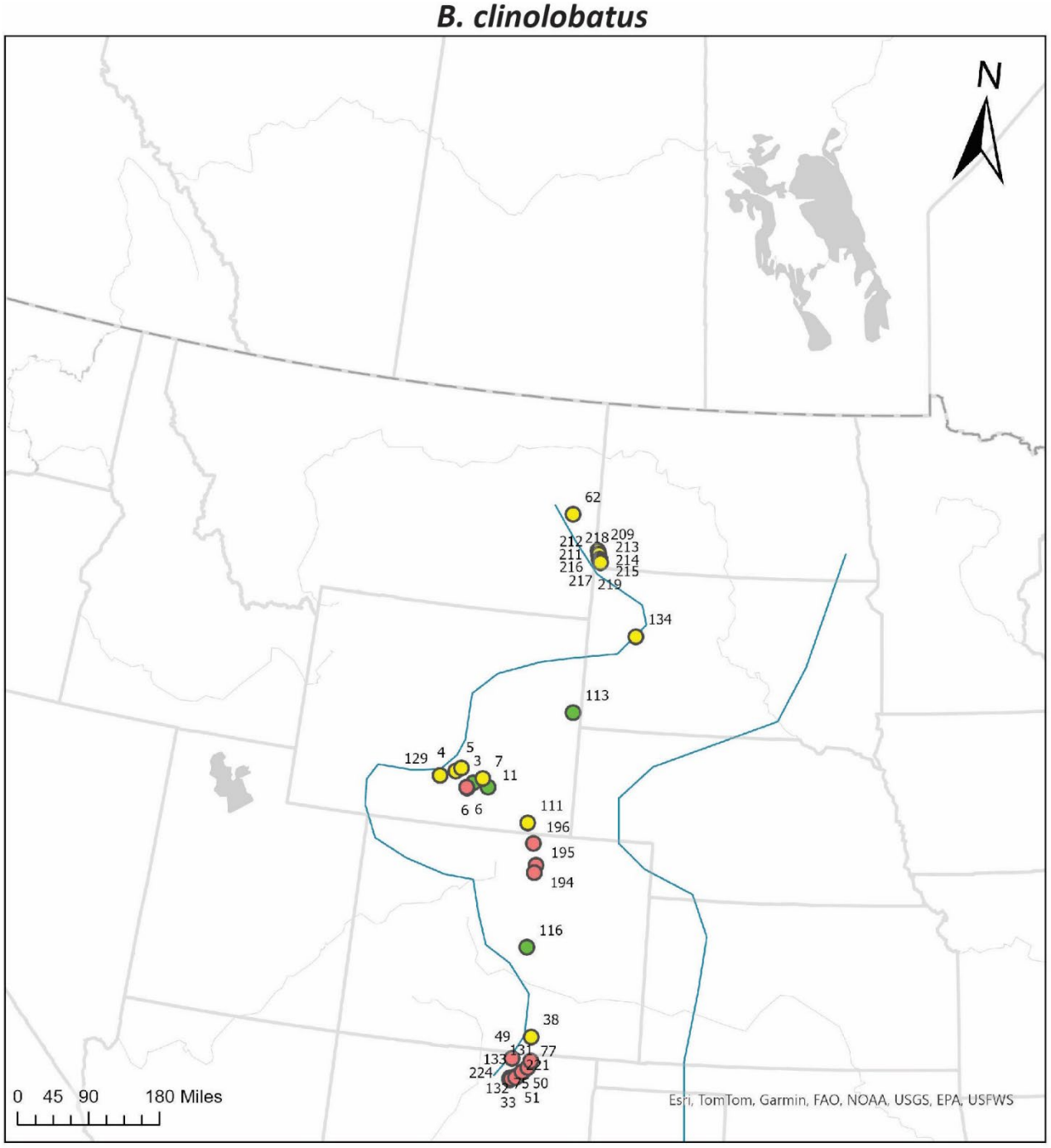

D. nebrascense- E. jenneyi
N
179
180
181
183
182
184
185
186
136
137
197
192
108
105
113
117
229
225
111
161
228
30
18
88
27
9
24
20
13
2
85
170
157
224
53
33
0 45 90 180 Miles
Esri, TomTom, Garmin, FAO, NOAA, USGS, EPA, USFWS


D. cheyennense
N
179
182
181
183
184
185
186
137
192
108
161
76
78
55
151
221
77
0 45 90 180 Miles
Esri, TomTom, Garmin, FAO, NOAA, USGS, EPA, USFWS


B. compressus - B. cuneatus
N
37
125
45
79
222
57
58
61
150
223
107
56
197
105
67
161
201
89
38
199
190
200
207
157
49
77
80
221
0 45 90 180 Miles
Esri, TomTom, Garmin, FAO, NOAA, USGS, EPA, USFWS


B. reesidei - B. eliasi
N
37
125
45
79
222
58
58
222
191
64
108
73
110
60
112
113
113
113
129
127
11
128
111
161
231
44
230
124
38
50
131
33
77
132
133
221
51
0 45 90 180 Miles
Esri, TomTom, Garmin, FAO, NOAA, USGS, EPA, USFWS


B. baculus - B. grandis
N
58
206
192
205
191
62
208
218
216
210
217
64
70
110
68
70
112
113
129
4
5
3
11
166
118
111
196
195
194
38
77
132
75
33
131
133
221
51
0 45 90 180 Miles
Esri, TomTom, Garmin, FAO, NOAA, USGS, EPA, USFWS


B. clinolobatus
N
62
212 218 209
211 213
216 214
217 219 215
134
113
129 4 5
3 7
11
6 6
111
196
195
194
116
38
49
133 131 77
224 221
132 75 50
33 51
0 45 90 180 Miles
Esri, TomTom, Garmin, FAO, NOAA, USGS, EPA, USFWS

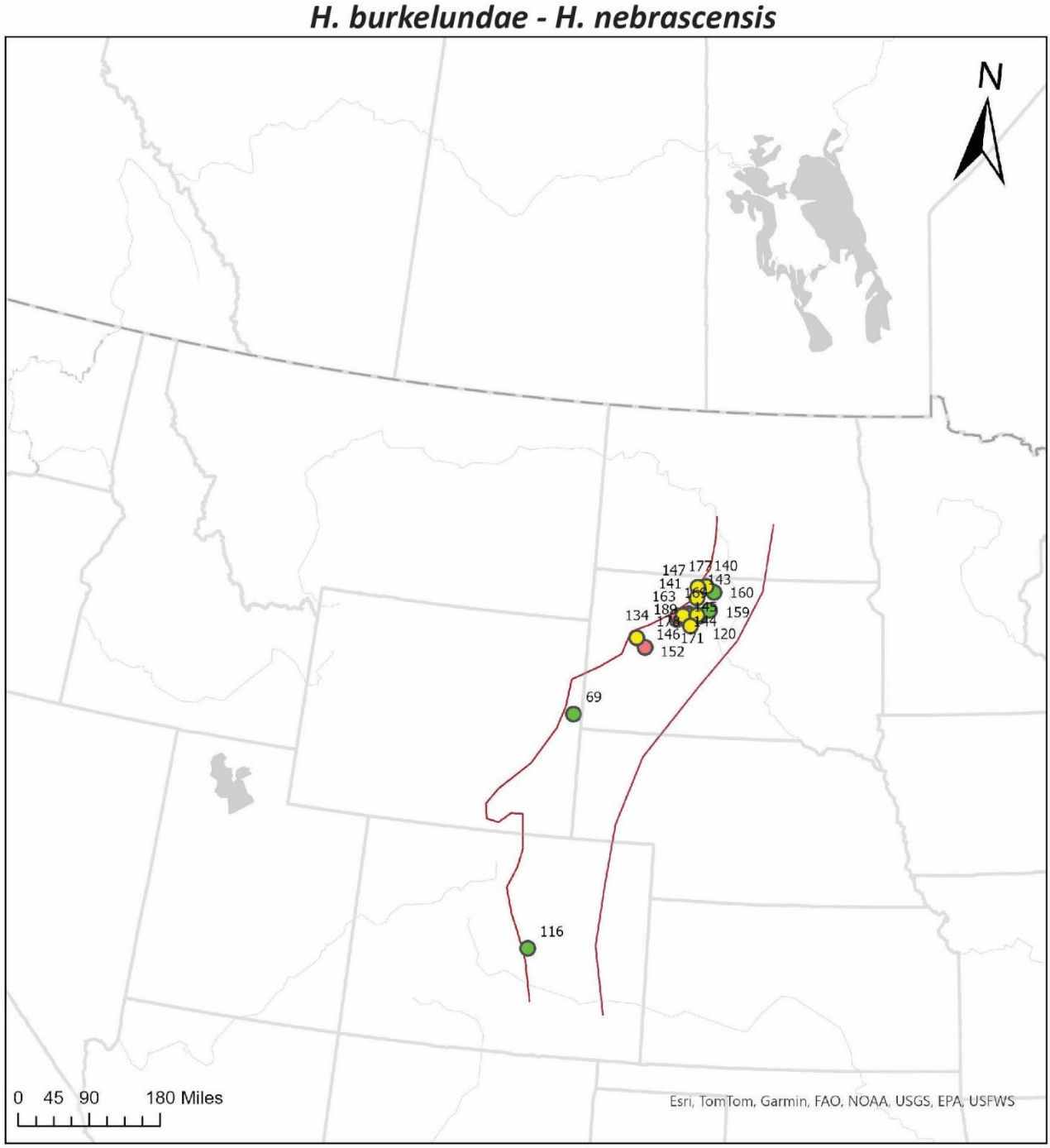

H. burkelundae - H. nebrascensis
N
147
177
140
141
143
163
160
134
159
146
171
120
152
69
116
0 45 90 180 Miles
Esri, TomTom, Garmin, FAO, NOAA, USGS, EPA, USFWS

Table S4. Sedimentary data references table, including numeric ID values for map reference. See Table S3 for full interval names.

| Interval | Interval Confidence | State | Reference | ID Number |
|---|---|---|---|---|
| B. baculus - B. grandis | High | WY | Gill and Cobban 1966 | 113 |
| B. baculus - B. grandis | High | WY | Robinson et al. 1964 | 112 |
| B. baculus - B. grandis | High | WY | Robinson et al. 1964 | 110 |
| B. baculus - B. grandis | Medium | MT | Robinson et al. 1964 | 70 |
| B. baculus - B. grandis | Medium | WY | Robinson et al. 1964 | 68 |
| B. baculus - B. grandis | Medium | WY | Robinson et al. 1964 | 70 |
| B. baculus - B. grandis | Medium | WY | Gill et al. 1970 | 118 |
| B. baculus - B. grandis | Medium | WY | Gill et al. 1970 | 129 |
| B. baculus - B. grandis | Medium | WY | Kiteley 1979 | 111 |
| B. baculus - B. grandis | High | WY | Roehler 1990 | 166 |
| B. baculus - B. grandis | Medium | WY | Gill 1974 | 3 |
| B. baculus - B. grandis | Medium | WY | Gill 1974 | 4 |
| B. baculus - B. grandis | High | WY | Gill 1974 | 5 |
| B. baculus - B. grandis | High | WY | Gill 1974 | 11 |
| B. baculus - B. grandis | Medium | WY | Johnson et al 2005 | 64 |
| B. baculus - B. grandis | Medium | MT | Daly 1984 | 208 |
| B. baculus - B. grandis | Medium | MT | Daly 1984 | 210 |
| B. baculus - B. grandis | Medium | ND | Daly 1984 | 216 |
| B. baculus - B. grandis | Medium | ND | Daly 1984 | 217 |
| B. baculus - B. grandis | Medium | ND | Daly 1984 | 218 |
| B. baculus - B. grandis | Low | NM | Sealey and Lucas 2022 | 33 |
| B. baculus - B. grandis | Low | NM | Sealey and Lucas 2022 | 221 |
| B. baculus - B. grandis | Medium | NM | Sealey and Lucas 2022 | 51 |
| B. baculus - B. grandis | Low | NM | Sealey and Lucas 2022 | 131 |
| B. baculus - B. grandis | Medium | NM | Sealey and Lucas 2022 | 132 |
| B. baculus - B. grandis | Medium | NM | Sealey and Lucas 2022 | 133 |
| B. baculus - B. grandis | High | NM | Sealey and Lucas 2022 | 77 |
| B. baculus - B. grandis | Medium | NM | Sealey and Lucas 2022 | 75 |
| B. baculus - B. grandis | High | NM | Sealey and Lucas 2022 | 38 |
| B. baculus - B. grandis | High | MT | Bishop 1973 | 62 |
| B. baculus - B. grandis | Low | CO | Mather 1928 | 195 |
| B. baculus - B. grandis | Low | CO | Mather 1928 | 194 |
| B. baculus - B. grandis | Low | CO | Mather 1928 | 196 |
| B. baculus - B. grandis | High | ALB | Tsujita 1955 | 58 |
| B. baculus - B. grandis | Low | MT | Johnson and Smith 1964 | 192 |
| B. baculus - B. grandis | Low | MT | Johnson and Smith 1964 | 206 |
| B. baculus - B. grandis | Low | MT | Johnson and Smith 1964 | 205 |
| B. baculus - B. grandis | Low | MT | Heald 1926 | 191 |
| B. clinobatus | High | WY | Gill and Cobban 1966 | 113 |
| B. clinobatus | Medium | WY | Gill et al. 1970 | 129 |

| Interval | Interval Confidence | State | Reference | ID Number |
|---|---|---|---|---|
| B. clinobatus | Medium | WY | Kiteley 1979 | 111 |
| B. clinobatus | High | WY | Gill 1974 | 3 |
| B. clinobatus | Medium | WY | Gill 1974 | 4 |
| B. clinobatus | Medium | WY | Gill 1974 | 5 |
| B. clinobatus | Low | WY | Gill 1974 | 6 |
| B. clinobatus | High | WY | Gill 1974 | 6 |
| B. clinobatus | Medium | WY | Gill 1974 | 7 |
| B. clinobatus | High | WY | Gill 1974 | 11 |
| B. clinobatus | Medium | MT | Daly 1984 | 209 |
| B. clinobatus | Medium | MT | Daly 1984 | 211 |
| B. clinobatus | Medium | MT | Daly 1984 | 212 |
| B. clinobatus | Medium | MT | Daly 1984 | 213 |
| B. clinobatus | Medium | ND | Daly 1984 | 214 |
| B. clinobatus | Medium | ND | Daly 1984 | 215 |
| B. clinobatus | Medium | ND | Daly 1984 | 216 |
| B. clinobatus | Medium | ND | Daly 1984 | 217 |
| B. clinobatus | Medium | ND | Daly 1984 | 218 |
| B. clinobatus | Medium | ND | Daly 1984 | 219 |
| B. clinobatus | Medium | ND | Daly 1984 | 220 |
| B. clinobatus | Low | NM | Sealey and Lucas 2022 | 33 |
| B. clinobatus | Low | NM | Sealey and Lucas 2022 | 224 |
| B. clinobatus | Low | NM | Sealey and Lucas 2022 | 221 |
| B. clinobatus | Low | NM | Sealey and Lucas 2022 | 50 |
| B. clinobatus | Low | NM | Sealey and Lucas 2022 | 51 |
| B. clinobatus | Low | NM | Sealey and Lucas 2022 | 131 |
| B. clinobatus | Low | NM | Sealey and Lucas 2022 | 132 |
| B. clinobatus | Low | NM | Sealey and Lucas 2022 | 133 |
| B. clinobatus | Low | NM | Sealey and Lucas 2022 | 77 |
| B. clinobatus | Low | NM | Sealey and Lucas 2022 | 75 |
| B. clinobatus | Low | NM | Sealey and Lucas 2022 | 49 |
| B. clinobatus | Medium | NM | Sealey and Lucas 2022 | 38 |
| B. clinobatus | High | CO | Nwangwu 1977 | 116 |
| B. clinobatus | Medium | SD | Pettyjohn 1967 | 134 |
| B. clinobatus | Medium | MT | Bishop 1973 | 62 |
| B. clinobatus | Low | CO | Mather 1928 | 195 |
| B. clinobatus | Low | CO | Mather 1928 | 194 |
| B. clinobatus | Low | CO | Mather 1928 | 196 |
| B. compressus - B. cuneatus | High | CO | Roehler 1990 | 161 |
| B. compressus - B. cuneatus | High | CO | Porter 1940 | 89 |
| B. compressus - B. cuneatus | High | NM | Sealey and Lucas 2022 | 221 |
| B. compressus - B. cuneatus | Low | NM | Sealey and Lucas 2022 | 77 |
| B. compressus - B. cuneatus | High | NM | Sealey and Lucas 2022 | 49 |
| B. compressus - B. cuneatus | High | NM | Sealey and Lucas 2022 | 38 |

| Interval | Interval Confidence | State | Reference | ID Number |
|---|---|---|---|---|
| B. compressus - B. cuneatus | Medium | NM | Baltz 1967 | 80 |
| B. compressus - B. cuneatus | Medium | NM | Reeside, 1924 | 190 |
| B. compressus - B. cuneatus | Medium | NM | Reeside, 1924 | 199 |
| B. compressus - B. cuneatus | Low | NM | Reeside, 1924 | 200 |
| B. compressus - B. cuneatus | Low | NM | Reeside, 1924 | 207 |
| B. compressus - B. cuneatus | Medium | SD | Hanczaryk and Gallagher 2007 | 67 |
| B. compressus - B. cuneatus | High | SD | Hanczaryk and Gallagher 2007 | 105 |
| B. compressus - B. cuneatus | Medium | CO | Mather 1928 | 201 |
| B. compressus - B. cuneatus | Medium | NM | Fitter 1958 | 157 |
| B. compressus - B. cuneatus | High | ALB | Tsujita 1955 | 58 |
| B. compressus - B. cuneatus | High | ALB | Tsujita 1955 | 222 |
| B. compressus - B. cuneatus | Medium | ALB | Tsujita 1955 | 57 |
| B. compressus - B. cuneatus | High | ALB | Tsujita 1955 | 61 |
| B. compressus - B. cuneatus | High | ALB | Tsujita 1955 | 79 |
| B. compressus - B. cuneatus | High | ALB | Tsujita 1955 | 45 |
| B. compressus - B. cuneatus | High | ALB | Tsujita 1955 | 37 |
| B. compressus - B. cuneatus | High | ALB | Tsujita 1955 | 125 |
| B. compressus - B. cuneatus | Low | MT | Johnson and Smith 1964 | 197 |
| B. compressus - B. cuneatus | Low | MT | Fuentes et al 2011 | 150 |
| B. compressus - B. cuneatus | Low | MT | Braun 1983 | 223 |
| B. compressus - B. cuneatus | Low | MT | Braun 1983 | 107 |
| B. compressus - B. cuneatus | Low | MT | Braun 1983 | 56 |
| B. maclearni - B. sp. (smooth) | Medium | WY | Gill and Cobban 1966 | 113 |
| B. maclearni - B. sp. (smooth) | High | WY | Gill and Cobban 1966 | 113 |
| B. maclearni - B. sp. (smooth) | Medium | WY | Robinson et al. 1964 | 109 |
| B. maclearni - B. sp. (smooth) | High | WY | Gill et al. 1970 | 226 |
| B. maclearni - B. sp. (smooth) | High | WY | Gill et al. 1970 | 130 |
| B. maclearni - B. sp. (smooth) | Medium | WY | Gill et al. 1970 | 119 |
| B. maclearni - B. sp. (smooth) | High | WY | Gill et al. 1970 | 117 |
| B. maclearni - B. sp. (smooth) | High | WY | Roehler 1990 | 162 |
| B. maclearni - B. sp. (smooth) | High | WY | Roehler 1990 | 172 |
| B. maclearni - B. sp. (smooth) | High | WY | Roehler 1990 | 166 |
| B. maclearni - B. sp. (smooth) | High | CO | Roehler 1990 | 161 |
| B. maclearni - B. sp. (smooth) | Medium | WY | Gill 1974 | 8 |
| B. maclearni - B. sp. (smooth) | Medium | WY | Gill 1974 | 12 |
| B. maclearni - B. sp. (smooth) | Medium | WY | Gill 1974 | 19 |
| B. maclearni - B. sp. (smooth) | Low | WY | Gill 1974 | 22 |
| B. maclearni - B. sp. (smooth) | High | WY | Gill 1974 | 22 |
| B. maclearni - B. sp. (smooth) | High | WY | Gill 1974 | 25 |
| B. maclearni - B. sp. (smooth) | High | WY | Gill 1974 | 28 |
| B. maclearni - B. sp. (smooth) | Low | WY | Johnson et al 2005 | 60 |
| B. maclearni - B. sp. (smooth) | Low | WY | Johnson et al 2005 | 64 |
| B. maclearni - B. sp. (smooth) | High | WY | Minor et al 2022 | 240 |

| Interval | Interval Confidence | State | Reference | ID Number |
|---|---|---|---|---|
| B. maclearni - B. sp. (smooth) | High | WY | Minor et al 2022 | 242 |
| B. maclearni - B. sp. (smooth) | High | WY | Minor et al 2022 | 241 |
| B. maclearni - B. sp. (smooth) | High | WY | Minor et al 2022 | 243 |
| B. maclearni - B. sp. (smooth) | Medium | UT | Seymour 2012 | 40 |
| B. maclearni - B. sp. (smooth) | High | UT | Seymour 2012 | 39 |
| B. maclearni - B. sp. (smooth) | Low | UT | Fisher et al 1960 | 99 |
| B. maclearni - B. sp. (smooth) | Low | UT | Fisher et al 1960 | 99 |
| B. maclearni - B. sp. (smooth) | Low | UT | Fisher et al 1960 | 100 |
| B. maclearni - B. sp. (smooth) | Low | UT | Fisher et al 1960 | 101 |
| B. maclearni - B. sp. (smooth) | Low | UT | Fisher et al 1960 | 102 |
| B. maclearni - B. sp. (smooth) | Low | UT | Fisher et al 1960 | 103 |
| B. maclearni - B. sp. (smooth) | Low | UT | Fisher et al 1960 | 90 |
| B. maclearni - B. sp. (smooth) | Low | UT | Fisher et al 1960 | 90 |
| B. maclearni - B. sp. (smooth) | Low | UT | Fisher et al 1960 | 91 |
| B. maclearni - B. sp. (smooth) | Low | UT | Fisher et al 1960 | 92 |
| B. maclearni - B. sp. (smooth) | Low | UT | Fisher et al 1960 | 104 |
| B. maclearni - B. sp. (smooth) | Low | UT | Fisher et al 1960 | 93 |
| B. maclearni - B. sp. (smooth) | Low | UT | Fisher et al 1960 | 94 |
| B. maclearni - B. sp. (smooth) | Low | UT | Fisher et al 1960 | 95 |
| B. maclearni - B. sp. (smooth) | Low | UT | Fisher et al 1960 | 96 |
| B. maclearni - B. sp. (smooth) | Medium | UT | Fisher et al 1960 | 97 |
| B. maclearni - B. sp. (smooth) | High | UT | Chan and Newman 1991 | 1 |
| B. maclearni - B. sp. (smooth) | High | UT | Chan and Newman 1991 | 10 |
| B. maclearni - B. sp. (smooth) | High | UT | Chan and Newman 1991 | 14 |
| B. maclearni - B. sp. (smooth) | High | UT | Chan and Newman 1991 | 17 |
| B. maclearni - B. sp. (smooth) | High | UT | Chan and Newman 1991 | 21 |
| B. maclearni - B. sp. (smooth) | High | UT | Chan and Newman 1991 | 23 |
| B. maclearni - B. sp. (smooth) | Medium | UT | Chan and Newman 1991 | 26 |
| B. maclearni - B. sp. (smooth) | High | UT | Chan and Newman 1991 | 29 |
| B. maclearni - B. sp. (smooth) | Low | NM | Reeside, 1924 | 190 |
| B. maclearni - B. sp. (smooth) | Low | NM | Reeside, 1924 | 198 |
| B. maclearni - B. sp. (smooth) | Low | NM | Reeside, 1924 | 207 |
| B. maclearni - B. sp. (smooth) | High | SD | Martin et al. 2007 in Martin & Parris | 35 |
| B. maclearni - B. sp. (smooth) | Medium | SD | Martin et al. 2007 in Martin & Parris | 36 |
| B. maclearni - B. sp. (smooth) | Low | SD | Bertog et al. 2007 | 52 |
| B. maclearni - B. sp. (smooth) | Low | KS | Bertog et al. 2007 | 84 |
| B. maclearni - B. sp. (smooth) | Medium | SD | Bertog et al. 2007 | 248 |
| B. maclearni - B. sp. (smooth) | Medium | SD | Bertog et al. 2007 | 249 |
| B. maclearni - B. sp. (smooth) | Medium | SD | Bertog et al. 2007 | 47 |
| B. maclearni - B. sp. (smooth) | Low | NM | Hutchinson 1974 | 151 |
| B. maclearni - B. sp. (smooth) | Low | NM | Hutchinson 1974 | 158 |
| B. maclearni - B. sp. (smooth) | Low | NM | Hutchinson 1974 | 188 |
| B. maclearni - B. sp. (smooth) | Low | NM | Hutchinson 1974 | 139 |

| Interval | Interval Confidence | State | Reference | ID Number |
|---|---|---|---|---|
| B. maclearni - B. sp. (smooth) | Low | SAS | Crockford 1949 | 149 |
| B. maclearni - B. sp. (smooth) | Low | SAS | Crockford 1949 | 142 |
| B. maclearni - B. sp. (smooth) | Medium | ALB | Stott 1967 | 164 |
| B. maclearni - B. sp. (smooth) | Medium | ALB | Stott 1967 | 165 |
| B. maclearni - B. sp. (smooth) | Medium | ALB | Stott 1967 | 168 |
| B. maclearni - B. sp. (smooth) | Low | MT | Johnson and Smith 1964 | 193 |
| B. obtusus | High | WY | Gill and Cobban 1966 | 113 |
| B. obtusus | Medium | WY | Robinson et al. 1964 | 109 |
| B. obtusus | High | WY | Gill et al. 1970 | 226 |
| B. obtusus | High | WY | Gill et al. 1970 | 130 |
| B. obtusus | Medium | WY | Gill et al. 1970 | 119 |
| B. obtusus | High | WY | Gill et al. 1970 | 117 |
| B. obtusus | Medium | WY | Kiteley 1979 | 71 |
| B. obtusus | Medium | CO | Scott 1969 | 85 |
| B. obtusus | High | WY | Roehler 1990 | 162 |
| B. obtusus | High | WY | Roehler 1990 | 172 |
| B. obtusus | High | WY | Roehler 1990 | 166 |
| B. obtusus | High | WY | Gill 1974 | 15 |
| B. obtusus | Low | WY | Gill 1974 | 19 |
| B. obtusus | Low | WY | Gill 1974 | 16 |
| B. obtusus | Low | WY | Johnson et al 2005 | 64 |
| B. obtusus | Low | WY | Minor et al 2022 | 234 |
| B. obtusus | Low | WY | Minor et al 2022 | 235 |
| B. obtusus | Low | WY | Minor et al 2022 | 236 |
| B. obtusus | Low | WY | Minor et al 2022 | 237 |
| B. obtusus | Low | WY | Minor et al 2022 | 238 |
| B. obtusus | Low | WY | Minor et al 2022 | 239 |
| B. obtusus | Medium | UT | Seymour 2012 | 42 |
| B. obtusus | Low | UT | Fisher et al 1960 | 99 |
| B. obtusus | Low | UT | Fisher et al 1960 | 100 |
| B. obtusus | Low | UT | Fisher et al 1960 | 101 |
| B. obtusus | Low | UT | Fisher et al 1960 | 102 |
| B. obtusus | Low | UT | Fisher et al 1960 | 103 |
| B. obtusus | Low | UT | Fisher et al 1960 | 90 |
| B. obtusus | Low | UT | Fisher et al 1960 | 91 |
| B. obtusus | Low | UT | Fisher et al 1960 | 92 |
| B. obtusus | Low | UT | Fisher et al 1960 | 104 |
| B. obtusus | Low | UT | Fisher et al 1960 | 93 |
| B. obtusus | Low | UT | Fisher et al 1960 | 94 |
| B. obtusus | High | UT | Maberry 1971 | 121 |
| B. obtusus | Medium | UT | Chan and Newman 1991 | 10 |
| B. obtusus | High | UT | Chan and Newman 1991 | 14 |
| B. obtusus | High | UT | Chan and Newman 1991 | 17 |

| Interval | Interval Confidence | State | Reference | ID Number |
|---|---|---|---|---|
| B. obtusus | Low | UT | Chan and Newman 1991 | 21 |
| B. obtusus | Low | UT | Chan and Newman 1991 | 23 |
| B. obtusus | Low | NM | Reeside, 1924 | 190 |
| B. obtusus | Low | NM | Reeside, 1924 | 207 |
| B. obtusus | High | SD | Martin et al. 2007 in Martin & Parris | 34 |
| B. obtusus | High | SD | Martin et al. 2007 in Martin & Parris | 35 |
| B. obtusus | Low | SD | Bertog et al. 2007 | 52 |
| B. obtusus | High | KS | Bertog et al. 2007 | 83 |
| B. obtusus | Medium | SD | Bertog et al. 2007 | 248 |
| B. obtusus | Low | SD | Bertog et al. 2007 | 249 |
| B. obtusus | High | SD | Bertog et al. 2007 | 47 |
| B. obtusus | Low | MT | Johnson and Smith 1964 | 54 |
| B. obtusus | Low | MT | Johnson and Smith 1964 | 204 |
| B. obtusus | Low | MT | Johnson and Smith 1964 | 193 |
| B. obtusus | High | SAS | McNeil and Caldwell 1981 | 154 |
| B. obtusus | High | SAS | McNeil and Caldwell 1981 | 155 |
| B. obtusus | High | SAS | McNeil and Caldwell 1981 | 156 |
| B. obtusus | Medium | MAN | McNeil and Caldwell 1981 | 173 |
| B. obtusus | High | MAN | McNeil and Caldwell 1981 | 174 |
| B. obtusus | High | MAN | McNeil and Caldwell 1981 | 175 |
| B. obtusus | Medium | MAN | McNeil and Caldwell 1981 | 176 |
| B. obtusus | High | MAN | McNeil and Caldwell 1981 | 187 |
| B. obtusus | High | MAN | McNeil and Caldwell 1981 | 135 |
| B. obtusus | High | MAN | McNeil and Caldwell 1981 | 136 |
| B. obtusus | High | MAN | McNeil and Caldwell 1981 | 137 |
| B. obtusus | High | SD | McNeil and Caldwell 1981 | 138 |
| B. perplexus - B. gregoryensis | High | WY | Gill and Cobban 1966 | 113 |
| B. perplexus - B. gregoryensis | High | WY | Gill and Cobban 1966 | 113 |
| B. perplexus - B. gregoryensis | High | WY | Robinson et al. 1964 | 115 |
| B. perplexus - B. gregoryensis | Medium | WY | Robinson et al. 1964 | 114 |
| B. perplexus - B. gregoryensis | Medium | WY | Robinson et al. 1964 | 109 |
| B. perplexus - B. gregoryensis | High | WY | Gill et al. 1970 | 226 |
| B. perplexus - B. gregoryensis | High | WY | Gill et al. 1970 | 130 |
| B. perplexus - B. gregoryensis | Medium | WY | Gill et al. 1970 | 119 |
| B. perplexus - B. gregoryensis | High | WY | Gill et al. 1970 | 225 |
| B. perplexus - B. gregoryensis | High | WY | Gill et al. 1970 | 117 |
| B. perplexus - B. gregoryensis | High | WY | Kiteley 1979 | 111 |
| B. perplexus - B. gregoryensis | Medium | CO | Scott 1969 | 85 |
| B. perplexus - B. gregoryensis | High | WY | Roehler 1990 | 162 |
| B. perplexus - B. gregoryensis | High | WY | Roehler 1990 | 172 |
| B. perplexus - B. gregoryensis | High | WY | Roehler 1990 | 166 |
| B. perplexus - B. gregoryensis | High | CO | Roehler 1990 | 161 |
| B. perplexus - B. gregoryensis | Medium | WY | Gill 1974 | 8 |

| Interval | Interval Confidence | State | Reference | ID Number |
|---|---|---|---|---|
| B. perplexus - B. gregoryensis | Medium | WY | Gill 1974 | 12 |
| B. perplexus - B. gregoryensis | Medium | WY | Gill 1974 | 15 |
| B. perplexus - B. gregoryensis | High | WY | Gill 1974 | 22 |
| B. perplexus - B. gregoryensis | High | WY | Gill 1974 | 25 |
| B. perplexus - B. gregoryensis | High | WY | Gill 1974 | 28 |
| B. perplexus - B. gregoryensis | High | WY | Gill 1974 | 31 |
| B. perplexus - B. gregoryensis | Medium | WY | Gill 1974 | 32 |
| B. perplexus - B. gregoryensis | Low | WY | Johnson et al 2005 | 65 |
| B. perplexus - B. gregoryensis | Low | WY | Johnson et al 2005 | 74 |
| B. perplexus - B. gregoryensis | Low | WY | Johnson et al 2005 | 60 |
| B. perplexus - B. gregoryensis | Low | WY | Johnson et al 2005 | 73 |
| B. perplexus - B. gregoryensis | Low | WY | Johnson et al 2005 | 64 |
| B. perplexus - B. gregoryensis | High | WY | Minor et al 2022 | 244 |
| B. perplexus - B. gregoryensis | High | CO | Minor et al 2022 | 247 |
| B. perplexus - B. gregoryensis | High | CO | Minor et al 2022 | 246 |
| B. perplexus - B. gregoryensis | High | CO | Minor et al 2022 | 245 |
| B. perplexus - B. gregoryensis | Medium | UT | Seymour 2012 | 46 |
| B. perplexus - B. gregoryensis | High | CO | Porter 1940 | 44 |
| B. perplexus - B. gregoryensis | High | CO | Porter 1940 | 44 |
| B. perplexus - B. gregoryensis | High | CO | Porter 1940 | 123 |
| B. perplexus - B. gregoryensis | High | CO | Porter 1940 | 66 |
| B. perplexus - B. gregoryensis | High | CO | Porter 1940 | 66 |
| B. perplexus - B. gregoryensis | High | CO | Porter 1940 | 66 |
| B. perplexus - B. gregoryensis | High | CO | Porter 1940 | 59 |
| B. perplexus - B. gregoryensis | High | CO | Porter 1940 | 59 |
| B. perplexus - B. gregoryensis | High | CO | Porter 1940 | 89 |
| B. perplexus - B. gregoryensis | High | CO | Porter 1940 | 86 |
| B. perplexus - B. gregoryensis | High | CO | Porter 1940 | 63 |
| B. perplexus - B. gregoryensis | Low | UT | Fisher et al 1960 | 90 |
| B. perplexus - B. gregoryensis | Low | UT | Fisher et al 1960 | 91 |
| B. perplexus - B. gregoryensis | Low | UT | Fisher et al 1960 | 92 |
| B. perplexus - B. gregoryensis | Low | UT | Fisher et al 1960 | 104 |
| B. perplexus - B. gregoryensis | Low | UT | Fisher et al 1960 | 93 |
| B. perplexus - B. gregoryensis | Low | UT | Fisher et al 1960 | 94 |
| B. perplexus - B. gregoryensis | Low | UT | Fisher et al 1960 | 95 |
| B. perplexus - B. gregoryensis | Low | UT | Fisher et al 1960 | 96 |
| B. perplexus - B. gregoryensis | Low | UT | Fisher et al 1960 | 97 |
| B. perplexus - B. gregoryensis | Low | SD | Bertog et al. 2007 | 52 |
| B. perplexus - B. gregoryensis | Medium | SD | Hanczaryk and Gallagher 2007 | 67 |
| B. perplexus - B. gregoryensis | Low | NM | Fitter 1958 | 157 |
| B. perplexus - B. gregoryensis | Low | SAS | Crockford 1949 | 149 |
| B. perplexus - B. gregoryensis | Low | SAS | Crockford 1949 | 142 |
| B. perplexus - B. gregoryensis | High | SAS | McNeil and Caldwell 1981 | 153 |

| Interval | Interval Confidence | State | Reference | ID Number |
|---|---|---|---|---|
| B. perplexus - B. gregoryensis | High | MAN | McNeil and Caldwell 1981 | 173 |
| B. perplexus - B. gregoryensis | High | MAN | McNeil and Caldwell 1981 | 174 |
| B. perplexus - B. gregoryensis | Low | MAN | McNeil and Caldwell 1981 | 180 |
| B. perplexus - B. gregoryensis | High | MAN | McNeil and Caldwell 1981 | 180 |
| B. perplexus - B. gregoryensis | High | MAN | McNeil and Caldwell 1981 | 135 |
| B. perplexus - B. gregoryensis | High | MAN | McNeil and Caldwell 1981 | 136 |
| B. perplexus - B. gregoryensis | High | SD | McNeil and Caldwell 1981 | 138 |
| B. reduncus - B. scotti | High | WY | Gill and Cobban 1966 | 113 |
| B. reduncus - B. scotti | High | WY | Gill and Cobban 1966 | 113 |
| B. reduncus - B. scotti | Low | WY | Kiteley 1979 | 111 |
| B. reduncus - B. scotti | High | CO | Scott 1969 | 85 |
| B. reduncus - B. scotti | High | CO | Minor et al 2022 | 227 |
| B. reduncus - B. scotti | Medium | UT | Seymour 2012 | 46 |
| B. reduncus - B. scotti | High | CO | Porter 1940 | 44 |
| B. reduncus - B. scotti | High | CO | Porter 1940 | 66 |
| B. reduncus - B. scotti | High | CO | Porter 1940 | 59 |
| B. reduncus - B. scotti | High | CO | Porter 1940 | 86 |
| B. reduncus - B. scotti | High | CO | Porter 1940 | 63 |
| B. reduncus - B. scotti | High | NM | Sealey and Lucas 2022 | 53 |
| B. reduncus - B. scotti | Low | UT | Fisher et al 1960 | 90 |
| B. reduncus - B. scotti | Low | UT | Fisher et al 1960 | 91 |
| B. reduncus - B. scotti | Low | UT | Fisher et al 1960 | 92 |
| B. reduncus - B. scotti | Low | UT | Fisher et al 1960 | 104 |
| B. reduncus - B. scotti | Low | UT | Fisher et al 1960 | 94 |
| B. reduncus - B. scotti | Low | UT | Fisher et al 1960 | 95 |
| B. reduncus - B. scotti | Low | UT | Fisher et al 1960 | 96 |
| B. reduncus - B. scotti | Low | UT | Fisher et al 1960 | 98 |
| B. reduncus - B. scotti | Medium | CO | Warner 1964 | 2 |
| B. reduncus - B. scotti | High | CO | Warner 1964 | 9 |
| B. reduncus - B. scotti | High | CO | Warner 1964 | 13 |
| B. reduncus - B. scotti | High | CO | Warner 1964 | 18 |
| B. reduncus - B. scotti | Low | CO | Warner 1964 | 20 |
| B. reduncus - B. scotti | Medium | CO | Warner 1964 | 24 |
| B. reduncus - B. scotti | Medium | CO | Warner 1964 | 27 |
| B. reduncus - B. scotti | Low | CO | Warner 1964 | 30 |
| B. reduncus - B. scotti | Medium | SD | Hanczaryk and Gallagher 2007 | 67 |
| B. reduncus - B. scotti | Medium | SD | Hanczaryk and Gallagher 2007 | 105 |
| B. reduncus - B. scotti | Medium | MAN | McNeil and Caldwell 1981 | 180 |
| B. reesidei - B. eliasi | High | WY | Gill and Cobban 1966 | 113 |
| B. reesidei - B. eliasi | High | WY | Gill and Cobban 1966 | 113 |
| B. reesidei - B. eliasi | High | WY | Gill and Cobban 1966 | 113 |
| B. reesidei - B. eliasi | Medium | WY | Robinson et al. 1964 | 108 |
| B. reesidei - B. eliasi | High | WY | Robinson et al. 1964 | 112 |

| Interval | Interval Confidence | State | Reference | ID Number |
|---|---|---|---|---|
| B. reesidei - B. eliasi | High | WY | Robinson et al. 1964 | 110 |
| B. reesidei - B. eliasi | Medium | WY | Gill et al. 1970 | 127 |
| B. reesidei - B. eliasi | Medium | WY | Gill et al. 1970 | 128 |
| B. reesidei - B. eliasi | High | WY | Gill et al. 1970 | 129 |
| B. reesidei - B. eliasi | High | WY | Kiteley 1979 | 111 |
| B. reesidei - B. eliasi | High | CO | Roehler 1990 | 161 |
| B. reesidei - B. eliasi | High | WY | Gill 1974 | 11 |
| B. reesidei - B. eliasi | Low | WY | Johnson et al 2005 | 60 |
| B. reesidei - B. eliasi | Medium | WY | Johnson et al 2005 | 73 |
| B. reesidei - B. eliasi | Low | WY | Johnson et al 2005 | 64 |
| B. reesidei - B. eliasi | High | CO | Minor et al 2022 | 230 |
| B. reesidei - B. eliasi | High | UT | Minor et al 2022 | 231 |
| B. reesidei - B. eliasi | High | CO | Porter 1940 | 44 |
| B. reesidei - B. eliasi | High | CO | Porter 1940 | 124 |
| B. reesidei - B. eliasi | High | NM | Sealey and Lucas 2022 | 33 |
| B. reesidei - B. eliasi | High | NM | Sealey and Lucas 2022 | 221 |
| B. reesidei - B. eliasi | Medium | NM | Sealey and Lucas 2022 | 50 |
| B. reesidei - B. eliasi | Medium | NM | Sealey and Lucas 2022 | 51 |
| B. reesidei - B. eliasi | High | NM | Sealey and Lucas 2022 | 131 |
| B. reesidei - B. eliasi | High | NM | Sealey and Lucas 2022 | 132 |
| B. reesidei - B. eliasi | High | NM | Sealey and Lucas 2022 | 133 |
| B. reesidei - B. eliasi | High | NM | Sealey and Lucas 2022 | 77 |
| B. reesidei - B. eliasi | High | NM | Sealey and Lucas 2022 | 38 |
| B. reesidei - B. eliasi | High | ALB | Tsujita 1955 | 58 |
| B. reesidei - B. eliasi | High | ALB | Tsujita 1955 | 58 |
| B. reesidei - B. eliasi | High | ALB | Tsujita 1955 | 222 |
| B. reesidei - B. eliasi | High | ALB | Tsujita 1955 | 222 |
| B. reesidei - B. eliasi | High | ALB | Tsujita 1955 | 79 |
| B. reesidei - B. eliasi | High | ALB | Tsujita 1955 | 45 |
| B. reesidei - B. eliasi | High | ALB | Tsujita 1955 | 37 |
| B. reesidei - B. eliasi | High | ALB | Tsujita 1955 | 125 |
| B. reesidei - B. eliasi | Low | MT | Heald 1926 | 191 |
| D. cheyennense | Medium | WY | Robinson et al. 1964 | 108 |
| D. cheyennense | High | CO | Roehler 1990 | 161 |
| D. cheyennense | High | NM | Sealey and Lucas 2022 | 221 |
| D. cheyennense | High | NM | Sealey and Lucas 2022 | 77 |
| D. cheyennense | Low | NM | Caldwell 1953 | 78 |
| D. cheyennense | Low | NM | Caldwell 1953 | 76 |
| D. cheyennense | Low | NM | Caldwell 1953 | 55 |
| D. cheyennense | Low | NM | Hutchinson 1974 | 151 |
| D. cheyennense | Low | MT | Johnson and Smith 1964 | 192 |
| D. cheyennense | Medium | MAN | McNeil and Caldwell 1981 | 179 |
| D. cheyennense | Medium | MAN | McNeil and Caldwell 1981 | 181 |

| Interval | Interval Confidence | State | Reference | ID Number |
|---|---|---|---|---|
| D. cheyennense | Medium | MAN | McNeil and Caldwell 1981 | 182 |
| D. cheyennense | Medium | MAN | McNeil and Caldwell 1981 | 183 |
| D. cheyennense | Medium | MAN | McNeil and Caldwell 1981 | 184 |
| D. cheyennense | Medium | MAN | McNeil and Caldwell 1981 | 185 |
| D. cheyennense | Medium | MAN | McNeil and Caldwell 1981 | 186 |
| D. cheyennense | Medium | MAN | McNeil and Caldwell 1981 | 137 |
| D. nebrascense - E. jenneyi | High | WY | Gill and Cobban 1966 | 113 |
| D. nebrascense - E. jenneyi | Medium | WY | Robinson et al. 1964 | 108 |
| D. nebrascense - E. jenneyi | High | WY | Gill et al. 1970 | 225 |
| D. nebrascense - E. jenneyi | High | WY | Gill et al. 1970 | 117 |
| D. nebrascense - E. jenneyi | High | WY | Kiteley 1979 | 111 |
| D. nebrascense - E. jenneyi | High | CO | Scott 1969 | 85 |
| D. nebrascense - E. jenneyi | High | CO | Roehler 1990 | 161 |
| D. nebrascense - E. jenneyi | High | CO | Minor et al 2022 | 228 |
| D. nebrascense - E. jenneyi | High | WY | Minor et al 2022 | 229 |
| D. nebrascense - E. jenneyi | Medium | UT | Seymour 2012 | 88 |
| D. nebrascense - E. jenneyi | High | NM | Sealey and Lucas 2022 | 33 |
| D. nebrascense - E. jenneyi | High | NM | Sealey and Lucas 2022 | 53 |
| D. nebrascense - E. jenneyi | Medium | NM | Sealey and Lucas 2022 | 224 |
| D. nebrascense - E. jenneyi | Medium | CO | Warner 1964 | 2 |
| D. nebrascense - E. jenneyi | Medium | CO | Warner 1964 | 9 |
| D. nebrascense - E. jenneyi | Medium | CO | Warner 1964 | 13 |
| D. nebrascense - E. jenneyi | Medium | CO | Warner 1964 | 18 |
| D. nebrascense - E. jenneyi | Medium | CO | Warner 1964 | 20 |
| D. nebrascense - E. jenneyi | Medium | CO | Warner 1964 | 24 |
| D. nebrascense - E. jenneyi | Medium | CO | Warner 1964 | 27 |
| D. nebrascense - E. jenneyi | Medium | CO | Warner 1964 | 30 |
| D. nebrascense - E. jenneyi | Medium | SD | Hanczaryk and Gallagher 2007 | 105 |
| D. nebrascense - E. jenneyi | Medium | NM | Hutchinson 1974 | 170 |
| D. nebrascense - E. jenneyi | Low | NM | Fitter 1958 | 157 |
| D. nebrascense - E. jenneyi | Low | MT | Johnson and Smith 1964 | 197 |
| D. nebrascense - E. jenneyi | Low | MT | Johnson and Smith 1964 | 192 |
| D. nebrascense - E. jenneyi | High | MAN | McNeil and Caldwell 1981 | 179 |
| D. nebrascense - E. jenneyi | Medium | MAN | McNeil and Caldwell 1981 | 180 |
| D. nebrascense - E. jenneyi | Medium | MAN | McNeil and Caldwell 1981 | 181 |
| D. nebrascense - E. jenneyi | Medium | MAN | McNeil and Caldwell 1981 | 182 |
| D. nebrascense - E. jenneyi | Medium | MAN | McNeil and Caldwell 1981 | 183 |
| D. nebrascense - E. jenneyi | Medium | MAN | McNeil and Caldwell 1981 | 184 |
| D. nebrascense - E. jenneyi | Medium | MAN | McNeil and Caldwell 1981 | 185 |
| D. nebrascense - E. jenneyi | Medium | MAN | McNeil and Caldwell 1981 | 186 |
| D. nebrascense - E. jenneyi | Medium | MAN | McNeil and Caldwell 1981 | 136 |
| D. nebrascense - E. jenneyi | High | MAN | McNeil and Caldwell 1981 | 137 |
| H. birkelundae - H. nebrascensis | High | WY | Gill and Cobban 1966 | 69 |

| Interval | Interval Confidence | State | Reference | ID Number |
|---|---|---|---|---|
| H. birkelundae - H. nebrascensis | High | CO | Nwangwu 1977 | 116 |
| H. birkelundae - H. nebrascensis | Medium | SD | Pettyjohn 1967 | 134 |
| H. birkelundae - H. nebrascensis | Low | SD | Pettyjohn 1967 | 152 |
| H. birkelundae - H. nebrascensis | High | SD | Waage 1968 | 120 |
| H. birkelundae - H. nebrascensis | High | SD | Waage 1968 | 159 |
| H. birkelundae - H. nebrascensis | High | SD | Waage 1968 | 160 |
| H. birkelundae - H. nebrascensis | Medium | SD | Waage 1968 | 163 |
| H. birkelundae - H. nebrascensis | Medium | SD | Waage 1968 | 169 |
| H. birkelundae - H. nebrascensis | Medium | SD | Waage 1968 | 171 |
| H. birkelundae - H. nebrascensis | Medium | SD | Waage 1968 | 177 |
| H. birkelundae - H. nebrascensis | Medium | SD | Waage 1968 | 178 |
| H. birkelundae - H. nebrascensis | Medium | SD | Waage 1968 | 189 |
| H. birkelundae - H. nebrascensis | Medium | SD | Waage 1968 | 140 |
| H. birkelundae - H. nebrascensis | Medium | SD | Waage 1968 | 141 |
| H. birkelundae - H. nebrascensis | Medium | SD | Waage 1968 | 143 |
| H. birkelundae - H. nebrascensis | Low | SD | Waage 1968 | 144 |
| H. birkelundae - H. nebrascensis | Low | SD | Waage 1968 | 145 |
| H. birkelundae - H. nebrascensis | Low | SD | Waage 1968 | 146 |
| H. birkelundae - H. nebrascensis | Low | SD | Waage 1968 | 147 |
| S. leei - B. sp. (weak flank ribs) | Medium | WY | Gill and Cobban 1966 | 113 |
| S. leei - B. sp. (weak flank ribs) | Medium | WY | Robinson et al. 1964 | 115 |
| S. leei - B. sp. (weak flank ribs) | Medium | WY | Robinson et al. 1964 | 114 |
| S. leei - B. sp. (weak flank ribs) | High | WY | Gill et al. 1970 | 226 |
| S. leei - B. sp. (weak flank ribs) | High | WY | Gill et al. 1970 | 130 |
| S. leei - B. sp. (weak flank ribs) | Medium | WY | Kiteley 1979 | 71 |
| S. leei - B. sp. (weak flank ribs) | High | CO | Scott 1969 | 85 |
| S. leei - B. sp. (weak flank ribs) | High | WY | Roehler 1990 | 162 |
| S. leei - B. sp. (weak flank ribs) | High | WY | Roehler 1990 | 148 |
| S. leei - B. sp. (weak flank ribs) | Low | WY | Gill 1974 | 12 |
| S. leei - B. sp. (weak flank ribs) | High | WY | Gill 1974 | 12 |
| S. leei - B. sp. (weak flank ribs) | High | WY | Gill 1974 | 19 |
| S. leei - B. sp. (weak flank ribs) | High | WY | Gill 1974 | 16 |
| S. leei - B. sp. (weak flank ribs) | Low | WY | Johnson et al 2005 | 106 |
| S. leei - B. sp. (weak flank ribs) | Medium | WY | Johnson et al 2005 | 65 |
| S. leei - B. sp. (weak flank ribs) | Low | WY | Johnson et al 2005 | 74 |
| S. leei - B. sp. (weak flank ribs) | Low | WY | Johnson et al 2005 | 72 |
| S. leei - B. sp. (weak flank ribs) | Low | WY | Johnson et al 2005 | 73 |
| S. leei - B. sp. (weak flank ribs) | Low | WY | Johnson et al 2005 | 64 |
| S. leei - B. sp. (weak flank ribs) | High | WY | Minor et al 2022 | 232 |
| S. leei - B. sp. (weak flank ribs) | High | WY | Minor et al 2022 | 233 |
| S. leei - B. sp. (weak flank ribs) | High | UT | Seymour 2012 | 82 |
| S. leei - B. sp. (weak flank ribs) | Medium | UT | Seymour 2012 | 87 |
| S. leei - B. sp. (weak flank ribs) | Medium | UT | Seymour 2012 | 43 |

| Interval | Interval Confidence | State | Reference | ID Number |
|---|---|---|---|---|
| S. leei - B. sp. (weak flank ribs) | Medium | UT | Seymour 2012 | 41 |
| S. leei - B. sp. (weak flank ribs) | Low | UT | Fisher et al 1960 | 99 |
| S. leei - B. sp. (weak flank ribs) | Medium | UT | Fisher et al 1960 | 99 |
| S. leei - B. sp. (weak flank ribs) | Medium | UT | Fisher et al 1960 | 99 |
| S. leei - B. sp. (weak flank ribs) | Low | UT | Fisher et al 1960 | 100 |
| S. leei - B. sp. (weak flank ribs) | Low | UT | Fisher et al 1960 | 101 |
| S. leei - B. sp. (weak flank ribs) | Low | UT | Fisher et al 1960 | 102 |
| S. leei - B. sp. (weak flank ribs) | High | UT | Fisher et al 1960 | 102 |
| S. leei - B. sp. (weak flank ribs) | Low | UT | Fisher et al 1960 | 103 |
| S. leei - B. sp. (weak flank ribs) | Low | UT | Fisher et al 1960 | 90 |
| S. leei - B. sp. (weak flank ribs) | Low | UT | Fisher et al 1960 | 91 |
| S. leei - B. sp. (weak flank ribs) | Medium | CO | Leckie et al 1997 | 81 |
| S. leei - B. sp. (weak flank ribs) | Medium | UT | Maberry 1971 | 122 |
| S. leei - B. sp. (weak flank ribs) | High | UT | Chan and Newman 1991 | 10 |
| S. leei - B. sp. (weak flank ribs) | High | UT | Chan and Newman 1991 | 14 |
| S. leei - B. sp. (weak flank ribs) | High | UT | Chan and Newman 1991 | 17 |
| S. leei - B. sp. (weak flank ribs) | Medium | UT | Chan and Newman 1991 | 21 |
| S. leei - B. sp. (weak flank ribs) | Low | NM | Reeside, 1924 | 190 |
| S. leei - B. sp. (weak flank ribs) | Low | NM | Reeside, 1924 | 207 |
| S. leei - B. sp. (weak flank ribs) | High | KS | Bertog et al. 2007 | 83 |
| S. leei - B. sp. (weak flank ribs) | High | KS | Bertog et al. 2007 | 48 |
| S. leei - B. sp. (weak flank ribs) | High | SD | Bertog et al. 2007 | 248 |
| S. leei - B. sp. (weak flank ribs) | High | WY | Bertog et al. 2007 | 126 |
| S. leei - B. sp. (weak flank ribs) | Medium | SD | Bertog et al. 2007 | 249 |
| S. leei - B. sp. (weak flank ribs) | Medium | ALB | Stott 1967 | 164 |
| S. leei - B. sp. (weak flank ribs) | Medium | ALB | Stott 1967 | 165 |
| S. leei - B. sp. (weak flank ribs) | Medium | ALB | Stott 1967 | 168 |
| S. leei - B. sp. (weak flank ribs) | Medium | ALB | Stott 1967 | 167 |
| S. leei - B. sp. (weak flank ribs) | Low | MT | Johnson and Smith 1964 | 203 |
| S. leei - B. sp. (weak flank ribs) | Low | MT | Johnson and Smith 1964 | 202 |
| S. leei - B. sp. (weak flank ribs) | Low | MT | Johnson and Smith 1964 | 54 |
| S. leei - B. sp. (weak flank ribs) | Low | MT | Johnson and Smith 1964 | 54 |
| S. leei - B. sp. (weak flank ribs) | Low | MT | Johnson and Smith 1964 | 204 |
| S. leei - B. sp. (weak flank ribs) | Low | MT | Fuentes et al 2011 | 150 |

## Sedimentary Variable Interpolation

Interpolations were performed using ArcGIS Pro (ArcGIS Pro 3.2.0). Cross validation comparisons of root mean squared error (RMSE) values were conducted on two example variable layers to determine the most appropriate method for the highly clustered, low sample-size datasets. Only two variable layers were compared in this way to save time on computation: percent sand in the *Hoploscaphites birkelundae-Hoploscaphites nebrasensis* and *Baculites obtusus* intervals (Table S5). Kriging was not utilized given the scarcity and non-normal distribution of values within the datasets (Webster and Oliver 2007; Wang et al. 2017; Zhao et al. 2018). The *H. birkelundae-H. nebrasensis* and *B. obtusus* percent sand variables were selected for this comparison because they contain one of the lowest and highest numbers of sedimentary localities (20 and 63, respectively) and both are highly clustered, therefore potentially representing the extremes in potential sources of bias (Figure S2). Inverse Distance Weighting (IDW) interpolation of raw values was used to create raster surfaces of these test variables, using a cell size of 30km, a power of 2, and the Standard search neighborhood and all other default values from the ArcGIS Pro IDW tool.

Table S5. RMSE results for raw data and averaged/extrapolated data.

| Interval | RMSE | |
|---|---|---|
| | Raw Data | Smoothed Data |
| *H. birkelundae-H. nebrascensis* | 0.268 | 0.099 |
| *S. leei-B. sp. (weak flank ribs)* | 0.260 | 0.064 |
| *B. obtusus* | 0.211 | 0.066 |

Given that many sedimentary localities are well within 30km of each other and contain potentially variable values even within these distances, average values within a 30km range were also calculated using the Point Statistic tool in ArcGIS Pro (calculating the mean value of 30km cells within a square neighborhood of two cells). This tool was used to calculate the mean of point values within a specified neighborhood of two 30km cells and convert those mean values to a 30km raster. Cells in the raster with no point values are given NA values. This raster was then converted to vector point values

using the Raster to Point tool, and those values were then interpolated using the IDW method as described above. Surfaces interpolated using this spatial averaging technique had much lower RMSE values than those produced using the raw data, though this result is considered highly biased given that values are smoothed prior to interpolation. However, the raster surfaces created using these mean values were considered more reasonable based on visual inspection and contain fewer interpolation artifacts caused by the raw data, such as bulls-eye patterns and sudden sharp changes in value (Figure S2). Both distance from seeps and distance from nearest shoreline were interpolated without smoothing the data first.

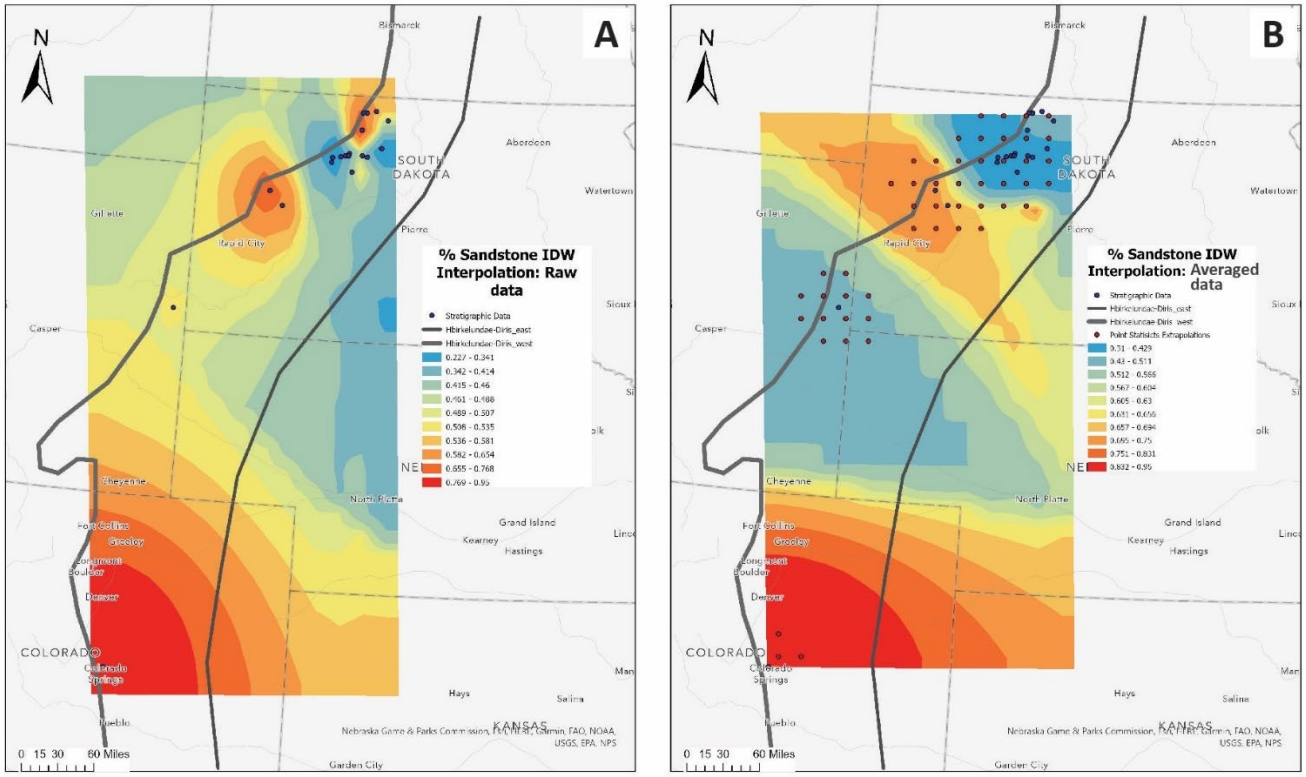


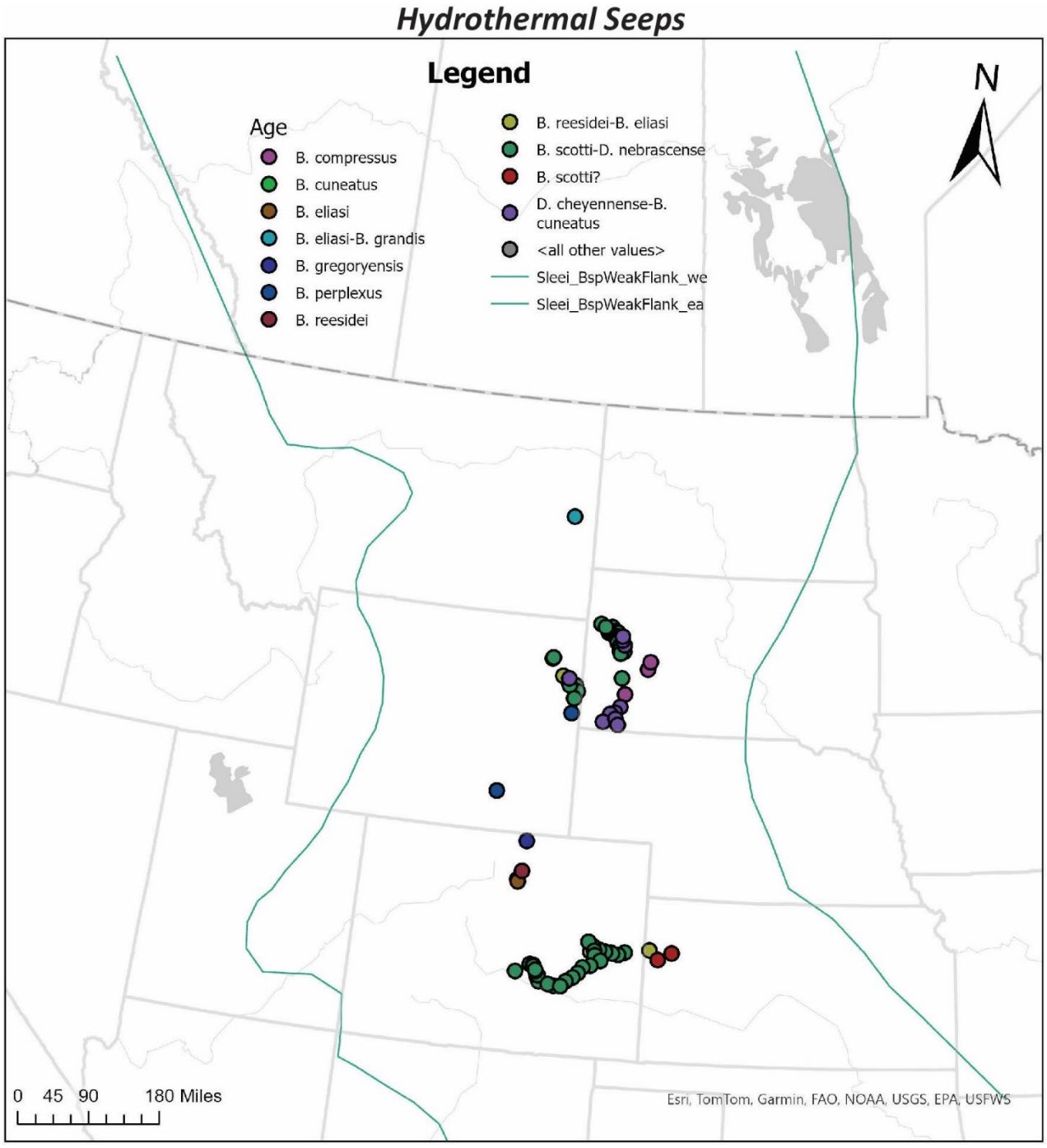

Figure S2. Example interpolation results for percent sand in the *H. birkelundae-H. nebrascensis* interval based on raw data values and averaged/extrapolated data values. Blue points represent stratigraphic data locations, red points represent averaged values calculated using the Point Statistics tool, and dark grey lines represent shorelines. Note the lack of artifacts in the averaged/extrapolated interpolation compared to the raw data interpolation. Red colors indicate higher percent sand, and cool colors indicate lower percent sand.

## Distance from Seeps Interpolation Methods:

Distance from seep locations was included as a variable in the analysis. This variable was calculated by first documenting general seep locations based on literature sources (see reference section). These locations did not attempt to include individual seeps, but instead mark the general location of seeps within a specific interval based on literature descriptions and map locations (Figure S3). Since seeps are often highly clustered in space, within a 10 km from each other (see Metz 2010 for discussion of seep distributions), and the raster cell size selected for this analysis is 30km based on taxonomic occurrence error allowance, it was not considered necessary to include higher resolution locality information. All seep locations were given distance from seep values of zero.

Distance from seeps was determined by plotting general seep occurrence localities and creating buffer distances around these points at 60km intervals from 60 to 480km (and an additional 2000km distance) using the Buffer tool in ArcGIS Pro (using geodesic distance and dissolving features). Points were then created along that polygon's boundary using the Generate Points Along Line tool (using the percentage option to create points every 10% of the polygon's boarder. These points were given the value of the initial buffer distance used to create them in meters using the Calculate Fields tool. Once all buffer distance polygon boundaries were converted into point distributions, these points were collected into a single feature class using the Merge tool. Merged points were then used to interpolate a new surface of distance from seeps. This analysis and all others performed in ArcGIS Pro were performed using the USA Contiguous Albers Equal Area Conic projected coordinate system which preserves area values.

Figure S3. Map of hydrothermal seep locations across the WIS, colored by age. Strandlines are from the S. leei-B. sp. (weak flank ribs), which provides the larges seaway extent.



## Distance from Strandlines Interpolation

Strandlines for each biostratigraphic interval were produced based on relevant publications (Gill and Cobban, 1973; Slattery et al., 2015; Cobban et al. 1994; Roberts and Krischbaum 1995) and adjusted based on the taxonomic occurrence data as needed. Vector line feature classes were produced in ArcGIS Pro and converted to point features using the Generate Points along Line tool every 120km along the shoreline, including endpoints. The Near tool was used to calculate the angle and distance from each point on the western shoreline to the nearest point along the eastern shoreline. These values were then used to calculate latitude and longitude values along the same angle at 5%, 25%, 50%, 75% and 95% of the total distance from a point on the western shoreline to the nearest point along the eastern shoreline. Vector points were created at these locations and given their distance from the closest shoreline as an integer attribute, i.e., a point created at 75% of the distance from a western shoreline point to an eastern shoreline point were given a value of 25% of the total distance between the two points (see Appendix 1-2 for code). These points were used to interpolate raster surfaces of distance from the nearest shoreline using the IDW method as described above.

## Masking by Shoreline/Outcrop Polygons

All variable raster surfaces were clipped based on relevant Upper Cretaceous outcrop polygons and strandlines boundaries (see above for discussion strandline production). USA State geologic outcrop polygons with relevant formation polygons were extracted based on age (CDC 1996; Green 1992; Green et al. 1994; Green et al. 1997; Hintze et al. 2000; North Dakota Geological Survey 2001; Ross and Jorgina 1992; Tomhave and Schulz 2004; Vuke et al 2007). Outcrop polygons were given a buffer of 30km distance using the Buffer tool, to account for taxonomic locality error and provide slightly more area for analysis on the assumption that some outcrop features will not be accounted for by the outcrop polygons, especially when geographically small. The buffered outcrop polygon was merged with biozone strandline boundary polygons to create a mask layer for each interval. Interpolation raster layers for each variable in each interval were then clipped to the mask extents of their respective interval using the Extract by Mask tool.

## Environmental Variables Correlation

All interpolated layers were exported as TIF files and input into R to run analysis (Figures S4). These layers were analyzed for autocorrelation using the *pairs* function in R (Figures S5). Each interval was evaluated for correlation with variables removed that showed a Pearson correlation coefficient of 0.8 or greater. All intervals included a different combination of variables following removal of variables to reduce correlation (Table S6 and S7). Shale, though present in all intervals, strongly correlated with other variables in all intervals, and was therefore removed from all analysis. All other variables were present in at least five intervals.

## Comparisons with Substage and Stage

Aggregations were also performed at the substage and stage level, when reasonable, creating five total substage time bins (lower, middle, and upper Campanian, and lower and upper Maastrichtian) and a single stage-level aggregation for the Maastrichtian. A Campanian stage bin was not created, given that the Campanian spans ~ 11.5 Ma, which was not considered to represent a meaningful level of environmental aggregation.

Aggregation and interpolations for Substages and the Maastrichtian were created using the same process used for biozone intervals. When possible, the same initial interpolation layers used for biozone intervals were recycled for use in the larger temporal bins, including the *B. baculus*-*B. grandis* distance to seep and distance to shore raster layers for the lower Maastrichtian and Maastrichtian and all *S. leei*-*B. sp.* (weak flank ribs) raster layers for the lower Campanian. To make shorelines for lower temporal resolution bins, shoreline polygons were merged and boundaries were dissolved to create a single, new polygon. In the case of the upper Campanian, where this new merged polygon showed highly irregular boundaries including right angles, shorelines were smoothed manually to reduce harsh angle artifacts (Figure S4). Raster layers for each lower resolution temporal bin were then clipped to these new shoreline boundaries using the smoothed polygons.

Table S6. Table of variables that had Pearson's correlation coefficient values less than 0.80 in each biozone. Y=Yes, N=No, and NA=not present

| Interval | Dist. to Shore | Dist. to Seeps | Percent Pebble | Percent Sand | Percent Shale | Percent Silt | Percent Mud | Percent Limestone | Bedding Style | Bioturbation | Bed Thickness |
|---|---|---|---|---|---|---|---|---|---|---|---|
| *H. birkelundae-H. nebrascensis* | Y | NA | Y | N | N | Y | N | Y | N | N | Y |
| *B. clinolobatus* | Y | NA | Y | Y | N | N | Y | N | N | N | Y |
| *B. baculus-B. grandis* | Y | Y | Y | Y | N | Y | Y | Y | N | Y | Y |
| *B. reesidei-B. eliasi* | Y | N | Y | Y | N | Y | Y | Y | Y | Y | Y |
| *B. compressus-B. cuneatus* | Y | Y | NA | Y | N | Y | Y | N | N | Y | N |
| *D. cheyennense* | Y | N | NA | Y | N | Y | Y | Y | N | N | N |
| *D. nebrascense-E. jenneyi* | Y | Y | NA | Y | N | Y | Y | Y | Y | Y | Y |
| *B. reduncus-B. scotti* | Y | Y | NA | Y | N | Y | Y | Y | Y | Y | N |
| *B. perplexus-B. gregoryensis* | Y | Y | Y | Y | N | Y | Y | Y | Y | Y | Y |
| *B. maclearni-B. sp.* (smooth) | Y | NA | N | Y | Y | N | Y | Y | N | Y | Y |
| *B. obtusus* | Y | NA | NA | Y | N | Y | Y | Y | Y | N | Y |
| *S. leei-B. sp.* (weak flank ribs) | Y | NA | Y | Y | N | Y | Y | Y | Y | N | N |

Table S7. Table of variables used in temporal comparison of niche overlap, indicated with an asterisk. Y=less than 0.80 Spearman's correlation with another variable, N=greater than or equal to 0.80 Spearman's correlation with another variable, and NA=not present in interval.

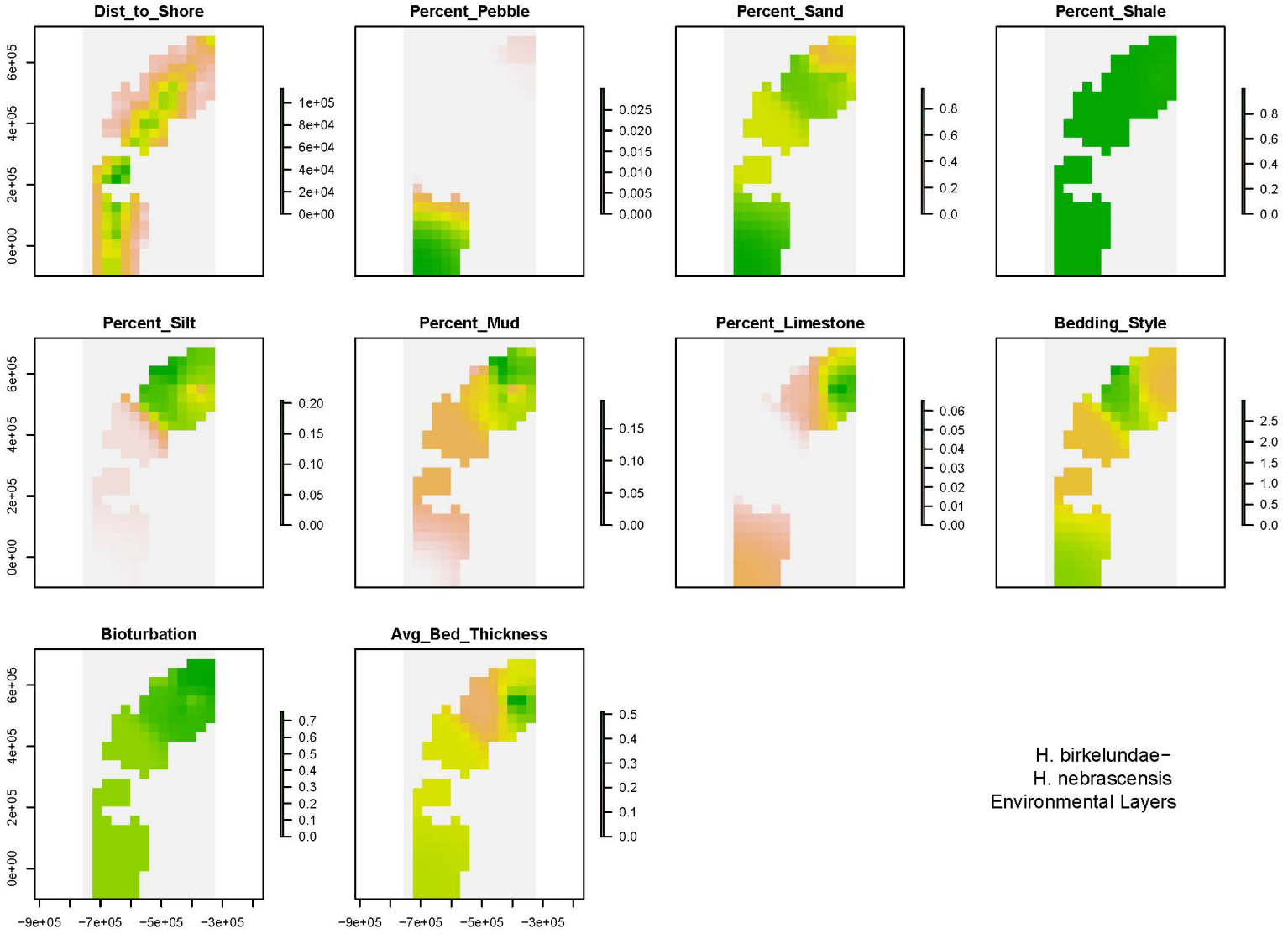
| | Dist. to Shore | Dist. to Seeps | Percent Pebble | Percent Sand | Percent Shale | Percent Silt | Percent Mud | Percent Limestone | Bedding Style | Bioturbation | Bed Thickness |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Upper Campanian** | Y* | Y | Y | Y* | N | Y* | Y* | Y | Y | Y* | Y |
| *B. reesidei-B. eliasi* | Y* | N | Y | Y* | N | Y* | Y* | Y | Y | Y* | Y |
| *B. compressus-B. cuneatus* | Y* | Y | NA | Y* | N | Y* | Y* | N | N | Y* | N |
| *D. nebrascense-E. jenneyi* | Y* | Y | NA | Y* | N | Y* | Y* | Y | Y | Y* | Y |
| **Middle Campanian** | Y* | Y | Y | Y* | N | Y* | Y* | Y* | N | Y* | Y |
| *B. reduncus-B. scotti* | Y* | Y | NA | Y* | N | Y* | Y* | Y* | Y | Y* | N |
| *B. maclearni-B. sp. (smooth)* | Y* | NA | N | Y* | N | Y* | Y* | Y* | N | Y* | Y |

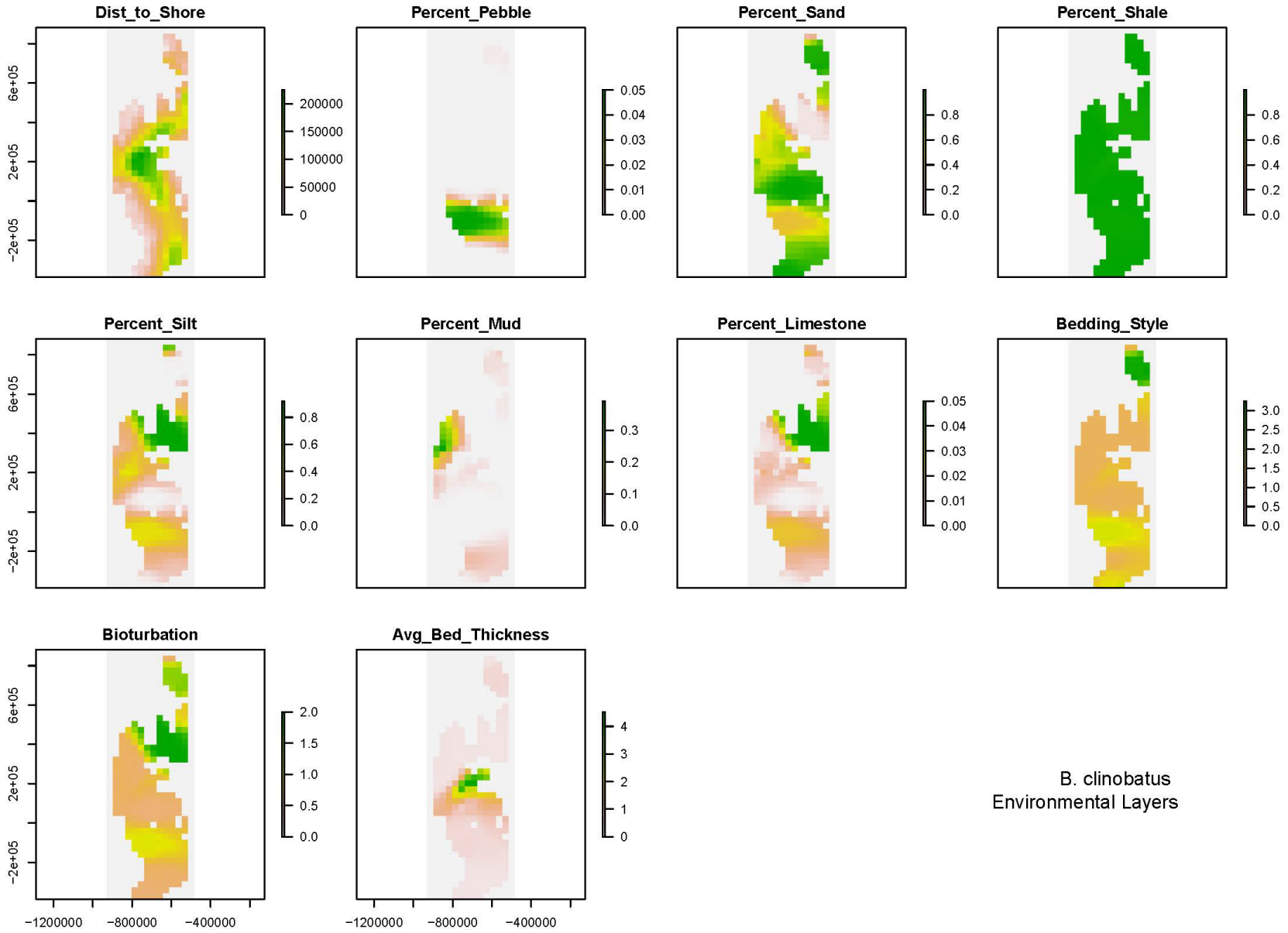
Stratigraphic Database:

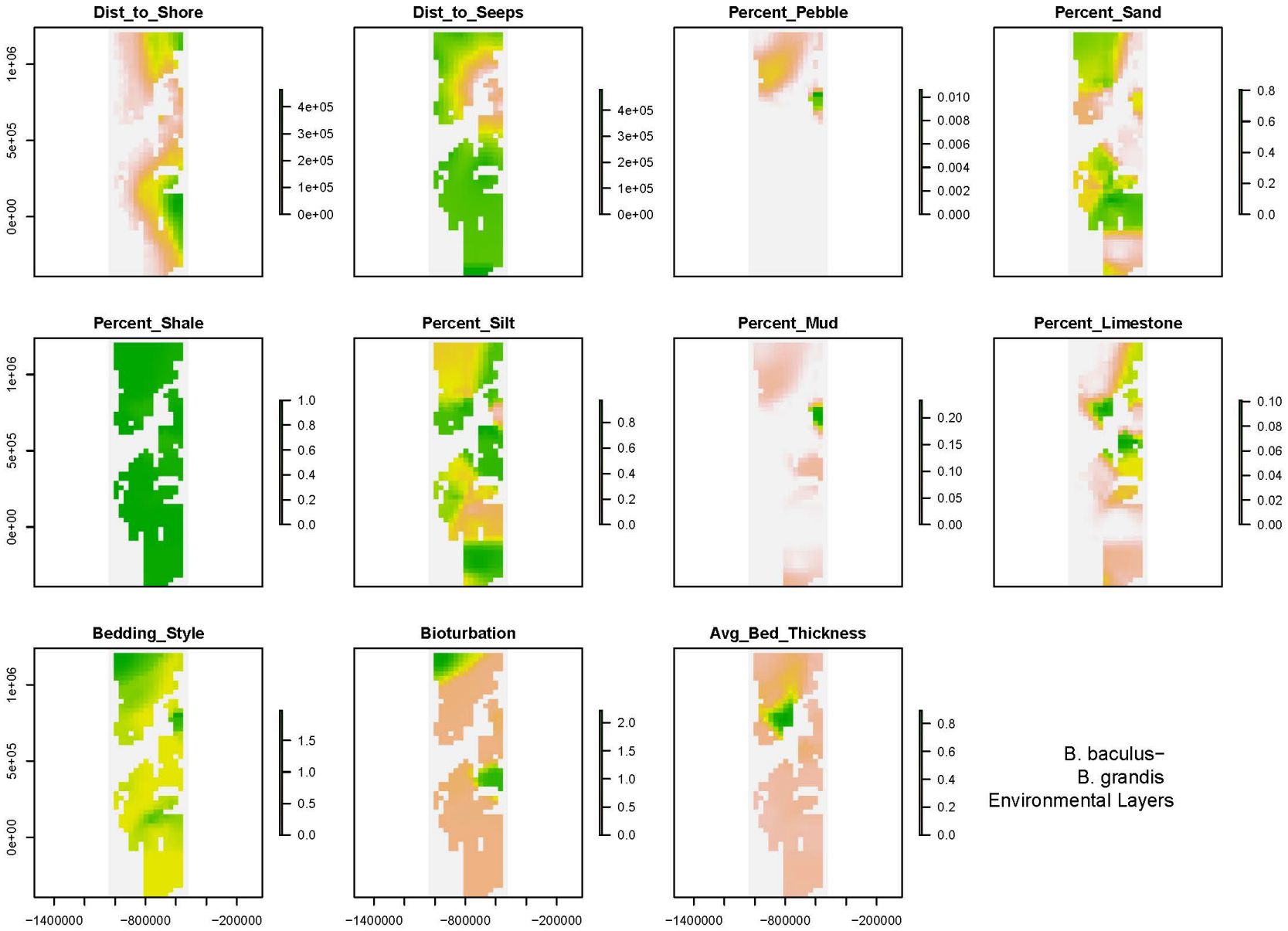
Raw stratigraphic data, collected using the methods described above, can be found in the separate excel table (Table S28).

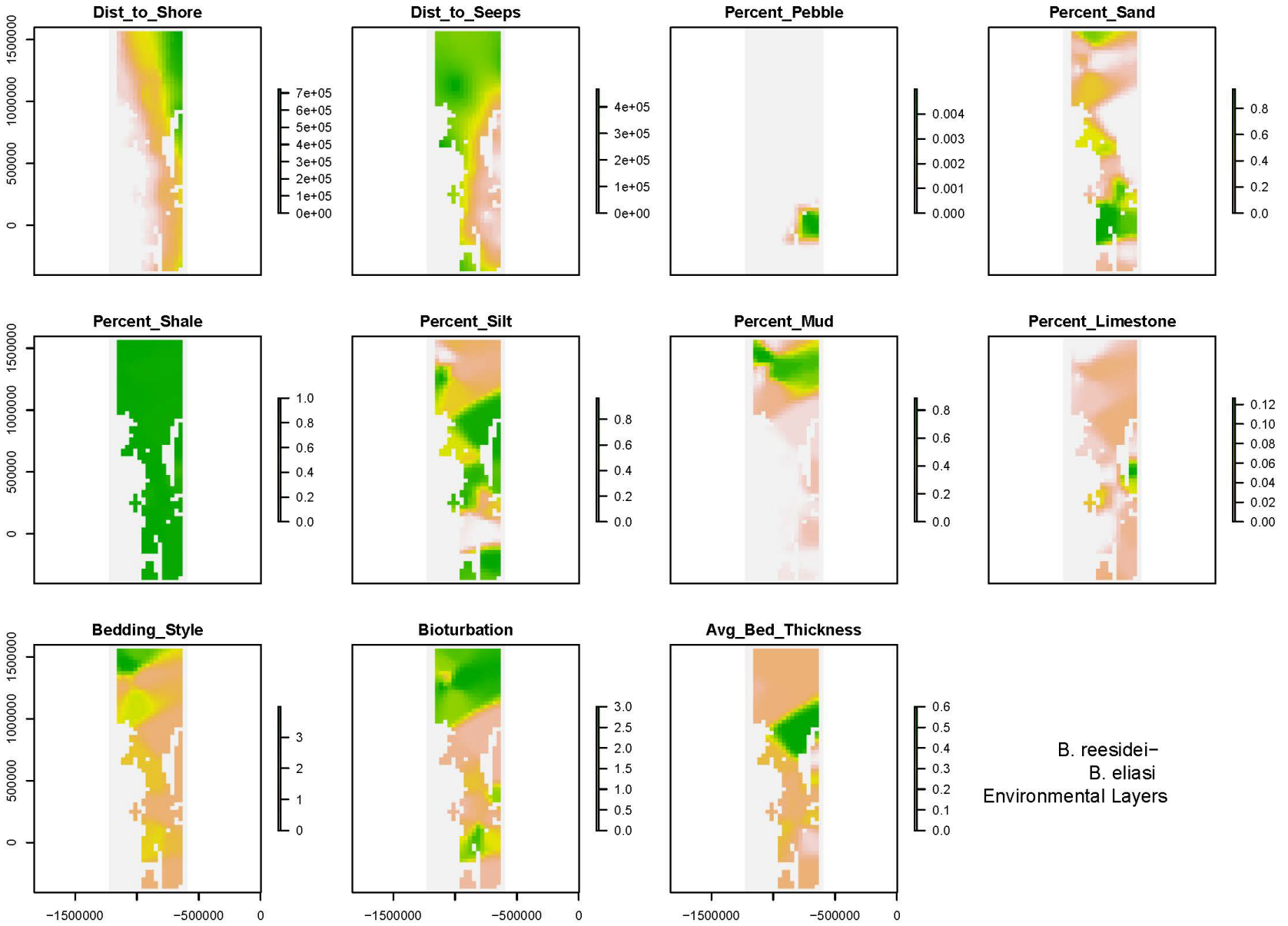
## Results

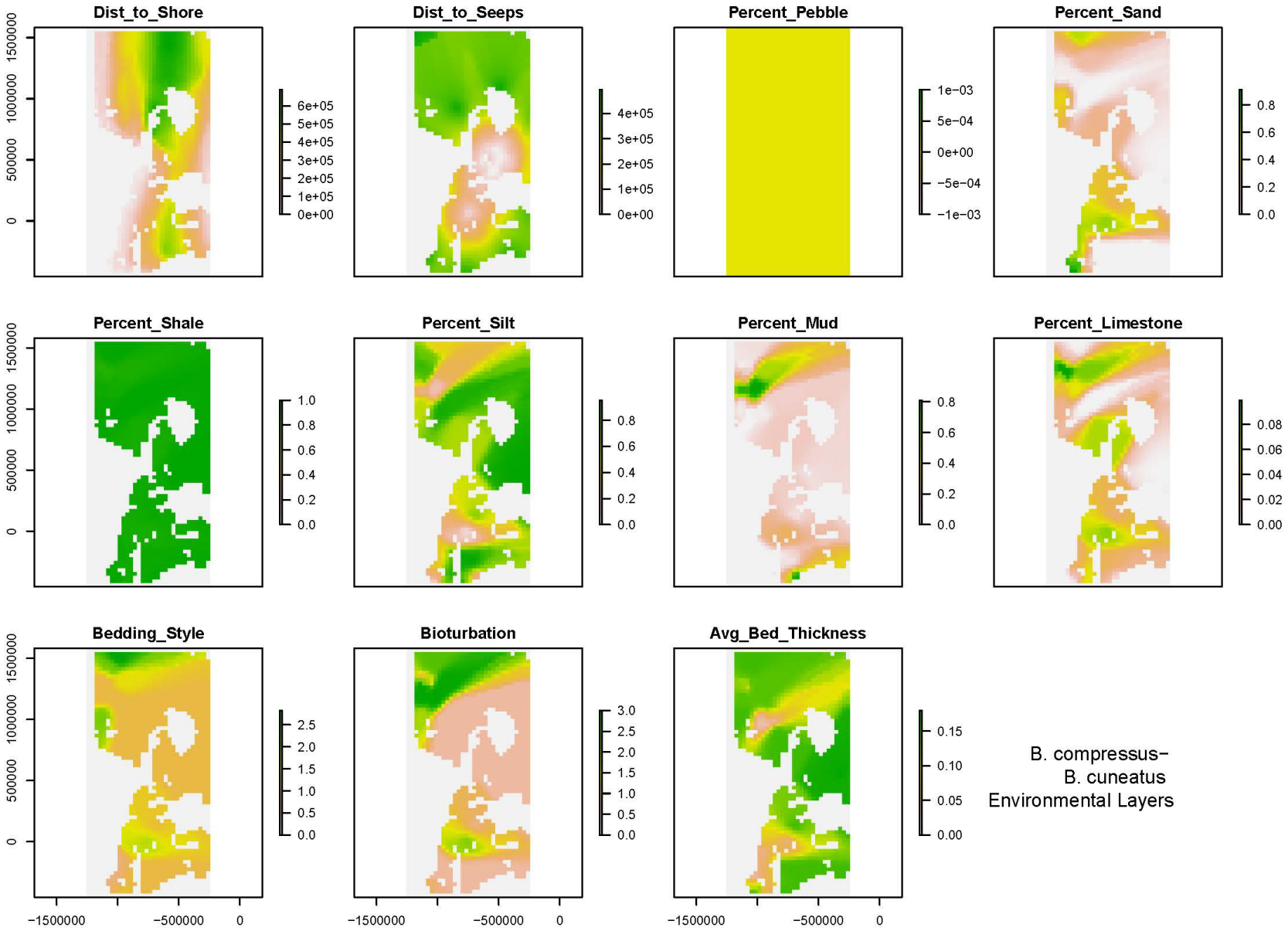
**Figure S4.** Raster map layers for each interval.

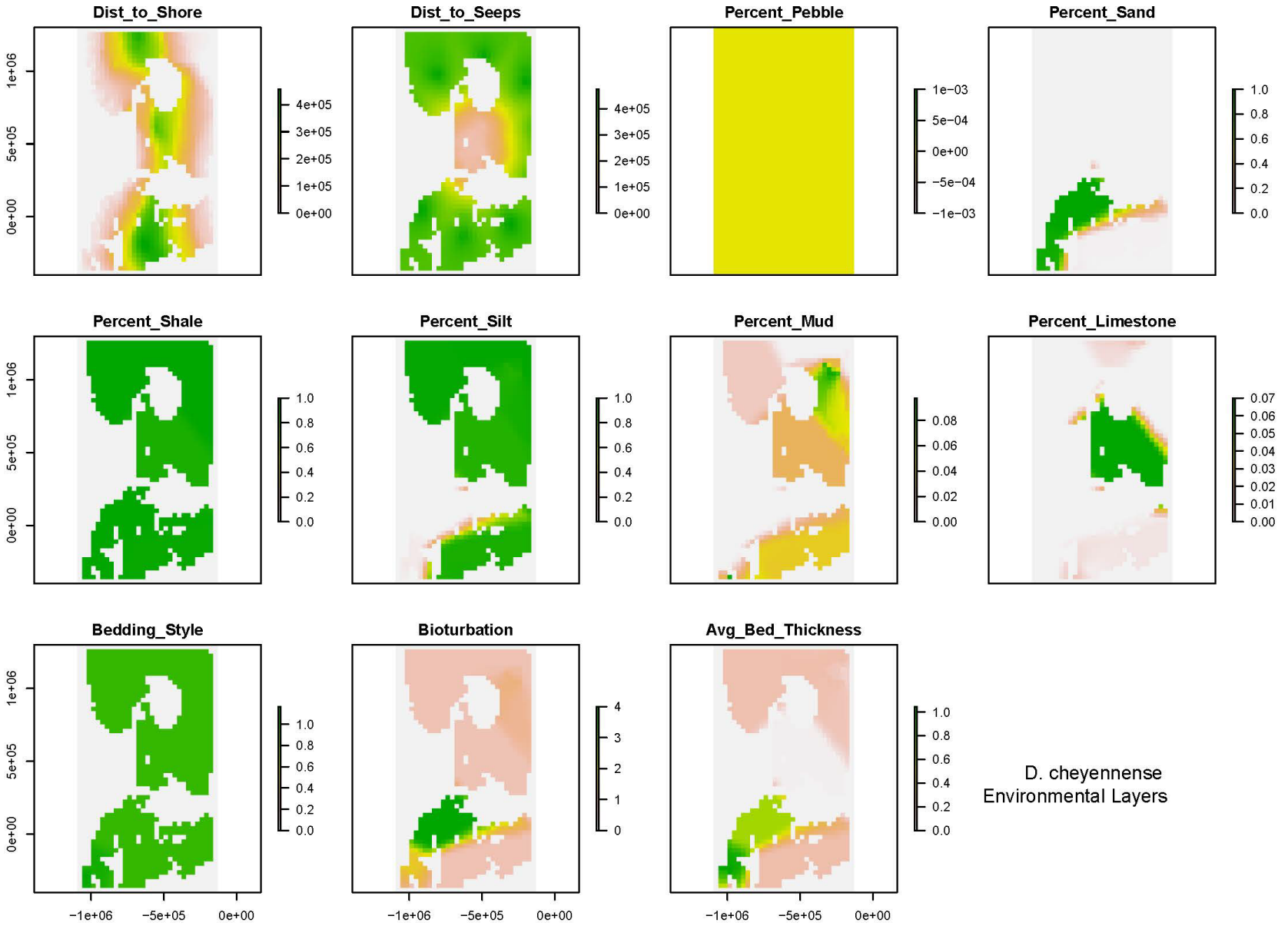



Dist_to_Shore
Percent_Pebble
Percent_Sand
Percent_Shale
Percent_Silt
Percent_Mud
Percent_Limestone
Bedding_Style
Bioturbation
Avg_Bed_Thickness
6e+05
2e+05
−2e+05
−1200000
−800000
−400000
B. clinobatus
Environmental Layers

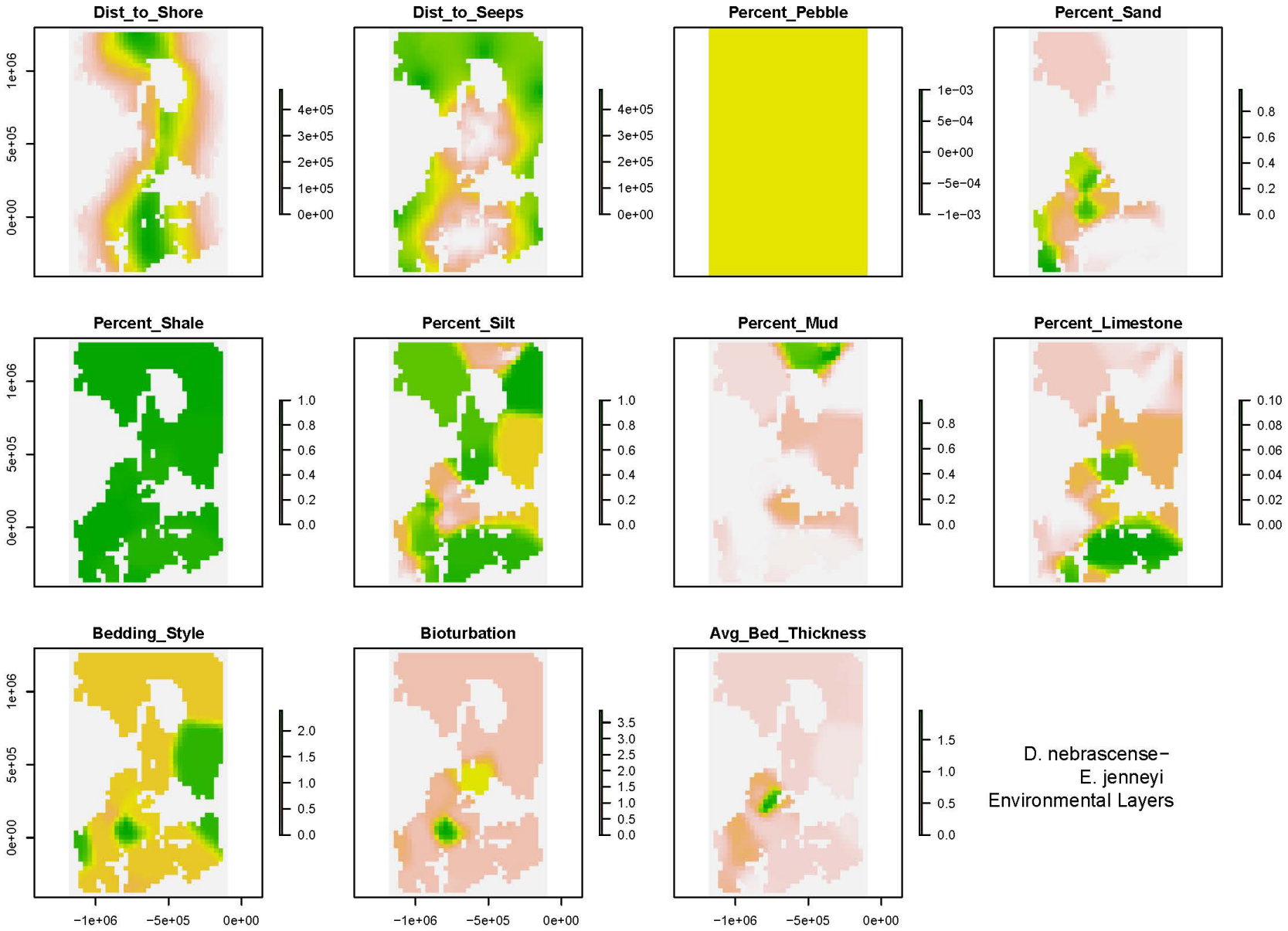

Dist_to_Shore
Dist_to_Seeps
Percent_Pebble
Percent_Sand
Percent_Shale
Percent_Silt
Percent_Mud
Percent_Limestone
Bedding_Style
Bioturbation
Avg_Bed_Thickness
0e+00
5e+05
1e+06
−1400000
−800000
−200000
B. baculus−
B. grandis
Environmental Layers

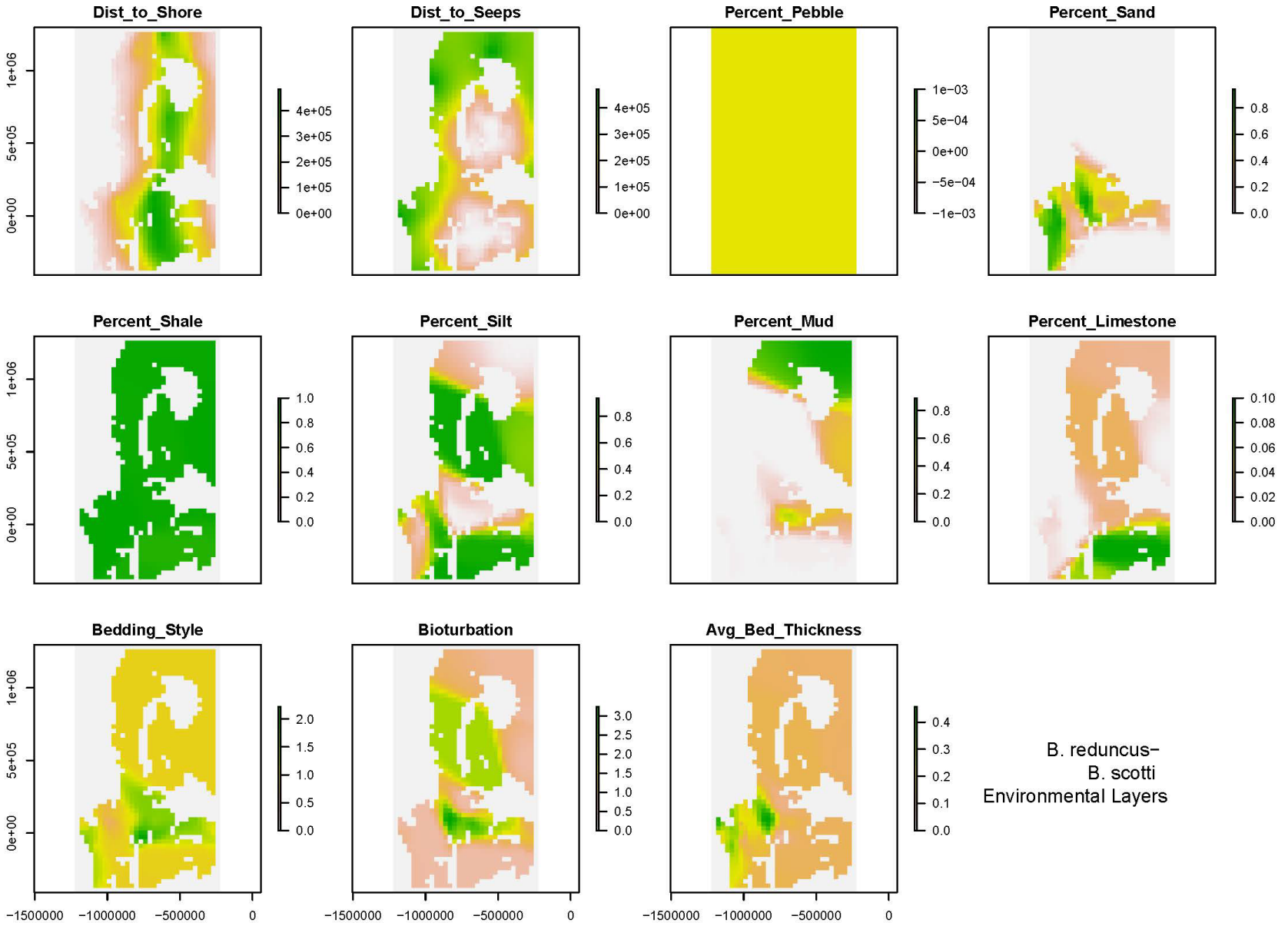

Dist_to_Shore
Dist_to_Seeps
Percent_Pebble
Percent_Sand
Percent_Shale
Percent_Silt
Percent_Mud
Percent_Limestone
Bedding_Style
Bioturbation
Avg_Bed_Thickness
0
500000
1000000
1500000
−1500000
−500000
0
B. reesidei−
B. eliasi
Environmental Layers

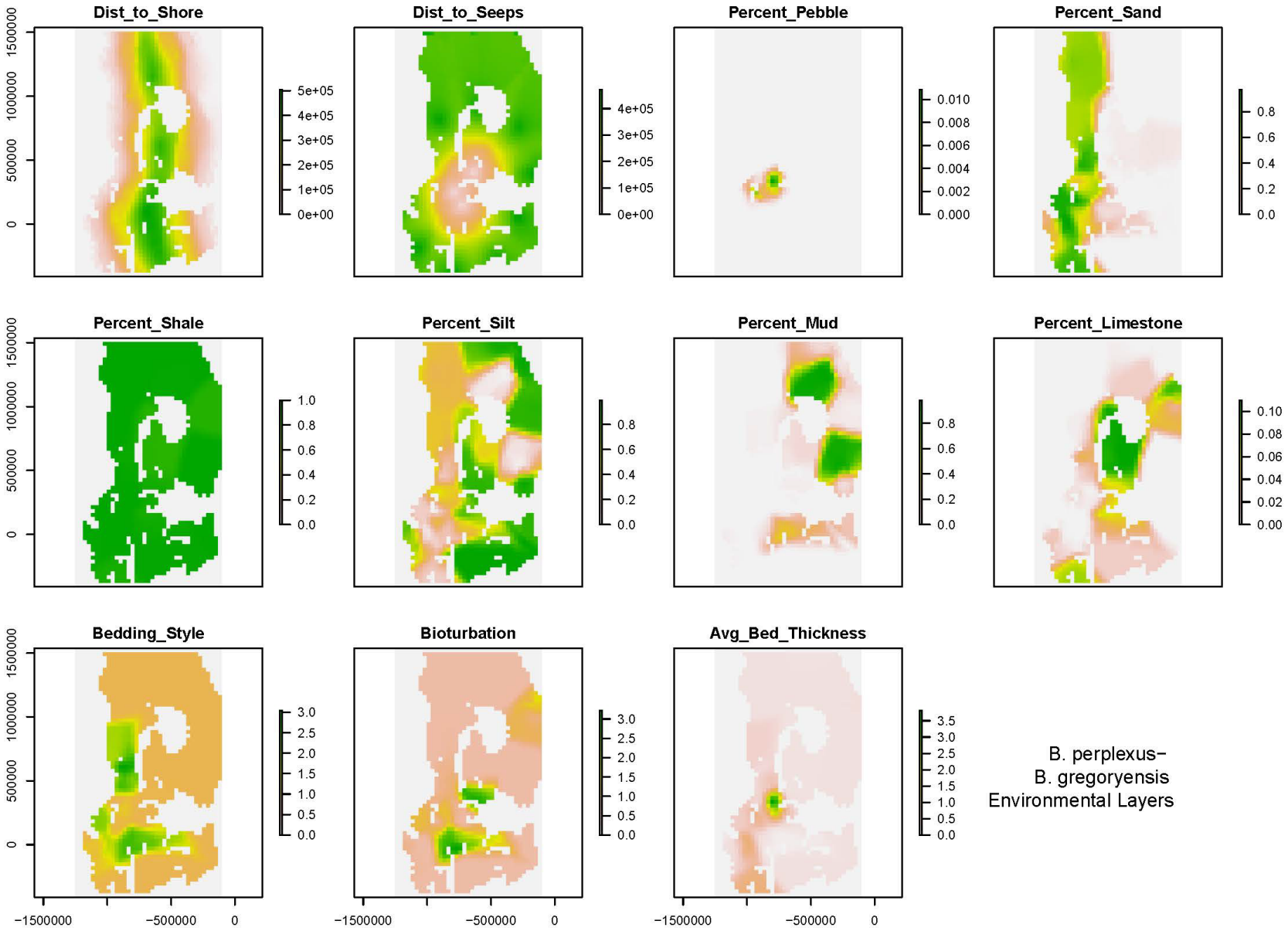

Dist_to_Shore
Dist_to_Seeps
Percent_Pebble
Percent_Sand
Percent_Shale
Percent_Silt
Percent_Mud
Percent_Limestone
Bedding_Style
Bioturbation
Avg_Bed_Thickness
B. compressus–
B. cuneatus
Environmental Layers

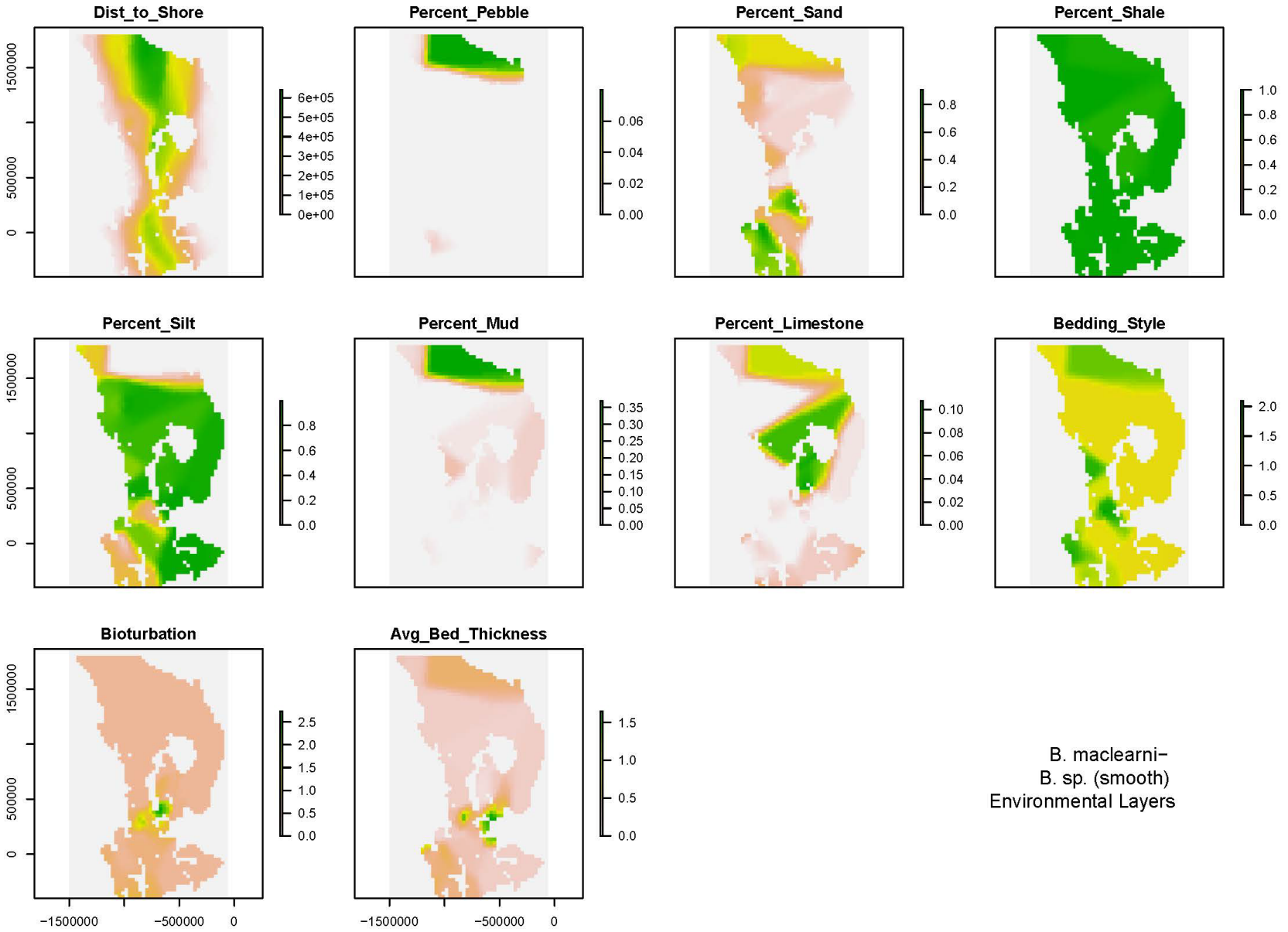



D. cheyennense
Environmental Layers

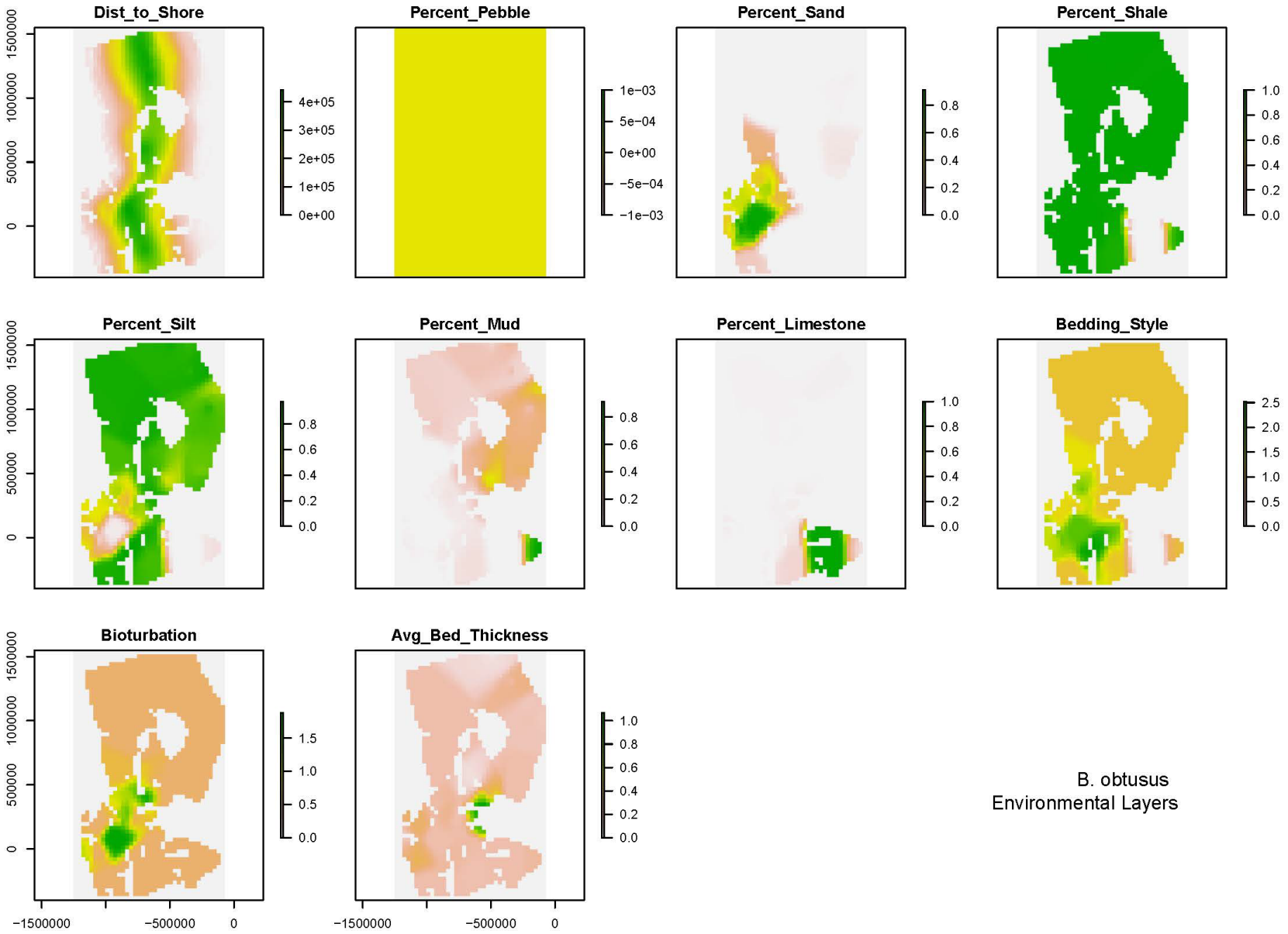

Dist_to_Shore
Dist_to_Seeps
Percent_Pebble
Percent_Sand
Percent_Shale
Percent_Silt
Percent_Mud
Percent_Limestone
Bedding_Style
Bioturbation
Avg_Bed_Thickness
D. nebrascense−
E. jenneyi
Environmental Layers


Dist_to_Shore
Dist_to_Seeps
Percent_Pebble
Percent_Sand
Percent_Shale
Percent_Silt
Percent_Mud
Percent_Limestone
Bedding_Style
Bioturbation
Avg_Bed_Thickness
0e+00
5e+05
1e+06
−1500000
−1000000
−500000
0
B. reduncus−
B. scotti
Environmental Layers

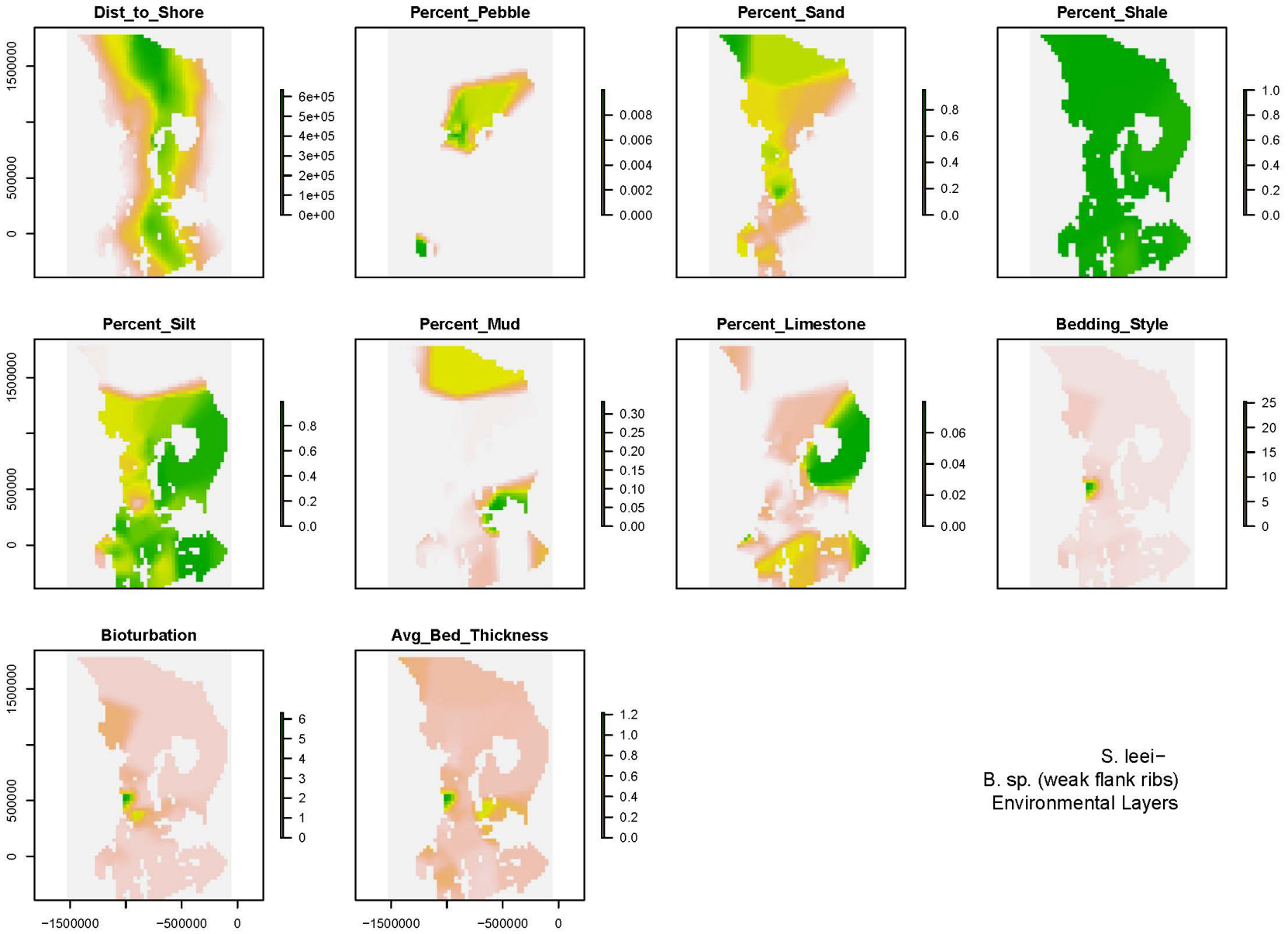

Dist_to_Shore
Dist_to_Seeps
Percent_Pebble
Percent_Sand
Percent_Shale
Percent_Silt
Percent_Mud
Percent_Limestone
Bedding_Style
Bioturbation
Avg_Bed_Thickness
B. perplexus–
B. gregoryensis
Environmental Layers

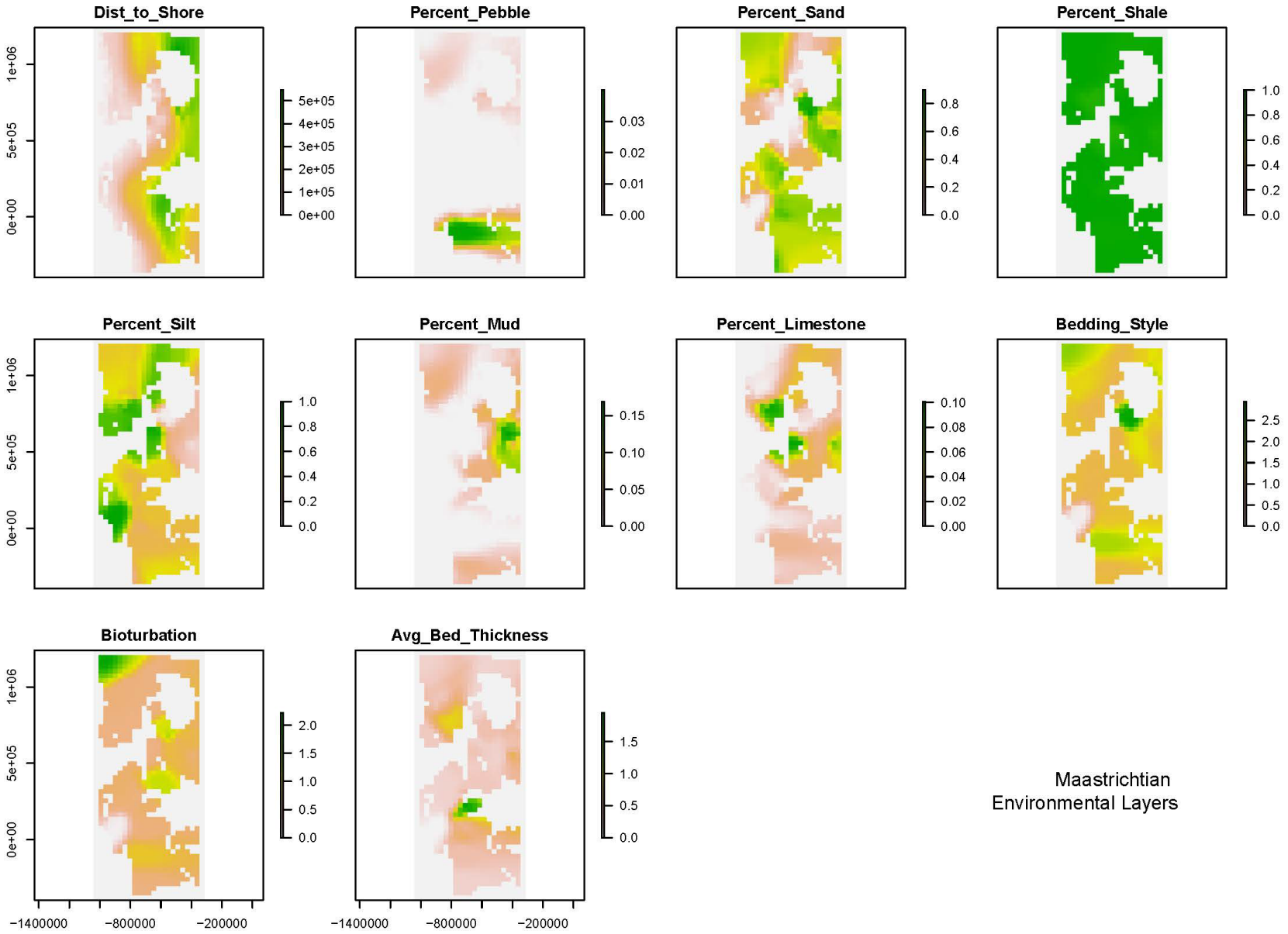


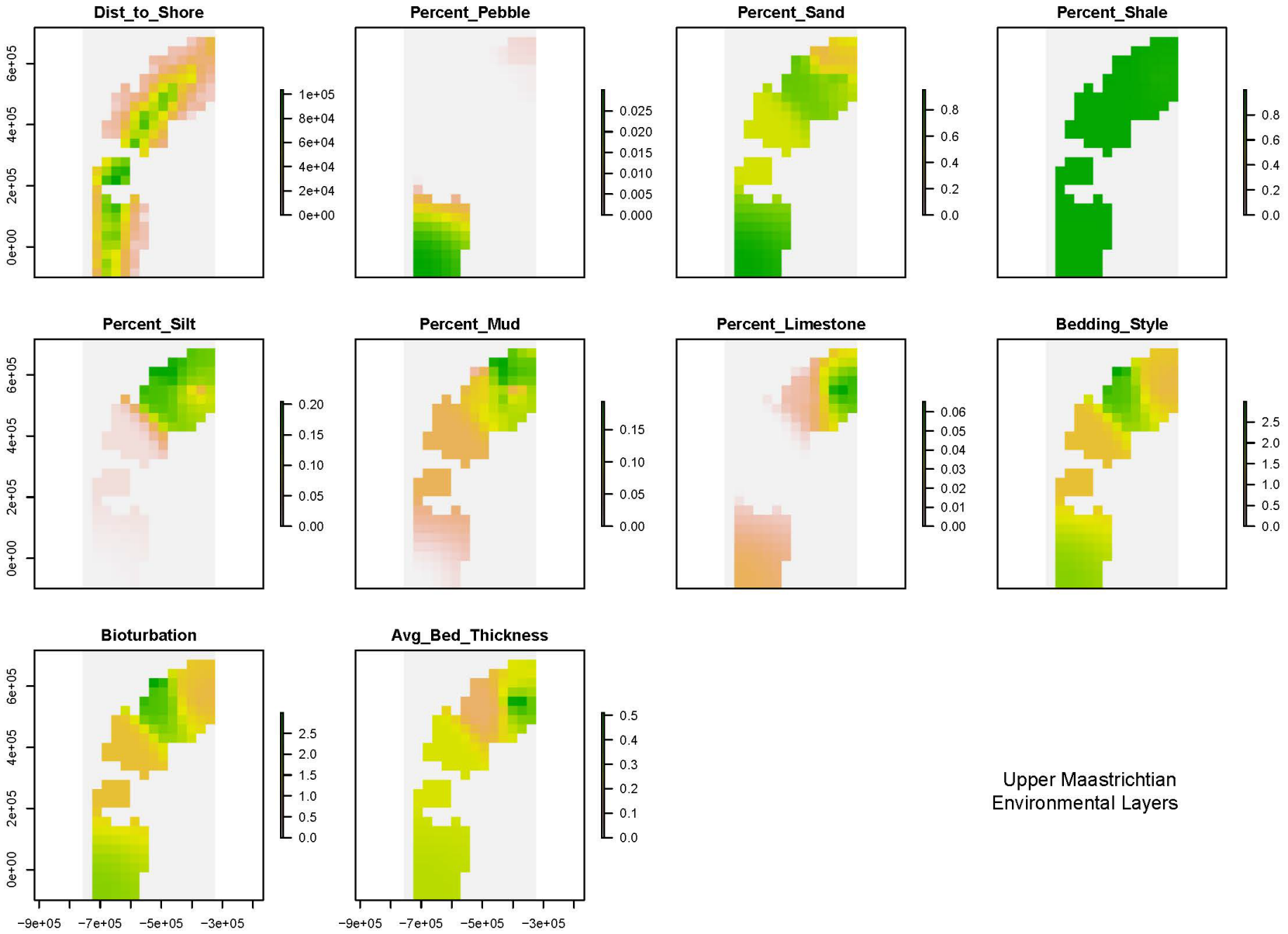

B. maclearni–
B. sp. (smooth)
Environmental Layers



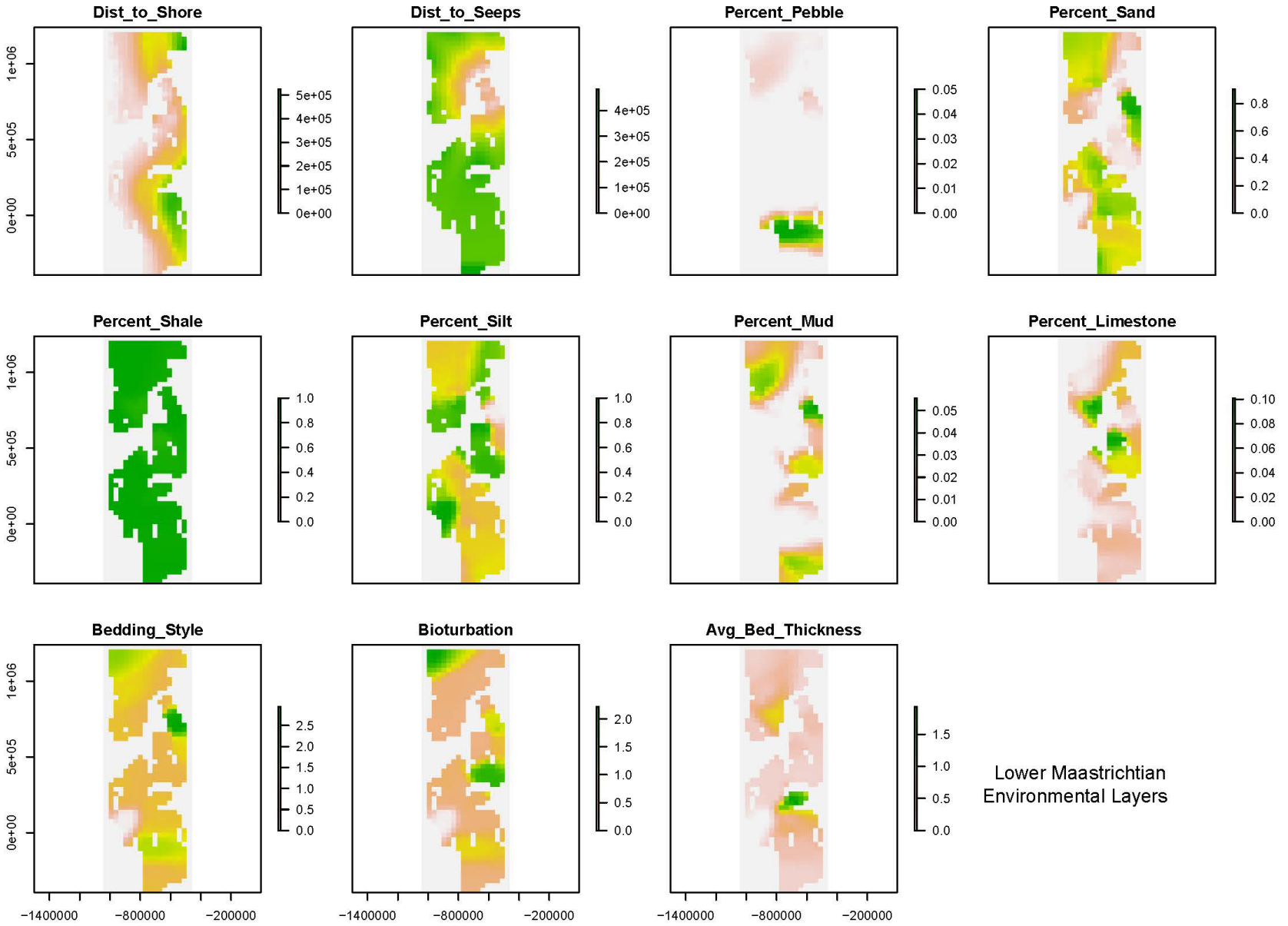

B. obtusus
Environmental Layers

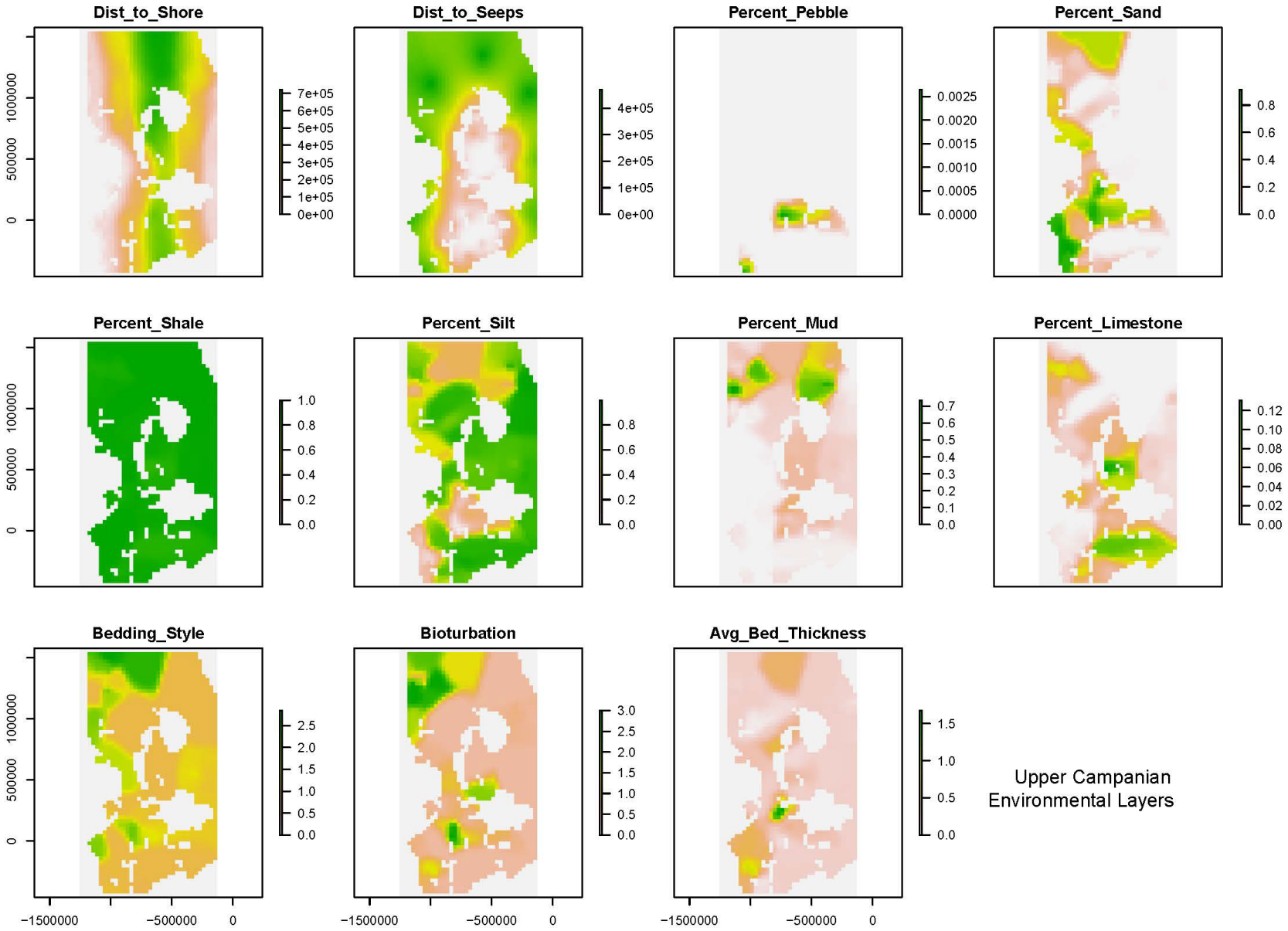

Dist_to_Shore
Percent_Pebble
Percent_Sand
Percent_Shale
Percent_Silt
Percent_Mud
Percent_Limestone
Bedding_Style
Bioturbation
Avg_Bed_Thickness
1500000
500000
0
−1500000
−500000
0
S. leei−
B. sp. (weak flank ribs)
Environmental Layers

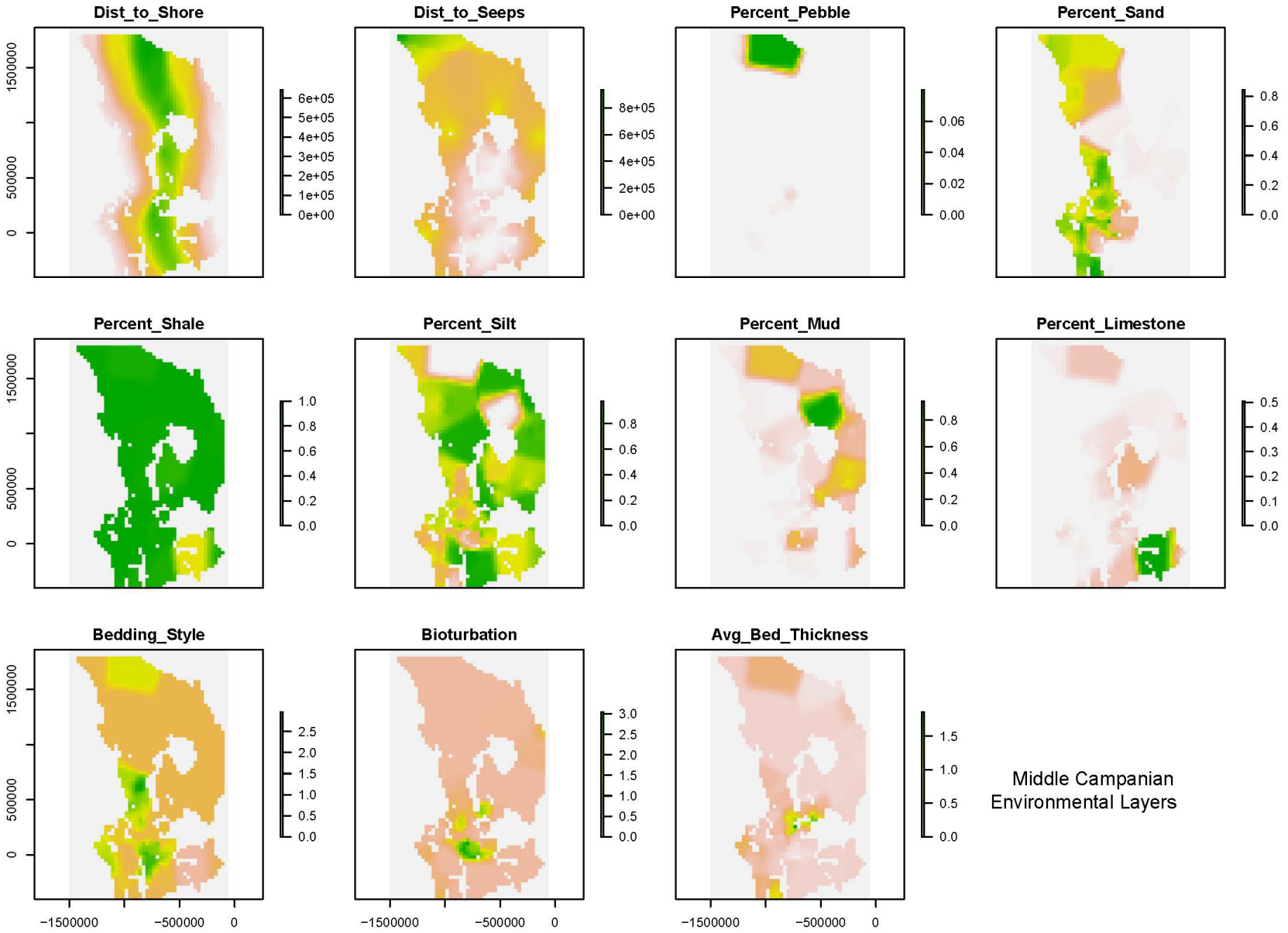

Dist_to_Shore
Percent_Pebble
Percent_Sand
Percent_Shale
Percent_Silt
Percent_Mud
Percent_Limestone
Bedding_Style
Bioturbation
Avg_Bed_Thickness
0e+00
5e+05
1e+06
−1400000
−800000
−200000
Maastrichtian
Environmental Layers

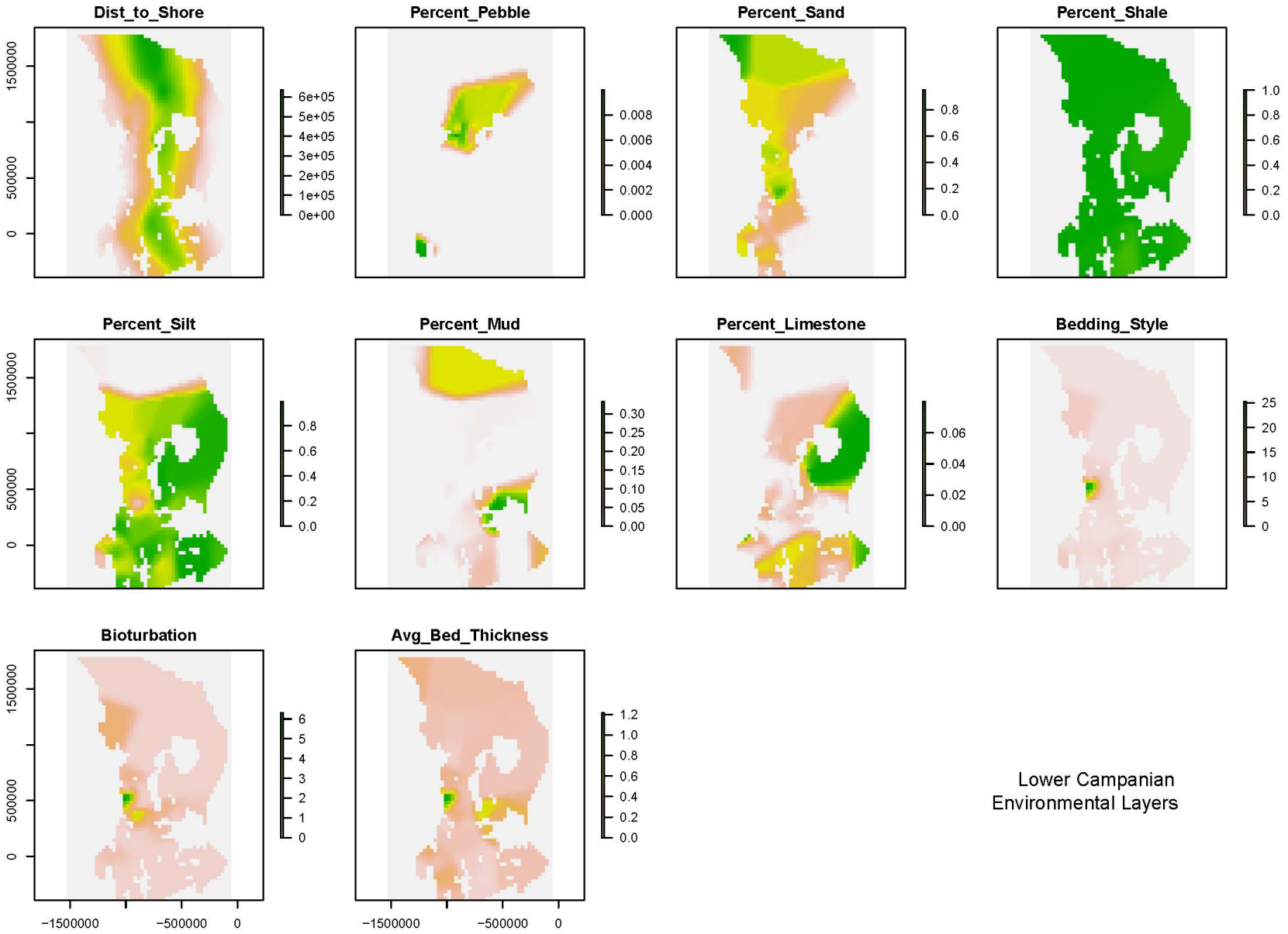



Upper Maastrichtian
Environmental Layers


Dist_to_Shore
Dist_to_Seeps
Percent_Pebble
Percent_Sand
Percent_Shale
Percent_Silt
Percent_Mud
Percent_Limestone
Bedding_Style
Bioturbation
Avg_Bed_Thickness
0e+00
5e+05
1e+06
−1400000
−800000
−200000
Lower Maastrichtian
Environmental Layers

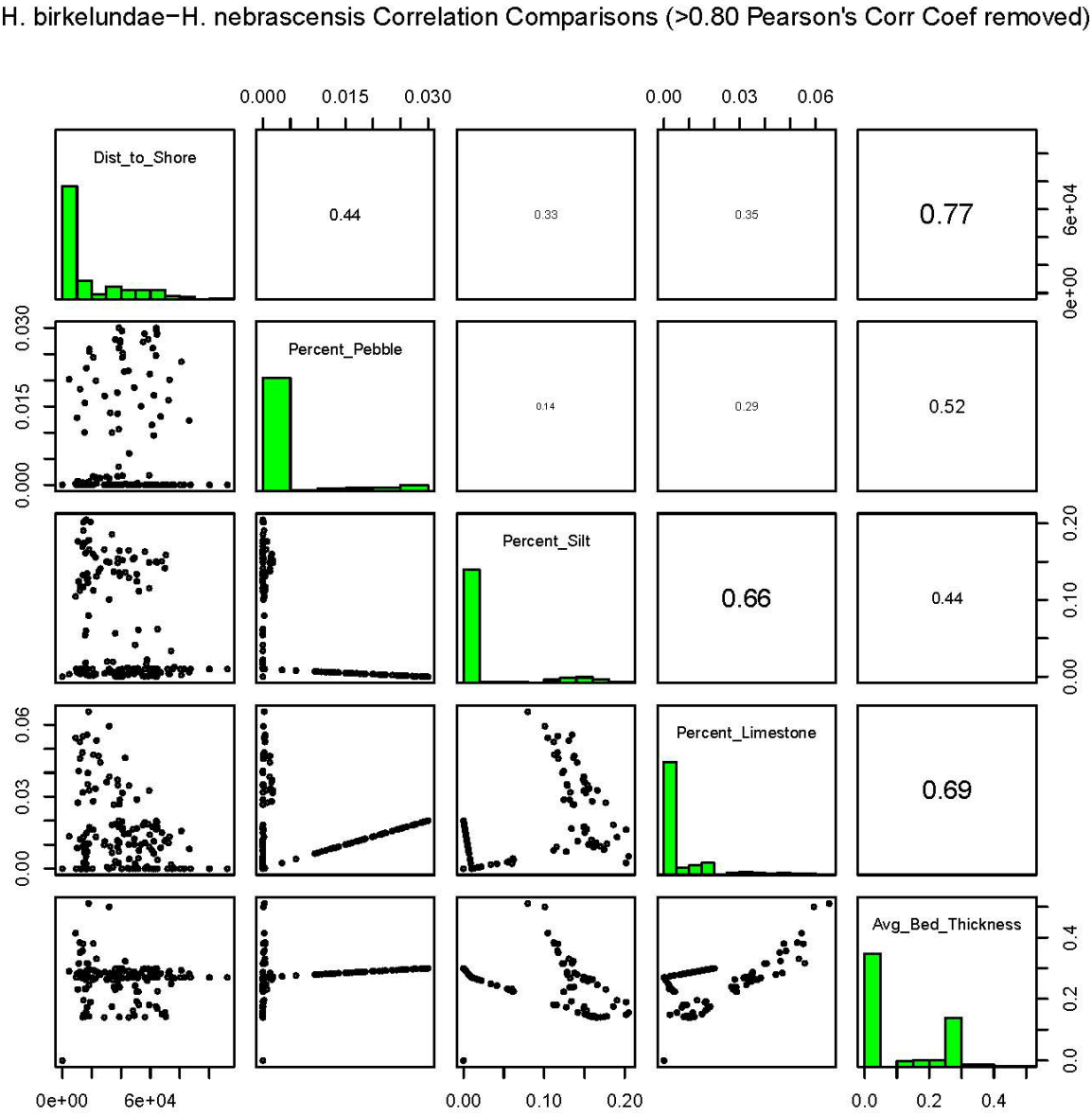

Dist_to_Shore
Dist_to_Seeps
Percent_Pebble
Percent_Sand
Percent_Shale
Percent_Silt
Percent_Mud
Percent_Limestone
Bedding_Style
Bioturbation
Avg_Bed_Thickness
0
500000
1000000
−1500000
−500000
0
Upper Campanian
Environmental Layers


Dist_to_Shore
Dist_to_Seeps
Percent_Pebble
Percent_Sand
Percent_Shale
Percent_Silt
Percent_Mud
Percent_Limestone
Bedding_Style
Bioturbation
Avg_Bed_Thickness
1500000
500000
0
−1500000
−500000
0
Middle Campanian
Environmental Layers

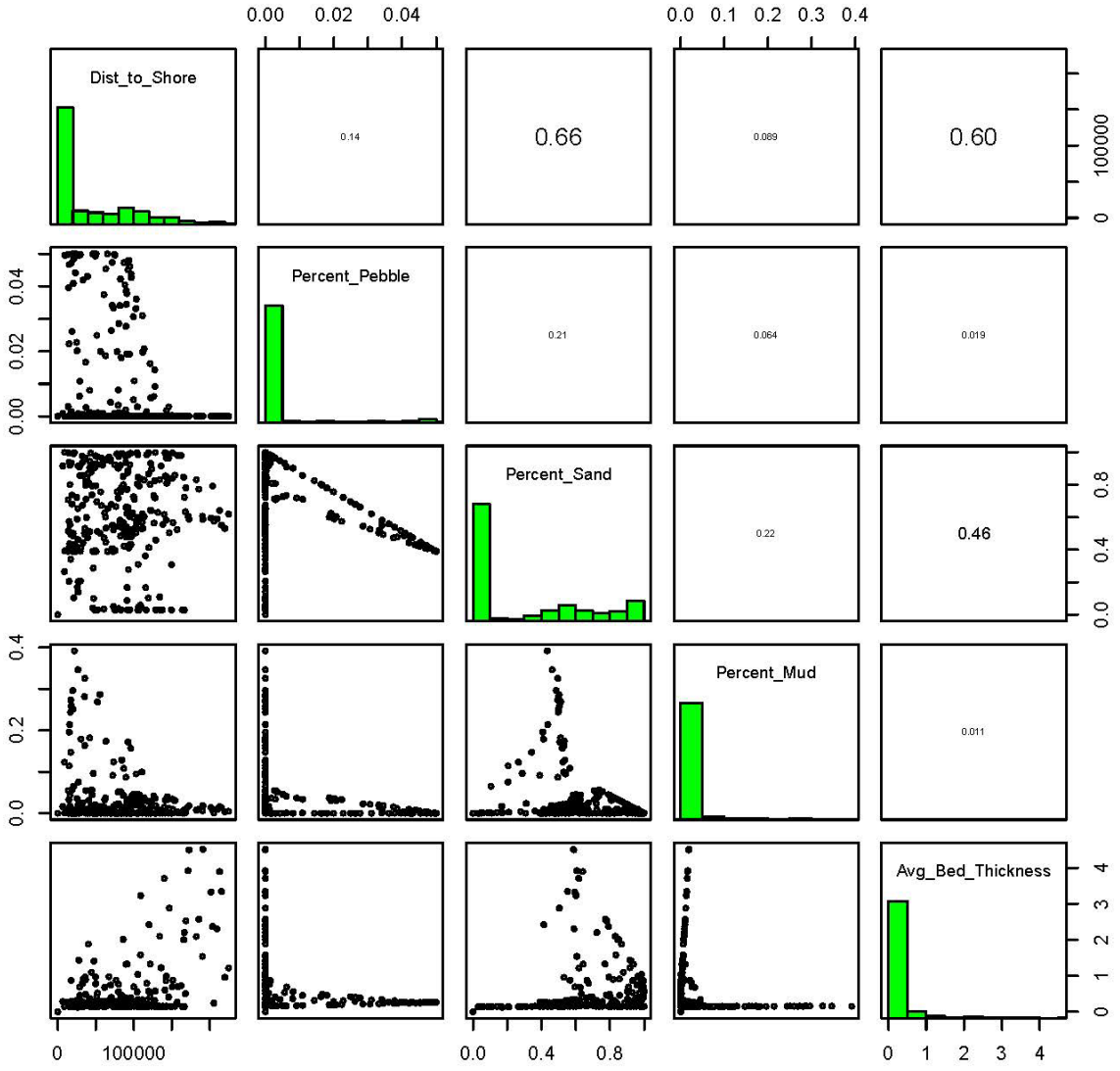



Lower Campanian
Environmental Layers

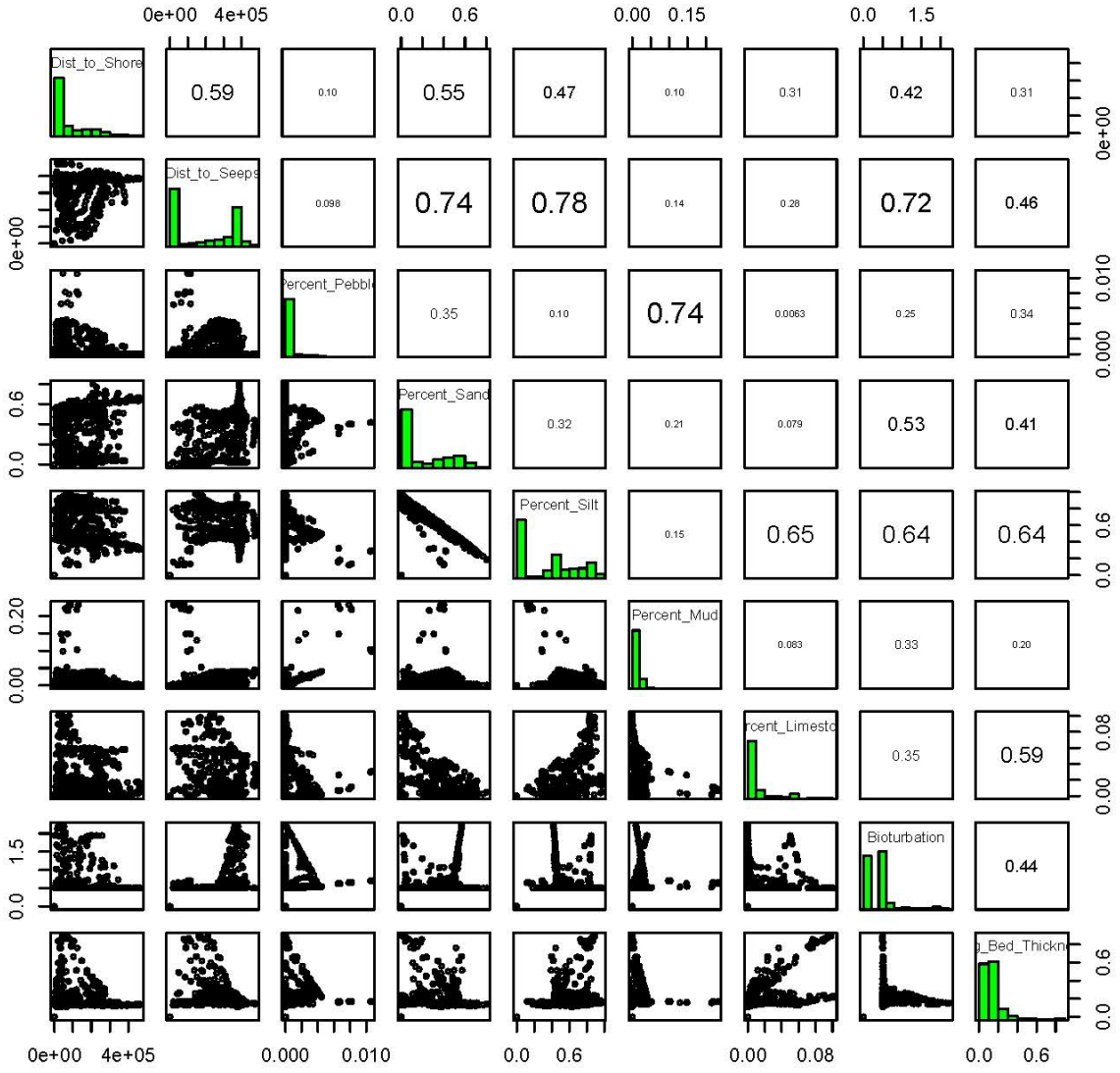
**Figure S5.** Correlation charts for environmental raster layers with variables removed to ensure all Pearson's Correlation Coefficients were less than 0.80.

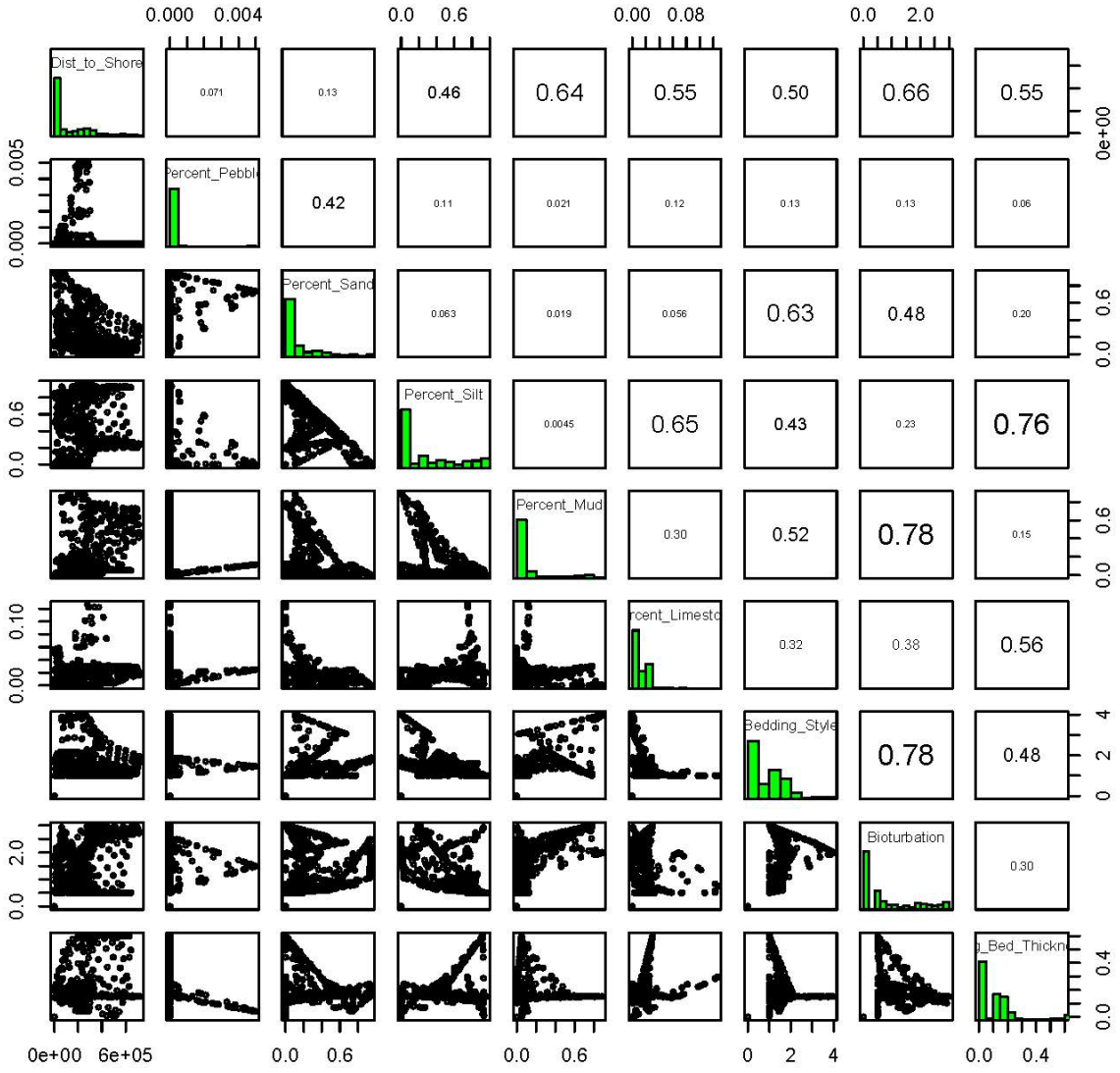


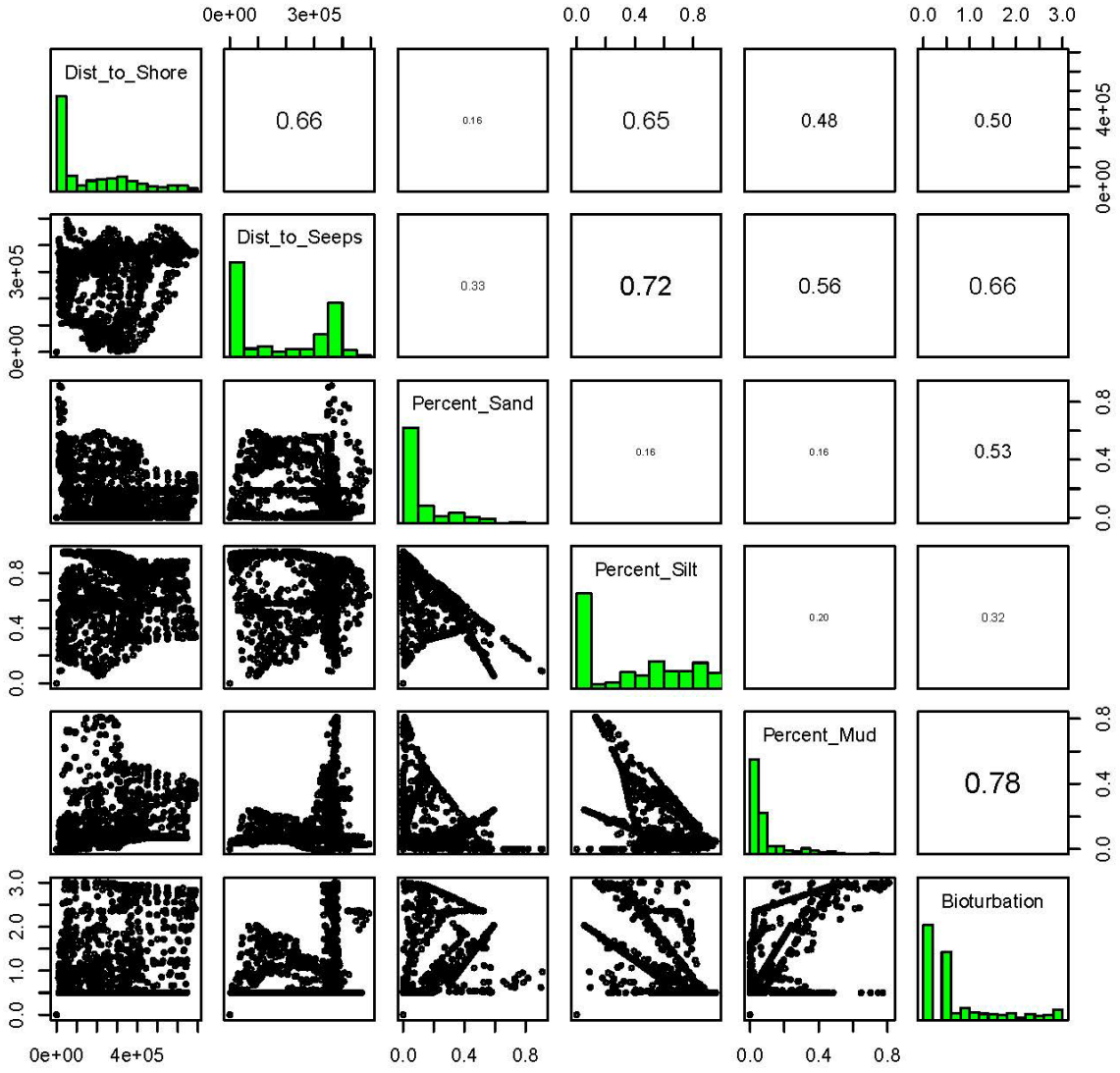
B. clinobatus Correlation Comparisons (>0.80 Pearson's Corr Coef removed)

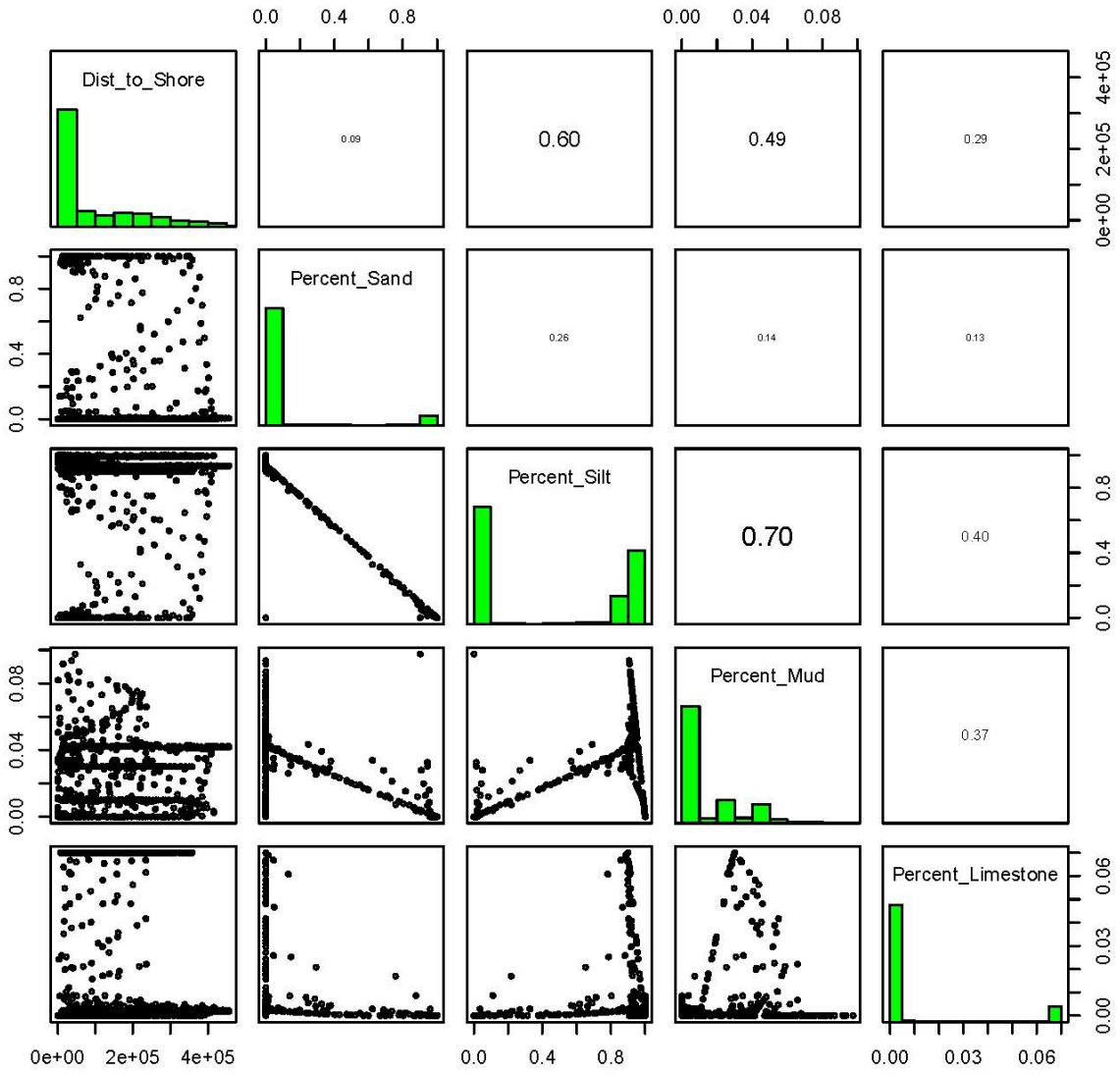


B. baculus−B. grandis Correlation Comparisons (>0.80 Pearson's Corr Coef removed)

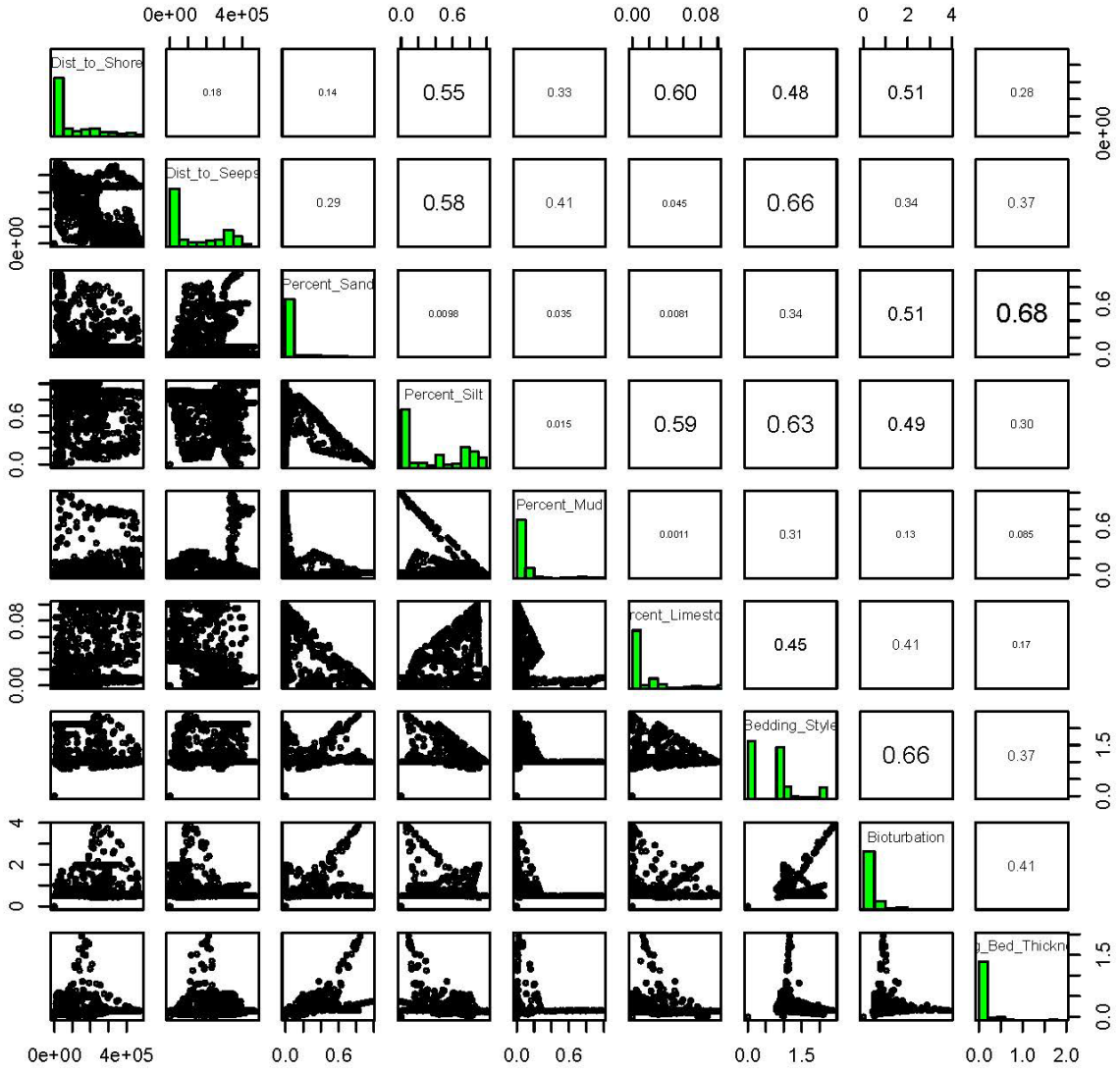


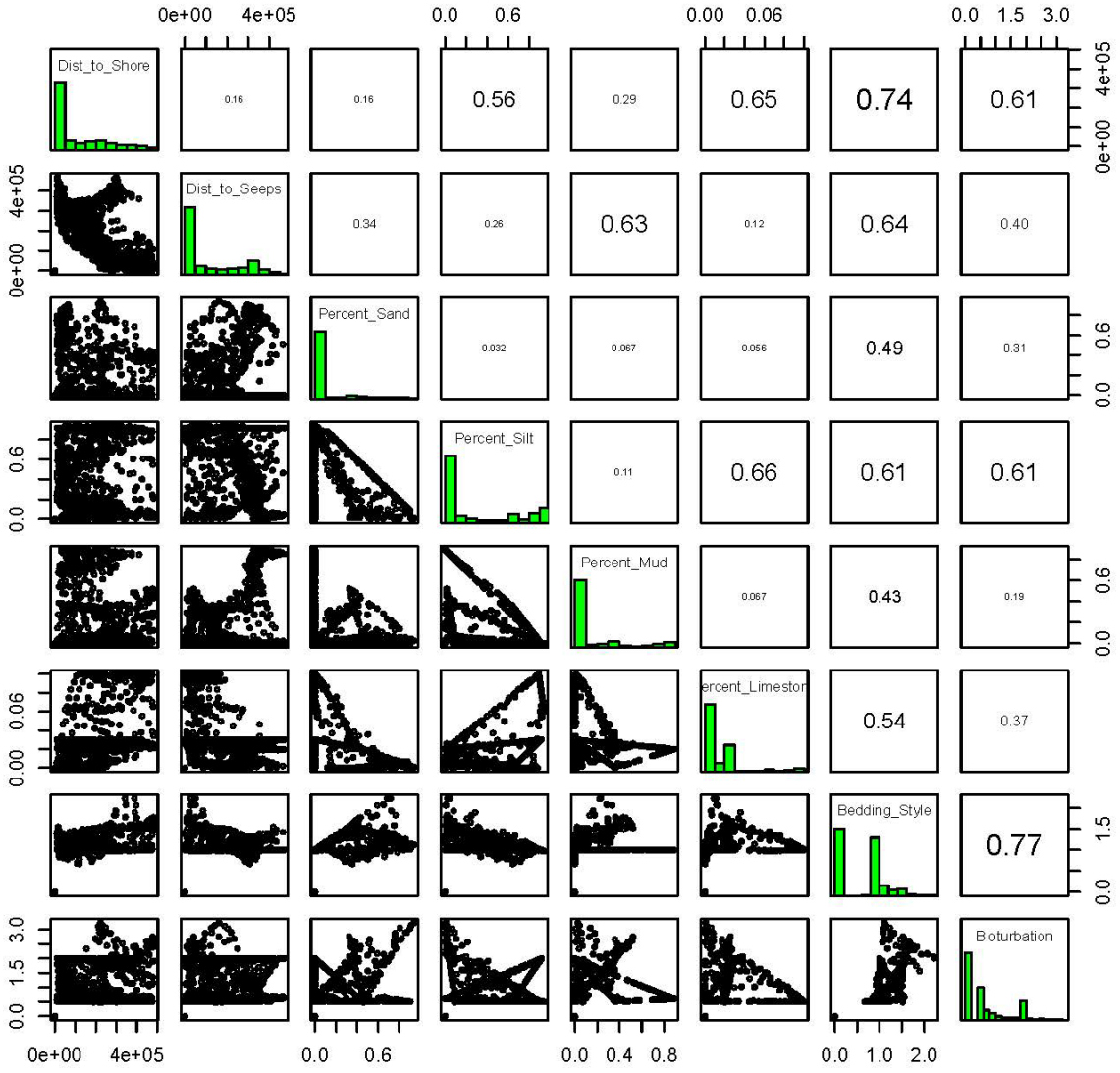

B. reesidei−B. eliasi Correlation Comparisons (>0.80 Pearson's Corr Coef removed)
Dist_to_Shore
Percent_Pebbl
Percent_Sand
Percent_Silt
Percent_Mud
rcent_Limesto
Bedding_Style
Bioturbation
g_Bed_Thickn
0.071
0.13
0.46
0.64
0.55
0.50
0.66
0.55
0.42
0.11
0.021
0.12
0.13
0.13
0.06
0.063
0.019
0.056
0.63
0.48
0.20
0.0045
0.65
0.43
0.23
0.76
0.30
0.52
0.78
0.15
0.32
0.38
0.56
0.78
0.48
0.30

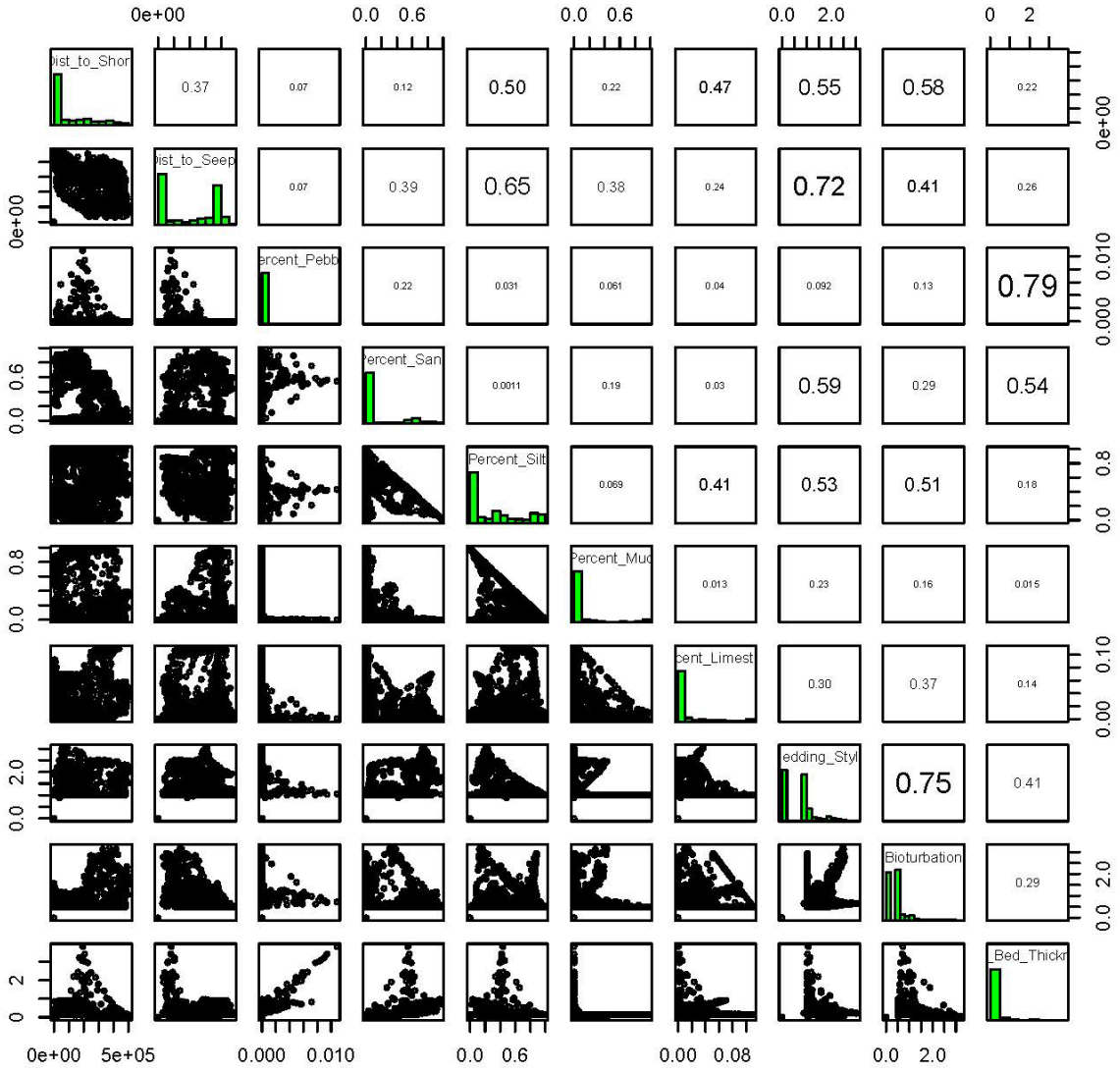

B. compressus−B. cuneatus Correlation Comparisons (>0.80 Pearson's Corr Coef removed)
Dist_to_Shore
Dist_to_Seeps
Percent_Sand
Percent_Silt
Percent_Mud
Bioturbation
0.66
0.16
0.65
0.48
0.50
0.33
0.72
0.56
0.66
0.16
0.16
0.53
0.20
0.32
0.78
0e+00
3e+05
4e+05
0.0
0.4
0.8
1.0
2.0
3.0

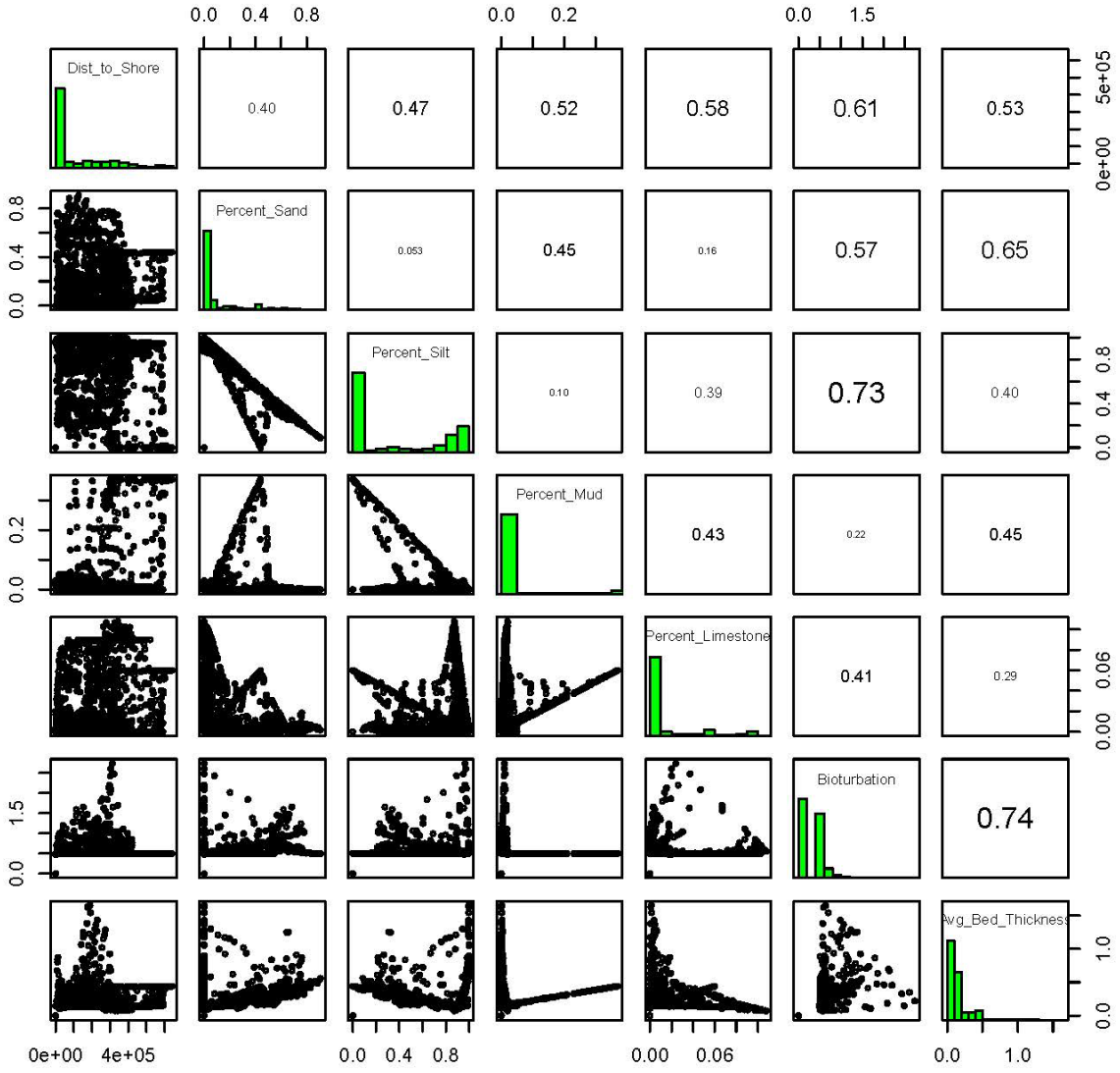

D. cheyennense Correlation Comparisons (>0.80 Pearson's Corr Coef removed)
Dist_to_Shore
Percent_Sand
Percent_Silt
Percent_Mud
Percent_Limestone
0.09
0.60
0.49
0.29
0.26
0.14
0.13
0.70
0.40
0.37
0.0
0.4
0.8
0.00
0.04
0.08
0e+00
2e+05
4e+05
0.03
0.06

D. nebrascense−E. jenneyi Correlation Comparisons (>0.80 Pearson's Corr Coef removed)

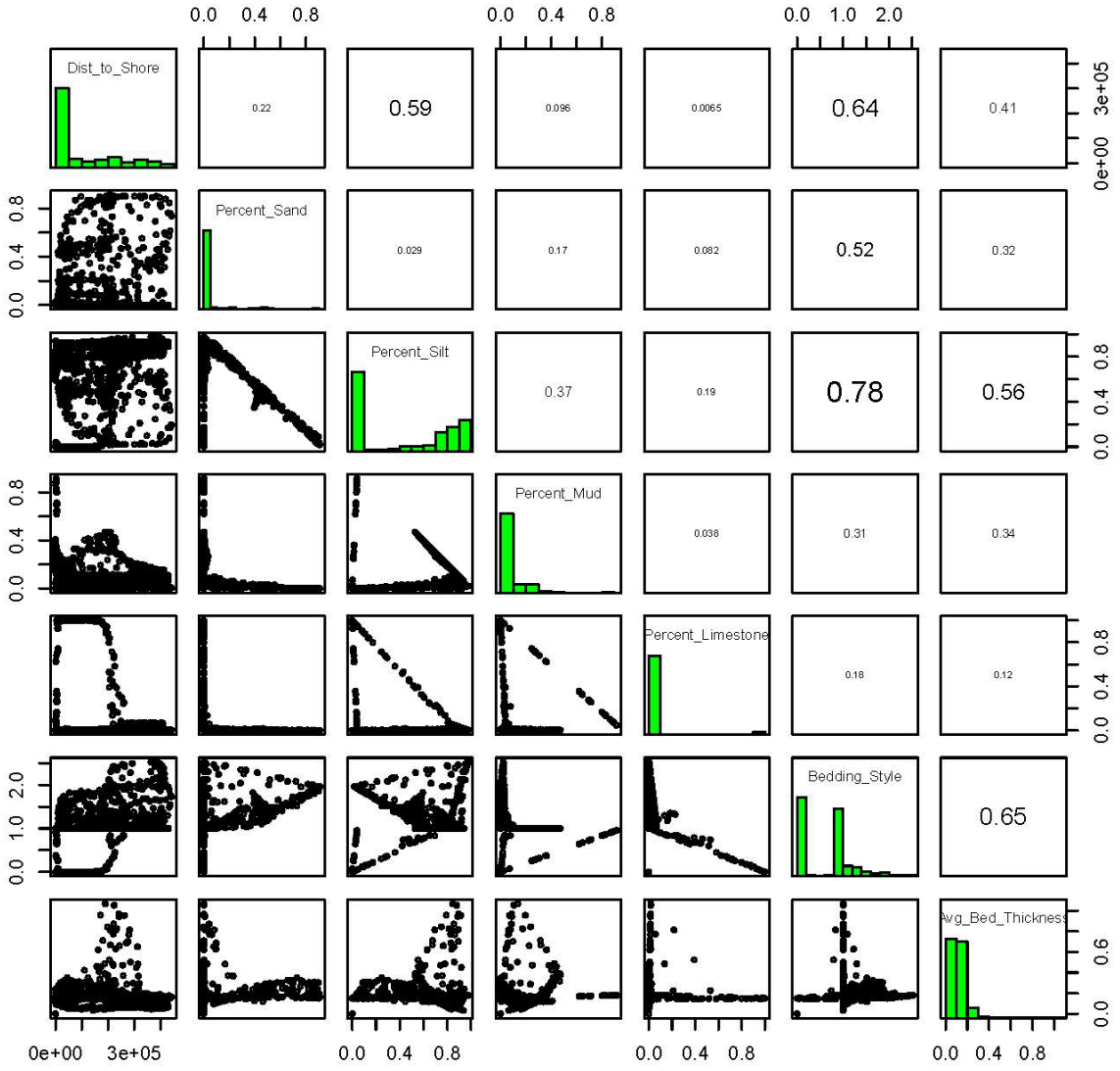


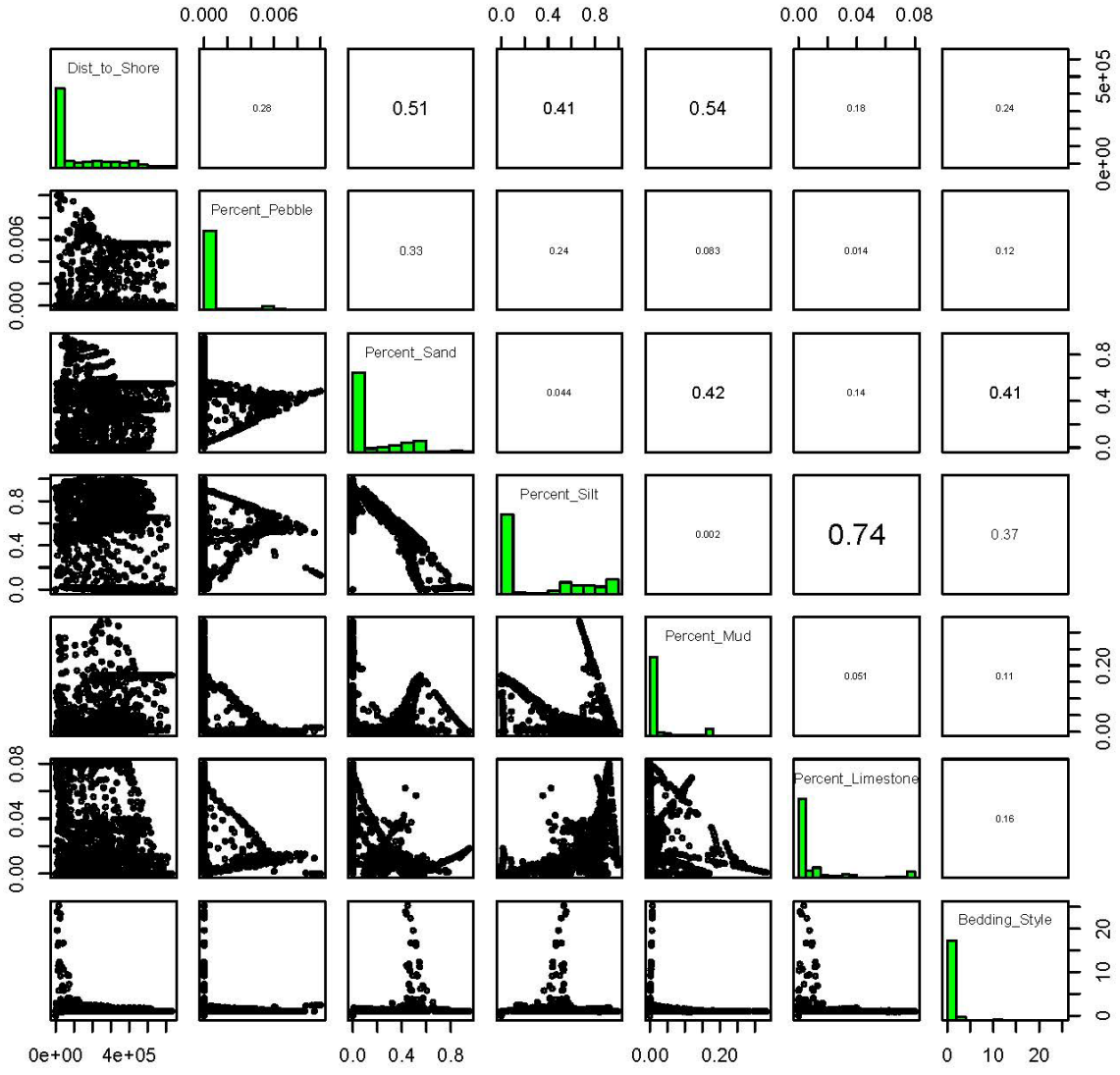

B. reduncus−B. scotti Correlation Comparisons (>0.80 Pearson's Corr Coef removed)
Dist_to_Shore
Dist_to_Seeps
Percent_Sand
Percent_Silt
Percent_Mud
ercent_Limestor
Bedding_Style
Bioturbation
0.16
0.16
0.56
0.29
0.65
0.74
0.61
0.34
0.26
0.63
0.12
0.64
0.40
0.032
0.067
0.056
0.49
0.31
0.11
0.66
0.61
0.61
0.067
0.43
0.19
0.54
0.37
0.77
0e+00
4e+05
0.0
0.6
0.00
0.06
0.0
1.5
3.0
0.4
0.8
1.0
2.0

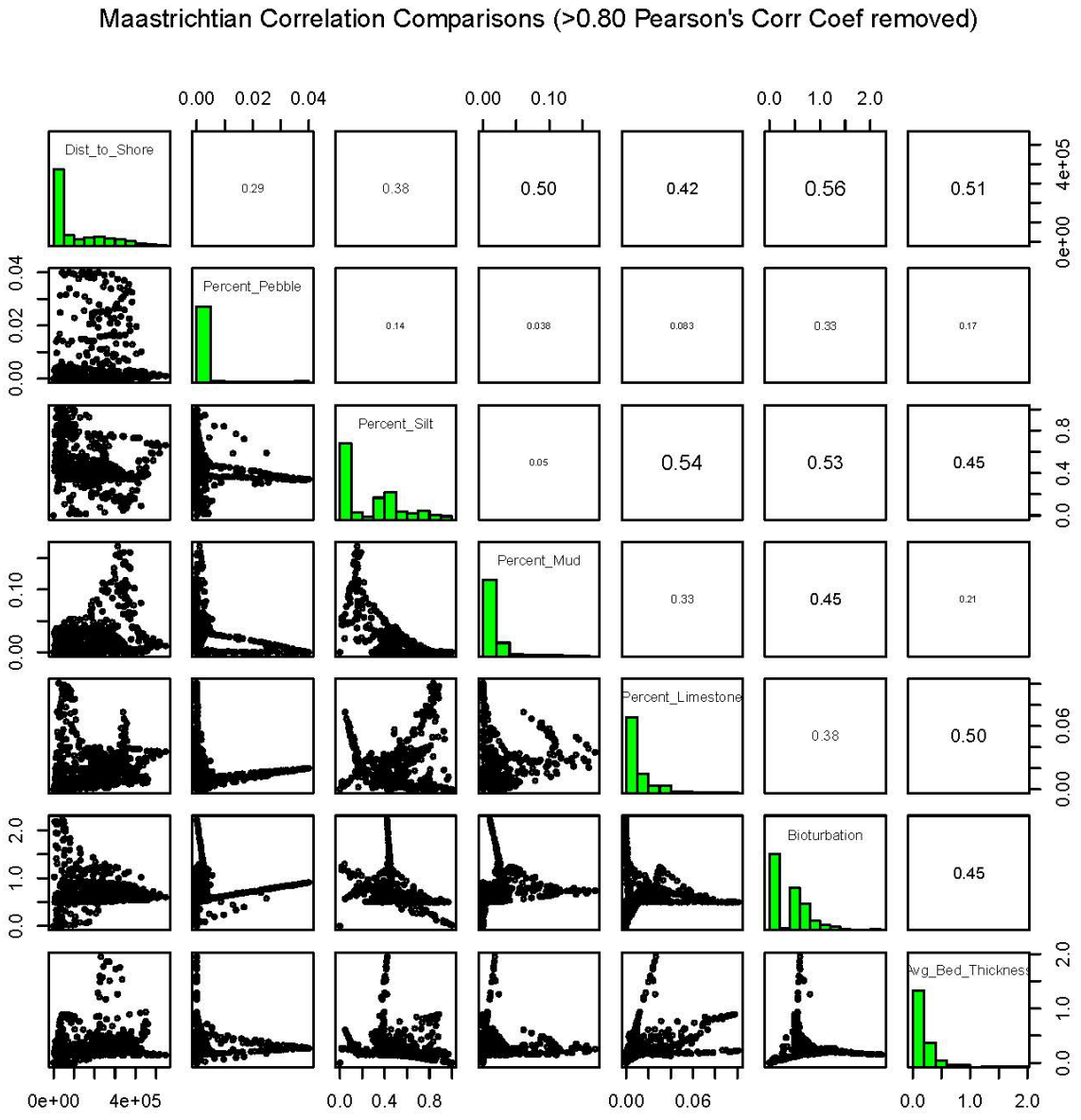

B. perplexus−B. gregoryensis Correlation Comparisons (>0.80 Pearson's Corr Coef removed)
Dist_to_Shor
Dist_to_Seep
Percent_Pebb
Percent_San
Percent_Silt
Percent_Mud
cent_Limest
edding_Styl
Bioturbation
Bed_Thickn
0.37
0.07
0.12
0.50
0.22
0.47
0.55
0.58
0.22
0.07
0.39
0.65
0.38
0.24
0.72
0.41
0.26
0.22
0.031
0.061
0.04
0.092
0.13
0.79
0.0011
0.19
0.03
0.59
0.29
0.54
0.069
0.41
0.53
0.51
0.18
0.013
0.23
0.16
0.015
0.30
0.37
0.14
0.75
0.41
0.29

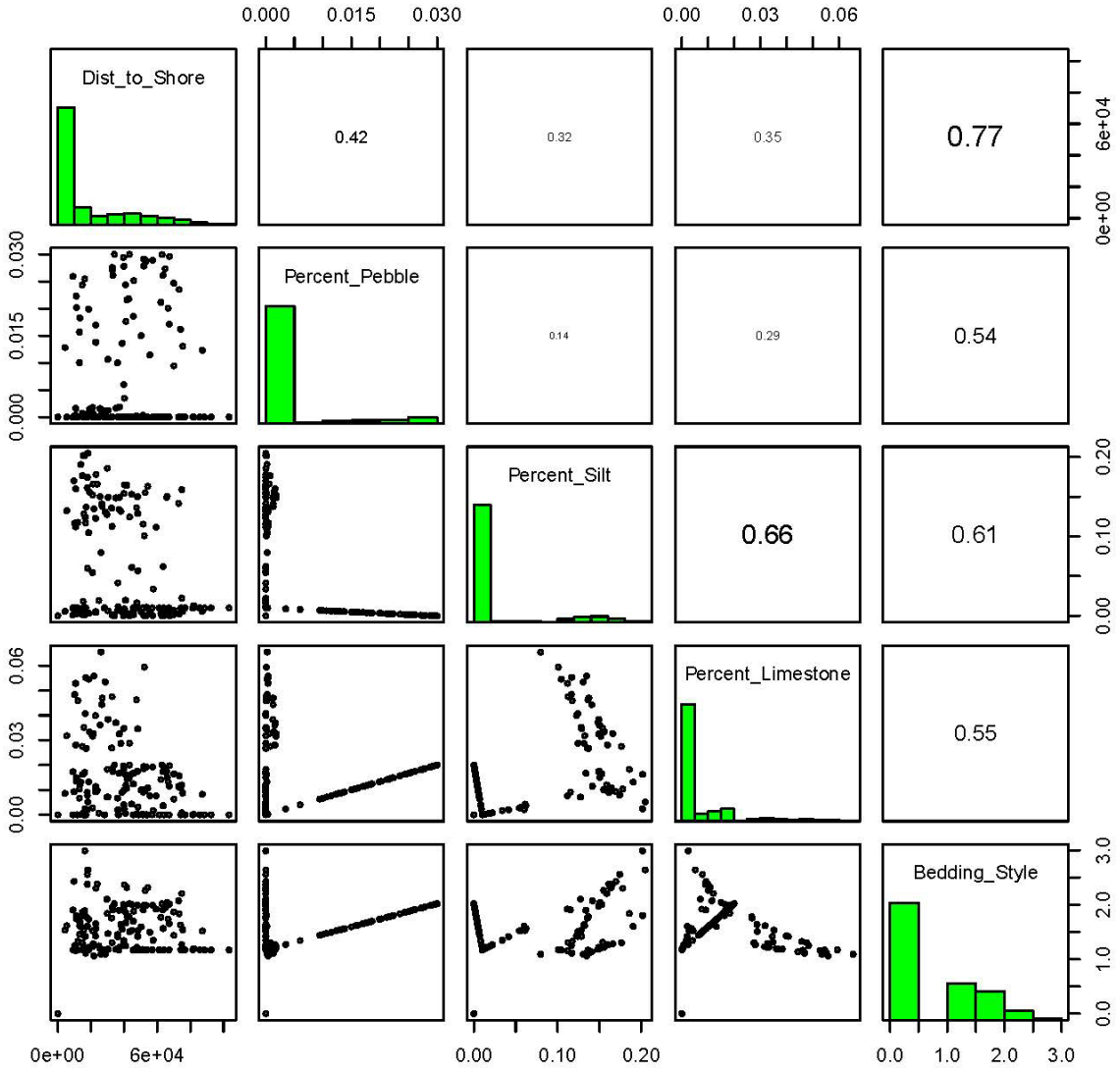

B. maclearni−B. sp. (smooth) Correlation Comparisons (>0.80 Pearson's Corr Coef removed)
Dist_to_Shore
Percent_Sand
Percent_Silt
Percent_Mud
Percent_Limestone
Bioturbation
Avg_Bed_Thickness
0.40
0.47
0.52
0.58
0.61
0.53
0.053
0.45
0.16
0.57
0.65
0.10
0.39
0.73
0.40
0.43
0.22
0.45
0.41
0.29
0.74

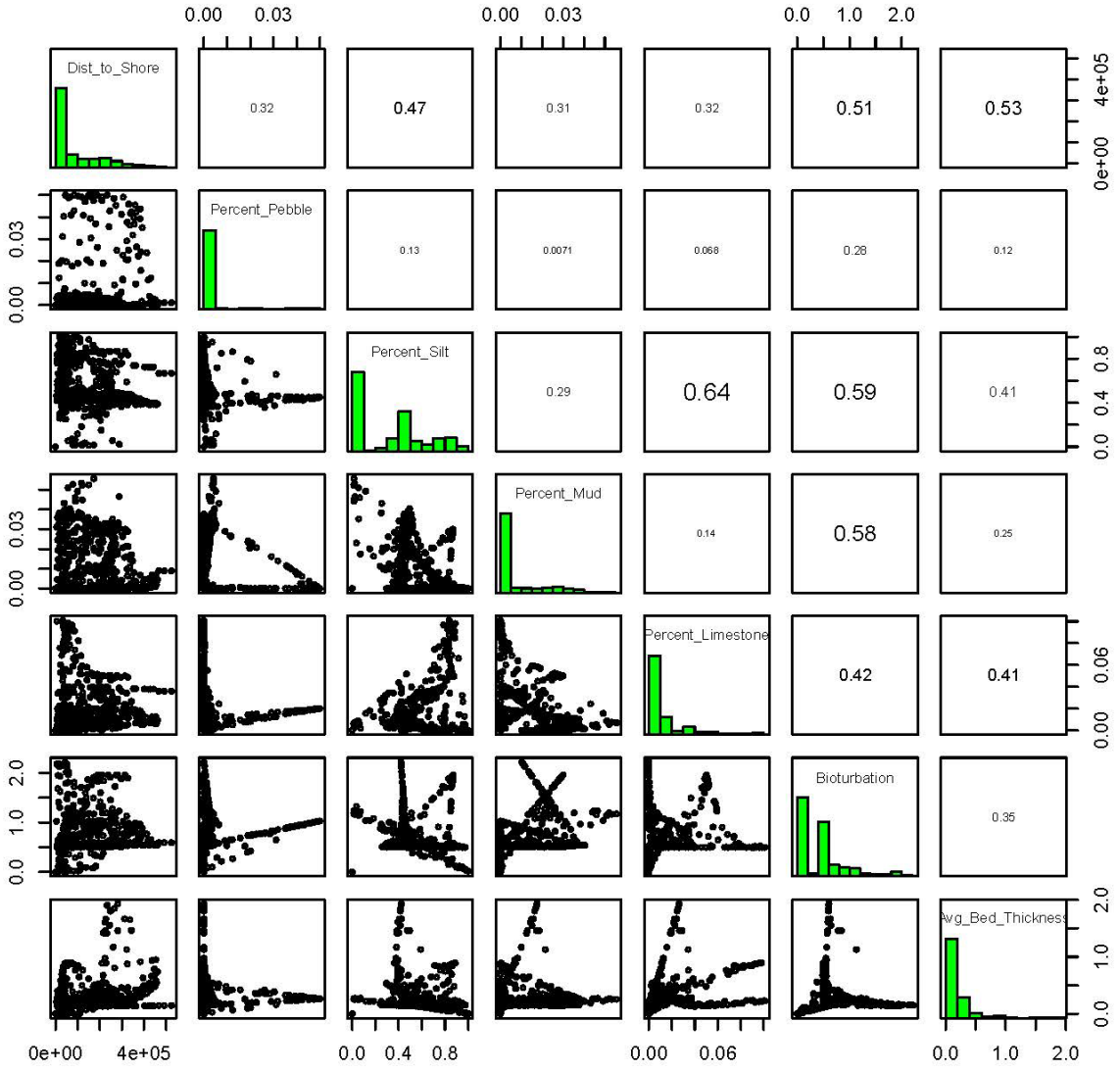
B. obtusus Correlation Comparisons (>0.80 Pearson's Corr Coef removed)

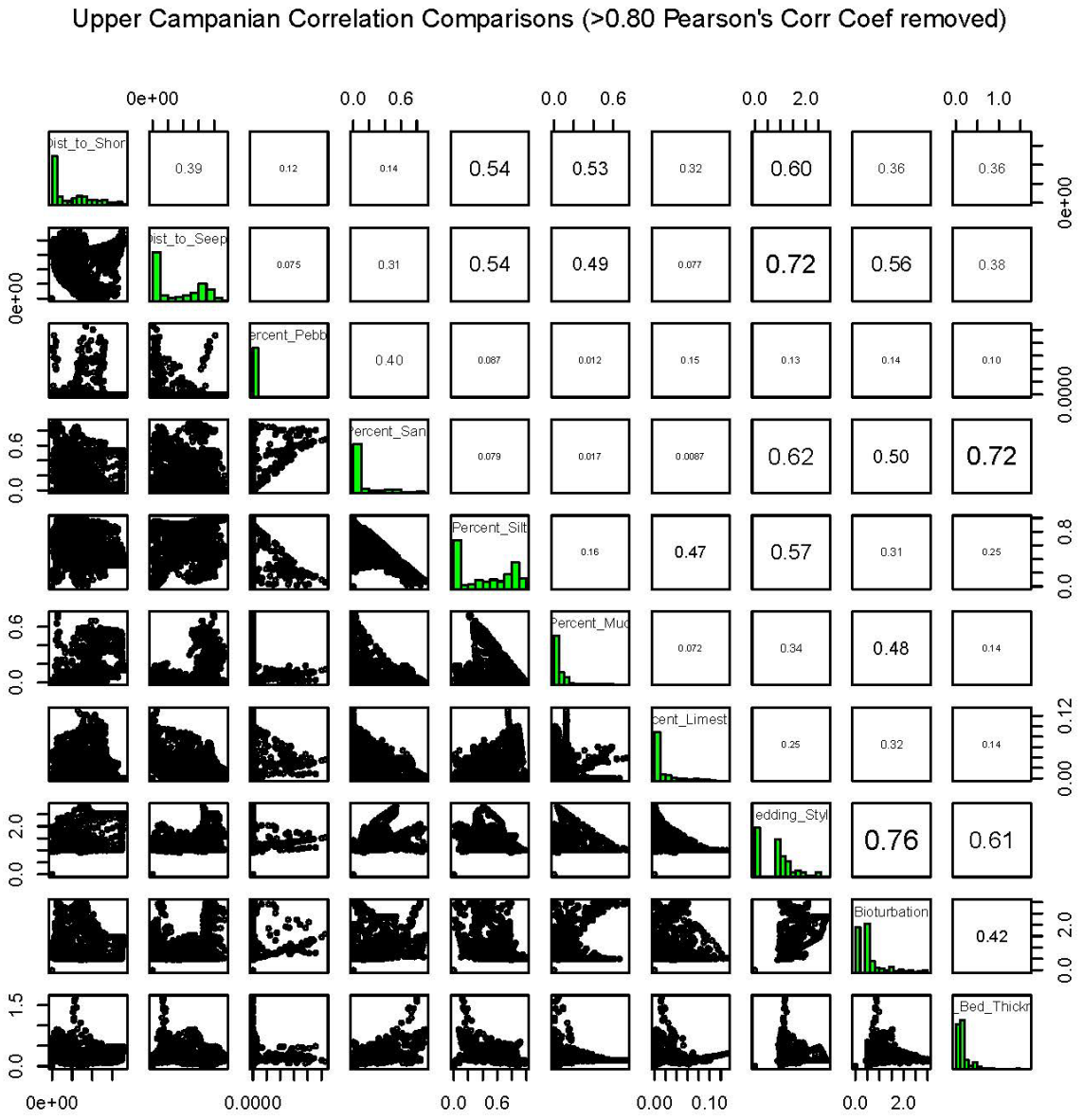


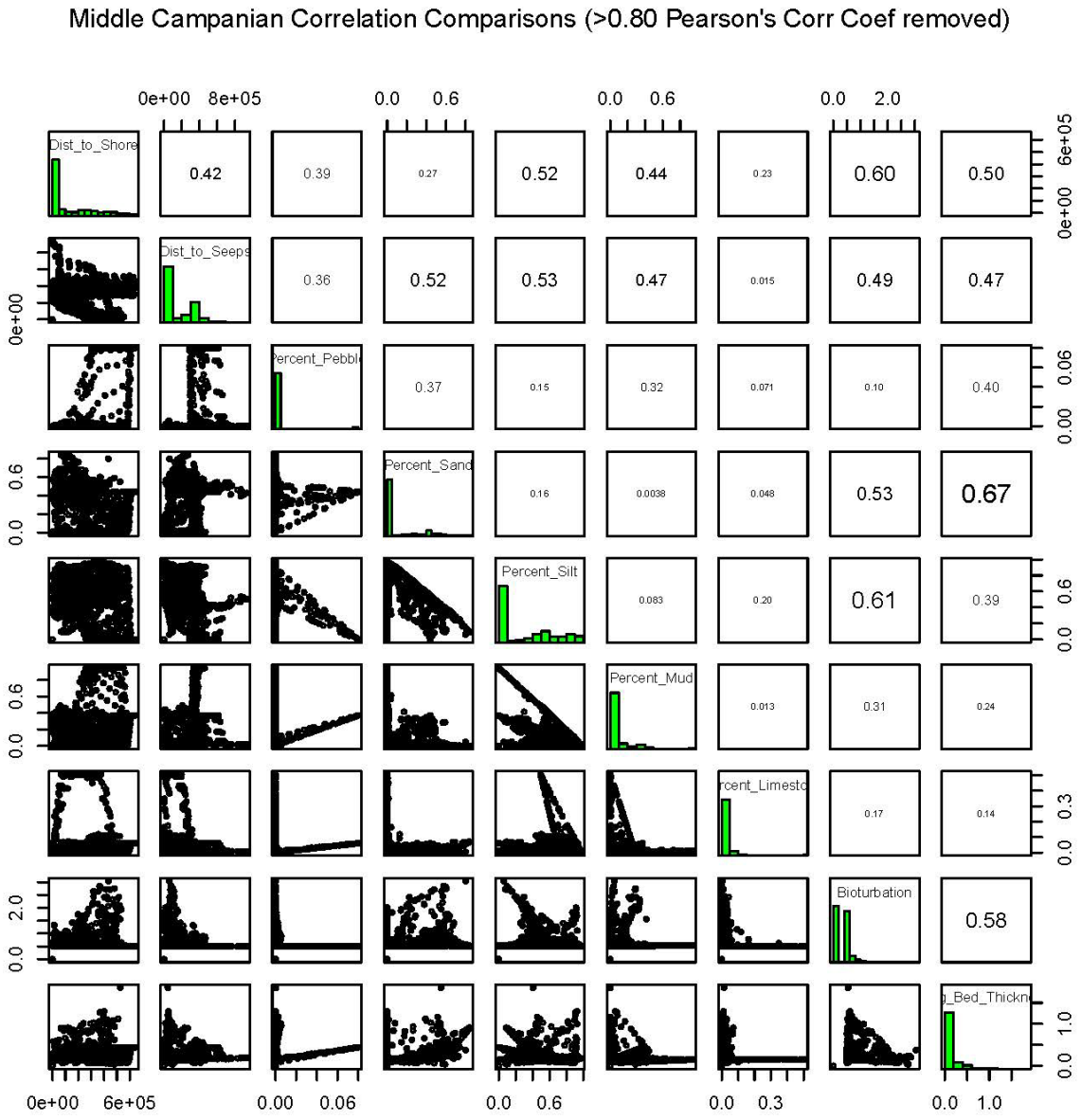

S. leei−B. sp. (weak flank ribs) Correlation Comparisons (>0.80 Pearson's Corr Coef removed)
Dist_to_Shore
Percent_Pebble
Percent_Sand
Percent_Silt
Percent_Mud
Percent_Limestone
Bedding_Style
0.28
0.51
0.41
0.54
0.18
0.24
0.33
0.24
0.083
0.014
0.12
0.044
0.42
0.14
0.41
0.002
0.74
0.37
0.051
0.11
0.16
0.000 0.006
0.0 0.4 0.8
0.00 0.04 0.08
0e+00 5e+05
0.0 0.4 0.8
0.00 0.20
0 10 20
0e+00 4e+05
0.0 0.4 0.8
0.00 0.20
0 10 20

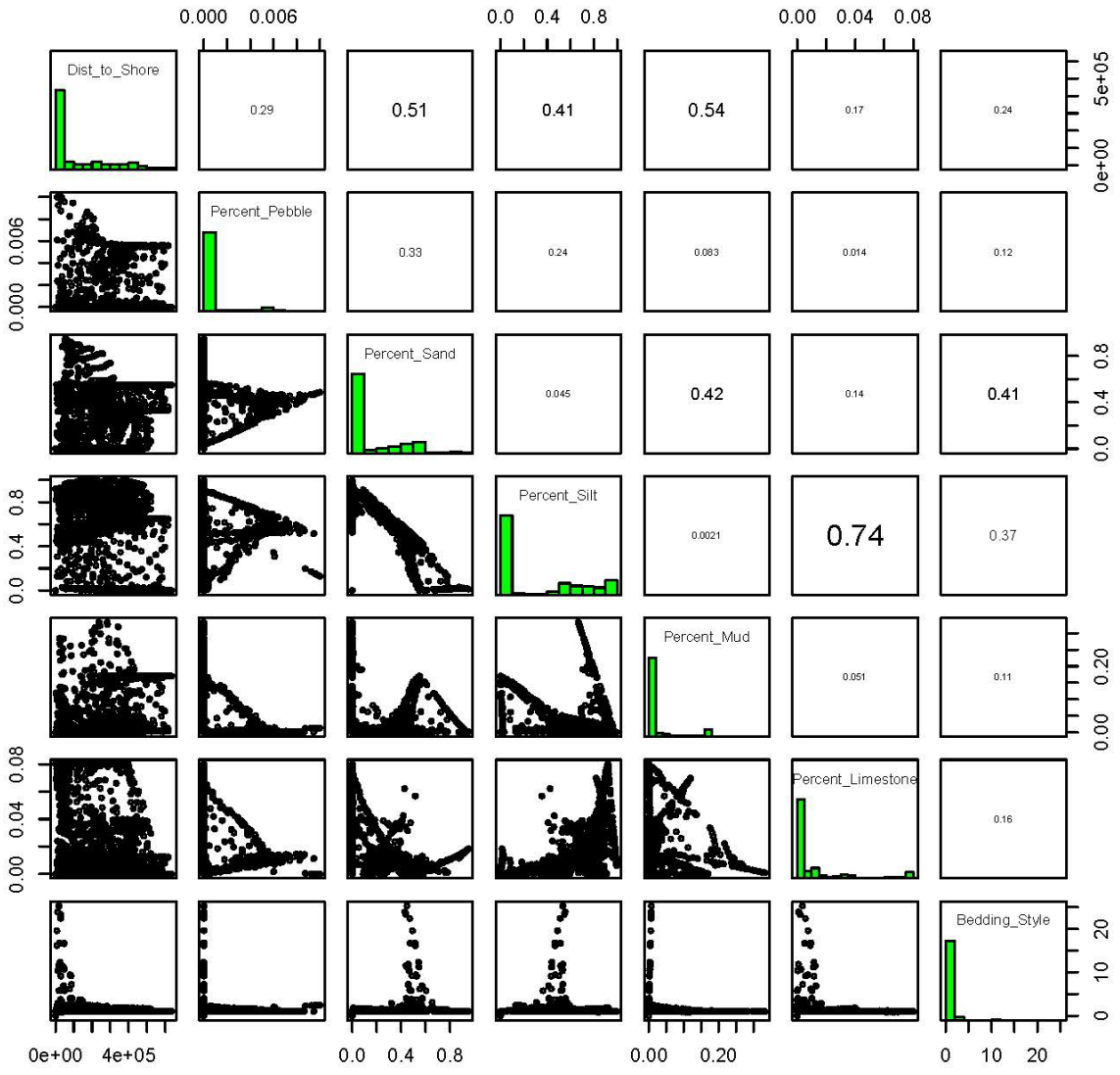

Maastrichtian Correlation Comparisons (>0.80 Pearson's Corr Coef removed)
Dist_to_Shore
Percent_Pebble
Percent_Silt
Percent_Mud
Percent_Limestone
Bioturbation
Avg_Bed_Thickness
0.29
0.38
0.50
0.42
0.56
0.51
0.14
0.038
0.083
0.33
0.17
0.05
0.54
0.53
0.45
0.33
0.45
0.21
0.38
0.50
0.45
0.00 0.02 0.04
0.00 0.10
0.0 1.0 2.0
0e+00 4e+05
0.0 0.4 0.8
0.00 0.06


Upper Maastrichtian Correlation Comparisons (>0.80 Pearson's Corr Coef removed)
Dist_to_Shore
Percent_Pebble
Percent_Silt
Percent_Limestone
Bedding_Style
0.42
0.32
0.35
0.77
0.14
0.29
0.54
0.66
0.61
0.55
0.000
0.015
0.030
0.00
0.03
0.06
0e+00
6e+04
0.00
0.10
0.20
0.0
1.0
2.0
3.0


Lower Maastrichtian Correlation Comparisons (>0.80 Pearson's Corr Coef removed)
Dist_to_Shore
Percent_Pebble
Percent_Silt
Percent_Mud
Percent_Limestone
Bioturbation
Avg_Bed_Thickness
0.32
0.47
0.31
0.32
0.51
0.53
0.13
0.0071
0.068
0.28
0.12
0.29
0.64
0.59
0.41
0.14
0.58
0.25
0.42
0.41
0.35
0.00
0.03
0.0
1.0
2.0
0e+00
4e+05
0.4
0.8
0.06


Upper Campanian Correlation Comparisons (>0.80 Pearson's Corr Coef removed)
Dist_to_Shor
Dist_to_Seep
ercent_Pebb
ercent_San
Percent_Silt
Percent_Mud
cent_Limest
edding_Styl
Bioturbation
_Bed_Thickn
0.39
0.12
0.14
0.54
0.53
0.32
0.60
0.36
0.36
0.075
0.31
0.54
0.49
0.077
0.72
0.56
0.38
0.40
0.087
0.012
0.15
0.13
0.14
0.10
0.079
0.017
0.0087
0.62
0.50
0.72
0.16
0.47
0.57
0.31
0.25
0.072
0.34
0.48
0.14
0.25
0.32
0.14
0.76
0.61
0.42


Middle Campanian Correlation Comparisons (>0.80 Pearson's Corr Coef removed)
Dist_to_Shore
Dist_to_Seeps
Percent_Pebbl
Percent_Sand
Percent_Silt
Percent_Mud
rcent_Limesto
Bioturbation
g_Bed_Thickn
0.42
0.39
0.27
0.52
0.44
0.23
0.60
0.50
0.36
0.52
0.53
0.47
0.015
0.49
0.47
0.37
0.15
0.32
0.071
0.10
0.40
0.16
0.0038
0.048
0.53
0.67
0.083
0.20
0.61
0.39
0.013
0.31
0.24
0.17
0.14
0.58


Lower Campanian Correlation Comparisons (>0.80 Pearson's Corr Coef removed)
Dist_to_Shore
Percent_Pebble
Percent_Sand
Percent_Silt
Percent_Mud
Percent_Limestone
Bedding_Style
0.29
0.51
0.41
0.54
0.17
0.24
0.33
0.24
0.083
0.014
0.12
0.045
0.42
0.14
0.41
0.0021
0.74
0.37
0.051
0.11
0.16
0.000 0.006
0.0 0.4 0.8
0.00 0.04 0.08
0e+00 5e+05
0.0 0.4 0.8
0.00 0.20
0 10 20
0e+00 4e+05
0.0 0.4 0.8
0.00 0.20
0 10 20

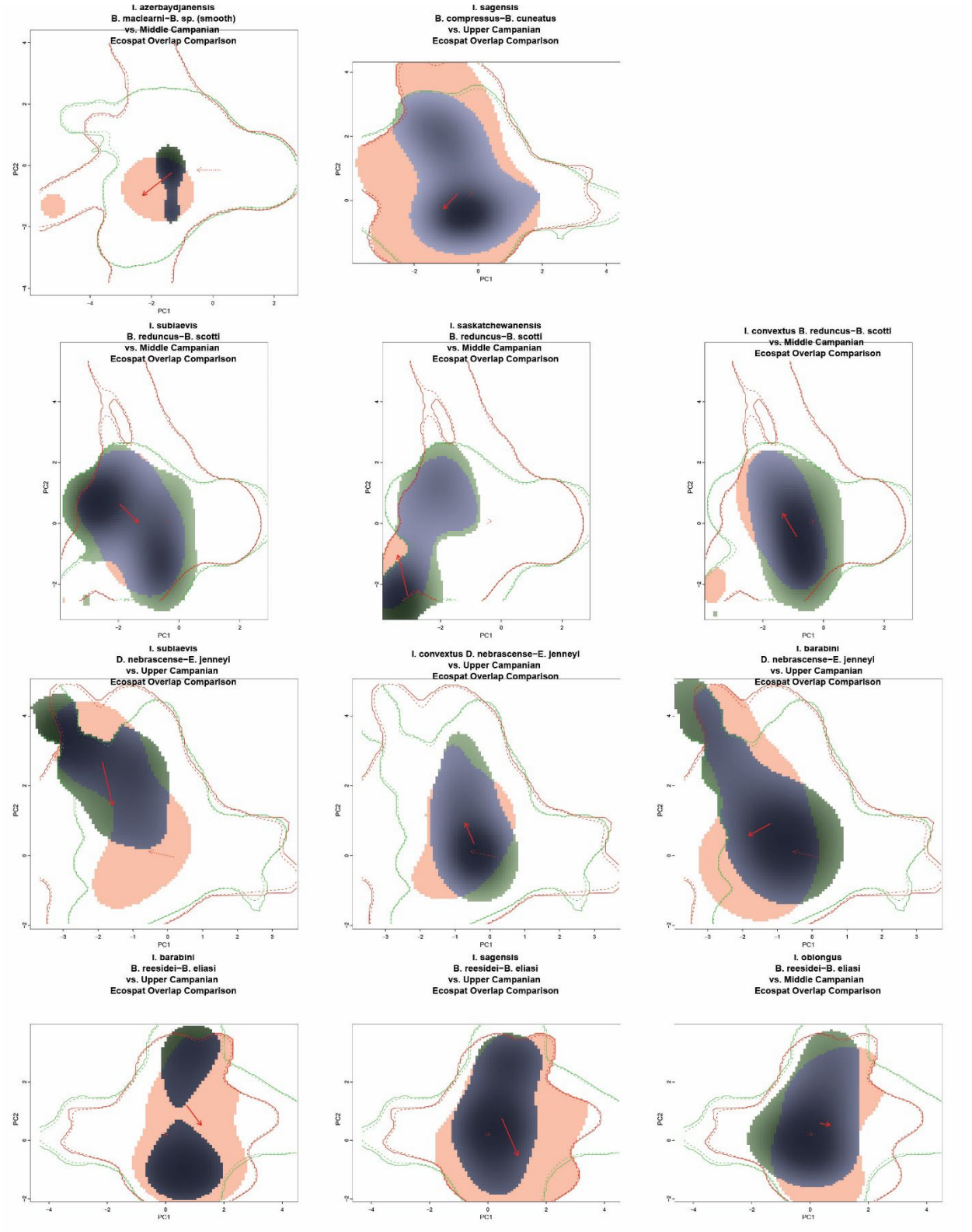



Figure S6. Overlap comparison plots of species' niches in different temporal bins. Blue represents regions of niche stability, red regions of niche expansion (i.e., where the longer interval niche has expanded beyond the biozone

interval niche), and green represents regions of niche unfilling (i.e., where the longer interval niche has moved away from the biozone interval niche). Red lines represent the extent of environmental space realized within the longer interval, green lines represent the extent of environmental space realized within the biozone interval, and the red arrow indicates the shift of the niche centroid between the two species. Non-analogous conditions are indicated by regions where the extent of environmental parameters in the different intervals do not overlap.

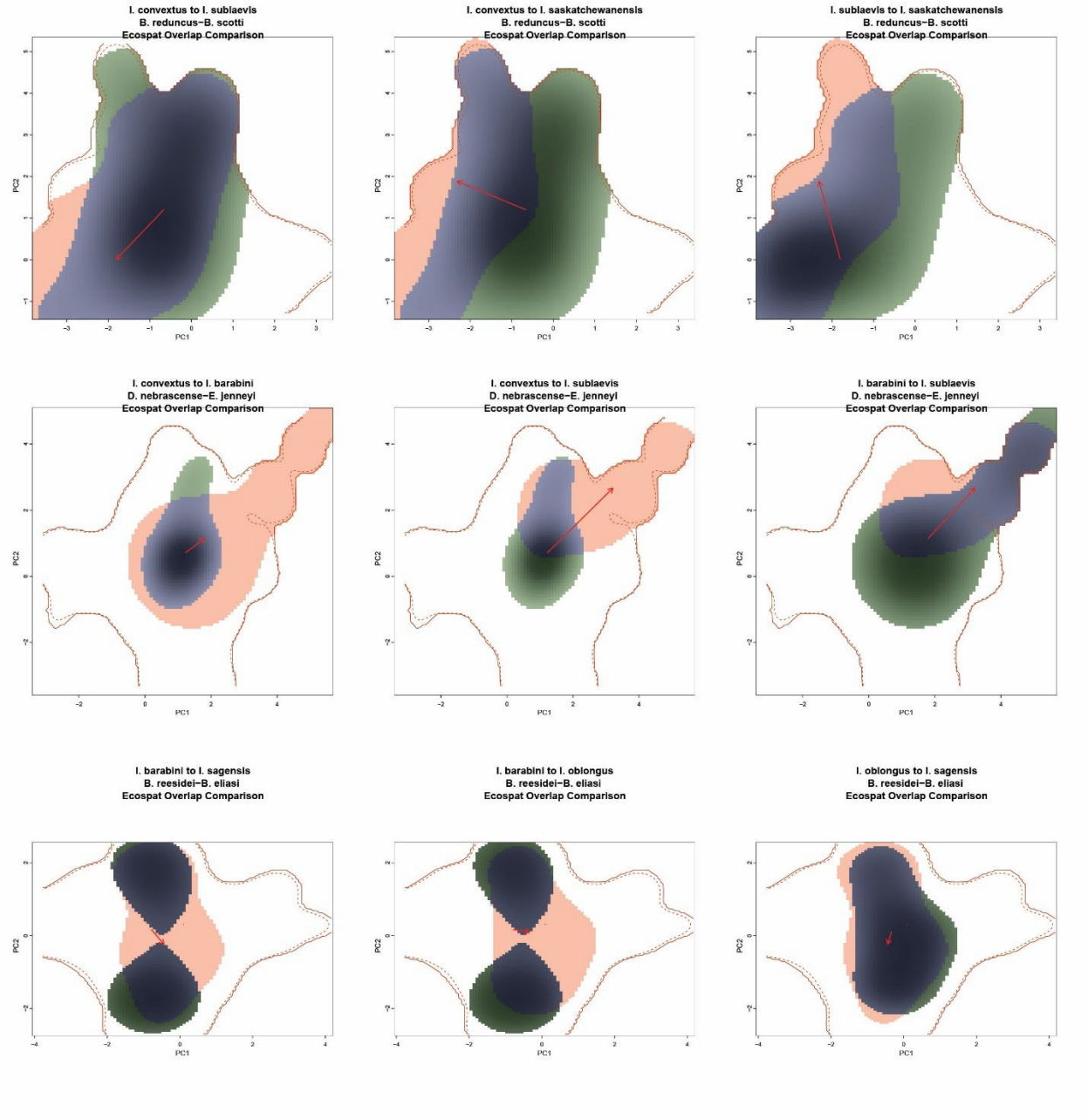



Figure S7. Overlap comparison plots of ingroup *Inoceramus* species' niches. Blue represents regions of niche stability, red regions of niche expansion (i.e., where the second listed species has expanded beyond the niche of the first listed species), and green represents regions of niche unfilling (i.e., where the second listed species niche has moved away from space occupied by the first listed species). Red lines represent the extent of environmental space realized within the interval and the red arrow indicates the shift of the niche centroid between the two species.

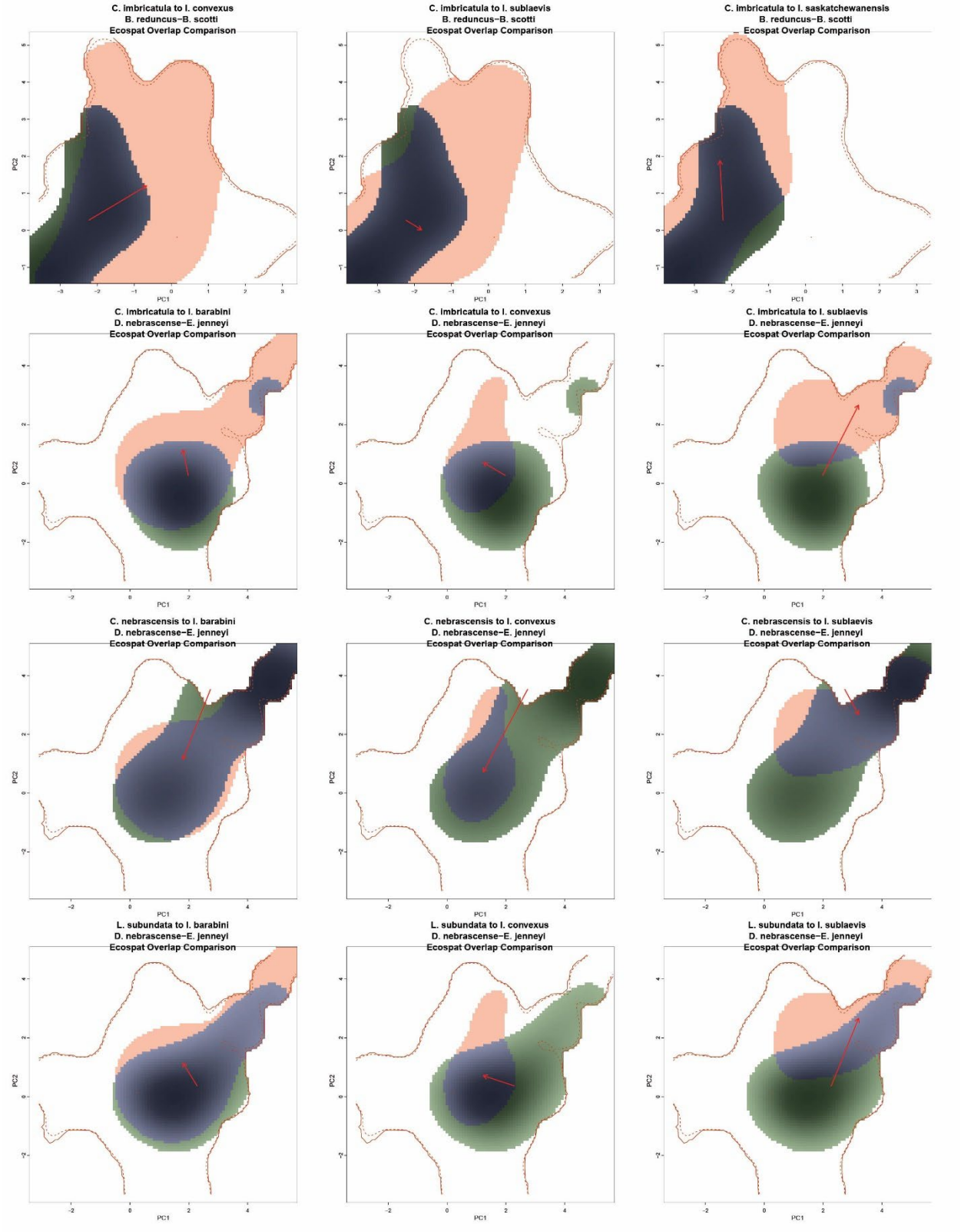



Figure S8a. Overlap comparison plots of comparison species' niches. Blue represents regions of niche stability, red regions of niche expansion (i.e., where the second listed species has expanded beyond the niche of the first listed

species), and green represents regions of niche unfilling (i.e., where the second listed species niche has moved away from space occupied by the first listed species). Red lines represent the extent of environmental space realized within the interval and the red arrow indicates the shift of the niche centroid between the two species.

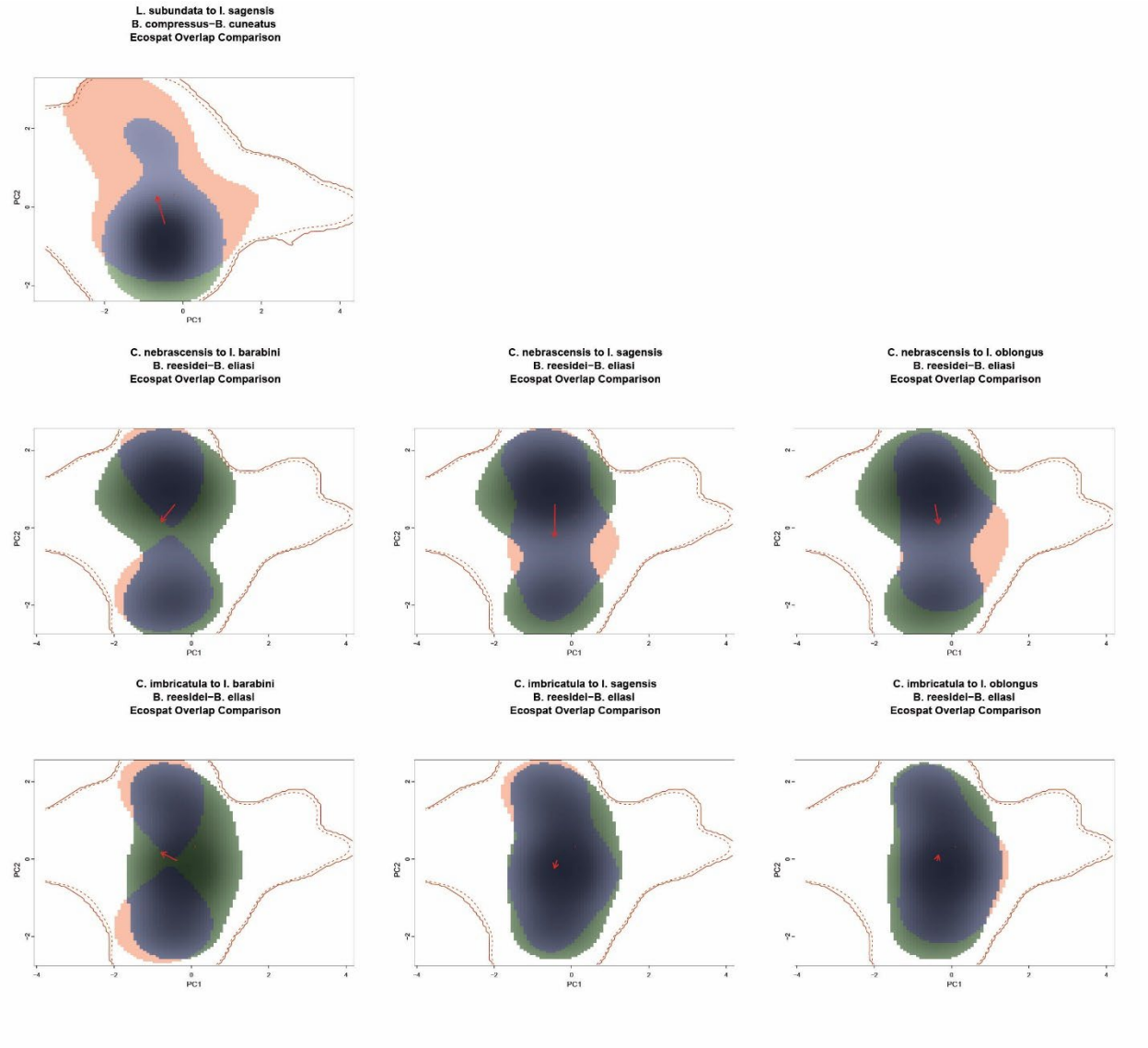



Figure S8b. Overlap comparison plots of comparison species' niches. Blue represents regions of niche stability, red regions of niche expansion (i.e., where the second listed species has expanded beyond the niche of the first listed species), and green represents regions of niche unfilling (i.e., where the second listed species niche has moved away from space occupied by the first listed species). Red lines represent the extent of environmental space realized within the interval and the red arrow indicates the shift of the niche centroid between the two species.

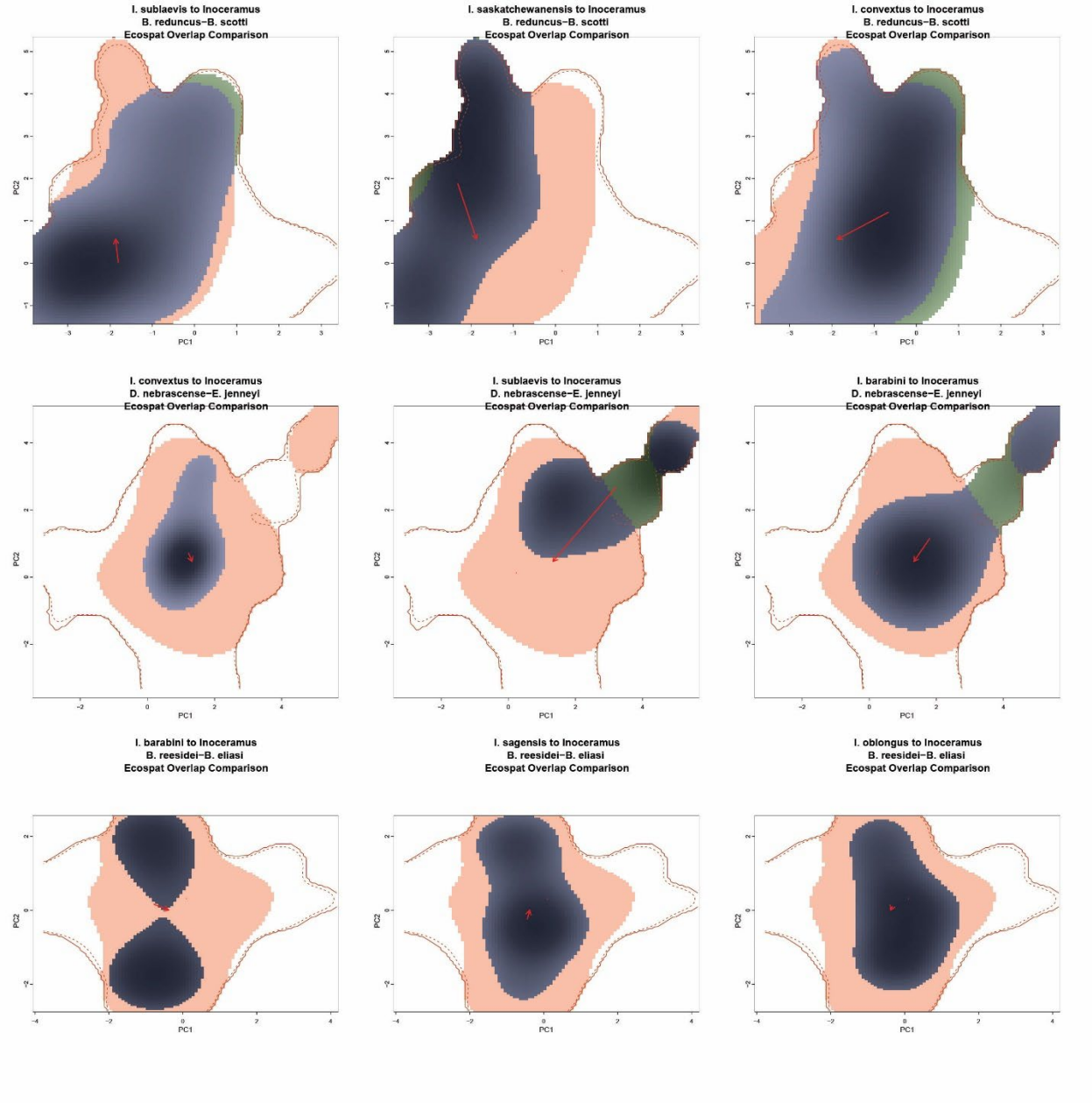



Figure S9. Overlap comparison plots of *Inoceramus* species' niches and the *Inoceramus* genus niche. Blue represents regions of niche stability, red regions of niche expansion (i.e., where the genus has expanded beyond the species), and green represents regions of niche unfilling (i.e., where the genus niche has moved away from space occupied by the species). Red lines represent the extent of environmental space realized within the interval and the red arrow indicates the shift of the niche centroid between the two species.

Table S8. Niche overlap values for temporal bin comparisons including I and D overlap values, and niche expansion, stability, and unfilling.

| Species | Intervals Compared | D value | I value | Expansion | Stability | Unfilling |
|---|---|---|---|---|---|---|
| I. convexus | B. reduncus-B. scotti to Mid Campanian | 0.31 | 0.44 | 0.09 | 0.91 | 0.24 |
| I. convexus | D. nebrascense-E. jenneyi to Upper Campanian | 0.68 | 0.84 | 0.12 | 0.88 | 0.15 |
| I. barabini | B. reesidei-B. eliasi to Upper Campanian | 0.41 | 0.58 | 0.42 | 0.58 | 0.05 |
| I. barabini | D. nebrascense-E. jenneyi to Upper Campanian | 0.43 | 0.59 | 0.18 | 0.82 | 0.16 |
| I. sagensis | B. reesidei-B. eliasi to Upper Campanian | 0.43 | 0.63 | 0.37 | 0.63 | 0.01 |
| I. sagensis | B. compressus-B. cuneatus to Upper Campanian | 0.43 | 0.62 | 0.29 | 0.71 | 0.00 |
| I. sublaevis | B. reduncus-B. scotti to Mid Campanian | 0.41 | 0.62 | 0.01 | 0.99 | 0.14 |
| I. sublaevis | D. nebrascense-E. jenneyi to Upper Campanian | 0.49 | 0.62 | 0.38 | 0.62 | 0.19 |
| I. azerbaydjanensis | B. maclearni-B. sp. (smooth) to Mid Campanian | 0.09 | 0.27 | 0.69 | 0.31 | 0.24 |
| I. oblongus | B. reesidei-B. eliasi to Upper Campanian | 0.55 | 0.70 | 0.19 | 0.81 | 0.16 |
| I. saskatchewanensis | B. reduncus-B. scotti to Mid Campanian | 0.46 | 0.63 | 0.00 | 1.00 | 0.15 |
| C. imbricatula | B. perplexus-B. gregoryensis to Mid Campanian | **0.018*** | 0.154 | 0.545 | 0.545 | 0.543 |
| C. imbricatula | B. reduncus-B. scotti to Mid Campanian | 0.224 | 0.253 | 0.756 | 0.756 | 0.547 |
| C. imbricatula | D. nebrascense-E. jenneyi to Upper Campanian | 0.297 | 0.307 | 0.589 | 0.589 | 0.527 |
| C. imbricatula | B. reesidei-B. eliasi to Upper Campanian | 0.273 | 0.535 | 0.381 | 0.381 | 0.932 |
| L. subundata | D. nebrascense-E. jenneyi to Upper Campanian | 0.267 | 0.267 | 0.339 | 0.339 | 0.489 |
| L. subundata | B. reesidei-B. eliasi to Upper Campanian | 0.896 | 0.854 | 0.870 | 0.870 | 0.154 |
| C. nebrascensis | D. nebrascense-E. jenneyi to Upper Campanian | 0.511 | 0.581 | 0.697 | 0.697 | 0.685 |
| C. nebrascensis | B. reesidei-B. eliasi to Upper Campanian | 0.321 | 0.367 | 0.158 | 0.158 | 0.790 |
| C. nebrascensis | B. baculus-B. grandis to Lower Maastrichtian | **0.014*** | **0.012*** | 0.212 | 0.212 | **0.006*** |
| C. nebrascensis | B. baculus-B. grandis to Maastrichtian | 0.275 | 0.579 | 0.798 | 0.798 | 0.429 |

Table S9. Results of greater niche equivalency test for temporal bin comparisons including p-values for I and D overlap (greater than random), and niche expansion (less than random), stability (greater than random), and unfilling (less than random).

| | | Greater Niche Equivalency Test | | | | |
|---|---|---|---|---|---|---|
| Species | Intervals Compared | D value | I Value | Expansion (lower) | Stability (higher) | Unfilling (lower) |
| I. convexus | B. reduncus-B. scotti vs. Mid Campanian | 0.413 | 0.463 | 0.190 | 0.190 | 0.812 |
| I. convexus | D. nebrascense-E. jenneyi vs. Upper Campanian | **0.008*** | **0.002*** | 0.208 | 0.208 | 0.275 |
| I. barabini | B. reesidei-B. eliasi vs. Upper Campanian | 0.495 | 0.549 | 0.729 | 0.729 | 0.220 |
| I. barabini | D. nebrascense-E. jenneyi vs Upper Campanian | 0.140 | 0.255 | 0.194 | 0.194 | 0.709 |
| I. sagensis | B. reesidei-B. eliasi vs. Upper Campanian | 0.493 | 0.495 | 0.739 | 0.739 | 0.166 |
| I. sagensis | B. compressus-B. cuneatus vs. Upper Campanian | 0.820 | 0.846 | 0.904 | 0.904 | 0.066 |
| I. sublaevis | B. reduncus-B. scotti vs. Mid Campanian | 0.064 | 0.114 | **0.040*** | **0.040*** | 0.709 |
| I. sublaevis | D. nebrascense-E. jenneyi vs. Upper Campanian | 0.054 | 0.140 | 0.415 | 0.415 | 0.397 |
| I. azerbaydjanensis | B. maclearni-B. sp. (smooth) vs. Mid Campanian | 0.966 | 0.970 | 0.984 | 0.984 | 0.774 |
| I. oblongus | B. reesidei-B. eliasi vs. Upper Campanian | 0.142 | 0.178 | 0.293 | 0.293 | 0.527 |
| I. saskatchewanensis | B. reduncus-B. scotti vs. Mid Campanian | 0.154 | 0.311 | 0.068 | 0.068 | 0.415 |
| C. imbricatula | B. perplexus-B. gregoryensis to Mid Campanian | **0.018*** | 0.154 | 0.545 | 0.545 | 0.543 |
| C. imbricatula | B. reduncus-B. scotti to Mid Campanian | 0.224 | 0.253 | 0.756 | 0.756 | 0.547 |
| C. imbricatula | D. nebrascense-E. jenneyi to Upper Campanian | 0.297 | 0.307 | 0.589 | 0.589 | 0.527 |
| C. imbricatula | B. reesidei-B. eliasi to Upper Campanian | 0.273 | 0.535 | 0.381 | 0.381 | 0.932 |
| L. subundata | D. nebrascense-E. jenneyi to Upper Campanian | 0.267 | 0.267 | 0.339 | 0.339 | 0.489 |
| L. subundata | B. reesidei-B. eliasi to Upper Campanian | 0.896 | 0.854 | 0.870 | 0.870 | 0.154 |
| C. nebrascensis | D. nebrascense-E. jenneyi to Upper Campanian | 0.511 | 0.581 | 0.697 | 0.697 | 0.685 |
| C. nebrascensis | B. reesidei-B. eliasi to Upper Campanian | 0.321 | 0.367 | 0.158 | 0.158 | 0.790 |
| C. nebrascensis | B. baculus-B. grandis to Lower Maastrichtian | **0.014*** | **0.012*** | 0.212 | 0.212 | **0.006*** |
| C. nebrascensis | B. baculus-B. grandis to Maastrichtian | 0.275 | 0.579 | 0.798 | 0.798 | 0.429 |

Table S10. Results of lower niche equivalency test for temporal bin comparisons including p-values for I and D overlap (less than random), and niche expansion (greater than random), stability (less than random), and unfilling (greater than random).

| Lower Niche Equivalency Test | | | | | | |
|---|---|---|---|---|---|---|
| Species | Intervals Compared | D value | I Value | Expansion (higher) | Stability (lower) | Unfilling (higher) |
| I. convexus | B. reduncus-B. scotti to Mid Campanian | 0.621 | 0.575 | 0.818 | 0.818 | 0.190 |
| I. convexus | D. nebrascense-E. jenneyi to Upper Campanian | 0.988 | 0.996 | 0.800 | 0.800 | 0.715 |
| I. barabini | B. reesidei-B. eliasi to Upper Campanian | 0.495 | 0.429 | 0.273 | 0.273 | 0.739 |
| I. barabini | D. nebrascense-E. jenneyi to Upper Campanian | 0.892 | 0.766 | 0.810 | 0.810 | 0.313 |
| I. sagensis | B. reesidei-B. eliasi to Upper Campanian | 0.499 | 0.497 | 0.261 | 0.261 | 0.870 |
| I. sagensis | B. compressus-B. cuneatus to Upper Campanian | 0.174 | 0.170 | 0.098 | 0.098 | 0.930 |
| I. sublaevis | B. reduncus-B. scotti to Mid Campanian | 0.940 | 0.886 | 0.968 | 0.968 | 0.295 |
| I. sublaevis | D. nebrascense-E. jenneyi to Upper Campanian | 0.944 | 0.832 | 0.565 | 0.565 | 0.607 |
| I. azerbaydjanensis | B. maclearni-B. sp. (smooth) to Mid Campanian | **0.034*** | **0.034*** | **0.020*** | **0.020*** | 0.244 |
| I. oblongus | B. reesidei-B. eliasi to Upper Campanian | 0.842 | 0.822 | 0.671 | 0.671 | 0.489 |
| I. saskatchewanensis | B. reduncus-B. scotti to Mid Campanian | 0.808 | 0.675 | 0.930 | 0.930 | 0.587 |
| C. imbricatula | B. perplexus-B. gregoryensis to Mid Campanian | 0.990 | 0.850 | 0.479 | 0.479 | 0.431 |
| C. imbricatula | B. reduncus-B. scotti to Mid Campanian | 0.754 | 0.709 | 0.196 | 0.196 | 0.421 |
| C. imbricatula | D. nebrascense-E. jenneyi to Upper Campanian | 0.701 | 0.697 | 0.435 | 0.435 | 0.481 |
| C. imbricatula | B. reesidei-B. eliasi to Upper Campanian | 0.750 | 0.461 | 0.647 | 0.647 | 0.058 |
| L. subundata | D. nebrascense-E. jenneyi to Upper Campanian | 0.693 | 0.691 | 0.639 | 0.639 | 0.411 |
| L. subundata | B. reesidei-B. eliasi to Upper Campanian | 0.104 | 0.148 | 0.134 | 0.134 | 1.000 |
| C. nebrascensis | D. nebrascense-E. jenneyi to Upper Campanian | 0.501 | 0.429 | 0.250 | 0.250 | 0.343 |
| C. nebrascensis | B. reesidei-B. eliasi to Upper Campanian | 0.749 | 0.689 | 0.832 | 0.832 | 0.232 |
| C. nebrascensis | B. baculus-B. grandis to Lower Maastrichtian | 0.982 | 0.990 | 0.806 | 0.806 | 0.994 |
| C. nebrascensis | B. baculus-B. grandis to Maastrichtian | 0.719 | 0.423 | 0.204 | 0.204 | 0.581 |

Table S11. Results of greater niche similarity test for temporal bin comparisons including p-values for I and D overlap (greater than random), and niche expansion (less than random), stability (greater than random), and unfilling (less than random).

| | | Greater Niche Similarity Test | | | | |
|---|---|---|---|---|---|---|
| Species | Intervals Compared | D value | I Value | Expansion (lower) | Stability (higher) | Unfilling (lower) |
| I. convexus | B. reduncus-B. scotti vs. Mid Campanian | 0.243 | 0.241 | 0.171 | 0.171 | 0.147 |
| I. convexus | D. nebrascense-E. jenneyi vs. Upper Campanian | **0.001*** | **0.001*** | **0.022*** | **0.022*** | 0.069 |
| I. barabini | B. reesidei-B. eliasi vs. Upper Campanian | **0.034*** | **0.021*** | 0.080 | 0.080 | 0.067 |
| I. barabini | D. nebrascense-E. jenneyi vs Upper Campanian | 0.059 | **0.045*** | 0.133 | 0.133 | 0.078 |
| I. sagensis | B. reesidei-B. eliasi vs. Upper Campanian | 0.113 | 0.063 | 0.205 | 0.205 | **0.013*** |
| I. sagensis | B. compressus-B. cuneatus vs. Upper Campanian | 0.133 | 0.113 | 0.166 | 0.166 | **0.008*** |
| I. sublaevis | B. reduncus-B. scotti vs. Mid Campanian | 0.105 | **0.023*** | **0.003*** | **0.003*** | 0.190 |
| I. sublaevis | D. nebrascense-E. jenneyi vs. Upper Campanian | **0.003*** | **0.001*** | 0.126 | 0.126 | 0.177 |
| I. azerbaydjanensis | B. maclearni-B. sp. (smooth) vs. Mid Campanian | 0.165 | 0.122 | 0.161 | 0.161 | 0.190 |
| I. oblongus | B. reesidei-B. eliasi vs. Upper Campanian | 0.014 | 0.016 | 0.191 | 0.191 | **0.016*** |
| I. saskatchewanensis | B. reduncus-B. scotti vs. Mid Campanian | 0.028 | 0.017 | 0.024 | 0.024 | 0.136 |
| C. imbricatula | B. perplexus-B. gregoryensis to Mid Campanian | **0.002*** | **0.006*** | 0.191 | 0.191 | 0.204 |
| C. imbricatula | B. reduncus-B. scotti to Mid Campanian | 0.102 | 0.065 | 0.239 | 0.239 | 0.146 |
| C. imbricatula | D. nebrascense-E. jenneyi to Upper Campanian | 0.093 | 0.081 | 0.153 | 0.153 | 0.094 |
| C. imbricatula | B. reesidei-B. eliasi to Upper Campanian | 0.085 | 0.127 | 0.168 | 0.168 | 0.114 |
| L. subundata | D. nebrascense-E. jenneyi to Upper Campanian | 0.076 | 0.065 | 0.133 | 0.133 | 0.076 |
| L. subundata | B. reesidei-B. eliasi to Upper Campanian | 0.337 | 0.241 | 0.139 | 0.139 | 0.058 |
| C. nebrascensis | D. nebrascense-E. jenneyi to Upper Campanian | **0.003*** | **0.004*** | **0.016*** | **0.016*** | 0.282 |
| C. nebrascensis | B. reesidei-B. eliasi to Upper Campanian | **0.040*** | **0.027*** | **0.021*** | **0.021*** | 0.082 |
| C. nebrascensis | B. baculus-B. grandis to Lower Maastrichtian | **0.001*** | **0.001*** | **0.005*** | **0.005*** | **0.003*** |
| C. nebrascensis | B. baculus-B. grandis to Maastrichtian | 0.105 | 0.127 | 0.118 | 0.118 | 0.153 |

Table S12. Results of lower niche similarity test of temporal bin sizes including p-values for I and D overlap (less than random), and niche expansion (greater than random), stability (less than random), and unfilling (greater than random).

| Lower Niche Similarity Test | | | | | | |
|---|---|---|---|---|---|---|
| Species | Intervals Compared | D value | I Value | Expansion (higher) | Stability (lower) | Unfilling (higher) |
| I. convexus | B. reduncus-B. scotti to Mid Campanian | 0.765 | 0.764 | 0.838 | 0.838 | 0.847 |
| I. convexus | D. nebrascense-E. jenneyi to Upper Campanian | 1.000 | 1.000 | 0.984 | 0.984 | 0.936 |
| I. barabini | B. reesidei-B. eliasi to Upper Campanian | 0.961 | 0.979 | 0.910 | 0.910 | 0.929 |
| I. barabini | D. nebrascense-E. jenneyi to Upper Campanian | 0.958 | 0.971 | 0.899 | 0.899 | 0.934 |
| I. sagensis | B. reesidei-B. eliasi to Upper Campanian | 0.889 | 0.940 | 0.792 | 0.792 | 0.984 |
| I. sagensis | B. compressus-B. cuneatus to Upper Campanian | 0.844 | 0.857 | 0.813 | 0.813 | 0.986 |
| I. sublaevis | B. reduncus-B. scotti to Mid Campanian | 0.876 | 0.976 | 0.999 | 0.999 | 0.819 |
| I. sublaevis | D. nebrascense-E. jenneyi to Upper Campanian | 0.993 | 0.996 | 0.870 | 0.870 | 0.830 |
| I. azerbaydjanensis | B. maclearni-B. sp. (smooth) to Mid Campanian | 0.871 | 0.908 | 0.867 | 0.867 | 0.828 |
| I. oblongus | B. reesidei-B. eliasi to Upper Campanian | 0.994 | 0.990 | 0.827 | 0.827 | 0.992 |
| I. saskatchewanensis | B. reduncus-B. scotti to Mid Campanian | 0.976 | 0.982 | 0.977 | 0.977 | 0.862 |
| C. imbricatula | B. perplexus-B. gregoryensis to Mid Campanian | 1.000 | 1.000 | 0.834 | 0.834 | 0.812 |
| C. imbricatula | B. reduncus-B. scotti to Mid Campanian | 0.886 | 0.907 | 0.747 | 0.747 | 0.826 |
| C. imbricatula | D. nebrascense-E. jenneyi to Upper Campanian | 0.916 | 0.930 | 0.878 | 0.878 | 0.924 |
| C. imbricatula | B. reesidei-B. eliasi to Upper Campanian | 0.930 | 0.898 | 0.866 | 0.866 | 0.909 |
| L. subundata | D. nebrascense-E. jenneyi to Upper Campanian | 0.910 | 0.933 | 0.877 | 0.877 | 0.926 |
| L. subundata | B. reesidei-B. eliasi to Upper Campanian | 0.630 | 0.762 | 0.867 | 0.867 | 1.000 |
| C. nebrascensis | D. nebrascense-E. jenneyi to Upper Campanian | 0.995 | 0.994 | 0.991 | 0.991 | 0.705 |
| C. nebrascensis | B. reesidei-B. eliasi to Upper Campanian | 0.957 | 0.973 | 0.981 | 0.981 | 0.933 |
| C. nebrascensis | B. baculus-B. grandis to Lower Maastrichtian | 1.000 | 1.000 | 0.989 | 0.989 | 0.997 |
| C. nebrascensis | B. baculus-B. grandis to Maastrichtian | 0.915 | 0.881 | 0.899 | 0.899 | 0.867 |

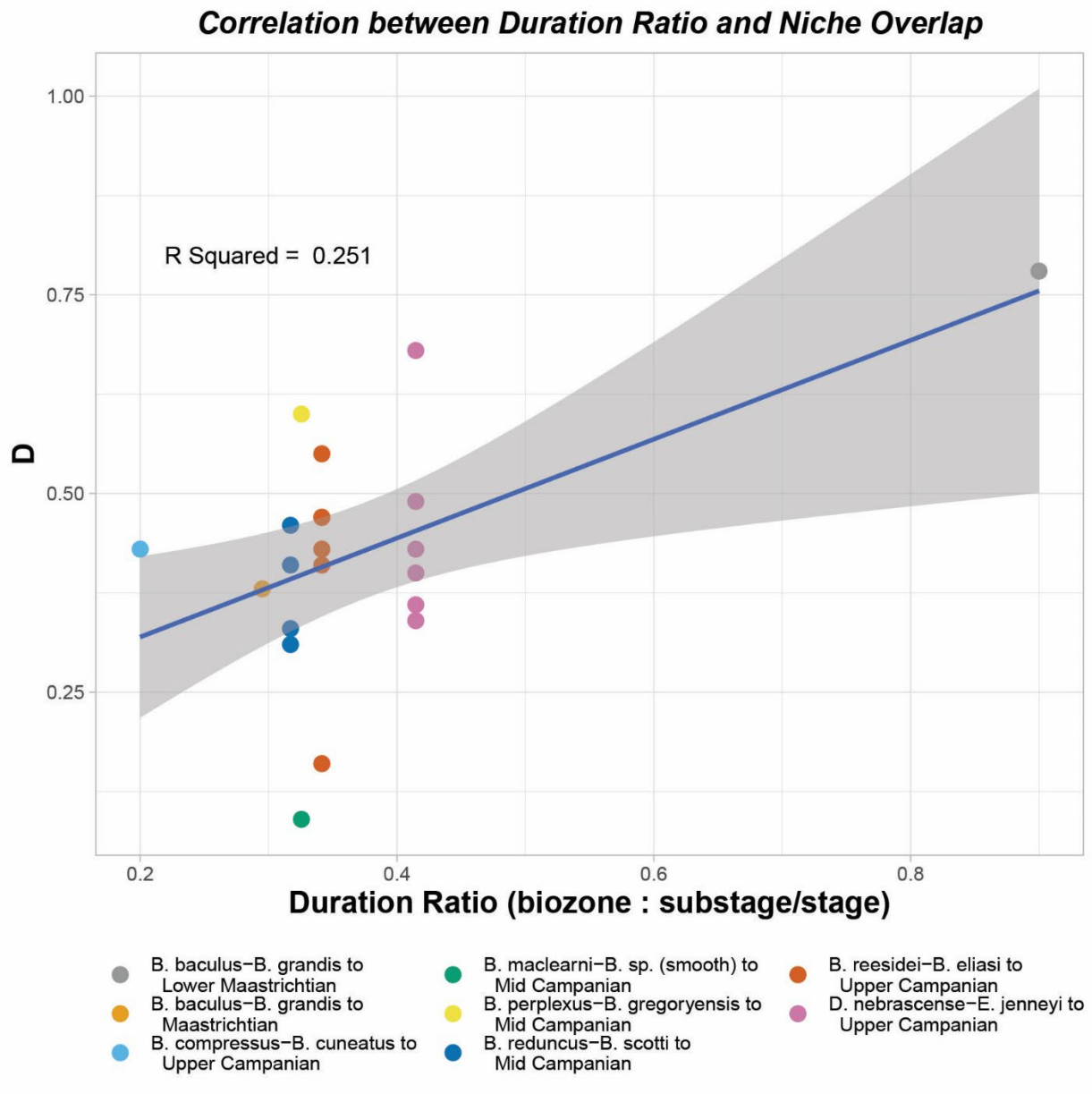



Figure S10. Correlation plot of duration ratio to niche overlap (D value). This plot represents non-averaged correlation data.

## Ingroup Comparison:

Table S13. Niche overlap values for *Inoceramus* species comparisons including I and D overlap values, and niche expansion, stability, and unfilling.

| Species | Interval | D value | I value | Expansion | Stability | Unfilling |
|---|---|---|---|---|---|---|
| *I. convexus vs. I. sublaevis* | *B. reduncus-B. scotti* | 0.721 | 0.779 | 0.674 | 0.326 | 0.505 |
| *I. convexus vs. I. saskatchewanensis* | *B. reduncus-B. scotti* | 0.518 | 0.633 | 0.217 | 0.783 | 0.540 |
| *I. sublaevis vs. I. saskatchewanensis* | *B. reduncus-B. scotti* | 0.491 | 0.585 | 0.231 | 0.769 | 0.383 |
| *I. convexus vs. I. barabini* | *D. nebrascense-E. jenneyi* | 0.106 | 0.289 | 0.589 | 0.411 | 0.101 |
| *I. convexus vs. I. sublaevis* | *B. reesidei-B. eliasi* | 0.088 | 0.254 | 0.674 | 0.326 | 0.505 |
| *I. barabini vs. I. sublaevis* | *D. nebrascense-E. jenneyi* | 0.796 | 0.851 | 0.186 | 0.814 | 0.495 |
| *I. barabini vs. I. sagensis* | *B. reesidei-B. eliasi* | 0.675 | 0.719 | 0.344 | 0.656 | 0.181 |
| *I. barabini vs. I. oblongus* | *B. reesidei-B. eliasi* | 0.545 | 0.621 | 0.406 | 0.594 | 0.295 |
| *I. oblongus vs. I. sagensis* | *B. reesidei-B. eliasi* | 0.777 | 0.845 | 0.125 | 0.875 | 0.087 |

Table S14. Results of greater niche equivalency test for *Inoceramus* species comparisons including p-values for I and D overlap (greater than random), and niche expansion (less than random), stability (greater than random), and unfilling (less than random).

| | | Greater Niche Equivalency Test | | | | |
|---|---|---|---|---|---|---|
| Species | Intervals Compared | D value | I Value | Expansion (lower) | Stability (higher) | Unfilling (lower) |
| *I. convexus vs. I. sublaevis* | *B. reduncus-B. scotti* | 0.912 | 0.882 | 0.926 | 0.926 | 0.810 |
| *I. convexus vs. I. saskatchewanensis* | *B. reduncus-B. scotti* | 0.695 | 0.663 | 0.381 | 0.381 | 0.884 |
| *I. sublaevis vs. I. saskatchewanensis* | *B. reduncus-B. scotti* | 0.677 | 0.741 | 0.617 | 0.617 | 0.749 |
| *I. convexus vs. I. barabini* | *D. nebrascense-E. jenneyi* | 0.866 | 0.884 | 0.934 | 0.934 | 0.303 |
| *I. convexus vs. I. sublaevis* | *D. nebrascense-E. jenneyi* | 0.912 | 0.882 | 0.926 | 0.926 | 0.810 |
| *I. barabini vs. I. sublaevis* | *D. nebrascense-E. jenneyi* | 0.066 | 0.092 | 0.202 | 0.202 | 0.591 |
| *I. barabini vs. I. sagensis* | *B. reesidei-B. eliasi* | 0.052 | 0.072 | 0.511 | 0.511 | 0.152 |
| *I. barabini vs. I. oblongus* | *B. reesidei-B. eliasi* | 0.130 | 0.186 | 0.589 | 0.589 | 0.357 |
| *I. oblongus vs. I. sagensis* | *B. reesidei-B. eliasi* | **0.012*** | **0.018*** | 0.158 | 0.158 | 0.132 |

Table S15. Results of lower niche equivalency test for *Inoceramus* species comparisons including p-values for I and D overlap (less than random), and niche expansion (greater than random), stability (less than random), and unfilling (greater than random).

Lower Niche Equivalency Test

| Species | Intervals Compared | D value | I Value | Expansion (higher) | Stability (lower) | Unfilling (higher) |
|---|---|---|---|---|---|---|
| *I. convexus vs. I. sublaevis* | *B. reduncus-B. scotti* | 0.068 | 0.112 | 0.060 | 0.060 | 0.162 |
| *I. convexus vs. I. saskatchewanensis* | *B. reduncus-B. scotti* | 0.317 | 0.319 | 0.689 | 0.689 | 0.140 |
| *I. sublaevis vs. I. saskatchewanensis* | *B. reduncus-B. scotti* | 0.281 | 0.253 | 0.347 | 0.347 | 0.240 |
| *I. convexus vs. I. barabini* | *D. nebrascense-E. jenneyi* | 0.122 | 0.110 | 0.076 | 0.076 | 0.669 |
| *I. convexus vs. I. sublaevis* | *D. nebrascense-E. jenneyi* | 0.068 | 0.112 | 0.060 | 0.060 | 0.162 |
| *I. barabini vs. I. sublaevis* | *D. nebrascense-E. jenneyi* | 0.930 | 0.910 | 0.798 | 0.798 | 0.443 |
| *I. barabini vs. I. sagensis* | *B. reesidei-B. eliasi* | 0.966 | 0.934 | 0.543 | 0.543 | 0.850 |
| *I. barabini vs. I. oblongus* | *B. reesidei-B. eliasi* | 0.862 | 0.828 | 0.371 | 0.371 | 0.629 |
| *I. oblongus vs. I. sagensis* | *B. reesidei-B. eliasi* | 0.986 | 0.980 | 0.832 | 0.832 | 0.866 |

Table S16. Results of greater niche similarity test for *Inoceramus* species comparisons including p-values for I and D overlap (greater than random), and niche expansion (less than random), stability (greater than random), and unfilling (less than random).

Greater Niche Similarity Test

| Species | Intervals Compared | D value | I value | Expansion (lower) | Stability (higher) | Unfilling (lower) |
|---|---|---|---|---|---|---|
| *I. convexus vs. I. sublaevis* | *B. reduncus-B. scotti* | 0.431 | 0.276 | 0.392 | 0.392 | 0.247 |
| *I. convexus vs. I. saskatchewanensis* | *B. reduncus-B. scotti* | 0.112 | 0.193 | 0.381 | 0.381 | 0.246 |
| *I. sublaevis vs. I. saskatchewanensis* | *B. reduncus-B. scotti* | 0.221 | 0.239 | 0.353 | 0.353 | 0.179 |
| *I. convexus vs. I. barabini* | *D. nebrascense-E. jenneyi* | 0.429 | 0.361 | 0.209 | 0.209 | 0.232 |
| *I. convexus vs. I. sublaevis* | *D. nebrascense-E. jenneyi* | 0.431 | 0.276 | 0.392 | 0.392 | 0.247 |
| *I. barabini vs. I. sublaevis* | *D. nebrascense-E. jenneyi* | **0.001*** | **0.001*** | 0.251 | 0.251 | **0.045*** |
| *I. barabini vs. I. sagensis* | *B. reesidei-B. eliasi* | **0.015*** | **0.028*** | **0.026*** | **0.026*** | 0.164 |
| *I. barabini vs. I. oblongus* | *B. reesidei-B. eliasi* | **0.038*** | 0.068 | 0.072 | 0.072 | 0.206 |
| *I. oblongus vs. I. sagensis* | *B. reesidei-B. eliasi* | **0.048*** | **0.043*** | 0.058 | 0.058 | **0.037*** |

Table S17. Results of lower niche similarity test for *Inoceramus* species comparisons including p-values for I and D overlap (less than random), and niche expansion (greater than random), stability (less than random), and unfilling (greater than random).

| Lower Niche Similarity Test | | | | | | |
|---|---|---|---|---|---|---|
| Species | Intervals Compared | D value | I value | Expansion (higher) | Stability (lower) | Unfilling (higher) |
| *I. convexus vs. I. sublaevis* | *B. reduncus-B. scotti* | 0.554 | 0.704 | 0.618 | 0.618 | 0.777 |
| *I. convexus vs. I. saskatchewanensis* | *B. reduncus-B. scotti* | 0.890 | 0.817 | 0.616 | 0.616 | 0.764 |
| *I. sublaevis vs. I. saskatchewanensis* | *B. reduncus-B. scotti* | 0.780 | 0.772 | 0.661 | 0.661 | 0.822 |
| *I. convexus vs. I. barabini* | *D. nebrascense-E. jenneyi* | 0.592 | 0.672 | 0.806 | 0.806 | 0.787 |
| *I. convexus vs. I. sublaevis* | *D. nebrascense-E. jenneyi* | 0.554 | 0.704 | 0.618 | 0.618 | 0.777 |
| *I. barabini vs. I. sublaevis* | *D. nebrascense-E. jenneyi* | 1.000 | 0.999 | 0.769 | 0.769 | 0.947 |
| *I. barabini vs. I. sagensis* | *B. reesidei-B. eliasi* | 0.985 | 0.977 | 0.975 | 0.975 | 0.864 |
| *I. barabini vs. I. oblongus* | *B. reesidei-B. eliasi* | 0.959 | 0.924 | 0.916 | 0.916 | 0.785 |
| *I. oblongus vs. I. sagensis* | *B. reesidei-B. eliasi* | 0.945 | 0.952 | 0.937 | 0.937 | 0.955 |

## Comparison taxa Comparisons:

Table S18. Niche overlap values for *Inoceramus* species to the genus comparisons including I and D overlap values, and niche expansion, stability, and unfilling.

| Species Compared | Interval | D value | I value | Expansion | Stability | Unfilling |
|---|---|---|---|---|---|---|
| *C. imbricatula vs. I. convexus* | *B. reduncus-B. scotti* | 0.247 | 0.425 | 0.665 | 0.335 | 0.164 |
| *C. imbricatula vs. I. sublaevis* | *B. reduncus-B. scotti* | 0.383 | 0.525 | 0.473 | 0.527 | 0.076 |
| *C. imbricatula vs. I. saskatchewanensis* | *B. reduncus-B. scotti* | 0.443 | 0.654 | 0.368 | 0.632 | 0.085 |
| *C. nebrascensis vs. I. barabini* | *D. nebrascense-E. jenneyi* | 0.832 | 0.918 | 0.069 | 0.931 | 0.092 |
| *C. imbricatula vs. I. convexus* | *D. nebrascense-E. jenneyi* | 0.147 | 0.267 | 0.280 | 0.720 | 0.582 |
| *C. imbricatula vs. I. sublaevis* | *D. nebrascense-E. jenneyi* | 0.207 | 0.290 | 0.784 | 0.216 | 0.834 |
| *C. nebrascensis vs. I. barabini* | *D. nebrascense-E. jenneyi* | 0.832 | 0.918 | 0.069 | 0.931 | 0.092 |
| *C. nebrascensis vs. I. convexus* | *D. nebrascense-E. jenneyi* | 0.072 | 0.248 | 0.088 | 0.912 | 0.668 |
| *C. nebrascensis vs. I. sagensis* | *D. nebrascense-E. jenneyi* | 0.854 | 0.925 | 0.116 | 0.884 | 0.431 |
| *L. subundata vs. I. barabini* | *D. nebrascense-E. jenneyi* | 0.700 | 0.791 | 0.167 | 0.833 | 0.093 |
| *L. subundata vs. I. convexus* | *D. nebrascense-E. jenneyi* | 0.104 | 0.260 | 0.187 | 0.813 | 0.604 |
| *L. subundata vs. I. sublaevis* | *D. nebrascense-E. jenneyi* | 0.571 | 0.661 | 0.426 | 0.574 | 0.658 |
| *L. subundata vs. I. sagensis* | *B. compressus-B. cuneatus* | 0.294 | 0.467 | 0.362 | 0.638 | 0.114 |
| *C. nebrascensis vs. I. barabini* | *B. reesidei-B. eliasi* | 0.611 | 0.709 | 0.062 | 0.938 | 0.371 |
| *C. nebrascensis vs. I. sagensis* | *B. reesidei-B. eliasi* | 0.655 | 0.771 | 0.084 | 0.916 | 0.248 |
| *C. nebrascensis vs. I. oblongus* | *B. reesidei-B. eliasi* | 0.547 | 0.700 | 0.126 | 0.874 | 0.315 |
| *C. imbricatula vs. I. barabini* | *B. reesidei-B. eliasi* | 0.599 | 0.659 | 0.137 | 0.863 | 0.396 |
| *C. imbricatula vs. I. sagensis* | *B. reesidei-B. eliasi* | 0.786 | 0.852 | 0.034 | 0.966 | 0.133 |
| *C. imbricatula vs. I. oblongus* | *B. reesidei-B. eliasi* | 0.710 | 0.834 | 0.021 | 0.979 | 0.160 |

Table S19. Results of greater niche equivalency test for *Inoceramus* species to comparison species comparisons including p-values for I and D overlap (greater than random), and niche expansion (less than random), stability (greater than random), and unfilling (less than random).

| Greater Equivalency Test (p-values) | | | | | | |
|---|---|---|---|---|---|---|
| Species | Interval | D value | I value | Expansion (lower) | Stability (higher) | Unfilling (lower) |
| *C. imbricatula vs. I. convexus* | *B. reduncus-B. scotti* | 0.800 | 0.743 | 0.884 | 0.884 | 0.345 |
| *C. imbricatula vs. I. sublaevis* | *B. reduncus-B. scotti* | 0.549 | 0.593 | 0.667 | 0.667 | 0.263 |
| *C. imbricatula vs. I. saskatchewanensis* | *B. reduncus-B. scotti* | 0.669 | 0.607 | 0.495 | 0.495 | 0.226 |
| *C. nebrascensis vs. I. barabini* | *D. nebrascense-E. jenneyi* | **0.012*** | **0.006*** | 0.116 | 0.116 | 0.182 |
| *C. imbricatula vs. I. convexus* | *D. nebrascense-E. jenneyi* | 0.788 | 0.888 | 0.717 | 0.717 | 0.908 |
| *C. imbricatula vs. I. sublaevis* | *D. nebrascense-E. jenneyi* | 0.782 | 0.800 | 0.946 | 0.946 | 0.958 |
| *C. nebrascensis vs. I. barabini* | *D. nebrascense-E. jenneyi* | **0.012*** | **0.006*** | 0.116 | 0.116 | 0.182 |
| *C. nebrascensis vs. I. convexus* | *D. nebrascense-E. jenneyi* | 0.900 | 0.814 | 0.070 | 0.070 | 0.944 |
| *C. nebrascensis vs. I. sagensis* | *D. nebrascense-E. jenneyi* | 0.068 | 0.054 | **0.042*** | **0.042*** | 0.545 |
| *L. subundata vs. I. barabini* | *D. nebrascense-E. jenneyi* | **0.026*** | **0.040*** | 0.337 | 0.337 | 0.182 |
| *L. subundata vs. I. convexus* | *D. nebrascense-E. jenneyi* | 0.850 | 0.886 | 0.451 | 0.451 | 0.906 |
| *L. subundata vs. I. sublaevis* | *D. nebrascense-E. jenneyi* | 0.311 | 0.359 | 0.561 | 0.561 | 0.794 |
| *L. subundata vs. I. sagensis* | *B. compressus-B. cuneatus* | 0.599 | 0.693 | 0.597 | 0.597 | 0.591 |
| *C. nebrascensis vs. I. barabini* | *B. reesidei-B. eliasi* | 0.070 | 0.096 | 0.112 | 0.112 | 0.539 |
| *C. nebrascensis vs. I. sagensis* | *B. reesidei-B. eliasi* | 0.060 | **0.042*** | 0.178 | 0.178 | 0.443 |
| *C. nebrascensis vs. I. oblongus* | *B. reesidei-B. eliasi* | 0.259 | 0.214 | 0.303 | 0.303 | 0.463 |
| *C. imbricatula vs. I. barabini* | *B. reesidei-B. eliasi* | 0.112 | 0.164 | 0.148 | 0.148 | 0.593 |
| *C. imbricatula vs. I. sagensis* | *B. reesidei-B. eliasi* | **0.014*** | **0.010*** | **0.030*** | **0.030*** | 0.190 |
| *C. imbricatula vs. I. oblongus* | *B. reesidei-B. eliasi* | **0.046*** | **0.020*** | **0.026*** | **0.026*** | 0.246 |

Table S20. Results of lower niche equivalency test for *Inoceramus* species to comparison species comparisons including p-values for I and D overlap (less than random), and niche expansion (greater than random), stability (less than random), and unfilling (greater than random).

| | | Lower Equivalency Test (p-values) | | | | |
|---|---|---|---|---|---|---|
| Species | Interval | D value | I value | Expansion (higher) | Stability (lower) | Unfilling (higher) |
| *C. imbricatula vs. I. convexus* | *B. reduncus-B. scotti* | 0.226 | 0.277 | 0.138 | 0.138 | 0.679 |
| *C. imbricatula vs. I. sublaevis* | *B. reduncus-B. scotti* | 0.503 | 0.489 | 0.345 | 0.345 | 0.715 |
| *C. imbricatula vs. I. saskatchewanensis* | *B. reduncus-B. scotti* | 0.323 | 0.425 | 0.525 | 0.525 | 0.764 |
| *C. nebrascensis vs. I. barabini* | *D. nebrascense-E. jenneyi* | 0.994 | 1.000 | 0.880 | 0.880 | 0.854 |
| *C. imbricatula vs. I. convexus* | *D. nebrascense-E. jenneyi* | 0.192 | 0.098 | 0.303 | 0.303 | 0.082 |
| *C. imbricatula vs. I. sublaevis* | *D. nebrascense-E. jenneyi* | 0.214 | 0.176 | 0.054 | 0.054 | **0.034*** |
| *C. nebrascensis vs. I. barabini* | *D. nebrascense-E. jenneyi* | 0.994 | 1.000 | 0.880 | 0.880 | 0.854 |
| *C. nebrascensis vs. I. convexus* | *D. nebrascense-E. jenneyi* | 0.068 | 0.148 | 0.932 | 0.932 | **0.044*** |
| *C. nebrascensis vs. I. sagensis* | *D. nebrascense-E. jenneyi* | 0.932 | 0.942 | 0.964 | 0.964 | 0.511 |
| *L. subundata vs. I. barabini* | *D. nebrascense-E. jenneyi* | 0.982 | 0.952 | 0.689 | 0.689 | 0.802 |
| *L. subundata vs. I. convexus* | *D. nebrascense-E. jenneyi* | 0.184 | 0.122 | 0.567 | 0.567 | 0.098 |
| *L. subundata vs. I. sublaevis* | *D. nebrascense-E. jenneyi* | 0.641 | 0.601 | 0.443 | 0.443 | 0.196 |
| *L. subundata vs. I. sagensis* | *B. compressus-B. cuneatus* | 0.421 | 0.357 | 0.417 | 0.417 | 0.429 |
| *C. nebrascensis vs. I. barabini* | *B. reesidei-B. eliasi* | 0.932 | 0.908 | 0.866 | 0.866 | 0.421 |
| *C. nebrascensis vs. I. sagensis* | *B. reesidei-B. eliasi* | 0.932 | 0.940 | 0.848 | 0.848 | 0.581 |
| *C. nebrascensis vs. I. oblongus* | *B. reesidei-B. eliasi* | 0.735 | 0.810 | 0.737 | 0.737 | 0.549 |
| *C. imbricatula vs. I. barabini* | *B. reesidei-B. eliasi* | 0.900 | 0.858 | 0.848 | 0.848 | 0.423 |
| *C. imbricatula vs. I. sagensis* | *B. reesidei-B. eliasi* | 0.988 | 0.990 | 0.974 | 0.974 | 0.794 |
| *C. imbricatula vs. I. oblongus* | *B. reesidei-B. eliasi* | 0.946 | 0.982 | 0.982 | 0.982 | 0.794 |

Table S21. Results of greater niche similarity test for *Inoceramus* species to comparison species comparisons including p-values for I and D overlap (greater than random), and niche expansion (less than random), stability (greater than random), and unfilling (less than random).

| Greater Similarity Test (p-values) | | | | | | |
|---|---|---|---|---|---|---|
| Species | Interval | D value | I value | Expansion (lower) | Stability (higher) | Unfilling (lower) |
| *C. imbricatula vs. I. convexus* | *B. reduncus-B. scotti* | 0.411 | 0.399 | 0.295 | 0.295 | 0.325 |
| *C. imbricatula vs. I. sublaevis* | *B. reduncus-B. scotti* | 0.306 | 0.362 | 0.147 | 0.147 | 0.294 |
| *C. imbricatula vs. I. saskatchewanensis* | *B. reduncus-B. scotti* | 0.112 | 0.054 | 0.074 | 0.074 | **0.038*** |
| *C. nebrascensis vs. I. barabini* | *D. nebrascense-E. jenneyi* | **0.001*** | **0.001*** | **0.001*** | **0.001*** | 0.199 |
| *C. imbricatula vs. I. convexus* | *D. nebrascense-E. jenneyi* | 0.398 | 0.349 | 0.305 | 0.305 | 0.304 |
| *C. imbricatula vs. I. sublaevis* | *D. nebrascense-E. jenneyi* | 0.356 | 0.332 | 0.626 | 0.626 | 0.536 |
| *C. nebrascensis vs. I. barabini* | *D. nebrascense-E. jenneyi* | **0.001*** | **0.001*** | **0.001*** | **0.001*** | 0.199 |
| *C. nebrascensis vs. I. convexus* | *D. nebrascense-E. jenneyi* | 0.449 | 0.300 | **0.015*** | **0.015*** | 0.350 |
| *C. nebrascensis vs. I. sublaevis* | *D. nebrascense-E. jenneyi* | **0.015*** | **0.006*** | **0.045*** | **0.045*** | 0.267 |
| *L. subundata vs. I. barabini* | *D. nebrascense-E. jenneyi* | 0.101 | 0.099 | 0.138 | 0.138 | 0.115 |
| *L. subundata vs. I. convexus* | *D. nebrascense-E. jenneyi* | 0.501 | 0.422 | 0.356 | 0.356 | 0.305 |
| *L. subundata vs. I. sublaevis* | *D. nebrascense-E. jenneyi* | **0.028*** | **0.050*** | 0.448 | 0.448 | 0.232 |
| *L. subundata vs. I. sagensis* | *B. compressus-B. cuneatus* | 0.265 | 0.234 | 0.158 | 0.158 | 0.167 |
| *C. nebrascensis vs. I. barabini* | *B. reesidei-B. eliasi* | **0.006*** | **0.008*** | 0.098 | 0.098 | **0.029*** |
| *C. nebrascensis vs. I. sagensis* | *B. reesidei-B. eliasi* | 0.067 | **0.047*** | **0.045*** | **0.045*** | 0.122 |
| *C. nebrascensis vs. I. oblongus* | *B. reesidei-B. eliasi* | 0.101 | 0.062 | 0.083 | 0.083 | 0.149 |
| *C. imbricatula vs. I. barabini* | *B. reesidei-B. eliasi* | **0.017*** | **0.039*** | 0.172 | 0.172 | **0.028*** |
| *C. imbricatula vs. I. sagensis* | *B. reesidei-B. eliasi* | **0.044*** | **0.049*** | **0.035*** | **0.035*** | **0.038*** |
| *C. imbricatula vs. I. oblongus* | *B. reesidei-B. eliasi* | 0.076 | 0.053 | **0.018*** | **0.018*** | 0.054 |

Table S22. Results of lower niche similarity test for *Inoceramus* species to comparison species comparisons including p-values for I and D overlap (less than random), and niche expansion (greater than random), stability (less than random), and unfilling (greater than random).

| | | Lower Similarity Test (p-values) | | | | |
|---|---|---|---|---|---|---|
| Species | Interval | D value | I value | Expansion (higher) | Stability (lower) | Unfilling (higher) |
| *C. imbricatula vs. I. convexus* | *B. reduncus-B. scotti* | 0.601 | 0.608 | 0.727 | 0.727 | 0.667 |
| *C. imbricatula vs. I. sublaevis* | *B. reduncus-B. scotti* | 0.716 | 0.683 | 0.844 | 0.844 | 0.742 |
| *C. imbricatula vs. I. saskatchewanensis* | *B. reduncus-B. scotti* | 0.901 | 0.953 | 0.937 | 0.937 | 0.954 |
| *C. nebrascensis vs. I. barabini* | *D. nebrascense-E. jenneyi* | 1.000 | 1.000 | 0.999 | 0.999 | 0.771 |
| *C. imbricatula vs. I. convexus* | *D. nebrascense-E. jenneyi* | 0.615 | 0.678 | 0.725 | 0.725 | 0.707 |
| *C. imbricatula vs. I. sublaevis* | *D. nebrascense-E. jenneyi* | 0.628 | 0.637 | 0.358 | 0.358 | 0.435 |
| *C. nebrascensis vs. I. barabini* | *D. nebrascense-E. jenneyi* | 1.000 | 1.000 | 0.999 | 0.999 | 0.771 |
| *C. nebrascensis vs. I. convexus* | *D. nebrascense-E. jenneyi* | 0.536 | 0.725 | 0.985 | 0.985 | 0.623 |
| *C. nebrascensis vs. I. sublaevis* | *D. nebrascense-E. jenneyi* | 0.993 | 0.998 | 0.952 | 0.952 | 0.719 |
| *L. subundata vs. I. barabini* | *D. nebrascense-E. jenneyi* | 0.910 | 0.907 | 0.863 | 0.863 | 0.907 |
| *L. subundata vs. I. convexus* | *D. nebrascense-E. jenneyi* | 0.517 | 0.566 | 0.650 | 0.650 | 0.698 |
| *L. subundata vs. I. sublaevis* | *D. nebrascense-E. jenneyi* | 0.980 | 0.942 | 0.535 | 0.535 | 0.762 |
| *L. subundata vs. I. sagensis* | *B. compressus-B. cuneatus* | 0.707 | 0.727 | 0.822 | 0.822 | 0.824 |
| *C. nebrascensis vs. I. barabini* | *B. reesidei-B. eliasi* | 0.988 | 0.989 | 0.875 | 0.875 | 0.976 |
| *C. nebrascensis vs. I. sagensis* | *B. reesidei-B. eliasi* | 0.939 | 0.955 | 0.948 | 0.948 | 0.907 |
| *C. nebrascensis vs. I. oblongus* | *B. reesidei-B. eliasi* | 0.875 | 0.923 | 0.923 | 0.923 | 0.846 |
| *C. imbricatula vs. I. barabini* | *B. reesidei-B. eliasi* | 0.972 | 0.956 | 0.827 | 0.827 | 0.969 |
| *C. imbricatula vs. I. sagensis* | *B. reesidei-B. eliasi* | 0.927 | 0.924 | 0.939 | 0.939 | 0.950 |
| *C. imbricatula vs. I. oblongus* | *B. reesidei-B. eliasi* | 0.933 | 0.949 | 0.967 | 0.967 | 0.956 |

## Species to Genus (all occurrences) Comparison

Table S23. Niche overlap values for *Inoceramus* species to the genus comparisons including I and D overlap values, and niche expansion, stability, and unfilling.

| Species | Interval | D value | I value | Expansion | Stability | Unfilling |
|---|---|---|---|---|---|---|
| *I. convexus* | *B. reduncus-B. scotti* | 0.721 | 0.826 | 0.112 | 0.888 | 0.078 |
| *I. sublaevis* | *B. reduncus-B. scotti* | 0.740 | 0.832 | 0.087 | 0.913 | 0.020 |
| *I. saskatchewanensis* | *B. reduncus-B. scotti* | 0.679 | 0.844 | 0.348 | 0.652 | 0.018 |
| *I. convexus* | *D. nebrascense-E. jenneyi* | 0.221 | 0.466 | 0.512 | 0.488 | 0.000 |
| *I. sublaevis* | *D. nebrascense-E. jenneyi* | 0.427 | 0.549 | 0.680 | 0.320 | 0.237 |
| *I. barabini* | *D. nebrascense-E. jenneyi* | 0.509 | 0.643 | 0.224 | 0.776 | 0.119 |
| *I. barabini* | *B. reesidei-B. eliasi* | 0.465 | 0.676 | 0.481 | 0.519 | 0.000 |
| *I. sagensis* | *B. reesidei-B. eliasi* | 0.477 | 0.687 | 0.302 | 0.698 | 0.000 |
| *I. oblongus* | *B. reesidei-B. eliasi* | 0.408 | 0.636 | 0.334 | 0.666 | 0.000 |

Table S24. Results of greater niche equivalency test for *Inoceramus* species to genus (all possible occurrences) comparisons including p-values for I and D overlap (greater than random), and niche expansion (less than random), stability (greater than random), and unfilling (less than random).

| | | Greater Niche Equivalency Test | | | | |
|---|---|---|---|---|---|---|
| Species | Intervals Compared | D value | I value | Expansion (lower) | Stability (higher) | Unfilling (lower) |
| *I. convexus* | *B. reduncus-B. scotti* | 0.132 | 0.152 | 0.160 | 0.160 | 0.766 |
| *I. sublaevis* | *B. reduncus-B. scotti* | 0.126 | 0.190 | 0.150 | 0.150 | 0.473 |
| *I. saskatchewanensis* | *B. reduncus-B. scotti* | 0.359 | 0.263 | 0.725 | 0.725 | 0.497 |
| *I. convexus* | *D. nebrascense-E. jenneyi* | 0.697 | 0.655 | 0.735 | 0.735 | 0.399 |
| *I. sublaevis* | *D. nebrascense-E. jenneyi* | 0.441 | 0.565 | 0.930 | 0.930 | 0.990 |
| *I. barabini* | *D. nebrascense-E. jenneyi* | 0.347 | 0.375 | 0.415 | 0.415 | 0.874 |
| *I. barabini* | *B. reesidei-B. eliasi* | 0.511 | 0.435 | 0.816 | 0.816 | 0.307 |
| *I. sagensis* | *B. reesidei-B. eliasi* | 0.527 | 0.467 | 0.535 | 0.535 | 0.295 |
| *I. oblongus* | *B. reesidei-B. eliasi* | 0.607 | 0.519 | 0.485 | 0.485 | 0.377 |

Table S25. Results of lower niche equivalency test for *Inoceramus* species to genus (all possible occurrences) comparisons including p-values for I and D overlap (less than random), and niche expansion (greater than random), stability (less than random), and unfilling (greater than random).

Lower Niche Equivalency Test

| Species | Intervals Compared | D value | I value | Expansion (higher) | Stability (lower) | Unfilling (higher) |
|---|---|---|---|---|---|---|
| *I. convexus* | *B. reduncus-B. scotti* | 0.886 | 0.860 | 0.854 | 0.854 | 0.246 |
| *I. sublaevis* | *B. reduncus-B. scotti* | 0.886 | 0.776 | 0.852 | 0.852 | 0.531 |
| *I. saskatchewanensis* | *B. reduncus-B. scotti* | 0.631 | 0.707 | 0.279 | 0.279 | 0.489 |
| *I. convexus* | *D. nebrascense-E. jenneyi* | 0.269 | 0.307 | 0.246 | 0.246 | 1.000 |
| *I. sublaevis* | *D. nebrascense-E. jenneyi* | 0.595 | 0.485 | 0.056 | 0.056 | **0.006*** |
| *I. barabini* | *D. nebrascense-E. jenneyi* | 0.687 | 0.673 | 0.611 | 0.611 | 0.134 |
| *I. barabini* | *B. reesidei-B. eliasi* | 0.469 | 0.561 | 0.188 | 0.188 | 1.000 |
| *I. sagensis* | *B. reesidei-B. eliasi* | 0.491 | 0.567 | 0.443 | 0.443 | 1.000 |
| *I. oblongus* | *B. reesidei-B. eliasi* | 0.447 | 0.533 | 0.533 | 0.533 | 1.000 |

Table S26. Results of greater niche similarity test for *Inoceramus* species to genus (all possible occurrences) comparisons including p-values for I and D overlap (greater than random), and niche expansion (less than random), stability (greater than random), and unfilling (less than random).

Greater Niche Similarity Test

| Species | Intervals Compared | D value | I value | Expansion (lower) | Stability (higher) | Unfilling (lower) |
|---|---|---|---|---|---|---|
| *I. convexus* | *B. reduncus-B. scotti* | 0.114 | 0.141 | 0.225 | 0.225 | 0.100 |
| *I. sublaevis* | *B. reduncus-B. scotti* | 0.165 | 0.194 | 0.144 | 0.144 | 0.125 |
| *I. saskatchewanensis* | *B. reduncus-B. scotti* | 0.060 | **0.047*** | 0.145 | 0.145 | 0.148 |
| *I. convexus* | *D. nebrascense-E. jenneyi* | 0.292 | 0.262 | 0.202 | 0.202 | 0.238 |
| *I. sublaevis* | *D. nebrascense-E. jenneyi* | 0.137 | 0.215 | 0.347 | 0.347 | 0.442 |
| *I. barabini* | *D. nebrascense-E. jenneyi* | 0.152 | 0.154 | 0.170 | 0.170 | **0.047*** |
| *I. barabini* | *B. reesidei-B. eliasi* | **0.033*** | **0.022*** | **0.041*** | **0.041*** | 0.157 |
| *I. sagensis* | *B. reesidei-B. eliasi* | 0.206 | 0.164 | 0.089 | 0.089 | 0.097 |
| *I. oblongus* | *B. reesidei-B. eliasi* | 0.212 | 0.174 | 0.076 | 0.076 | 0.068 |

Table S27. Results of lower niche similarity test for *Inoceramus* species to genus (all possible occurrences) comparisons including p-values for I and D overlap (less than random), and niche expansion (greater than random), stability (less than random), and unfilling (greater than random).

| Lower Niche Similarity Test | | | | | | |
|---|---|---|---|---|---|---|
| Species | Intervals Compared | D value | I value | Expansion (higher) | Stability (lower) | Unfilling (higher) |
| *I. convexus* | *B. reduncus-B. scotti* | 0.907 | 0.880 | 0.764 | 0.764 | 0.923 |
| *I. sublaevis* | *B. reduncus-B. scotti* | 0.858 | 0.829 | 0.869 | 0.869 | 0.877 |
| *I. saskatchewanensis* | *B. reduncus-B. scotti* | 0.936 | 0.958 | 0.831 | 0.831 | 0.825 |
| *I. convexus* | *D. nebrascense-E. jenneyi* | 0.729 | 0.763 | 0.836 | 0.836 | 1.000 |
| *I. sublaevis* | *D. nebrascense-E. jenneyi* | 0.883 | 0.804 | 0.672 | 0.672 | 0.558 |
| *I. barabini* | *D. nebrascense-E. jenneyi* | 0.840 | 0.849 | 0.831 | 0.831 | 0.952 |
| *I. barabini* | *B. reesidei-B. eliasi* | 0.972 | 0.981 | 0.965 | 0.965 | 1.000 |
| *I. sagensis* | *B. reesidei-B. eliasi* | 0.827 | 0.860 | 0.906 | 0.906 | 1.000 |
| *I. oblongus* | *B. reesidei-B. eliasi* | 0.779 | 0.827 | 0.900 | 0.900 | 1.000 |

References

Wang, Y., O. V. Akeju, and T. Zhao. 2017: Interpolation of spatially varying but sparsely measured geo-data: A comparative study. Engineering Geology 231:200–217.

Webster, R., and M. A. Oliver. 2007: Geostatistics for Environmental Scientists. John Wiley & Sons, Hoboken, New York., p.

Zhao, T., Y. Hu, and Y. Wang. 2018: Statistical interpretation of spatially varying 2D geo-data from sparse measurements using Bayesian compressive sampling. Engineering Geology 246:162–175.

Appendix C-3. Detailed References for Environmental Raster Layers

Sedimentological Literature References:

Baltz, J.E.H., 1962, Stratigraphy and geologic structure of uppermost Cretaceous and Tertiary rocks of the east-central part of the San Juan Basin, New Mexico: 1–294 p.

Baltz, E.H., 1967, Stratigraphy and regional tectonic implications of part of Upper Cretaceous and Tertiary rocks East-Central San Juan Basin New Mexico: Geological Survey Professional Paper, v. 552, p. 1–101.

Bertog, J., Huff, W., and Maritin, J.E., 2007, Geochemical and mineralogical recognition of the bentonites in the lower Pierre Shale Group and their use in regional stratigraphic correlation: The Geological Society of America Special Papers, v. 427, p. 23–50.

Bishop, G.A., 1973, Geology, stratigraphy, and biostratigraphy of the north end of the Cedar Creek Anticline, Dawson County, Montana: Montana Bureau of Mines and Geology Special Publication, v. 61.

Braun, R.E., 1983, The transition from the Judith River Formation to the Bearpaw Shale (Campanian), north-central Montana: iii–66 p.

Caldwell, J.W., 1953, Stratigraphic analysis of the Upper Cretaceous Pictured Cliffs Sandstone in Northeastern San Juan County, New Mexico: 1–59 p.

Chan, M.A., Newman, S.L., and May, F.E., 1991, Deltaic and shelf deposits in the Cretaceous Blackhawk Formation and Mancos Shale, Grand County, Utah: Utah Geological Survey Miscellaneous Publication, v. 91–6, p. 1–83.

Crockford, M.B.B., 1949, Oldman and Foremost Formations of Southern Alberta: Bulletin of the American Association of Petroleum Geologists, v. 33, p. 500–510, http://pubs.geoscienceworld.org/aapgbull/article-pdf/33/4/500/4368701/aapg_1949_0033_0004_0500.pdf.

Daly, D.J., 1984, Stratigraphy and depositional environments of the Fox Hills Formation (Upper Cretaceous), Bowman County, North Dakota: ii–288 p., https://commons.und.edu/theses/67.

Fisher, D.J., Erdmann, C.E., and Reeside, J.B., 1960, Cretaceous and Tertiary Formations of the Book Cliffs Carbon, Emery, and Grand Counties, Utah, and Garfield and Mesa Counties, Colorado: Geological Survey Professional Paper, v. 332, p. 1–80.

Fitter, F.L., 1958, Stratigraphy and Structure of the French Mesa Area, Rio Arriba County, New Mexico: 1–66 p., https://digitalrepository.unm.edu/eps_etds/148.

Fuentes, F., DeCelles, P.G., Constenius, K.N., and Gehrels, G.E., 2011a, Evolution of the cordilleran foreland basin system in northwestern montana, U.S.A.: Bulletin of the Geological Society of America, v. 123, p. 507–533, doi:10.1130/B30204.1.

Fuentes, F., DeCelles, P.G., Constenius, K.N., and Gehrels, G.E., 2011b, Evolution of the cordilleran foreland basin system in northwestern montana, U.S.A.: Bulletin of the Geological Society of America, v. 123, p. 507–533, doi:10.1130/B30204.1.

Gilbert, M.M., Bamforth, E.L., Buatois, L.A., and Renaut, R.W., 2018, Paleoecology and sedimentology of a vertebrate microfossil assemblage from the easternmost Dinosaur Park Formation (Late Cretaceous, Upper Campanian,) Saskatchewan, Canada: Reconstructing diversity in a coastal ecosystem: Palaeogeography, Palaeoclimatology, Palaeoecology, v. 495, p. 227–244, doi:10.1016/j.palaeo.2018.01.016.

Gilbert, M.M., Buatois, L.A., and Renaut, R.W., 2019, Ichnology and depositional environments of the Upper Cretaceous Dinosaur Park – Bearpaw formation transition in the Cypress Hills region of Southwestern Saskatchewan, Canada: Cretaceous Research, v. 98, p. 189–210, doi:10.1016/j.cretres.2018.12.017.

Gill, J.R., 1974, Stratigraphic sections of the Mesaverde Group, Lewis Shale, Fox Hills Formation, and Medicine Bow Formation, Carbon County, Wyoming: U.S. Geological Survey Open File Report, v. 74–1038, p. Plate I-Plate IV.

Gill, J.R., and Cobban, W.A., 1966, The Red Section of the Upper Cretaceous Pierre Shale in Wyoming: Geological Survey Professional Paper, v. 393-A, p. 1–76.

Gill, J.R., Merewether, E.A., and Cobban, W.A., 1970, Stratigraphy and Nomenclature of some Upper Cretaceous and Lower Tertiary rocks in South-Central Wyoming: Geological Survey Professional Paper, v. 667, p. 1–53.

Hanczarykt, P.A., and Gallagher, W.B., 2007, Stratigraphy and paleoecology of the middle Pierre Shale along the Missouri River, central South Dakota, in The Geology and Paleontology of the Late Cretaceous Marine Deposits of the Dakotas: Geological Society of America Special Paper, Geological Society, v. 427, p. 51–69.

Heald, K.C., 1926, The geology of the Ingomar Anticline, Treasure and Rosebud Counties, Montana: Contributions to Economic Geology, p. 1–38.

Hutchinson, J., 1974, Stratigraphy and paleontology of the Bisti Badlands Area, San Juan County, New Mexico: iv–219 p.

Johnson, R.C., Keefer, W.R., Keighin, C.W., and Finn, T.M., 1998, Detailed Outcrop Studies of the Upper Part of the Upper Cretaceous Cody Shale and the Upper Cretaceous Mesaverde, Meeteetse, and Lance Formations, Bighorn Basin, Wyoming: Cretaceous and Lower Tertiary Rocks of the Bighorn Basin, Wyoming and Montana; 49th Annual Field Conference Guidebook, p. 59–78.

Johnson, W.D., and Smith, H.R., 1964, Geology of the Winnett-Mosby area, Petroleum, Garfield, Rosebud, and Fergus Counties, Montana: Geological Survey Bulletin, v. 1149, p. 1–91.

Kiteley, L.W., 1976, Marine Shales and Sandstones in the Upper Cretaceous Pierre Shale at the Francis Ranch, Laramie County, Wyoming: The Mountain Geologist, v. 13, p. 1–19.

Kupsch, W.O., 1956, Geology of eastern Cypress Hills (Knollys and Dollard Quadrangles), Saskatchewan: Saskatchewan Department of Mineral Resources, Saskatchewan Geological Survey, Saskatchewan Industry and Resources (SIR) Report, v. 20, p. 1–30.

Lawson, R.J., 2019, The stratigraphy and paleontology of the Eastend Formation in Saskatchewan, Canada: i–179 p.

Leckie, R.M., Kirkland, J.I., and Elder, W.P., 2022, Stratigraphic framework and correlation of a principal reference section of the Mancos Shale (Upper Cretaceous), Mesa Verde, Colorado: Mesozoic Geology and Paleontology of the Four Corners Area. Anderson, O.; Kues, B.; Lucas, S; [eds.], New Mexico Geological Society 48th Annual Fall Field Conference Guidebook , p. 163–216, doi:10.56577/ffc-48.163.

Maberry, J.O., 1971, Sedimentary Features of the Blackhawk Formation (Cretaceous) in the Sunnyside District, Carbon County, Utah: Geological Survey Professional Paper, v. 688, p. 1–44.

Martin, J.E., Bertog, J.L., and Parris, D.C., 2007, Revised lithostratigraphy of the lower Pierre Shale Group (Campanian) of central South Dakota, including newly designated members: Special Paper of the Geological Society of America, v. 427, p. 9–21, doi:10.1130/2007.2427(02).

Mather, K.F., Gilluly, J., and Lusk, R.G., 1927, Geology and oil and gas prospects of Northeastern Colorado: Contributions to Economic Geology Part II, p. 65–124.

McNeil, D.H., and Caldwell, W.G.E., 1981, Cretaceous rocks and their Foraminifera in the Manitoba Escarpment: The Geological Association of Canada Special Paper, v. 21, p. 1–439.

Minor, K.P., Wroblewski, A., Steel, R.J., Olariu, C., and Crabaugh, J.P., 2022, Facies partitioning of fluvial, wave, and tidal influences across the shoreline-to-shelf architecture in the Western Interior Campanian Seaway, USA: Geological Society of London, Special Publications, v. 523, p. 487–523, doi:10.6084/m9.figshare.c.6237323.

Nwangwu, U., 1977, Depositional environments, Upper Pierre Shale, Denver Basin, Colorado, in Veal, H.K. ed., Exploration Frontiers of the Central and Southern Rockies, Denver, CO, Rocky Mountain Association of Geologists, p. 213–233.

Pettyjohn, W.A., 1967, New members of Upper Cretaceous Fox Hills Formation in South Dakota, representing delta deposits: AAPG Bulletin, v. 51, p. 1361–1367,

http://pubs.geoscienceworld.org/aapgbull/article-pdf/51/7/1361/4406986/aapg_1967_0051_0007_1361.pdf?casa_token=MK1cWLOl2VAAAAAA:90D_KoFaWlg.

Porter, K.W., 1976, Stratigraphic model for the Upper Cretaceous (Campanian) Hygiene Member, Pierre Shale West Denver Basin, Colorado [PhD]: Colorado School of Mines.

Reeside, J.B., and Knowlton, F.H., 1924, Upper Cretaceous and Tertiary Formations of the western part of the San Juan Basin Colorado and New Mexico: USGS Professional Paper, v. 134, p. 1–70.

Robinson, C.S., Mapel, W.J., and Bergendahl, M.H., 1964, Stratigraphy and structure of the northern and western flanks of The Black Hills Uplift, Wyoming Montana, and South Dakota: Geological Survey Professional Paper, v. 404, p. 1–135.

Roehler, H.W., Lujan, M., and Peck, D.L., 1990, Stratigraphy of the Mesaverde Group in the Central and Eastern Greater Green River Basin, Wyoming, Colorado, and Utah: U.S. Geological Survey Professional Paper, v. 1508, p. 1–52.

Scott, G.R., 1964, Geologic Map of the Northwest and Northeast Pueblo Quadrangles, Colorado.:

Sealey, P.L., and Lucas, S.G., 2022, Late Cretaceous (Campanian-Maastrichtian) Ammonites from the Pierre Shale, Raton Basin, Northeastern New Mexico, and Southeastern Colorado: Bulletin of the New Mexico Museum of Natural History and Science, v. 19, p. 1–183, www.nmnaturalhistory.org.

Seymour, D.L., 2012, High-resolution correlation of the Upper Cretaceous stratigraphy between the Book cliffs and the Western Henry Mountains Syncline, Utah, U.S.A.: 1–135 p., http://digitalcommons.unl.edu/geoscidisshttp://digitalcommons.unl.edu/geoscidiss/88.

Slipper, S.E., and Hunter, H.M., 1931, Stratigraphy of the Foremost, Pakowki, and Milk River Formations of Southern Plains of Alberta: AAPG Bulletin, v. 15, p. 1181, http://pubs.geoscienceworld.org/aapgbull/article-pdf/15/10/1181/4358904/aapg_1931_0015_0010_1181.pdf.

Stott, D.F., 1967, The Cretaceous Smoky Group, Rocky Mountain Foothills, Alberta and British Columbia: Geological Survey of Canada Bulletin, v. 132, p. 1.

Tsujita, C.J., 1995, Stratigraphy, taphonomy, and paleoecology of the Upper Cretaceous Bearpaw Formation in southern Alberta: xxii–375 p.

Waage, K.M., 1968, The type Fox Hills Formation, Cretaceous (Maestrichtian), South Dakota Part 1. Stratigraphy and paleoenvironments: Bulletin of the Peabody Museum of Natural History, Yale University, v. 27, p. 1–175, https://bioone.org/.

Warner, D.L., 1964, Mancos-Mesaverde (Upper Cretaceous) intertonguing relations Southeast Piceance Basin, Colorado: Bulletin of the American Association of Petroleum Geologists, v. 48, p. 1091–1107.

Seep Localities Literature Sources:

Dane, C.H., Pierce, W.G., and Reeside, J.B., 1936, The stratigraphy of the Upper Cretaceous rocks north of the Arkansas River in eastern Colorado: U.S. Geological Survey Professional Paper, v. 186-K, p. 207–232.

Darton, N.H., 1904, Newcastle folio, Wyoming-South Dakota (US Geological Atlas of the United States, Vol. 107).:

Darton, N.H., 1919, Newell folio, South Dakota (USGS Geological Atlas of the United States, Vol. 209).:

Darton, N.H., 1902, Oelrichs folio, South Dakota–Nebraska (USGS Geologic Atlas of the United States, Vol. 85).:

Darton, G.K., 1897, Pueblo folio, Colorado (USGS Geological Atlas of the United States, Vol. 36).:

Darton, N.H., and O'Harra, C.C., 1909, Belle Fourche folio, South Dakota (USGS Geologic Atlas of the United States, Vol. 164).:

Darton, N.H., and Paige, S., 1925, Central Black Hills folio, South Dakota (USGS Geologic Atlas of the United States, Vol. 219).:

Elias, M.K., 1931, The geology of the Wallace County Kansas: State Geological Survey of Kansas Bulletin, v. 18, p. 1–254.

Finlay, G.J., 1916, Colorado Spring folio, Colorado (USGS Geological Atlas of the United States, Vol. 203).:

Gilbert, G.K., 1897, Pueblo folio, Colorado (USGS Geologic Atlas of the United States, Vol. 36).:

Gilbert, G.K., and Gulliver, F.P., 1895, Tepee Buttes: Bulletin of the Geological Society of America, v. 6, p. 333–342, http://pubs.geoscienceworld.org/gsa/gsabulletin/article-pdf/6/1/333/3441334/BUL6-0333.pdf.

Gill, J.R., and Cobban, W.A., 1966, The Red Section of the Upper Cretaceous Pierre Shale in Wyoming: Geological Survey Professional Paper, v. 393-A, p. 1–76.

Gill, J.R., Cobban, W.A., and Schultz, L.G., 1972, Stratigraphy and Composition of the Sharon Springs Member of the Pierre Shale in Western Kansas: Geological Survey Professional Paper, v. 728, p. 1–50.

Griffitts, M.O., 1949, Zones of Pierre Formation of Colorado: Bulletin of the American Association of Petroleum Geologists, v. 33, p. 2011–2028, http://pubs.geoscienceworld.org/aapgbull/article-pdf/33/12/2011/4375210/aapg_1949_0033_0012_2011.pdf?casa_token=1EYLxoh3IVcAAAAA:2sKZWIARCI3.

Landman, N.H., Jicha, B.R., Cochran, J.K., Garb, M.P., Brophy, S.K., Larson, N.L., and Brezina, J., 2018a, 40Ar/39Ar date of a bentonite associated with a methane seep deposit in the upper Campanian Baculites compressus Zone, Pierre Shale, South Dakota: Cretaceous Research,.

Landman, N.H., Kirk Cochran, J., Slovacek, M., Larson, N.L., Garb, M.P., Brezina, J., and Witts, J.D., 2018b, Isotope sclerochronology of ammonites (Baculites compressus) from methane seep and non-seep sites in the late cretaceous Western Interior Seaway, USA: Implications for ammonite habitat and mode of life: American Journal of Science, v. 318, p. 603–639, doi:10.2475/06.2018.01.

Lavington, C.S., 1933, Montana Group in Eastern Colorado: Bulletin of the American Association of Petroleum Geologists, v. 17, p. 397–410.

Robinson, C.S., Mapel, W.J., Bergendahl, M.H., Udall, S.L., and Nolan, T.B., 19641, Stratigraphy and Structure of the Northern and Western Flanks of the Black Hills Uplift, Wyoming Montana, and South Dakota: Geological Survey Professional Paper, v. 404, p. 1.

Robinson, C.S., Mapel, W.J., and Cobban, W.A., 1959, Pierre Shale along Western and Northern flanks of Black Hills, Wyoming and Montana: Bulletin of the American Association of Petroleum Geologists, v. 43, p. 101–123, http://pubs.geoscienceworld.org/aapgbull/article-pdf/43/1/101/4379769/aapg_1959_0043_0001_0101.pdf.

Ryan, D.R., Witts, J.D., and Landman, N.H., 2021, Palaeoecological analysis of a methane seep deposit from the Upper Cretaceous (Maastrichtian) of the U.S. Western Interior: Lethaia, v. 54, p. 185–203, doi:10.1111/let.12396.

Scott, G.R., 1969, General and engineering geology of the northern part of Pueblo, Colorado.: Geological Survey Bulletin, v. 1262, p. 1–131.

Scott, G.R., and Cobban, W.A., 1965, Geologic and biostratigraphic map of the Pierre Shale between Jarre Creek and Loveland, Colorado: U.S. Geological Survey Miscellaneous Investigation Series, v. I–439.

Scott, G.R., and Cobban, W.A., 1986a, Geologic and biostratigraphic map of the Pierre Shale in the Colorado Springs–Pueblo area, Colorado: U.S. Geological Survey Miscellaneous Investigation Series, v. I–1627.

Scott, G.R., and Cobban, W.A., 1975, Geology and biostratigraphic map of the Pierre Shale in the Canon City–Florence Basin and the Twelvemile Park area, south-central Colorado: Geological Survey Miscellaneous Investigations Series, v. I–937.

Scott, G.R., and Cobban, W.A., 1986b, Geology, biostratigraphic, and structure map of the Pierre Shale between Loveland and Round Butte, Colorado: U.S. Geological Survey Miscellaneous Investigating Series, v. I–1700.

Outcrop Polygon References:

Conservation and Survey Division, University of Nebraska-Lincoln (CSD), 1996, bedrock.e00: Nebraska Geological Survey, Conservation and Survey Division, University of Nebraska-Lincoln, Lincoln, Nebraska, scale 1:1,000,000.

Green, G.N., 1992, The Digital Geologic Map of Colorado in ARC/INFO Format: U.S. Geological Survey Open-File Report 92-0507, 9 p., scale 1:500,000.

Green, Gregory N., and Drouillard, Patricia H., 1994, The Digital Geologic Map of Wyoming in ARC/INFO Format: U.S. Geological Survey Open-File Report 94-0425, scale 1:500,000.

Green, G.N., Jones, G.E., and Anderson, O.J., 1997, The Digital Geologic Map of New Mexico in ARC/INFO Format: U.S. Geological Survey Open-File Report 97-0052, 9 p., scale 1:500,000.

Hintze, L.F., Willis, G.C., Laes, D.Y.M., Sprinkel, D.A., and Brown, K.D., 2000, Digital Geologic Map of Utah: Utah Geological Survey, Map 179DM, scale 1:500,000.

North Dakota Geological Survey, 2001, Geologic Bedrock Map of North Dakota based on Bluemle, John P., 1983, Geologic and Topographic Bedrock Map of North Dakota, NDGS Miscellaneous Map 25, 1:670,000 scale.

Ross, Jorgina A., 1992, A digital representation of the Geological map of Kansas: Kansas Geological Survey, Map M-23, scale 1:500,000.

Tomhave, D.W., and Schulz, L.D., 2004, Bedrock Geologic Map Showing Configuration of the Bedrock Surface in South Dakota East of the Missouri River, South Dakota Geological Survey, GM 9, scale 1:500,000.

Vuke, S.M., Porter, K.W., Lonn, J.D., and Lopez, D.A., 2007, Geologic Map of Montana - Compact Disc: Montana Bureau of Mines and Geology: Geologic Map 62-C, 73 p., 2 sheets, scale 1:500,000. This map was digitized in 2012 as a result of a contract betwe

Appendix C-4. R Code

```
library(ecospat)
library(raster)
library(rgbif)
library(terra)
library(geodata)

library(ade4)
library(rlist)

library(raster) # for raster analysis
library(sf) # for spatial data analysis
library(dplyr) #
library(scales)
library(dismo) # a collection of ENM/SDM tools
library(tidyr)
library(colorRamps)
library(ENMeval) # for a few new tools in ENM/SDM
library(ggplot2)
# library(GGally) # Use for ggpairs fuction (not used...)
library(spThin)
library(sdm)
library(rasterVis)



######### READ IN ENVRIONMENTAL DATA #########

setwd("C:/Users/Ceara/Documents/Province Project/ENM_Analysis_Code-and-Files")

# Function to get stacked raster layers of environmental proxies (with proper names for plotting)
get_rasters<- function(interval_abrev_strg,seep_presence){
```

```
 if (seep_presence== TRUE){
   sr <- raster::raster(paste0("Env Raster Files/",interval_abrev_strg,"_sr_av_idw_clip.tif")) # Dist to Shore
   se <- raster::raster(paste0("Env Raster Files/",interval_abrev_strg,"_se_av_idw_clip.tif")) # Dist to Seeps
   p <- raster::raster(paste0("Env Raster Files/",interval_abrev_strg,"_p_av_idw_clip.tif")) # % Pebble
   s <- raster::raster(paste0("Env Raster Files/",interval_abrev_strg,"_s_av_idw_clip.tif")) # % Sand
   sh <- raster::raster(paste0("Env Raster Files/",interval_abrev_strg,"_sh_av_idw_clip.tif")) # % Shale
   sl <- raster::raster(paste0("Env Raster Files/",interval_abrev_strg,"_sl_av_idw_clip.tif")) # % Silt
   m <- raster::raster(paste0("Env Raster Files/",interval_abrev_strg,"_m_av_idw_clip.tif")) # % Mud
   l <- raster::raster(paste0("Env Raster Files/",interval_abrev_strg,"_l_av_idw_clip.tif")) # % LS
   bs <- raster::raster(paste0("Env Raster Files/",interval_abrev_strg,"_bs_av_idw_clip.tif")) # Bed Style
   b <- raster::raster(paste0("Env Raster Files/",interval_abrev_strg,"_b_av_idw_clip.tif")) # Bioturbation
   t <- raster::raster(paste0("Env Raster Files/",interval_abrev_strg,"_t_av_idw_clip.tif")) # Bed Thickness

   ### Stack env parameters
   stacked_env = stack(sr,se,p,s,sl,sh,m,l,bs,b,t)

   names(stacked_env) <- c("Dist_to_Shore","Dist_to_Seeps",
"Percent_Pebble","Percent_Sand","Percent_Shale","Percent_Silt","Percent_Mud",
                           "Percent_Limestone","Bedding_Style","Bioturbation","Avg_Bed_Thickness")

 } else {
   sr <- raster::raster(paste0("Env Raster Files/",interval_abrev_strg,"_sr_av_idw_clip.tif")) # Dist to Shore
   p <- raster::raster(paste0("Env Raster Files/",interval_abrev_strg,"_p_av_idw_clip.tif")) # % Pebble
   s <- raster::raster(paste0("Env Raster Files/",interval_abrev_strg,"_s_av_idw_clip.tif")) # % Sand
   sh <- raster::raster(paste0("Env Raster Files/",interval_abrev_strg,"_sh_av_idw_clip.tif")) # % Shale
   sl <- raster::raster(paste0("Env Raster Files/",interval_abrev_strg,"_sl_av_idw_clip.tif")) # % Silt
   m <- raster::raster(paste0("Env Raster Files/",interval_abrev_strg,"_m_av_idw_clip.tif")) # % Mud
   l <- raster::raster(paste0("Env Raster Files/",interval_abrev_strg,"_l_av_idw_clip.tif")) # % LS
   bs <- raster::raster(paste0("Env Raster Files/",interval_abrev_strg,"_bs_av_idw_clip.tif")) # Bed Style
   b <- raster::raster(paste0("Env Raster Files/",interval_abrev_strg,"_b_av_idw_clip.tif")) # Bioturbation
   t <- raster::raster(paste0("Env Raster Files/",interval_abrev_strg,"_t_av_idw_clip.tif")) # Bed Thickness
```

```
   ### Stack env parameters
   stacked_env = stack(sr,p,s,sl,sh,m,l,bs,b,t)

   names(stacked_env) <-
c("Dist_to_Shore","Percent_Pebble","Percent_Sand","Percent_Shale","Percent_Silt","Percent_Mud",
                    "Percent_Limestone","Bedding_Style","Bioturbation","Avg_Bed_Thickness")
  }

  return(stacked_env)
}



# Use above function to get stacked rasters for each time bin

maastr_env <- get_rasters("maastr",FALSE)
camplow_env <- get_rasters("camplow",FALSE)
campmid_env <- get_rasters("campmid",TRUE)
campup_env <- get_rasters("campup",TRUE)
maastrlow_env <-  get_rasters("maastrlow",TRUE)
maastrup_env <-  get_rasters("maastrup",FALSE)



birk_env <- get_rasters("birk",FALSE)
clin_env <- get_rasters("clin",FALSE)
bacu_env <- get_rasters("bacu",TRUE)
rees_env <-  get_rasters("rees",TRUE)
comp_env <-  get_rasters("comp",TRUE)
chey_env <- get_rasters("chey",TRUE)
nebr_env <- get_rasters("nebr",TRUE)
redu_env <- get_rasters("redu",TRUE)
```

```
perp_env <-  get_rasters("perp",TRUE)
macl_env <-  get_rasters("macl",FALSE)
obtu_env <- get_rasters("obtu",FALSE)
leei_env <-  get_rasters("leei",FALSE)



#### Remove the variables that are NA or that correlate more than 0.8 Pearson's Corr Coef from raster stacks
####

# H. birkelundae- H. nebrascensis
birk_env_clean <- birk_env[[c(1:2,5,7,10)]]
# Removed all but Dist to shore, Pebble, Silt, Limestone, and Bed thickness

# B. clinobatus
clin_env_clean <- clin_env[[c(1:3,6,10)]]
# Removed all but Dist to shore, Pebble, Sand, Mud, and Bed thickness

# B. baculus-B. grandis
bacu_env_clean <- bacu_env[[c(1:4,6:8,10:11)]]
# Removed all but Dist to shore, Dist to seeps, Pebble, Sand, Silt, Mud, LS, Bioturb, and Bed thickness

# B. reesidei-B. eliasi
rees_env_clean <- rees_env[[c(1,3:4,6:11)]]
# Removed all but Dist to shore, Pebble, Sand, Silt, Mud, LS, Bed Style, Bioturb, and Bed thickness

# B. compressus-B. cuneatus
comp_env_clean <- comp_env[[c(1:2,4,6:7,10)]]
# Removed all but Dist to shore, Dist to Seeps, Sand, Silt, Mud, and Bioturbation

# D. cheyennense
chey_env_clean <- chey_env[[c(1,4,6:8)]]
# Removed all but Dist to shore, Sand, Silt, Mud, and LS
```

```
# D. nebrascense-E. jenneyi
nebr_env_clean <- nebr_env[[c(1:2,4,6:11)]]
# Removed all but Dist to shore, dist to seeps, Sand, Silt, Mud, LS, Bed style, Bioturb, and Bed thickness


# B. reduncus-B. scotti
redu_env_clean <- redu_env[[c(1:2,4,6:10)]]
# Removed all but Dist to shore, dist to seeps, Sand, Silt, Mud, LS, Bed style, Bioturb

# B. perplexus-B. gregoryensis
perp_env_clean <- perp_env[[c(1:4,6:11)]]
# Removed all but Dist to shore, dist to seeps, Sand, Silt, Mud, LS, Bed style, Bioturb

# B. maclearni-B. sp. (smooth)
macl_env_clean <- macl_env[[c(1,3,5:7,9:10)]]
# Removed all but Dist to shore, dist to seeps, Sand, Silt, Mud, LS, Bed style, Bioturb

# B. obtusus
obtu_env_clean <- obtu_env[[c(1,3,5:8,10)]]
# Removed all but Dist to shore, Sand, Silt, Mud, LS, Bed style, and Bed Thickness

# S. leei-B. sp. (weak flank ribs)
leei_env_clean <- leei_env[[c(1:3,5:8)]]
# Removed all but Dist to shore, Pebble, Sand, Silt, Mud, LS, Bed style

pairs(obtu_env_clean)


#### Remove the variables that are not shared between intervals (substage-biozone) ####
```

```
# B. baculus-B. grandis
bacu_env_shared <-bacu_env[[c(1,3,4,6:8,11)]]

# B. reesidei-B. eliasi
rees_env_shared <- rees_env[[c(1,4,6:7,10)]]
# Removed all but Dist to shore, Pebble, Sand, Silt, Mud, LS, Bed Style, Bioturb, and Bed thickness

# B. compressus-B. cuneatus
comp_env_shared <- comp_env[[c(1,4,6:7,10)]]
# Removed all but Dist to shore, Dist to Seeps, Sand, Silt, Mud, and Bioturbation

# D. nebrascense-E. jenneyi
nebr_env_shared <- nebr_env[[c(1,4,6:7,10)]]
# Removed all but Dist to shore, dist to seeps, Sand, Silt, Mud, LS, Bed style, Bioturb, and Bed thickness

# B. reduncus-B. scotti
redu_env_shared <- redu_env[[c(1,4,6:8,10)]]
# Removed all but Dist to shore, Sand, Silt, Mud, LS, Bed style, Bioturb

# B. maclearni-B. sp. (smooth)
macl_env_shared <- macl_env[[c(1,3,5,6:7,9)]]

# B. perplexus-B. gregoryensis
perp_env_shared <- perp_env[[c(1,4,6:8,10)]]

# Upper Campanian
campup_env_shared <- campup_env[[c(1,4,6:7,10)]]
# Removed all but Dist to shore, Dist to seeps, Pebble, Sand, Silt, Mud, LS, Bed style, Bioturb, Bed thickness

# Middle Campanian
```

```
campmid_env_shared <- campmid_env[[c(1,4,6:8,10)]]

# Lower Maastrichtian
maastrlow_env_shared <- maastrlow_env[[c(1,3,4,6:8,11)]]

# Maastrichtian
maastr_env_shared <- maastr_env[[c(1,2,3,5:7,10)]]

#### Read in species occurrence data for the taxa at each relevant interval of comparison ####

# set name to use in calling/naming files
bara_nebr_name = "Ibarabini_nebr"

bara_rees_name = "Ibarabini_rees"

bara_campup_name = "Ibarabini_campup"


conv_redu_name = "Iconvexus_redu"

conv_nebr_name = "Iconvexus_nebr"

conv_campmid_name = "Iconvexus_campmid"

conv_campup_name = "Iconvexus_campup"


sage_comp_name = "Isagensis_comp"

sage_rees_name = "Isagensis_rees"
```

```
sage_campup_name = "Isagensis_campup"


subl_redu_name = "Isublaevis_redu"

subl_nebr_name = "Isublaevis_nebr"

subl_campmid_name = "Isublaevis_campmid"

subl_campup_name = "Isublaevis_campup"



azer_macl_name <- "Iazerbaydjanensis_macl"

azer_campmid_name <- "Iazerbaydjanensis_campmid"


oblo_rees_name <- "Ioblongus_rees"

oblo_campup_name <- "Ioblongus_campup"

sask_redu_name <- "Isaskatchewanensis_redu"

sask_campmid_name <- "Isaskatchewanensis_campmid"

# Outgroup Taxa:

chla_nebr_name <- "Cnebrascensis_nebr"
chla_rees_name <- "Cnebrascensis_rees"
chla_bacu_name <- "Cnebrascensis_bacu"
```

```
chla_campup_name <- "Cnebrascensis_campup"
chla_maastrlow_name <- "Cnebrascensis_maastrlow"
chla_maastr_name <- "Cnebrascensis_maa"

luci_nebr_name <- "Lsubundata_nebr"
luci_comp_name <- "Lsubundata_comp"
luci_campup_name <- "Lsubundata_campup"



cten_perp_name <- "Cimbricatula_perp"
cten_redu_name <- "Cimbricatula_redu"
cten_nebr_name <- "Cimbricatula_nebr"
cten_rees_name <- "Cimbricatula_rees"
cten_campup_name <- "Cimbricatula_campup"
cten_campmid_name <- "Cimbricatula_campmid"



# Additional taxa for analysis

cmon_macl_name <- "Cmontanensis_macl"
cmon_perp_name <- "Cmontanensis_perp"
cmon_campmid_name <- "Cmontanensis_campmid"
cmon_redu_name <- "Cmontanensis_redu"
cmon_comp_name <- "Cmontanensis_comp"
cmon_campup_name <- "Cmontanensis_campup"

hnod_comp_name <- "Hnodosus_comp"
hnod_rees_name <- "Hnodosus_rees"
hnod_campup_name <- "Hnodosus_campup"

ifib_clin_name <- "Ifibrosus_clin"
```

```
ifib_maastrlow_name <- "Ifibrosus_maastrlow"
ifib_maastr_name <- "Ifibrosus_maastr"



pmee_nebr_name <- "Pmeeki_nebr"
pmee_chey_name <- "Pmeeki_chey"
pmee_comp_name <- "Pmeeki_comp"
pmee_campup_name <- "Pmeeki_campup"

psyr_leei_name <- "Psyrtale_leei"
psyr_camplow_name <- "Psyrtale_camplow"

# Read in the occurrence data for the relevant intervals
conv_redu <- read.csv(as.character(paste0("Thinned_taxa_csvs/",conv_redu_name,"_thin1.csv")))
conv_nebr <- read.csv(as.character(paste0("Thinned_taxa_csvs/",conv_nebr_name,"_thin1.csv")))
conv_campmid <- read.csv(as.character(paste0("Thinned_taxa_csvs/",conv_campmid_name,"_thin1.csv")))
conv_campup <- read.csv(as.character(paste0("Thinned_taxa_csvs/",conv_campup_name,"_thin1.csv")))



bara_rees <- read.csv(as.character(paste0("Thinned_taxa_csvs/",bara_rees_name,"_thin1.csv")))
bara_nebr <- read.csv(as.character(paste0("Thinned_taxa_csvs/",bara_nebr_name,"_thin1.csv")))
bara_campup <- read.csv(as.character(paste0("Thinned_taxa_csvs/",bara_campup_name,"_thin1.csv")))



sage_comp <- read.csv(as.character(paste0("Thinned_taxa_csvs/",sage_comp_name,"_thin1.csv")))
sage_rees <- read.csv(as.character(paste0("Thinned_taxa_csvs/",sage_rees_name,"_thin1.csv")))
sage_campup <- read.csv(as.character(paste0("Thinned_taxa_csvs/",sage_campup_name,"_thin1.csv")))



subl_redu <- read.csv(as.character(paste0("Thinned_taxa_csvs/",subl_redu_name,"_thin1.csv")))
subl_nebr <- read.csv(as.character(paste0("Thinned_taxa_csvs/",subl_nebr_name,"_thin1.csv")))
```

```
subl_campmid <- read.csv(as.character(paste0("Thinned_taxa_csvs/",subl_campmid_name,"_thin1.csv")))
subl_campup <- read.csv(as.character(paste0("Thinned_taxa_csvs/",subl_campup_name,"_thin1.csv")))



azer_macl <- read.csv(as.character(paste0("Thinned_taxa_csvs/",azer_macl_name,"_thin1.csv")))
azer_campmid <- read.csv(as.character(paste0("Thinned_taxa_csvs/",azer_campmid_name,"_thin1.csv")))



oblo_rees <- read.csv(as.character(paste0("Thinned_taxa_csvs/",oblo_rees_name,"_thin1.csv")))
oblo_campup <- read.csv(as.character(paste0("Thinned_taxa_csvs/",oblo_campup_name,"_thin1.csv")))



sask_redu <- read.csv(as.character(paste0("Thinned_taxa_csvs/",sask_redu_name,"_thin1.csv")))
sask_campmid <- read.csv(as.character(paste0("Thinned_taxa_csvs/",sask_campmid_name,"_thin1.csv")))



# Outgroup Taxa:

chla_nebr <- read.csv(as.character(paste0("Thinned_taxa_csvs/",chla_nebr_name,"_thin1.csv")))
chla_rees <- read.csv(as.character(paste0("Thinned_taxa_csvs/",chla_rees_name,"_thin1.csv")))

chla_bacu <- read.csv(as.character(paste0("Thinned_taxa_csvs/",chla_bacu_name,"_thin1.csv")))
chla_campup <- read.csv(as.character(paste0("Thinned_taxa_csvs/",chla_campup_name,"_thin1.csv")))
chla_maastrlow <- read.csv(as.character(paste0("Thinned_taxa_csvs/",chla_maastrlow_name,"_thin1.csv")))
chla_maastr <- read.csv(as.character(paste0("Thinned_taxa_csvs/",chla_maastr_name,"_thin1.csv")))



luci_nebr <- read.csv(as.character(paste0("Thinned_taxa_csvs/",luci_nebr_name,"_thin1.csv")))
luci_comp <- read.csv(as.character(paste0("Thinned_taxa_csvs/",luci_comp_name,"_thin1.csv")))
```

```
luci_campup <- read.csv(as.character(paste0("Thinned_taxa_csvs/",luci_campup_name,"_thin1.csv")))


cten_perp <- read.csv(as.character(paste0("Thinned_taxa_csvs/",cten_perp_name,"_thin1.csv")))
cten_redu <- read.csv(as.character(paste0("Thinned_taxa_csvs/",cten_redu_name,"_thin1.csv")))
cten_nebr <- read.csv(as.character(paste0("Thinned_taxa_csvs/",cten_nebr_name,"_thin1.csv")))
cten_rees <- read.csv(as.character(paste0("Thinned_taxa_csvs/",cten_rees_name,"_thin1.csv")))

cten_campup <- read.csv(as.character(paste0("Thinned_taxa_csvs/",cten_campup_name,"_thin1.csv")))
cten_campmid <- read.csv(as.character(paste0("Thinned_taxa_csvs/",cten_campmid_name,"_thin1.csv")))

# Additional Taxa for an analysis
cmon_macl <- read.csv(as.character(paste0("Thinned_taxa_csvs/",cmon_macl_name,"_thin1.csv")))
cmon_perp <- read.csv(as.character(paste0("Thinned_taxa_csvs/",cmon_perp_name,"_thin1.csv")))
cmon_campmid <- read.csv(as.character(paste0("Thinned_taxa_csvs/",cmon_campmid_name,"_thin1.csv")))
cmon_redu <- read.csv(as.character(paste0("Thinned_taxa_csvs/",cmon_redu_name,"_thin1.csv")))
cmon_comp <- read.csv(as.character(paste0("Thinned_taxa_csvs/",cmon_comp_name,"_thin1.csv")))
cmon_campup <- read.csv(as.character(paste0("Thinned_taxa_csvs/",cmon_campup_name,"_thin1.csv")))


hnod_comp <- read.csv(as.character(paste0("Thinned_taxa_csvs/",hnod_comp_name,"_thin1.csv")))
hnod_rees <- read.csv(as.character(paste0("Thinned_taxa_csvs/",hnod_rees_name,"_thin1.csv")))
hnod_campup <- read.csv(as.character(paste0("Thinned_taxa_csvs/",hnod_campup_name,"_thin1.csv")))


ifib_clin <- read.csv(as.character(paste0("Thinned_taxa_csvs/",ifib_clin_name,"_thin1.csv")))
ifib_maastrlow <- read.csv(as.character(paste0("Thinned_taxa_csvs/",ifib_maastrlow_name,"_thin1.csv")))
ifib_maastr <- read.csv(as.character(paste0("Thinned_taxa_csvs/","Ifibrosus_maa","_thin1.csv")))
```

```
pmee_macl <- read.csv(as.character(paste0("Thinned_taxa_csvs/",pmee_nebr_name,"_thin1.csv")))
pmee_perp <- read.csv(as.character(paste0("Thinned_taxa_csvs/",pmee_chey_name,"_thin1.csv")))
pmee_comp <- read.csv(as.character(paste0("Thinned_taxa_csvs/",pmee_comp_name,"_thin1.csv")))
pmee_campup <- read.csv(as.character(paste0("Thinned_taxa_csvs/",pmee_campup_name,"_thin1.csv")))


psyr_leei <- read.csv(as.character(paste0("Thinned_taxa_csvs/",psyr_leei_name,"_thin1.csv")))
psyr_camplow <- read.csv(as.character(paste0("Thinned_taxa_csvs/",psyr_camplow_name,"_thin1.csv")))


#### Thin Genus level information (by selecting all relevant spp occ from database) ####
# Function to thin spp occ by 30km in intervals if >6 initial occurrences, creates csvs and outputs summary table
thin_taxa_by_interv<- function(taxa_names,occ_data,tax_abrev){
 # Subset out species from occ table
 spp_dat <- subset(occ_data,Taxa_name %in% taxa_names, select=c(Taxa_name,Lat,Long,interval_biozone))
 # remove duplicate locations
 spp_dat <- unique(spp_dat)
 # Get list of number of occ in each interval
 unique_interv<- as.data.frame(table(spp_dat$interval_biozone))
 unique_interv_7great <- subset(unique_interv, Freq >6)
 unique_interv_7great <- unique_interv_7great$Var1
 # Get list of df of occ for each interval with >6 occ
 inter_dat_list <- list()  # create empty list to hold df for each relevant interval
 for (i in 1:length(unique_interv_7great)){
  spp_dat_inter <- subset(spp_dat,interval_biozone==unique_interv_7great[i]) # subset out by interval name
  inter_dat_list[[i]] <-spp_dat_inter[,c("Taxa_name","Lat","Long")] # get just the lat/lon and taxa name
 }

 # Run thinning on each df in list:
 for (i in 1:nrow(as.data.frame(unique_interv_7great))){
  # Thinning function:
```

```
  thin_tax_in_loop <- function(tax,file_name){
    tax <- as.data.frame(tax)
    thin(tax, long.col="Long", lat.col="Lat", spec.col="Taxa_name", thin.par=30, reps=1,
        out.dir="C:/Users/ceara/Documents/Province Project/ENM_Analysis_Code-and-
Files/Thinned_taxa_csvs", out.base=file_name, write.log.file=FALSE, verbose=FALSE)
  }
  taxa_interval_name=as.character(paste(tax_abrev,unique(unique_interv_7great[i]), sep="_")) # Set the name
of output file
  thin_tax_in_loop(inter_dat_list[[i]],taxa_interval_name) # run thinning function to output files
 }
 output_list<-list()
 for (i in 1:length(unique_interv_7great)){
  taxa_interval_name=as.character(paste(tax_abrev,unique(unique_interv_7great[i]), sep="_")) # Set the name
of output file
  output_list[[i]] <- read.csv(as.character(paste0("Thinned_taxa_csvs/",taxa_interval_name,"_thin1.csv")))
  output_list[[i]]$interval <- taxa_interval_name
 }
 # Summarize the output CSV files into a vector of occ counts (frequency)
 summarize<- function (df){as.data.frame(table(df$interval))} # function to get frequency counts of each
interval
 output_list<-sapply(output_list,summarize) # apply the above function to each df
 rownames(output_list) <- c("Taxa-Interval","Freq") # renamed the summary table row names

 return(output_list)
}



# Make lists of names to use for genus-level analysis in relevant intervals
redu_tax_names <- c("Inoceramus convexus","Inoceramus sublaevis", "Inoceramus saskatchewanensis")
nebr_tax_names <- c("Inoceramus barabini","Inoceramus convexus","Inoceramus sublaevis")
rees_tax_names <- c("Inoceramus barabini","Inoceramus sagensis","Inoceramus oblongus")

#### Thin Genus level inforamtion (using all occ, including Inoceramus w/out spp ID)
```

```
# Read in occ database (with spp-genus combine in column Taxa_name)
tax_sp_inter <- read.csv('FE_data_compress_fe_3-7-23_basins_Fm_update_gen-sp_name_interval_update.csv')

# Extract all relevant spp names to use for collecting Inoceramus genus data
ino_spp_df <- unique(subset(tax_sp_inter,Updated_Genus == "Inoceramus", select=c(Taxa_name)))

ino_spp_vec <- as.character(ino_spp_df[,1]) # make into a character vector

# Make lists of names to use for genus-level analysis in relevant intervals
redu_tax_names <- ino_spp_vec
nebr_tax_names <- ino_spp_vec
rees_tax_names <- ino_spp_vec

# Use above function to run thinning on genus-level data in relevant intervals
# Inoceramus_redu<- thin_taxa_by_interv(redu_tax_names,tax_sp_inter,"Inoceramus_all_redu")
#
# Inoceramus_nebr<- thin_taxa_by_interv(nebr_tax_names,tax_sp_inter,"Inoceramus_all_nebr")
#
# Inoceramus_nebr<- thin_taxa_by_interv(rees_tax_names,tax_sp_inter,"Inoceramus_all_rees")

#### Read in Genus Occurrence data for the taxa at each relevant interval of comparison  ####
# Read in the genus occurrence data for the relevant intervals
Inoceramus_redu <- read.csv(as.character(paste0("Thinned_taxa_csvs/","Inoceramus_redu_B. reduncus-B. scotti","_thin1.csv")))
Inoceramus_nebr <- read.csv(as.character(paste0("Thinned_taxa_csvs/","Inoceramus_nebr_D. nebrascense-E. jenneyi","_thin1.csv")))
Inoceramus_rees <- read.csv(as.character(paste0("Thinned_taxa_csvs/","Inoceramus_rees_B. reesidei-B. eliasi","_thin1.csv")))

Inoceramus_all_redu <- read.csv(as.character(paste0("Thinned_taxa_csvs/","Inoceramus_all_redu_B. reduncus-B. scotti","_thin1.csv")))
Inoceramus_all_nebr <- read.csv(as.character(paste0("Thinned_taxa_csvs/","Inoceramus_all_nebr_D. nebrascense-E. jenneyi","_thin1.csv")))
```

```
Inoceramus_all_rees <- read.csv(as.character(paste0("Thinned_taxa_csvs/","Inoceramus_all_rees_B. reesidei-B. eliasi","_thin1.csv")))

#### MAKE MAPS? Using the directions from this site: https://plantarum.ca/2021/07/29/ecospat/ ####

# Tell what columns are lat/long
coordinates(conv_redu) <- c("Long",
                            "Lat")

# Download the maps of North American countries to plot
NorthAmerica <- gadm(country = country_codes("North America")$ISO3,
                     level = 0, resolution = 2,
                     path = "maps_basemaps")
us <- gadm(country = "USA", level = 1, resolution = 2,
           path = "maps_basemaps")
canada <- gadm(country = "CAN", level = 1, resolution = 2,
               path = "maps_basemaps")

# bind the maps into a single feature to map
CanUS <- rbind(NorthAmerica,us, canada)

# Plot the maps
plot(CanUS, xlim = c(-115, -95),ylim = c(55, 23),
     border = "gray")
points(conv_redu,pch=16,col = "red")

#### Run function to prep data for use in ecospat (clean NAs, make into spatial object, extract values, etc) ####
run_data_prep <- function(taxa_inter_name, interval_raster_stack){

  # Call the thinned data file
  tax_test <- read.csv(as.character(paste0("Thinned_taxa_csvs/",taxa_inter_name,"_thin1.csv")))
```

```
# Convert occ pts for taxa to spatial points
coordinates(tax_test) = ~ Long + Lat


# make sure to set spatial coor system (WGS 84)
myCRS1 <- CRS("+init=epsg:4326")
crs(tax_test) <- myCRS1

# project raster stack back to WGS 84 coor system
interval_raster_proj <- projectRaster(interval_raster_stack, crs=myCRS1)

# look at conditions at train occ locations (find nulls, or pts that are not loc on variable values)
conditions_occ <- raster::extract(interval_raster_proj,tax_test) # Use the thinned, cleaned taxa data and the env stacked rasters to extract variable vals
bad_rec <- is.na(conditions_occ[,1]) # Get the locations of those extracted pts with NA (no vals)
table(bad_rec) # Look at the list of NA locations
conditions_occ[bad_rec,] # Another way to look at it

# remove null values from occ
occ <- tax_test[!bad_rec,] # take out the pts that are over null vals
occ # Check result

# extract env values for occurrences
occ <- cbind(occ, extract(interval_raster_proj, occ))

#Plot final data:
par(mfrow=c(1,1))
plot(interval_raster_proj[[1]]) # plot env layer
plot(occ, add=T, col="red") # Plot occ pts

output_list <- list(occ, interval_raster_proj)
```

```
  return(output_list)

}

# Use function to prep the data for further analysis: output is list of occ then env data
conv_redu_prep <- run_data_prep(conv_redu_name,redu_env_shared)
conv_campmid_prep <- run_data_prep(conv_campmid_name,campmid_env_shared)
conv_nebr_prep <- run_data_prep(conv_nebr_name,nebr_env_shared)
conv_campup_prep <- run_data_prep(conv_campup_name,campup_env_shared)

bara_rees_prep <- run_data_prep(bara_rees_name,rees_env_shared)
bara_nebr_prep <- run_data_prep(bara_nebr_name,nebr_env_shared)
bara_campup_prep <- run_data_prep(bara_campup_name,campup_env_shared)

sage_comp_prep <- run_data_prep(sage_comp_name,comp_env_shared)
sage_rees_prep <- run_data_prep(sage_rees_name,rees_env_shared)
sage_campup_prep <- run_data_prep(sage_campup_name,campup_env_shared)


subl_redu_prep <- run_data_prep(subl_redu_name,redu_env_shared)
subl_campmid_prep <- run_data_prep(subl_campmid_name,campmid_env_shared)
subl_nebr_prep <- run_data_prep(subl_nebr_name,nebr_env_shared)
subl_campup_prep <- run_data_prep(subl_campup_name,campup_env_shared)


azer_macl_prep <- run_data_prep(azer_macl_name,macl_env_shared)
azer_campmid_prep <- run_data_prep(azer_campmid_name,campmid_env_shared)


oblo_rees_prep <- run_data_prep(oblo_rees_name,rees_env_shared)
oblo_campup_prep <- run_data_prep(oblo_campup_name,campup_env_shared)
```

```
sask_redu_prep <- run_data_prep(sask_redu_name,redu_env_shared)
sask_campmid_prep <- run_data_prep(sask_campmid_name,campmid_env_shared)

# Outgroup Taxa:

chla_nebr_prep <- run_data_prep(chla_nebr_name,nebr_env_shared)
chla_rees_prep <- run_data_prep(chla_rees_name,rees_env_shared)
chla_bacu_prep <- run_data_prep(chla_bacu_name,bacu_env_shared)

chla_campup_prep <- run_data_prep(chla_campup_name,campup_env_shared)
chla_maastrlow_prep <- run_data_prep(chla_maastrlow_name,maastrlow_env_shared)
chla_maastr_prep <- run_data_prep(chla_maastr_name,maastr_env_shared)


luci_nebr_prep <- run_data_prep(luci_nebr_name,nebr_env_shared)
luci_comp_prep <- run_data_prep(luci_comp_name,comp_env_shared)

luci_campup_prep <- run_data_prep(luci_campup_name,campup_env_shared)


cten_campmid_prep <- run_data_prep(cten_campmid_name,campmid_env_shared)
cten_campup_prep <- run_data_prep(cten_campup_name,campup_env_shared)

cten_perp_prep <- run_data_prep(cten_perp_name,perp_env_shared)
cten_redu_prep <- run_data_prep(cten_redu_name,redu_env_shared)
cten_nebr_prep <- run_data_prep(cten_nebr_name,nebr_env_shared)
cten_rees_prep <- run_data_prep(cten_rees_name,rees_env_shared)
```

```
# Additional taxa for analysis
cmon_rees_prep <- run_data_prep(cmon_rees_name,rees_env_shared)
cmon_rees_prep <- run_data_prep(cmon_rees_name,rees_env_shared)
cmon_rees_prep <- run_data_prep(cmon_rees_name,rees_env_shared)
cmon_rees_prep <- run_data_prep(cmon_rees_name,rees_env_shared)


cmon_rees_prep <- run_data_prep(cmon_rees_name,rees_env_shared)
cmon_rees_prep <- run_data_prep(cmon_rees_name,rees_env_shared)


cmon_rees_prep <- run_data_prep(cmon_rees_name,rees_env_shared)


cmon_rees_prep <- run_data_prep(cmon_rees_name,rees_env_shared)
cmon_rees_prep <- run_data_prep(cmon_rees_name,rees_env_shared)
cmon_rees_prep <- run_data_prep(cmon_rees_name,rees_env_shared)


cmon_rees_prep <- run_data_prep(cmon_rees_name,rees_env_shared)

# Function to prep genus level data for analysis
prep_genus <- function(gen_data,interval_raster_stack){
  # Convert occ pts for taxa to spatial points
  coordinates(gen_data) = ~ Long + Lat


  # make sure to set spatial coor system (WGS 84)
  myCRS1 <- CRS("+init=epsg:4326")
  crs(gen_data) <- myCRS1
```

```
  # project raster stack back to WGS 84 coor system
  interval_raster_proj <- projectRaster(interval_raster_stack, crs=myCRS1)

  # look at conditions at train occ locations (find nulls, or pts that are not loc on variable values)
  conditions_occ <- raster::extract(interval_raster_proj,gen_data) # Use the thinned, cleaned taxa data and the
env stacked rasters to extract variable vals
  bad_rec <- is.na(conditions_occ[,1]) # Get the locations of those extracted pts with NA (no vals)
  table(bad_rec) # Look at the list of NA locations
  conditions_occ[bad_rec,] # Another way to look at it

  # remove null values from occ
  occ <- gen_data[!bad_rec,] # take out the pts that are over null vals
  occ # Check result

  # extract env values for occurrences
  occ <- cbind(occ, extract(interval_raster_proj, occ))

  #Plot final data:
  par(mfrow=c(1,1))
  plot(interval_raster_proj[[1]]) # plot env layer
  plot(occ, add=T, col="red") # Plot occ pts

  output_list <- list(occ, interval_raster_proj)

  return(output_list)

}

# Use above function to prep genus data
Inoceramus_redu_prep <- prep_genus(Inoceramus_redu,redu_env_shared)
Inoceramus_nebr_prep <- prep_genus(Inoceramus_nebr,nebr_env_shared)
Inoceramus_rees_prep <- prep_genus(Inoceramus_rees,rees_env_shared)
```

```
# Use above function to prep genus including no spp ID data
Inoceramus_all_redu_prep <- prep_genus(Inoceramus_all_redu,redu_env_shared)
Inoceramus_all_nebr_prep <- prep_genus(Inoceramus_all_nebr,nebr_env_shared)
Inoceramus_all_rees_prep <- prep_genus(Inoceramus_all_rees,rees_env_shared)

#### Run Ecospat analysis to compare overlap (make plots and tables) ####
# Function to run ecospat overlap comparison, calc overlap and other vals, and create PCA summary and
overlap plots
run_ecospat_dyn <- function(first_occ_env_list,second_occ_env_list,spp_comp_title,spp_comp_pdf_name){
  ## Extract values to matrix for each raster stack of env proxy variables:
  first_bg<- getValues(first_occ_env_list[[2]])
  second_bg <- getValues(second_occ_env_list[[2]])

  ## Clean out missing values:
  first_bg <- first_bg[complete.cases(first_bg), ]
  second_bg <- second_bg[complete.cases(second_bg), ]

  ## Combined global environment:
  joined_bg <- rbind(first_bg, second_bg)



  # Run PCA analysis on combine BG pts (all env data)
  pca_joined_bg <- dudi.pca(joined_bg, center = TRUE,
                            scale = TRUE, scannf = FALSE, nf = 2)
  joined_bg_scores <- pca_joined_bg$li # get the pca scores

  # Start saving figures to pdf file
  pdf(paste("Figures/",spp_comp_pdf_name, "ecospat_overlap_comparison.pdf"))

  # look at variable contribution
```

```
ecospat.plot.contrib(contrib=pca_joined_bg$co, eigen=pca_joined_bg$eig)


# Map occurrence data into the 2d ordination (have to coerce spatialPointsDataFrame inot data.frame)
# Explicitly match the colnames so only using the right ones
# Only selected the li element (doesn't include others)
first_occ_scores <-
  suprow(pca_joined_bg,
         data.frame(first_occ_env_list[[1]])[, colnames(joined_bg)])$li
second_occ_scores <-
  suprow(pca_joined_bg,
         data.frame(second_occ_env_list[[1]])[, colnames(joined_bg)])$li

# Map BG data to the 2d ordination
first_bg_scores <- suprow(pca_joined_bg, first_bg)$li
second_bg_scores <- suprow(pca_joined_bg, second_bg)$li


# density distribution for first interval
grid_first <- ecospat.grid.clim.dyn(
  glob = joined_bg_scores,
  glob1 = first_bg_scores,
  sp = first_occ_scores,
  R = 100,
  th.sp = 0
)

# density distribution for second interval
grid_second <- ecospat.grid.clim.dyn(
  glob = joined_bg_scores,
  glob1 = second_bg_scores,
```

```
  sp = second_occ_scores,
  R = 100,
  th.sp = 0
)


D_overlap <- ecospat.niche.overlap (grid_first, grid_second, cor=T)
# Schoener's D metric and I metric

#Make summary table of D and I metrics
d_i_overlap <- matrix(nrow=1,ncol=2)
d_i_overlap[1,1] <- D_overlap[[1]]
d_i_overlap[1,2] <- D_overlap[[2]]
colnames(d_i_overlap) <- c("D_val","I_val")

# Finally we're ready to do the Niche Quantification/Comparisons. We'll use the
#PCA scores for the global environment, the native and invasive environments,
#and the native and invasive occurrence records

first_grid <- ecospat.grid.clim.dyn(joined_bg_scores,
                                    first_bg_scores,
                                    first_occ_scores)

second_grid <- ecospat.grid.clim.dyn(joined_bg_scores,
                                     second_bg_scores,
                                     second_occ_scores)


ecospat.plot.niche.dyn(first_grid,title=paste(spp_comp_title,"\nEcospat Overlap Comparison"), second_grid,
                       quant = 0.05, name.axis1="PC1",name.axis2="PC2")
ecospat.shift.centroids(first_occ_scores, second_occ_scores, first_bg_scores, second_bg_scores)
```

```
## Niche Equivalency Test
## Obs: observed overlaps, sim: similulated overlaps, p.D: pvalue of the test on D
## p.I: pvalue of the test on I
## Test for greater or lower equivalency
eq_testgr <- ecospat.niche.equivalency.test(first_grid, second_grid,
                                    rep=500, overlap.alternative = "higher",
                                    expansion.alternative = "lower",
                                    stability.alternative = "higher",
                                    unfilling.alternative = "lower",
                                    ncores=4) ##rep = 1000 recommended for operational runs
eq_testlw <- ecospat.niche.equivalency.test(first_grid, second_grid,
                                    rep=500, overlap.alternative = "lower",
                                    expansion.alternative = "higher",
                                    stability.alternative = "lower",
                                    unfilling.alternative = "higher",
                                    ncores=4) ##rep = 1000 recommended for operational runs

# write D and I values of equivalency test
EQ_DI = cbind("Obs_D"=eq_testgr$obs$D,"Obs_I"=eq_testgr$obs$I)

# write p values of equivalency test
p_EQ_DI =
cbind("p.D_GR"=eq_testgr$p.D,"p.I_GR"=eq_testgr$p.I,"p.D_LW"=eq_testlw$p.D,"p.I_LW"=eq_testlw$p.D)

# write p values of expansion, stability, and unfilling
p_EQ_dynam =
cbind("p.expan"=eq_testgr$p.expansion,"p.stab"=eq_testgr$p.stability,"p.unfill"=eq_testgr$p.unfilling)

## Test for greater (niche conservatism) or lower (niche divergence) similarity
sim_testgr <- ecospat.niche.similarity.test(first_grid, second_grid,
```

```
                                rep=1000, overlap.alternative = "higher",
                                expansion.alternative = "lower",
                                stability.alternative = "higher",
                                unfilling.alternative = "lower",
                                rand.type=1,ncores=4)
sim_testlw <- ecospat.niche.similarity.test(first_grid, second_grid,
                                rep=1000, overlap.alternative = "lower",
                                expansion.alternative = "higher",
                                stability.alternative = "lower",
                                unfilling.alternative = "higher",
                                rand.type=1,ncores=4)

# write D and I values of equivalency test
SIM_DI = cbind("Obs_D"=sim_testgr$obs$D,"Obs_I"=sim_testgr$obs$I)

# write p values of similarity test
p_SIM_DI =
cbind("p.D_GR"=sim_testgr$p.D,"p.I_GR"=sim_testgr$p.I,"p.D_LW"=sim_testlw$p.D,"p.I_LW"=sim_testlw$
p.D)

# write p values of expansion, stability, and unfilling
p_SIM_dynam =
cbind("p.expan"=sim_testgr$p.expansion,"p.stab"=sim_testgr$p.stability,"p.unfill"=sim_testgr$p.unfilling)

# Plot test distributions
ecospat.plot.overlap.test(eq_testgr, "D", "Greater Equivalency")
ecospat.plot.overlap.test(eq_testlw, "D", "Lower Equivalency")
ecospat.plot.overlap.test(sim_testgr, "D", "Greater Similarity")
ecospat.plot.overlap.test(sim_testlw, "D", "Lower Similarity")
```

```
  dev.off() # close the pdf so will stop adding plots


  ## Plot niche space for first (grid.clim1) and second (grid.clim2) intervals (NOT WORKING, THEY MIGHT
BE FIXING IT)
  # ecospat.plot.niche(grid_first, title=spp_first_inter_name, name.axis1='PC1', name.axis2='PC2', cor=FALSE)
  # ecospat.plot.niche(grid_second, title=spp_second_inter_name, name.axis1='PC1', name.axis2='PC2')



  ### Look at niche expansion, stability, and unfilling
  # NA=analysis on entire area, 0=analysis on only overlapping, 0.05=analysis on 5th quantile intersection
  dynam_allAreas = ecospat.niche.dyn.index (grid_first, grid_second, intersection=NA)
  dynam_overlap = ecospat.niche.dyn.index (grid_first, grid_second, intersection=0)


  # write csv of niche dynamic percentages
  Niche_Ex_St = cbind("AllAreas"=dynam_allAreas$dynamic.index.w,
                      "Overlapping"=dynam_overlap$dynamic.index.w)


  list_outputs <- list(d_i_overlap,eq_testgr,eq_testlw,sim_testgr,sim_testlw,Niche_Ex_St)
  names(list_outputs) <-
c("D_I_Overlap","Eq_Test_Great","Eq_Test_Low","Sim_Test_Great","Sim_Test_Low","Niche_dynamics")
  return(list_outputs)


}



#### TEMPORAL RESOLUTION COMPARISON ####
# Use above function to make plots of overlap comparison and produce list of overlap metric results
# (output is in the form of a list including 1-D/I values of overlap, 2-Greater Equiv results (also list),
# 3-Lower Equiv results, 4-Greater Similarity results, 5-Lower Similarity results, and 6-niche dynamics)
conv_redu_campmid_overlap <- run_ecospat_dyn(conv_redu_prep,conv_campmid_prep,
                                             "I. convextus B. reduncus-B. scotti\nvs. Middle
Campanian","Iconv_redu_to_midcamp")
```

```
conv_nebr_campup_overlap <- run_ecospat_dyn(conv_nebr_prep,conv_campup_prep,
                                            "I. convextus D. nebrascense-E. jenneyi\nvs. Upper
Campanian","Iconv_nebr_to_upcamp")

bara_rees_campup_overlap <- run_ecospat_dyn(bara_rees_prep,bara_campup_prep,
                                            "I. barabini\nB. reesidei-B. eliasi\nvs. Upper Campanian","Ibara_rees_to_upcamp")

bara_nebr_campup_overlap <- run_ecospat_dyn(bara_nebr_prep,bara_campup_prep,
                                            "I. barabini\nD. nebrascense-E. jenneyi\nvs. Upper
Campanian","Ibara_nebr_to_upcamp")

sage_rees_campup_overlap <- run_ecospat_dyn(sage_rees_prep,sage_campup_prep,
                                            "I. sagensis\nB. reesidei-B. eliasi\nvs. Upper Campanian","Isage_rees_to_upcamp")

sage_comp_campup_overlap <- run_ecospat_dyn(sage_comp_prep,sage_campup_prep,
                                            "I. sagensis\nB. compressus-B. cuneatus\nvs. Upper
Campanian","Isage_comp_to_upcamp")

subl_redu_campmid_overlap <- run_ecospat_dyn(subl_redu_prep,subl_campmid_prep,
                                             "I. sublaevis\nB. reduncus-B. scotti\nvs. Middle
Campanian","Isubl_redu_to_midcamp")

subl_nebr_campup_overlap <- run_ecospat_dyn(subl_nebr_prep,subl_campup_prep,
                                            "I. sublaevis\nD. nebrascense-E. jenneyi\nvs. Upper
Campanian","Isubl_nebr_to_upcamp")

azer_macl_campmid_overlap <- run_ecospat_dyn(azer_macl_prep,azer_campmid_prep,
                                             "I. azerbaydjanensis\nB. maclearni-B. sp. (smooth)\nvs. Middle
Campanian","Iazer_macl_to_midcamp")

sask_redu_campmid_overlap <- run_ecospat_dyn(sask_redu_prep,sask_campmid_prep,
                                             "I. saskatchewanensis\nB. reduncus-B. scotti\nvs. Middle
Campanian","Isask_redu_to_midcamp")
```

```
oblo_rees_campup_overlap <- run_ecospat_dyn(oblo_rees_prep,oblo_campup_prep,
                                "I. oblongus\nB. reesidei-B. eliasi\nvs. Upper
Campanian","Ioblo_rees_to_upcamp")

cten_perp_campmid_overlap <- run_ecospat_dyn(cten_perp_prep,cten_campmid_prep,
                                "C. imbricatula\nB. perplexus-B. gregoryensis\nvs. Middle
Campanian","Cimbr_perp_to_midcamp")

cten_redu_campmid_overlap <- run_ecospat_dyn(cten_redu_prep,cten_campmid_prep,
                                "C. imbricatula\nB. reduncus-B. scotti\nvs. Middle
Campanian","Cimbr_redu_to_midcamp")

cten_nebr_campup_overlap <- run_ecospat_dyn(cten_nebr_prep,cten_campup_prep,
                               "C. imbricatula\nD. nebrascense-E. jenneyi\nvs. Upper
Campanian","Cimbr_nebr_to_upcamp")

cten_rees_campup_overlap <- run_ecospat_dyn(cten_rees_prep,cten_campup_prep,
                                "C. imbricatula\nB. reesidei-B. eliasi\nvs. Upper
Campanian","Cimbr_rees_to_upcamp")

luci_nebr_campup_overlap <- run_ecospat_dyn(luci_nebr_prep,luci_campup_prep,
                                "L. subundata\nD. nebrascense-E. jenneyi\nvs. Upper
Campanian","Lsubu_nebr_to_upcamp")

luci_rees_campup_overlap <- run_ecospat_dyn(luci_comp_prep,luci_campup_prep,
                               "L. subundata\nB. compressus-B. cuneatus\nvs. Upper
Campanian","Lsubu_comp_to_upcamp")

chla_nebr_campup_overlap <- run_ecospat_dyn(chla_nebr_prep,chla_campup_prep,
```

```
                                "C. nebrascensis\nD. nebrascense-E. jenneyi\nvs. Upper
Campanian","Cnebr_nebr_to_upcamp")


chla_rees_campup_overlap <- run_ecospat_dyn(chla_rees_prep,chla_campup_prep,

                                "C. nebrascensis\nB. reesidei-B. eliasi\nvs. Upper
Campanian","Cnebr_rees_to_upcamp")


chla_bacu_maastrlow_overlap <- run_ecospat_dyn(chla_bacu_prep,chla_maastrlow_prep,

                                "C. nebrascensis\nB. baculus-B. grandis\nvs. Lower
Maastrichtian","Cnebr_bacu_to_lowmaas")


chla_bacu_maastr_overlap <- run_ecospat_dyn(chla_bacu_prep,chla_maastr_prep,

                                  "C. nebrascensis\nB. baculus-B. grandis\nvs.
Maastrichtian","Cnebr_bacu_to_maas")


#### Summarize Temporal Resolution Comparisons and Export Tables ####

# Combine the results of observed overlap

make_D_I_row <- function(data1,name,interval){

  named1 <- as.data.frame(cbind(name,interval,data1[[2]]$obs$D,data1[[2]]$obs$I))

  colnames(named1) <- c("species","Intervals_compared","D_val","I_val")

  named1[,c(3:4)] <- as.numeric(as.character(named1[,c(3:4)]))

  return(named1)

}


# Outgroup Tables


conv_redu_campmid_D_I_row <- make_D_I_row(conv_redu_campmid_overlap,"I. convexus","B. reduncus-B.
scotti to Mid Campanian")

conv_nebr_campup_D_I_row <- make_D_I_row(conv_nebr_campup_overlap,"I. convexus","D. nebrascense-E.
jenneyi to Upper Campanian")


bara_rees_campup_D_I_row <- make_D_I_row(bara_rees_campup_overlap,"I. barabini","B. reesidei-B. eliasi
to Upper Campanian")
```

```
bara_nebr_campup_D_I_row <- make_D_I_row(bara_nebr_campup_overlap,"I. barabini","D. nebrascense-E. jenneyi to Upper Campanian")


sage_rees_campup_D_I_row <- make_D_I_row(sage_rees_campup_overlap,"I. sagensis","B. reesidei-B. eliasi to Upper Campanian")

sage_comp_campup_D_I_row <- make_D_I_row(sage_comp_campup_overlap,"I. sagensis","B. compressus-B. cuneatus to Upper Campanian")


subl_redu_campmid_D_I_row <- make_D_I_row(subl_redu_campmid_overlap,"I. sublaevis","B. reduncus-B. scotti to Mid Campanian")

subl_nebr_campup_D_I_row <- make_D_I_row(subl_nebr_campup_overlap,"I. sublaevis","D. nebrascense-E. jenneyi to Upper Campanian")


azer_macl_campmid_D_I_row <- make_D_I_row(azer_macl_campmid_overlap,"I. azerbaydjanensis","B. maclearni-B. sp. (smooth) to Mid Campanian")


oblo_rees_campup_D_I_row <- make_D_I_row(oblo_rees_campup_overlap,"I. oblongus","B. reesidei-B. eliasi to Upper Campanian")


sask_redu_campmid_D_I_row <- make_D_I_row(sask_macl_campmid_overlap,"I. saskatchewanensis","B. reduncus-B. scotti to Mid Campanian")


temp_res_I_D_results <-
rbind(conv_redu_campmid_D_I_row,conv_nebr_campup_D_I_row,bara_rees_campup_D_I_row,
                bara_nebr_campup_D_I_row,sage_rees_campup_D_I_row,sage_comp_campup_D_I_row,
                subl_redu_campmid_D_I_row,subl_nebr_campup_D_I_row,azer_macl_campmid_D_I_row,
                oblo_rees_campup_D_I_row,sask_redu_campmid_D_I_row)


# Outgroup tables


conv_redu_campmid_D_I_row <- make_D_I_row(cten_perp_campmid_overlap,"C. imbricatula","B. perplexus-B. gregoryensis to Mid Campanian")

conv_nebr_campup_D_I_row <- make_D_I_row(cten_redu_campmid_overlap,"C. imbricatula","B. reduncus-B. scotti to Upper Campanian")
```

```
bara_rees_campup_D_I_row <- make_D_I_row(cten_nebr_campup_overlap,"C. imbricatula","D. nebrascense-E. jenneyi to Upper Campanian")

bara_nebr_campup_D_I_row <- make_D_I_row(cten_rees_campup_overlap,"C. imbricatula","B. reesidei-B. eliasi to Upper Campanian")


sage_rees_campup_D_I_row <- make_D_I_row(luci_nebr_campup_overlap,"L. subundata","D. nebrascense-E. jenneyi to Upper Campanian")

sage_comp_campup_D_I_row <- make_D_I_row(luci_rees_campup_overlap,"L. subundata","B. reesidei-B. eliasi to Upper Campanian")


subl_redu_campmid_D_I_row <- make_D_I_row(chla_nebr_campup_overlap,"C. nebrascensis","D. nebrascense-E. jenneyi to Upper Campanian")

subl_nebr_campup_D_I_row <- make_D_I_row(chla_rees_campup_overlap,"C. nebrascensis","B. reesidei-B. eliasi to Upper Campanian")


azer_macl_campmid_D_I_row <- make_D_I_row(chla_bacu_maastrlow_overlap,"C. nebrascensis","B. baculus-B. grandis to Lower Maastrichtian")


oblo_rees_campup_D_I_row <- make_D_I_row(chla_bacu_maastr_overlap,"C. nebrascensis","B. baculus-B. grandis to Maastrichtian")



temp_res_outgroup_I_D_results <-
rbind(conv_redu_campmid_D_I_row,conv_nebr_campup_D_I_row,bara_rees_campup_D_I_row,
                    bara_nebr_campup_D_I_row,sage_rees_campup_D_I_row,sage_comp_campup_D_I_row,
                    subl_redu_campmid_D_I_row,subl_nebr_campup_D_I_row,azer_macl_campmid_D_I_row,
                    oblo_rees_campup_D_I_row)


# Combine the results of equivalency tests

make_exp_p_row <- function(data1,name,interval){

  named1 <- as.data.frame(cbind(name,interval,data1[[2]]$p.D,data1[[2]]$p.I,data1[[3]]$p.D,data1[[3]]$p.I))

  colnames(named1) <-
c("species","Intervals_compared","Greater_Eq_p_D","Greater_Eq_p_I","Lower_Eq_p_D","Lower_Eq_p_I")

  named1[,c(3:4)] <- as.numeric(as.character(named1[,c(3:4)]))

  return(named1)
```

```
}

conv_redu_campmid_eq_row <- make_exp_p_row(conv_redu_campmid_overlap,"I. convexus","B. reduncus-B. scotti to Mid Campanian")

conv_nebr_campup_eq_row <- make_exp_p_row(conv_nebr_campup_overlap,"I. convexus","D. nebrascense-E. jenneyi to Upper Campanian")


bara_rees_campup_eq_row <- make_exp_p_row(bara_rees_campup_overlap,"I. barabini","B. reesidei-B. eliasi to Upper Campanian")

bara_nebr_campup_eq_row <- make_exp_p_row(bara_nebr_campup_overlap,"I. barabini","D. nebrascense-E. jenneyi to Upper Campanian")


sage_rees_campup_eq_row <- make_exp_p_row(sage_rees_campup_overlap,"I. sagensis","B. reesidei-B. eliasi to Upper Campanian")

sage_comp_campup_eq_row <- make_exp_p_row(sage_comp_campup_overlap,"I. sagensis","B. compressus-B. cuneatus to Upper Campanian")


subl_redu_campmid_eq_row <- make_exp_p_row(subl_redu_campmid_overlap,"I. sublaevis","B. reduncus-B. scotti to Mid Campanian")

subl_nebr_campup_eq_row <- make_exp_p_row(subl_nebr_campup_overlap,"I. sublaevis","D. nebrascense-E. jenneyi to Upper Campanian")


azer_macl_campmid_eq_row <- make_exp_p_row(azer_macl_campmid_overlap,"I. azerbaydjanensis","B. maclearni-B. sp. (smooth) to Mid Campanian")


oblo_rees_campup_eq_row <- make_exp_p_row(oblo_rees_campup_overlap,"I. oblongus","B. reesidei-B. eliasi to Upper Campanian")


sask_redu_campmid_eq_row <- make_exp_p_row(sask_macl_campmid_overlap,"I. saskatchewanensis","B. reduncus-B. scotti to Mid Campanian")


temp_res_eq_results <-
rbind(conv_redu_campmid_eq_row,conv_nebr_campup_eq_row,bara_rees_campup_eq_row,
    bara_nebr_campup_eq_row,sage_rees_campup_eq_row,sage_comp_campup_eq_row,
    subl_redu_campmid_eq_row,subl_nebr_campup_eq_row,azer_macl_campmid_eq_row,
    oblo_rees_campup_eq_row,sask_redu_campmid_eq_row)
```

```
# Outgroup tables


conv_redu_campmid_eq_row <- make_exp_p_row(cten_perp_campmid_overlap,"C. imbricatula","B.
perplexus-B. gregoryensis to Mid Campanian")

conv_nebr_campup_eq_row <- make_exp_p_row(cten_redu_campmid_overlap,"C. imbricatula","B. reduncus-
B. scotti to Upper Campanian")


bara_rees_campup_eq_row <- make_exp_p_row(cten_nebr_campup_overlap,"C. imbricatula","D. nebrascense-
E. jenneyi to Upper Campanian")

bara_nebr_campup_eq_row <- make_exp_p_row(cten_rees_campup_overlap,"C. imbricatula","B. reesidei-B.
eliasi to Upper Campanian")


sage_rees_campup_eq_row <- make_exp_p_row(luci_nebr_campup_overlap,"L. subundata","D. nebrascense-E.
jenneyi to Upper Campanian")

sage_comp_campup_eq_row <- make_exp_p_row(luci_rees_campup_overlap,"L. subundata","B. reesidei-B.
eliasi to Upper Campanian")


subl_redu_campmid_eq_row <- make_exp_p_row(chla_nebr_campup_overlap,"C. nebrascensis","D.
nebrascense-E. jenneyi to Upper Campanian")

subl_nebr_campup_eq_row <- make_exp_p_row(chla_rees_campup_overlap,"C. nebrascensis","B. reesidei-B.
eliasi to Upper Campanian")


azer_macl_campmid_eq_row <- make_exp_p_row(chla_bacu_maastrlow_overlap,"C. nebrascensis","B.
baculus-B. grandis to Lower Maastrichtian")


oblo_rees_campup_eq_row <- make_exp_p_row(chla_bacu_maastr_overlap,"C. nebrascensis","B. baculus-B.
grandis to Maastrichtian")



temp_res_outgroup_eq_results <-
rbind(conv_redu_campmid_eq_row,conv_nebr_campup_eq_row,bara_rees_campup_eq_row,
                        bara_nebr_campup_eq_row,sage_rees_campup_eq_row,sage_comp_campup_eq_row,
                        subl_redu_campmid_eq_row,subl_nebr_campup_eq_row,azer_macl_campmid_eq_row,
                        oblo_rees_campup_eq_row)
```

```
# Combine the results of equivalency tests of Dynamics (p-vals)

make_exp_dyn_p_row <- function(data1,name,interval){

  named1 <-
as.data.frame(cbind(name,interval,data1[[2]]$p.expansion,data1[[2]]$p.stability,data1[[2]]$p.unfilling,

                      data1[[3]]$p.expansion,data1[[3]]$p.stability,data1[[3]]$p.unfilling))

  colnames(named1) <-
c("species","Intervals_compared","Greater_Eq_p_expansion","Greater_Eq_p_stability","Greater_Eq_p_unfillin
g",

                "Lower_Eq_p_expansion","Lower_Eq_p_stability","Lower_Eq_p_unfilling")

  named1[,c(3:4)] <- as.numeric(as.character(named1[,c(3:4)]))

  return(named1)

}


conv_redu_campmid_eq_dyn_p_row <- make_exp_dyn_p_row(conv_redu_campmid_overlap,"I. convexus","B.
reduncus-B. scotti to Mid Campanian")

conv_nebr_campup_eq_dyn_p_row <- make_exp_dyn_p_row(conv_nebr_campup_overlap,"I. convexus","D.
nebrascense-E. jenneyi to Upper Campanian")


bara_rees_campup_eq_dyn_p_row <- make_exp_dyn_p_row(bara_rees_campup_overlap,"I. barabini","B.
reesidei-B. eliasi to Upper Campanian")

bara_nebr_campup_eq_dyn_p_row <- make_exp_dyn_p_row(bara_nebr_campup_overlap,"I. barabini","D.
nebrascense-E. jenneyi to Upper Campanian")


sage_rees_campup_eq_dyn_p_row <- make_exp_dyn_p_row(sage_rees_campup_overlap,"I. sagensis","B.
reesidei-B. eliasi to Upper Campanian")

sage_comp_campup_eq_dyn_p_row <- make_exp_dyn_p_row(sage_comp_campup_overlap,"I. sagensis","B.
compressus-B. cuneatus to Upper Campanian")


subl_redu_campmid_eq_dyn_p_row <- make_exp_dyn_p_row(subl_redu_campmid_overlap,"I. sublaevis","B.
reduncus-B. scotti to Mid Campanian")

subl_nebr_campup_eq_dyn_p_row <- make_exp_dyn_p_row(subl_nebr_campup_overlap,"I. sublaevis","D.
nebrascense-E. jenneyi to Upper Campanian")


azer_macl_campmid_eq_dyn_p_row <- make_exp_dyn_p_row(azer_macl_campmid_overlap,"I.
azerbaydjanensis","B. maclearni-B. sp. (smooth) to Mid Campanian")
```

```
oblo_rees_campup_eq_dyn_p_row <- make_exp_dyn_p_row(oblo_rees_campup_overlap,"I. oblongus","B. reesidei-B. eliasi to Upper Campanian")


sask_redu_campmid_eq_dyn_p_row <- make_exp_dyn_p_row(sask_macl_campmid_overlap,"I. saskatchewanensis","B. reduncus-B. scotti to Mid Campanian")




temp_res_eq_dyn_p_results <- rbind(conv_redu_campmid_eq_dyn_p_row,conv_nebr_campup_eq_dyn_p_row,bara_rees_campup_eq_dyn_p_row,

bara_nebr_campup_eq_dyn_p_row,sage_rees_campup_eq_dyn_p_row,sage_comp_campup_eq_dyn_p_row,


subl_redu_campmid_eq_dyn_p_row,subl_nebr_campup_eq_dyn_p_row,azer_macl_campmid_eq_dyn_p_row,
                    oblo_rees_campup_eq_dyn_p_row,sask_redu_campmid_eq_dyn_p_row)


# Outgroup tables


conv_redu_campmid_eq_dyn_p_row <- make_exp_dyn_p_row(cten_perp_campmid_overlap,"C. imbricatula","B. perplexus-B. gregoryensis to Mid Campanian")

conv_nebr_campup_eq_dyn_p_row <- make_exp_dyn_p_row(cten_redu_campmid_overlap,"C. imbricatula","B. reduncus-B. scotti to Upper Campanian")


bara_rees_campup_eq_dyn_p_row <- make_exp_dyn_p_row(cten_nebr_campup_overlap,"C. imbricatula","D. nebrascense-E. jenneyi to Upper Campanian")

bara_nebr_campup_eq_dyn_p_row <- make_exp_dyn_p_row(cten_rees_campup_overlap,"C. imbricatula","B. reesidei-B. eliasi to Upper Campanian")


sage_rees_campup_eq_dyn_p_row <- make_exp_dyn_p_row(luci_nebr_campup_overlap,"L. subundata","D. nebrascense-E. jenneyi to Upper Campanian")

sage_comp_campup_eq_dyn_p_row <- make_exp_dyn_p_row(luci_rees_campup_overlap,"L. subundata","B. reesidei-B. eliasi to Upper Campanian")


subl_redu_campmid_eq_dyn_p_row <- make_exp_dyn_p_row(chla_nebr_campup_overlap,"C. nebrascensis","D. nebrascense-E. jenneyi to Upper Campanian")

subl_nebr_campup_eq_dyn_p_row <- make_exp_dyn_p_row(chla_rees_campup_overlap,"C. nebrascensis","B. reesidei-B. eliasi to Upper Campanian")
```

```
azer_macl_campmid_eq_dyn_p_row <- make_exp_dyn_p_row(chla_bacu_maastrlow_overlap,"C. nebrascensis","B. baculus-B. grandis to Lower Maastrichtian")


oblo_rees_campup_eq_dyn_p_row <- make_exp_dyn_p_row(chla_bacu_maastr_overlap,"C. nebrascensis","B. baculus-B. grandis to Maastrichtian")



temp_res_outgroup_eq_dyn_p_results <- rbind(conv_redu_campmid_eq_dyn_p_row,conv_nebr_campup_eq_dyn_p_row,bara_rees_campup_eq_dyn_p_row,

bara_nebr_campup_eq_dyn_p_row,sage_rees_campup_eq_dyn_p_row,sage_comp_campup_eq_dyn_p_row,

subl_redu_campmid_eq_dyn_p_row,subl_nebr_campup_eq_dyn_p_row,azer_macl_campmid_eq_dyn_p_row,
                        oblo_rees_campup_eq_dyn_p_row)


# Combine the results of similarity tests
make_sim_p_row <- function(data1,name,interval){
  named1 <- as.data.frame(cbind(name,interval,data1[[4]]$p.D,data1[[4]]$p.I,data1[[5]]$p.D,data1[[5]]$p.I))
  colnames(named1) <- c("species","Intervals_compared","Greater_Sim_p_D","Greater_Sim_p_I","Lower_Sim_p_D","Lower_Sim_p_I")
  named1[,c(3:4)] <- as.numeric(as.character(named1[,c(3:4)]))
  return(named1)
}


conv_redu_campmid_sim_row <- make_sim_p_row(conv_redu_campmid_overlap,"I. convexus","B. reduncus-B. scotti to Mid Campanian")
conv_nebr_campup_sim_row <- make_sim_p_row(conv_nebr_campup_overlap,"I. convexus","D. nebrascense-E. jenneyi to Upper Campanian")


bara_rees_campup_sim_row <- make_sim_p_row(bara_rees_campup_overlap,"I. barabini","B. reesidei-B. eliasi to Upper Campanian")
bara_nebr_campup_sim_row <- make_sim_p_row(bara_nebr_campup_overlap,"I. barabini","D. nebrascense-E. jenneyi to Upper Campanian")
```

```
sage_rees_campup_sim_row <- make_sim_p_row(sage_rees_campup_overlap,"I. sagensis","B. reesidei-B. eliasi to Upper Campanian")

sage_comp_campup_sim_row <- make_sim_p_row(sage_comp_campup_overlap,"I. sagensis","B. compressus-B. cuneatus to Upper Campanian")


subl_redu_campmid_sim_row <- make_sim_p_row(subl_redu_campmid_overlap,"I. sublaevis","B. reduncus-B. scotti to Mid Campanian")

subl_nebr_campup_sim_row <- make_sim_p_row(subl_nebr_campup_overlap,"I. sublaevis","D. nebrascense-E. jenneyi to Upper Campanian")


azer_macl_campmid_sim_row <- make_sim_p_row(azer_macl_campmid_overlap,"I. azerbaydjanensis","B. maclearni-B. sp. (smooth) to Mid Campanian")


oblo_rees_campup_sim_row <- make_sim_p_row(oblo_rees_campup_overlap,"I. oblongus","B. reesidei-B. eliasi to Upper Campanian")


sask_redu_campmid_sim_row <- make_sim_p_row(sask_macl_campmid_overlap,"I. saskatchewanensis","B. reduncus-B. scotti to Mid Campanian")


temp_res_sim_results <-
rbind(conv_redu_campmid_sim_row,conv_nebr_campup_sim_row,bara_rees_campup_sim_row,

    bara_nebr_campup_sim_row,sage_rees_campup_sim_row,sage_comp_campup_sim_row,

    subl_redu_campmid_sim_row,subl_nebr_campup_sim_row,azer_macl_campmid_sim_row,

    oblo_rees_campup_sim_row,sask_redu_campmid_sim_row)


# Outgroup tables


conv_redu_campmid_sim_row <- make_sim_p_row(cten_perp_campmid_overlap,"C. imbricatula","B. perplexus-B. gregoryensis to Mid Campanian")

conv_nebr_campup_sim_row <- make_sim_p_row(cten_redu_campmid_overlap,"C. imbricatula","B. reduncus-B. scotti to Upper Campanian")


bara_rees_campup_sim_row <- make_sim_p_row(cten_nebr_campup_overlap,"C. imbricatula","D. nebrascense-E. jenneyi to Upper Campanian")
```

```
bara_nebr_campup_sim_row <- make_sim_p_row(cten_rees_campup_overlap,"C. imbricatula","B. reesidei-B. eliasi to Upper Campanian")


sage_rees_campup_sim_row <- make_sim_p_row(luci_nebr_campup_overlap,"L. subundata","D. nebrascense-E. jenneyi to Upper Campanian")

sage_comp_campup_sim_row <- make_sim_p_row(luci_rees_campup_overlap,"L. subundata","B. reesidei-B. eliasi to Upper Campanian")


subl_redu_campmid_sim_row <- make_sim_p_row(chla_nebr_campup_overlap,"C. nebrascensis","D. nebrascense-E. jenneyi to Upper Campanian")

subl_nebr_campup_sim_row <- make_sim_p_row(chla_rees_campup_overlap,"C. nebrascensis","B. reesidei-B. eliasi to Upper Campanian")


azer_macl_campmid_sim_row <- make_sim_p_row(chla_bacu_maastrlow_overlap,"C. nebrascensis","B. baculus-B. grandis to Lower Maastrichtian")


oblo_rees_campup_sim_row <- make_sim_p_row(chla_bacu_maastr_overlap,"C. nebrascensis","B. baculus-B. grandis to Maastrichtian")


temp_res_outgroup_sim_results <-
rbind(conv_redu_campmid_sim_row,conv_nebr_campup_sim_row,bara_rees_campup_sim_row,


bara_nebr_campup_sim_row,sage_rees_campup_sim_row,sage_comp_campup_sim_row,


subl_redu_campmid_sim_row,subl_nebr_campup_sim_row,azer_macl_campmid_sim_row,

                                oblo_rees_campup_sim_row)


# Combine the results of equivalency tests of Dynamics (p-vals)

make_sim_dyn_p_row <- function(data1,name,interval){

  named1 <-
as.data.frame(cbind(name,interval,data1[[4]]$p.expansion,data1[[4]]$p.stability,data1[[4]]$p.unfilling,

                        data1[[5]]$p.expansion,data1[[5]]$p.stability,data1[[5]]$p.unfilling))

  colnames(named1) <-
c("species","Intervals_compared","Greater_Sim_p_expansion","Greater_Sim_p_stability","Greater_Sim_p_unfilling",

                  "Lower_Sim_p_expansion","Lower_Sim_p_stability","Lower_Sim_p_unfilling")
```

```
  named1[,c(3:4)] <- as.numeric(as.character(named1[,c(3:4)]))
  return(named1)
}

conv_redu_campmid_sim_dyn_p_row <- make_sim_dyn_p_row(conv_redu_campmid_overlap,"I. convexus","B. reduncus-B. scotti to Mid Campanian")
conv_nebr_campup_sim_dyn_p_row <- make_sim_dyn_p_row(conv_nebr_campup_overlap,"I. convexus","D. nebrascense-E. jenneyi to Upper Campanian")

bara_rees_campup_sim_dyn_p_row <- make_sim_dyn_p_row(bara_rees_campup_overlap,"I. barabini","B. reesidei-B. eliasi to Upper Campanian")
bara_nebr_campup_sim_dyn_p_row <- make_sim_dyn_p_row(bara_nebr_campup_overlap,"I. barabini","D. nebrascense-E. jenneyi to Upper Campanian")

sage_rees_campup_sim_dyn_p_row <- make_sim_dyn_p_row(sage_rees_campup_overlap,"I. sagensis","B. reesidei-B. eliasi to Upper Campanian")
sage_comp_campup_sim_dyn_p_row <- make_sim_dyn_p_row(sage_comp_campup_overlap,"I. sagensis","B. compressus-B. cuneatus to Upper Campanian")

subl_redu_campmid_sim_dyn_p_row <- make_sim_dyn_p_row(subl_redu_campmid_overlap,"I. sublaevis","B. reduncus-B. scotti to Mid Campanian")
subl_nebr_campup_sim_dyn_p_row <- make_sim_dyn_p_row(subl_nebr_campup_overlap,"I. sublaevis","D. nebrascense-E. jenneyi to Upper Campanian")

azer_macl_campmid_sim_dyn_p_row <- make_sim_dyn_p_row(azer_macl_campmid_overlap,"I. azerbaydjanensis","B. maclearni-B. sp. (smooth) to Mid Campanian")

oblo_rees_campup_sim_dyn_p_row <- make_sim_dyn_p_row(oblo_rees_campup_overlap,"I. oblongus","B. reesidei-B. eliasi to Upper Campanian")

sask_redu_campmid_sim_dyn_p_row <- make_sim_dyn_p_row(sask_macl_campmid_overlap,"I. saskatchewanensis","B. reduncus-B. scotti to Mid Campanian")
```

```
temp_res_sim_dyn_p_results <-
rbind(conv_redu_campmid_sim_dyn_p_row,conv_nebr_campup_sim_dyn_p_row,bara_rees_campup_sim_dyn
_p_row,

bara_nebr_campup_sim_dyn_p_row,sage_rees_campup_sim_dyn_p_row,sage_comp_campup_sim_dyn_p_row,

subl_redu_campmid_sim_dyn_p_row,subl_nebr_campup_sim_dyn_p_row,azer_macl_campmid_sim_dyn_p_ro
w,
                    oblo_rees_campup_sim_dyn_p_row,sask_redu_campmid_sim_dyn_p_row)


# Outgroup tables


conv_redu_campmid_sim_dyn_p_row <- make_sim_dyn_p_row(cten_perp_campmid_overlap,"C.
imbricatula","B. perplexus-B. gregoryensis to Mid Campanian")

conv_nebr_campup_sim_dyn_p_row <- make_sim_dyn_p_row(cten_redu_campmid_overlap,"C.
imbricatula","B. reduncus-B. scotti to Upper Campanian")


bara_rees_campup_sim_dyn_p_row <- make_sim_dyn_p_row(cten_nebr_campup_overlap,"C. imbricatula","D.
nebrascense-E. jenneyi to Upper Campanian")

bara_nebr_campup_sim_dyn_p_row <- make_sim_dyn_p_row(cten_rees_campup_overlap,"C. imbricatula","B.
reesidei-B. eliasi to Upper Campanian")


sage_rees_campup_sim_dyn_p_row <- make_sim_dyn_p_row(luci_nebr_campup_overlap,"L. subundata","D.
nebrascense-E. jenneyi to Upper Campanian")

sage_comp_campup_sim_dyn_p_row <- make_sim_dyn_p_row(luci_rees_campup_overlap,"L. subundata","B.
reesidei-B. eliasi to Upper Campanian")


subl_redu_campmid_sim_dyn_p_row <- make_sim_dyn_p_row(chla_nebr_campup_overlap,"C.
nebrascensis","D. nebrascense-E. jenneyi to Upper Campanian")

subl_nebr_campup_sim_dyn_p_row <- make_sim_dyn_p_row(chla_rees_campup_overlap,"C.
nebrascensis","B. reesidei-B. eliasi to Upper Campanian")


azer_macl_campmid_sim_dyn_p_row <- make_sim_dyn_p_row(chla_bacu_maastrlow_overlap,"C.
nebrascensis","B. baculus-B. grandis to Lower Maastrichtian")


oblo_rees_campup_sim_dyn_p_row <- make_sim_dyn_p_row(chla_bacu_maastr_overlap,"C. nebrascensis","B.
baculus-B. grandis to Maastrichtian")
```

```
temp_res_outgroup_sim_dyn_p_results <-
rbind(conv_redu_campmid_sim_dyn_p_row,conv_nebr_campup_sim_dyn_p_row,bara_rees_campup_sim_dyn
_p_row,


bara_nebr_campup_sim_dyn_p_row,sage_rees_campup_sim_dyn_p_row,sage_comp_campup_sim_dyn_p_row,


subl_redu_campmid_sim_dyn_p_row,subl_nebr_campup_sim_dyn_p_row,azer_macl_campmid_sim_dyn_p_ro
w,
                              oblo_rees_campup_sim_dyn_p_row)


# Combine the results of niche dynamics

make_dynam_row <- function(data1,name,interval){

  data <- as.data.frame(data1[[6]])

  named1 <- as.data.frame(cbind(name,interval,t(data[,2])))

  colnames(named1) <- c("species","Intervals_compared","Expansion","Stability","Unfilling")

  named1[,c(3:5)] <- as.numeric(as.character(named1[,c(3:5)]))

  return(named1)

}


conv_redu_campmid_dyn_row <- make_dynam_row(conv_redu_campmid_overlap,"I. convexus","B. reduncus-
B. scotti to Mid Campanian")

conv_nebr_campup_dyn_row <- make_dynam_row(conv_nebr_campup_overlap,"I. convexus","D.
nebrascense-E. jenneyi to Upper Campanian")


bara_rees_campup_dyn_row <- make_dynam_row(bara_rees_campup_overlap,"I. barabini","B. reesidei-B.
eliasi to Upper Campanian")

bara_nebr_campup_dyn_row <- make_dynam_row(bara_nebr_campup_overlap,"I. barabini","D. nebrascense-E.
jenneyi to Upper Campanian")


sage_rees_campup_dyn_row <- make_dynam_row(sage_rees_campup_overlap,"I. sagensis","B. reesidei-B.
eliasi to Upper Campanian")

sage_comp_campup_dyn_row <- make_dynam_row(sage_comp_campup_overlap,"I. sagensis","B. compressus-
B. cuneatus to Upper Campanian")
```

```
subl_redu_campmid_dyn_row <- make_dynam_row(subl_redu_campmid_overlap,"I. sublaevis","B. reduncus-B. scotti to Mid Campanian")

subl_nebr_campup_dyn_row <- make_dynam_row(subl_nebr_campup_overlap,"I. sublaevis","D. nebrascense-E. jenneyi to Upper Campanian")


azer_macl_campmid_dyn_row <- make_dynam_row(azer_macl_campmid_overlap,"I. azerbaydjanensis","B. maclearni-B. sp. (smooth) to Mid Campanian")


oblo_rees_campup_dyn_row <- make_dynam_row(oblo_rees_campup_overlap,"I. oblongus","B. reesidei-B. eliasi to Upper Campanian")


sask_redu_campmid_dyn_row <- make_dynam_row(sask_macl_campmid_overlap,"I. saskatchewanensis","B. reduncus-B. scotti to Mid Campanian")


temp_res_dyn_results <-
rbind(conv_redu_campmid_dyn_row,conv_nebr_campup_dyn_row,bara_rees_campup_dyn_row,
    bara_nebr_campup_dyn_row,sage_rees_campup_dyn_row,sage_comp_campup_dyn_row,
    subl_redu_campmid_dyn_row,subl_nebr_campup_dyn_row,azer_macl_campmid_dyn_row,
    oblo_rees_campup_dyn_row,sask_redu_campmid_dyn_row)


# Outgroup tables


conv_redu_campmid_dyn_row <- make_dynam_row(cten_perp_campmid_overlap,"C. imbricatula","B. perplexus-B. gregoryensis to Mid Campanian")

conv_nebr_campup_dyn_row <- make_dynam_row(cten_redu_campmid_overlap,"C. imbricatula","B. reduncus-B. scotti to Upper Campanian")


bara_rees_campup_dyn_row <- make_dynam_row(cten_nebr_campup_overlap,"C. imbricatula","D. nebrascense-E. jenneyi to Upper Campanian")

bara_nebr_campup_dyn_row <- make_dynam_row(cten_rees_campup_overlap,"C. imbricatula","B. reesidei-B. eliasi to Upper Campanian")


sage_rees_campup_dyn_row <- make_dynam_row(luci_nebr_campup_overlap,"L. subundata","D. nebrascense-E. jenneyi to Upper Campanian")

sage_comp_campup_dyn_row <- make_dynam_row(luci_rees_campup_overlap,"L. subundata","B. reesidei-B. eliasi to Upper Campanian")
```

```
subl_redu_campmid_dyn_row <- make_dynam_row(chla_nebr_campup_overlap,"C. nebrascensis","D. nebrascense-E. jenneyi to Upper Campanian")

subl_nebr_campup_dyn_row <- make_dynam_row(chla_rees_campup_overlap,"C. nebrascensis","B. reesidei-B. eliasi to Upper Campanian")


azer_macl_campmid_dyn_row <- make_dynam_row(chla_bacu_maastrlow_overlap,"C. nebrascensis","B. baculus-B. grandis to Lower Maastrichtian")


oblo_rees_campup_dyn_row <- make_dynam_row(chla_bacu_maastr_overlap,"C. nebrascensis","B. baculus-B. grandis to Maastrichtian")


temp_res_outgroup_dyn_results <-
rbind(conv_redu_campmid_dyn_row,conv_nebr_campup_dyn_row,bara_rees_campup_dyn_row,


bara_nebr_campup_dyn_row,sage_rees_campup_dyn_row,sage_comp_campup_dyn_row,


subl_redu_campmid_dyn_row,subl_nebr_campup_dyn_row,azer_macl_campmid_dyn_row,
                                oblo_rees_campup_dyn_row)


## Write csv files of results

write.csv(temp_res_outgroup_I_D_results,file="tables/temp_res_outgroup_I_D_results.csv")

write.csv(temp_res_outgroup_eq_results,file="tables/temp_res_outgroup_eq_results.csv")

write.csv(temp_res_outgroup_sim_results,file="tables/temp_res_outgroup_sim_results.csv")

write.csv(temp_res_outgroup_dyn_results,file="tables/temp_res_outgroup_dyn_results.csv")

write.csv(temp_res_outgroup_eq_dyn_p_results,file="tables/temp_res_outgroup_eq_dyn_p_results.csv")

write.csv(temp_res_outgroup_sim_dyn_p_results,file="tables/temp_res_outgroup_sim_dyn_p_results.csv")


#### SPECIES WITHIN SAME INTERVAL COMPARISON ####

# Use above function to make plots of overlap comparison between spp and produce list of overlap metric results


conv_to_subl_redu_overlap <- run_ecospat_dyn(conv_redu_prep,subl_redu_prep,
                                "I. convextus to I. sublaevis\nB. reduncus-B. scotti","Iconv_to_Isubl_redu_")
```

```
conv_to_sask_redu_overlap <- run_ecospat_dyn(conv_redu_prep,sask_redu_prep,
                                "I. convextus to I. saskatchewanensis\nB. reduncus-B.
scotti","Iconv_to_Isask_redu_")

subl_to_sask_redu_overlap <- run_ecospat_dyn(subl_redu_prep,sask_redu_prep,
                                "I. sublaevis to I. saskatchewanensis\nB. reduncus-B.
scotti","Isubl_to_Isask_redu_")

conv_to_bara_nebr_overlap <- run_ecospat_dyn(conv_nebr_prep,bara_nebr_prep,
                               "I. convextus to I. barabini\nD. nebrascense-E. jenneyi","Iconv_to_Ibara_nebr")

conv_to_subl_nebr_overlap <- run_ecospat_dyn(conv_nebr_prep,subl_nebr_prep,
                                "I. convextus to I. sublaevis\nD. nebrascense-E. jenneyi","Iconv_to_Isubl_nebr")

bara_to_subl_nebr_overlap <- run_ecospat_dyn(bara_nebr_prep,subl_nebr_prep,
                                "I. barabini to I. sublaevis\nD. nebrascense-E. jenneyi","Ibara_to_Isubl_nebr")

bara_to_sage_rees_overlap <- run_ecospat_dyn(bara_rees_prep,sage_rees_prep,
                                "I. barabini to I. sagensis\nB. reesidei-B. eliasi","Ibara_to_Isage_rees")

bara_to_oblo_rees_overlap <- run_ecospat_dyn(bara_rees_prep,oblo_rees_prep,
                                "I. barabini to I. oblongus\nB. reesidei-B. eliasi","Ibara_to_Ioblo_rees")

oblo_to_sage_rees_overlap <- run_ecospat_dyn(oblo_rees_prep,sage_rees_prep,
                                "I. oblongus to I. sagensis\nB. reesidei-B. eliasi","Ioblo_to_Isage_rees")

#### Summarize Species Comparisons and Export Tables ####
# Combine the results of observed overlap
make_D_I_row <- function(data1,name,interval){
  named1 <- as.data.frame(cbind(name,interval,data1[[2]]$obs$D,data1[[2]]$obs$I))
  colnames(named1) <- c("species","Intervals_compared","D_val","I_val")
```

```
  named1[,c(3:4)] <- as.numeric(as.character(named1[,c(3:4)]))
  return(named1)
}

conv_to_subl_redu_D_I_row <- make_D_I_row(conv_to_subl_redu_overlap,"I. convexus vs. I. sublaevis","B. reduncus-B. scotti")

conv_to_sask_redu_D_I_row <- make_D_I_row(conv_to_sask_redu_overlap,"I. convexus vs. I. saskatchewanensis","B. reduncus-B. scotti")

subl_to_sask_redu_D_I_row <- make_D_I_row(subl_to_sask_redu_overlap,"I. sublaevis vs. I. saskatchewanensis","B. reduncus-B. scotti")

conv_to_bara_nebr_D_I_row <- make_D_I_row(conv_to_bara_nebr_overlap,"I. convexus vs. I. barabini","D. nebrascense-E. jenneyi")

conv_to_subl_nebr_D_I_row <- make_D_I_row(conv_to_subl_nebr_overlap,"I. convexus vs. I. sublaevis","B. reesidei-B. eliasi")

bara_to_subl_nebr_D_I_row <- make_D_I_row(bara_to_subl_nebr_overlap,"I. barabini vs. I. sublaevis","D. nebrascense-E. jenneyi")

bara_to_sage_rees_D_I_row <- make_D_I_row(bara_to_sage_rees_overlap,"I. barabini vs. I. sagensis","B. reesidei-B. eliasi")

bara_to_oblo_rees_D_I_row <- make_D_I_row(bara_to_oblo_rees_overlap,"I. barabini vs. I. oblongus","B. reesidei-B. eliasi")

oblo_to_sage_rees_D_I_row <- make_D_I_row(oblo_to_sage_rees_overlap,"I. oblongus vs. I. sagensis","B. reesidei-B. eliasi")

species_compare_I_D_results <- rbind(conv_to_subl_redu_D_I_row,conv_to_sask_redu_D_I_row,
                                     subl_to_sask_redu_D_I_row,conv_to_bara_nebr_D_I_row,
                                     conv_to_subl_nebr_D_I_row,
```

```
                bara_to_subl_nebr_D_I_row,bara_to_sage_rees_D_I_row,
                bara_to_oblo_rees_D_I_row,oblo_to_sage_rees_D_I_row)


# Combine the results of equivalency tests
make_exp_p_row <- function(data1,name,interval){
  named1 <- as.data.frame(cbind(name,interval,data1[[2]]$p.D,data1[[2]]$p.I,data1[[3]]$p.D,data1[[3]]$p.I))
  colnames(named1) <-
c("species","Intervals_compared","Greater_Eq_p_D","Greater_Eq_p_I","Lower_Eq_p_D","Lower_Eq_p_I")
  named1[,c(3:4)] <- as.numeric(as.character(named1[,c(3:4)]))
  return(named1)
}


conv_to_subl_eq_row <- make_exp_p_row(conv_to_subl_redu_overlap,"I. convexus vs. I. sublaevis","B.
reduncus-B. scotti")


conv_to_sask_eq_row <- make_exp_p_row(conv_to_sask_redu_overlap,"I. convexus vs. I.
saskatchewanensis","B. reduncus-B. scotti")


subl_to_sask_eq_row <- make_exp_p_row(subl_to_sask_redu_overlap,"I. sublaevis vs. I.
saskatchewanensis","B. reduncus-B. scotti")


conv_to_bara_eq_row <- make_exp_p_row(conv_to_bara_nebr_overlap,"I. convexus vs. I. barabini","D.
nebrascense-E. jenneyi")


conv_to_subl_eq_row <- make_exp_p_row(conv_to_subl_nebr_overlap,"I. convexus vs. I. sublaevis","D.
nebrascense-E. jenneyi")


bara_to_subl_eq_row <- make_exp_p_row(bara_to_subl_nebr_overlap,"I. barabini vs. I. sublaevis","D.
nebrascense-E. jenneyi")


bara_to_sage_eq_row <- make_exp_p_row(bara_to_sage_rees_overlap,"I. barabini vs. I. sagensis","B. reesidei-
B. eliasi")


bara_to_oblo_eq_row <- make_exp_p_row(bara_to_oblo_rees_overlap,"I. barabini vs. I. oblongus","B. reesidei-
B. eliasi")
```

```
oblo_to_sage_eq_row <- make_exp_p_row(oblo_to_sage_rees_overlap,"I. oblongus vs. I. sagensis","B.
reesidei-B. eliasi")



species_compare_eq_results <- rbind(conv_to_subl_eq_row,conv_to_sask_eq_row,
                                    subl_to_sask_eq_row,conv_to_bara_eq_row,
                                    conv_to_subl_eq_row,
                                    bara_to_subl_eq_row,bara_to_sage_eq_row,
                                    bara_to_oblo_eq_row,oblo_to_sage_eq_row)


# Combine the results of equivalency tests of Dynamics (p-vals)
make_exp_dyn_p_row <- function(data1,name,interval){
  named1 <-
as.data.frame(cbind(name,interval,data1[[2]]$p.expansion,data1[[2]]$p.stability,data1[[2]]$p.unfilling,
                      data1[[3]]$p.expansion,data1[[3]]$p.stability,data1[[3]]$p.unfilling))
  colnames(named1) <-
c("species","Intervals_compared","Greater_Eq_p_expansion","Greater_Eq_p_stability","Greater_Eq_p_unfillin
g",
                  "Lower_Eq_p_expansion","Lower_Eq_p_stability","Lower_Eq_p_unfilling")
  named1[,c(3:4)] <- as.numeric(as.character(named1[,c(3:4)]))
  return(named1)
}
conv_to_subl_eq_dyn_p_row <- make_exp_dyn_p_row(conv_to_subl_redu_overlap,"I. convexus vs. I.
sublaevis","B. reduncus-B. scotti")


conv_to_sask_eq_dyn_p_row <- make_exp_dyn_p_row(conv_to_sask_redu_overlap,"I. convexus vs. I.
saskatchewanensis","B. reduncus-B. scotti")


subl_to_sask_eq_dyn_p_row <- make_exp_dyn_p_row(subl_to_sask_redu_overlap,"I. sublaevis vs. I.
saskatchewanensis","B. reduncus-B. scotti")


conv_to_bara_eq_dyn_p_row <- make_exp_dyn_p_row(conv_to_bara_nebr_overlap,"I. convexus vs. I.
barabini","D. nebrascense-E. jenneyi")
```

```
conv_to_subl_eq_dyn_p_row <- make_exp_dyn_p_row(conv_to_subl_nebr_overlap,"I. convexus vs. I. sublaevis","D. nebrascense-E. jenneyi")


bara_to_subl_eq_dyn_p_row <- make_exp_dyn_p_row(bara_to_subl_nebr_overlap,"I. barabini vs. I. sublaevis","D. nebrascense-E. jenneyi")


bara_to_sage_eq_dyn_p_row <- make_exp_dyn_p_row(bara_to_sage_rees_overlap,"I. barabini vs. I. sagensis","B. reesidei-B. eliasi")


bara_to_oblo_eq_dyn_p_row <- make_exp_dyn_p_row(bara_to_oblo_rees_overlap,"I. barabini vs. I. oblongus","B. reesidei-B. eliasi")


oblo_to_sage_eq_dyn_p_row <- make_exp_dyn_p_row(oblo_to_sage_rees_overlap,"I. oblongus vs. I. sagensis","B. reesidei-B. eliasi")



species_compare_eq_dyn_p_results <- rbind(conv_to_subl_eq_dyn_p_row,conv_to_sask_eq_dyn_p_row,
                                          subl_to_sask_eq_dyn_p_row,conv_to_bara_eq_dyn_p_row,
                                          conv_to_subl_eq_dyn_p_row,
                                          bara_to_subl_eq_dyn_p_row,bara_to_sage_eq_dyn_p_row,
                                          bara_to_oblo_eq_dyn_p_row,oblo_to_sage_eq_dyn_p_row)


# Combine the results of similarity tests
make_sim_p_row <- function(data1,name,interval){
  named1 <- as.data.frame(cbind(name,interval,data1[[4]]$p.D,data1[[4]]$p.I,data1[[5]]$p.D,data1[[5]]$p.I))
  colnames(named1) <- c("species","Intervals_compared","Greater_Sim_p_D","Greater_Sim_p_I","Lower_Sim_p_D","Lower_Sim_p_I")
  named1[,c(3:4)] <- as.numeric(as.character(named1[,c(3:4)]))
  return(named1)
}
```

```
conv_to_subl_sim_row <- make_sim_p_row(conv_to_subl_redu_overlap,"I. convexus vs. I. sublaevis","B.
reduncus-B. scotti")


conv_to_sask_sim_row <- make_sim_p_row(conv_to_sask_redu_overlap,"I. convexus vs. I.
saskatchewanensis","B. reduncus-B. scotti")


subl_to_sask_sim_row <- make_sim_p_row(subl_to_sask_redu_overlap,"I. sublaevis vs. I.
saskatchewanensis","B. reduncus-B. scotti")


conv_to_bara_sim_row <- make_sim_p_row(conv_to_bara_nebr_overlap,"I. convexus vs. I. barabini","D.
nebrascense-E. jenneyi")


conv_to_subl_sim_row <- make_sim_p_row(conv_to_subl_nebr_overlap,"I. convexus vs. I. sublaevis","D.
nebrascense-E. jenneyi")


bara_to_subl_sim_row <- make_sim_p_row(bara_to_subl_nebr_overlap,"I. barabini vs. I. sublaevis","D.
nebrascense-E. jenneyi")


bara_to_sage_sim_row <- make_sim_p_row(bara_to_sage_rees_overlap,"I. barabini vs. I. sagensis","B.
reesidei-B. eliasi")


bara_to_oblo_sim_row <- make_sim_p_row(bara_to_oblo_rees_overlap,"I. barabini vs. I. oblongus","B.
reesidei-B. eliasi")


oblo_to_sage_sim_row <- make_sim_p_row(oblo_to_sage_rees_overlap,"I. oblongus vs. I. sagensis","B.
reesidei-B. eliasi")




species_compare_sim_results <- rbind(conv_to_subl_sim_row,conv_to_sask_sim_row,
                                     subl_to_sask_sim_row,conv_to_bara_sim_row,
                                     conv_to_subl_sim_row,
                                     bara_to_subl_sim_row,bara_to_sage_sim_row,
                                     bara_to_oblo_sim_row,oblo_to_sage_sim_row)


# Combine the results of equivalency tests of Dynamics (p-vals)
```

```
make_sim_dyn_p_row <- function(data1,name,interval){

  named1 <-
as.data.frame(cbind(name,interval,data1[[4]]$p.expansion,data1[[4]]$p.stability,data1[[4]]$p.unfilling,

                      data1[[5]]$p.expansion,data1[[5]]$p.stability,data1[[5]]$p.unfilling))

  colnames(named1) <-
c("species","Intervals_compared","Greater_Sim_p_expansion","Greater_Sim_p_stability","Greater_Sim_p_unfi
lling",

                "Lower_Sim_p_expansion","Lower_Sim_p_stability","Lower_Sim_p_unfilling")

  named1[,c(3:4)] <- as.numeric(as.character(named1[,c(3:4)]))

  return(named1)

}


conv_to_subl_sim_dyn_p_row <- make_sim_dyn_p_row(conv_to_subl_redu_overlap,"I. convexus vs. I.
sublaevis","B. reduncus-B. scotti")


conv_to_sask_sim_dyn_p_row <- make_sim_dyn_p_row(conv_to_sask_redu_overlap,"I. convexus vs. I.
saskatchewanensis","B. reduncus-B. scotti")


subl_to_sask_sim_dyn_p_row <- make_sim_dyn_p_row(subl_to_sask_redu_overlap,"I. sublaevis vs. I.
saskatchewanensis","B. reduncus-B. scotti")


conv_to_bara_sim_dyn_p_row <- make_sim_dyn_p_row(conv_to_bara_nebr_overlap,"I. convexus vs. I.
barabini","D. nebrascense-E. jenneyi")


conv_to_subl_sim_dyn_p_row <- make_sim_dyn_p_row(conv_to_subl_nebr_overlap,"I. convexus vs. I.
sublaevis","D. nebrascense-E. jenneyi")


bara_to_subl_sim_dyn_p_row <- make_sim_dyn_p_row(bara_to_subl_nebr_overlap,"I. barabini vs. I.
sublaevis","D. nebrascense-E. jenneyi")


bara_to_sage_sim_dyn_p_row <- make_sim_dyn_p_row(bara_to_sage_rees_overlap,"I. barabini vs. I.
sagensis","B. reesidei-B. eliasi")


bara_to_oblo_sim_dyn_p_row <- make_sim_dyn_p_row(bara_to_oblo_rees_overlap,"I. barabini vs. I.
oblongus","B. reesidei-B. eliasi")
```

```
oblo_to_sage_sim_dyn_p_row <- make_sim_dyn_p_row(oblo_to_sage_rees_overlap,"I. oblongus vs. I. sagensis","B. reesidei-B. eliasi")



species_compare_sim_dyn_p_results <- rbind(conv_to_subl_sim_dyn_p_row,conv_to_sask_sim_dyn_p_row,
                                           subl_to_sask_sim_dyn_p_row,conv_to_bara_sim_dyn_p_row,
                                           conv_to_subl_sim_dyn_p_row,
                                           bara_to_subl_sim_dyn_p_row,bara_to_sage_sim_dyn_p_row,
                                           bara_to_oblo_sim_dyn_p_row,oblo_to_sage_sim_dyn_p_row)


# Combine the results of niche dynamics
make_dynam_row <- function(data1,name,interval){
  data <- as.data.frame(data1[[6]])
  named1 <- as.data.frame(cbind(name,interval,t(data[,2])))
  colnames(named1) <- c("species","Intervals_compared","Expansion","Stability","Unfilling")
  named1[,c(3:5)] <- as.numeric(as.character(named1[,c(3:5)]))
  return(named1)
}


conv_to_subl_dyn_row <- make_dynam_row(conv_to_subl_redu_overlap,"I. convexus vs. I. sublaevis","B. reduncus-B. scotti")


conv_to_sask_dyn_row <- make_dynam_row(conv_to_sask_redu_overlap,"I. convexus vs. I. saskatchewanensis","B. reduncus-B. scotti")


subl_to_sask_dyn_row <- make_dynam_row(subl_to_sask_redu_overlap,"I. sublaevis vs. I. saskatchewanensis","B. reduncus-B. scotti")


conv_to_bara_dyn_row <- make_dynam_row(conv_to_bara_nebr_overlap,"I. convexus vs. I. barabini","D. nebrascense-E. jenneyi")


conv_to_subl_dyn_row <- make_dynam_row(conv_to_subl_nebr_overlap,"I. convexus vs. I. sublaevis","D. nebrascense-E. jenneyi")
```

```
bara_to_subl_dyn_row <- make_dynam_row(bara_to_subl_nebr_overlap,"I. barabini vs. I. sublaevis","D. nebrascense-E. jenneyi")


bara_to_sage_dyn_row <- make_dynam_row(bara_to_sage_rees_overlap,"I. barabini vs. I. sagensis","B. reesidei-B. eliasi")


bara_to_oblo_dyn_row <- make_dynam_row(bara_to_oblo_rees_overlap,"I. barabini vs. I. oblongus","B. reesidei-B. eliasi")


oblo_to_sage_dyn_row <- make_dynam_row(oblo_to_sage_rees_overlap,"I. oblongus vs. I. sagensis","B. reesidei-B. eliasi")


species_compare_dyn_results <- rbind(conv_to_subl_dyn_row,conv_to_sask_dyn_row,
                        subl_to_sask_dyn_row,conv_to_bara_dyn_row,
                        conv_to_subl_dyn_row,
                        bara_to_subl_dyn_row,bara_to_sage_dyn_row,
                        bara_to_oblo_dyn_row,oblo_to_sage_dyn_row)


## Write csv files of results
write.csv(species_compare_I_D_results,file="tables/species_compare_I_D_results.csv")
write.csv(species_compare_eq_results,file="tables/species_compare_eq_results.csv")
write.csv(species_compare_sim_results,file="tables/species_compare_sim_results.csv")
write.csv(species_compare_dyn_results,file="tables/species_compare_dyn_results.csv")
write.csv(species_compare_eq_dyn_p_results,file="tables/species_compare_eq_dyn_p_results.csv")
write.csv(species_compare_sim_dyn_p_results,file="tables/species_compare_sim_dyn_p_results.csv")


#### INOCERAMUS TO OUTGROUP COMPARISONS ####


cten_to_conv_redu_overlap <- run_ecospat_dyn(cten_redu_prep,conv_redu_prep,
                        "C. imbricatula to I. convexus\nB. reduncus-B. scotti","Cimbr_to_Iconv_redu_")


cten_to_subl_redu_overlap <- run_ecospat_dyn(cten_redu_prep,subl_redu_prep,
                        "C. imbricatula to I. sublaevis\nB. reduncus-B. scotti","Cimbr_to_Isubl_redu_")
```

```
cten_to_sask_redu_overlap <- run_ecospat_dyn(cten_redu_prep,sask_redu_prep,
                                "C. imbricatula to I. saskatchewanensis\nB. reduncus-B.
scotti","Cimbr_to_Isask_redu_")

cten_to_bara_nebr_overlap <- run_ecospat_dyn(cten_nebr_prep,bara_nebr_prep,
                                "C. imbricatula to I. barabini\nD. nebrascense-E. jenneyi","Cimbr_to_Ibara_nebr_")

cten_to_conv_nebr_overlap <- run_ecospat_dyn(cten_nebr_prep,conv_nebr_prep,
                                "C. imbricatula to I. convexus\nD. nebrascense-E.
jenneyi","Cimbr_to_Iconv_nebr_")

cten_to_subl_nebr_overlap <- run_ecospat_dyn(cten_nebr_prep,subl_nebr_prep,
                                "C. imbricatula to I. sublaevis\nD. nebrascense-E.
jenneyi","Cimbr_to_Isubl_nebr_")

chla_to_bara_nebr_overlap <- run_ecospat_dyn(chla_nebr_prep,bara_nebr_prep,
                                "C. nebrascensis to I. barabini\nD. nebrascense-E.
jenneyi","Cnebr_to_Ibara_nebr_")

chla_to_conv_nebr_overlap <- run_ecospat_dyn(chla_nebr_prep,conv_nebr_prep,
                                "C. nebrascensis to I. convexus\nD. nebrascense-E.
jenneyi","Cnebr_to_Iconv_nebr_")

chla_to_subl_nebr_overlap <- run_ecospat_dyn(chla_nebr_prep,subl_nebr_prep,
                                "C. nebrascensis to I. sublaevis\nD. nebrascense-E.
jenneyi","Cnebr_to_Isubl_nebr_")

luci_to_bara_nebr_overlap <- run_ecospat_dyn(luci_nebr_prep,bara_nebr_prep,
                                "L. subundata to I. barabini\nD. nebrascense-E. jenneyi","Lsubu_to_Ibara_nebr_")

luci_to_conv_nebr_overlap <- run_ecospat_dyn(luci_nebr_prep,conv_nebr_prep,
                                "L. subundata to I. convexus\nD. nebrascense-E. jenneyi","Lsubu_to_Iconv_nebr_")
```

```
luci_to_subl_nebr_overlap <- run_ecospat_dyn(luci_nebr_prep,subl_nebr_prep,
                              "L. subundata to I. sublaevis\nD. nebrascense-E. jenneyi","Lsubu_to_Isubl_nebr_")

luci_to_sage_comp_overlap <- run_ecospat_dyn(luci_comp_prep,sage_comp_prep,
                              "L. subundata to I. sagensis\nB. compressus-B. cuneatus","Lsubu_to_Isage_comp")

chla_to_bara_rees_overlap <- run_ecospat_dyn(chla_rees_prep,bara_rees_prep,
                              "C. nebrascensis to I. barabini\nB. reesidei-B. eliasi","Cnebr_to_Ibara_rees")

chla_to_sage_rees_overlap <- run_ecospat_dyn(chla_rees_prep,sage_rees_prep,
                              "C. nebrascensis to I. sagensis\nB. reesidei-B. eliasi","Cnebr_to_Isage_rees")

chla_to_oblo_rees_overlap <- run_ecospat_dyn(chla_rees_prep,oblo_rees_prep,
                              "C. nebrascensis to I. oblongus\nB. reesidei-B. eliasi","Cnebr_to_Ioblo_rees")

cten_to_bara_rees_overlap <- run_ecospat_dyn(cten_rees_prep,bara_rees_prep,
                              "C. imbricatula to I. barabini\nB. reesidei-B. eliasi","Cimbr_to_Ibara_rees")

cten_to_sage_rees_overlap <- run_ecospat_dyn(cten_rees_prep,sage_rees_prep,
                              "C. imbricatula to I. sagensis\nB. reesidei-B. eliasi","Cimbr_to_Isage_rees")

cten_to_oblo_rees_overlap <- run_ecospat_dyn(cten_rees_prep,oblo_rees_prep,
                              "C. imbricatula to I. oblongus\nB. reesidei-B. eliasi","Cibmr_to_Ioblo_rees")

#### Summarize Outgroup Comparisons and Export Tables ####
# Combine the results of observed overlap
make_D_I_row <- function(data1,name,interval){
  named1 <- as.data.frame(cbind(name,interval,data1[[2]]$obs$D,data1[[2]]$obs$I))
  colnames(named1) <- c("species","Intervals_compared","D_val","I_val")
  named1[,c(3:4)] <- as.numeric(as.character(named1[,c(3:4)]))
  return(named1)
```

```
}

cten_to_conv_redu_D_I_row <- make_D_I_row(cten_to_conv_redu_overlap,"C. imbricatula vs. I. convexus","B. reduncus-B. scotti")

cten_to_subl_redu_D_I_row <- make_D_I_row(cten_to_subl_redu_overlap,"C. imbricatula vs. I. sublaevis","B. reduncus-B. scotti")

cten_to_sask_redu_D_I_row <- make_D_I_row(cten_to_sask_redu_overlap,"C. imbricatula vs. I. saskatchewanensis","B. reduncus-B. scotti")

cten_to_bara_nebr_D_I_row <- make_D_I_row(cten_to_bara_nebr_overlap,"C. imbricatula vs. I. barabini","D. nebrascense-E. jenneyi")

cten_to_conv_nebr_D_I_row <- make_D_I_row(cten_to_conv_nebr_overlap,"C. imbricatula vs. I. convexus","D. nebrascense-E. jenneyi")

cten_to_subl_nebr_D_I_row <- make_D_I_row(cten_to_subl_nebr_overlap,"C. imbricatula vs. I. sublaevis","D. nebrascense-E. jenneyi")

chla_to_bara_nebr_D_I_row <- make_D_I_row(chla_to_bara_nebr_overlap,"C. nebrascensis vs. I. barabini","D. nebrascense-E. jenneyi")

chla_to_conv_nebr_D_I_row <- make_D_I_row(chla_to_conv_nebr_overlap,"C. nebrascensis vs. I. convexus","D. nebrascense-E. jenneyi")

chla_to_subl_nebr_D_I_row <- make_D_I_row(chla_to_subl_nebr_overlap,"C. nebrascensis vs. I. sagensis","D. nebrascense-E. jenneyi")

luci_to_bara_nebr_D_I_row <- make_D_I_row(luci_to_bara_nebr_overlap,"L. subundata vs. I. barabini","D. nebrascense-E. jenneyi")

luci_to_conv_nebr_D_I_row <- make_D_I_row(luci_to_conv_nebr_overlap,"L. subundata vs. I. convexus","D. nebrascense-E. jenneyi")
```

```
luci_to_subl_nebr_D_I_row <- make_D_I_row(luci_to_subl_nebr_overlap,"L. subundata vs. I. sublaevis","D. nebrascense-E. jenneyi")


luci_to_sage_comp_D_I_row <- make_D_I_row(luci_to_sage_comp_overlap,"L. subundata vs. I. sagensis","B. compressus-B. cuneatus")


chla_to_bara_rees_D_I_row <- make_D_I_row(chla_to_bara_rees_overlap,"C. nebrascensis vs. I. barabini","B. reesidei-B. eliasi")


chla_to_sage_rees_D_I_row <- make_D_I_row(chla_to_sage_rees_overlap,"C. nebrascensis vs. I. sagensis","B. reesidei-B. eliasi")


chla_to_oblo_rees_D_I_row <- make_D_I_row(chla_to_oblo_rees_overlap,"C. nebrascensis vs. I. oblongus","B. reesidei-B. eliasi")


cten_to_bara_rees_D_I_row <- make_D_I_row(cten_to_bara_rees_overlap,"C. imbricatula vs. I. barabini","B. reesidei-B. eliasi")


cten_to_sage_rees_D_I_row <- make_D_I_row(cten_to_sage_rees_overlap,"C. imbricatula vs. I. sagensis","B. reesidei-B. eliasi")


cten_to_oblo_rees_D_I_row <- make_D_I_row(cten_to_oblo_rees_overlap,"C. imbricatula vs. I. oblongus","B. reesidei-B. eliasi")


outgroup_compare_I_D_results <-
rbind(cten_to_conv_redu_D_I_row,cten_to_subl_redu_D_I_row,cten_to_sask_redu_D_I_row,
                    chla_to_bara_nebr_D_I_row,cten_to_conv_nebr_D_I_row,cten_to_subl_nebr_D_I_row,
                    chla_to_bara_nebr_D_I_row,chla_to_conv_nebr_D_I_row,chla_to_subl_nebr_D_I_row,
                    luci_to_bara_nebr_D_I_row,luci_to_conv_nebr_D_I_row,luci_to_subl_nebr_D_I_row,
                    luci_to_sage_comp_D_I_row,
                    chla_to_bara_rees_D_I_row,chla_to_sage_rees_D_I_row,chla_to_oblo_rees_D_I_row,
                    cten_to_bara_rees_D_I_row,cten_to_sage_rees_D_I_row,cten_to_oblo_rees_D_I_row)


# Combine the results of equivalency tests

make_exp_p_row <- function(data1,name,interval){
```

```
  named1 <- as.data.frame(cbind(name,interval,data1[[2]]$p.D,data1[[2]]$p.I,data1[[3]]$p.D,data1[[3]]$p.I))
  colnames(named1) <-
c("species","Intervals_compared","Greater_Eq_p_D","Greater_Eq_p_I","Lower_Eq_p_D","Lower_Eq_p_I")
  named1[,c(3:4)] <- as.numeric(as.character(named1[,c(3:4)]))
  return(named1)
}

cten_to_conv_redu_eq_row <- make_exp_p_row(cten_to_conv_redu_overlap,"C. imbricatula vs. I.
convexus","B. reduncus-B. scotti")

cten_to_subl_redu_eq_row <- make_exp_p_row(cten_to_subl_redu_overlap,"C. imbricatula vs. I.
sublaevis","B. reduncus-B. scotti")

cten_to_sask_redu_eq_row <- make_exp_p_row(cten_to_sask_redu_overlap,"C. imbricatula vs. I.
saskatchewanensis","B. reduncus-B. scotti")

cten_to_bara_nebr_eq_row <- make_exp_p_row(cten_to_bara_nebr_overlap,"C. imbricatula vs. I. barabini","D.
nebrascense-E. jenneyi")

cten_to_conv_nebr_eq_row <- make_exp_p_row(cten_to_conv_nebr_overlap,"C. imbricatula vs. I.
convexus","D. nebrascense-E. jenneyi")

cten_to_subl_nebr_eq_row <- make_exp_p_row(cten_to_subl_nebr_overlap,"C. imbricatula vs. I.
sublaevis","D. nebrascense-E. jenneyi")

chla_to_bara_nebr_eq_row <- make_exp_p_row(chla_to_bara_nebr_overlap,"C. nebrascensis vs. I.
barabini","D. nebrascense-E. jenneyi")

chla_to_conv_nebr_eq_row <- make_exp_p_row(chla_to_conv_nebr_overlap,"C. nebrascensis vs. I.
convexus","D. nebrascense-E. jenneyi")

chla_to_subl_nebr_eq_row <- make_exp_p_row(chla_to_subl_nebr_overlap,"C. nebrascensis vs. I.
sagensis","D. nebrascense-E. jenneyi")

luci_to_bara_nebr_eq_row <- make_exp_p_row(luci_to_bara_nebr_overlap,"L. subundata vs. I. barabini","D.
nebrascense-E. jenneyi")
```

```
luci_to_conv_nebr_eq_row <- make_exp_p_row(luci_to_conv_nebr_overlap,"L. subundata vs. I. convexus","D. nebrascense-E. jenneyi")


luci_to_subl_nebr_eq_row <- make_exp_p_row(luci_to_subl_nebr_overlap,"L. subundata vs. I. sublaevis","D. nebrascense-E. jenneyi")


luci_to_sage_comp_eq_row <- make_exp_p_row(luci_to_sage_comp_overlap,"L. subundata vs. I. sagensis","B. compressus-B. cuneatus")


chla_to_bara_rees_eq_row <- make_exp_p_row(chla_to_bara_rees_overlap,"C. nebrascensis vs. I. barabini","B. reesidei-B. eliasi")


chla_to_sage_rees_eq_row <- make_exp_p_row(chla_to_sage_rees_overlap,"C. nebrascensis vs. I. sagensis","B. reesidei-B. eliasi")


chla_to_oblo_rees_eq_row <- make_exp_p_row(chla_to_oblo_rees_overlap,"C. nebrascensis vs. I. oblongus","B. reesidei-B. eliasi")


cten_to_bara_rees_eq_row <- make_exp_p_row(cten_to_bara_rees_overlap,"C. imbricatula vs. I. barabini","B. reesidei-B. eliasi")


cten_to_sage_rees_eq_row <- make_exp_p_row(cten_to_sage_rees_overlap,"C. imbricatula vs. I. sagensis","B. reesidei-B. eliasi")


cten_to_oblo_rees_eq_row <- make_exp_p_row(cten_to_oblo_rees_overlap,"C. imbricatula vs. I. oblongus","B. reesidei-B. eliasi")


outgroup_compare_eq_results <-
rbind(cten_to_conv_redu_eq_row,cten_to_subl_redu_eq_row,cten_to_sask_redu_eq_row,
                    chla_to_bara_nebr_eq_row,cten_to_conv_nebr_eq_row,cten_to_subl_nebr_eq_row,
                    chla_to_bara_nebr_eq_row,chla_to_conv_nebr_eq_row,chla_to_subl_nebr_eq_row,
                    luci_to_bara_nebr_eq_row,luci_to_conv_nebr_eq_row,luci_to_subl_nebr_eq_row,
                    luci_to_sage_comp_eq_row,
                    chla_to_bara_rees_eq_row,chla_to_sage_rees_eq_row,chla_to_oblo_rees_eq_row,
```

```
                    cten_to_bara_rees_eq_row,cten_to_sage_rees_eq_row,cten_to_oblo_rees_eq_row)


# Combine the results of equivalency tests of Dynamics (p-vals)

make_exp_dyn_p_row <- function(data1,name,interval){

  named1 <-
as.data.frame(cbind(name,interval,data1[[2]]$p.expansion,data1[[2]]$p.stability,data1[[2]]$p.unfilling,

                      data1[[3]]$p.expansion,data1[[3]]$p.stability,data1[[3]]$p.unfilling))

  colnames(named1) <-
c("species","Intervals_compared","Greater_Eq_p_expansion","Greater_Eq_p_stability","Greater_Eq_p_unfillin
g",

                "Lower_Eq_p_expansion","Lower_Eq_p_stability","Lower_Eq_p_unfilling")

  named1[,c(3:4)] <- as.numeric(as.character(named1[,c(3:4)]))

  return(named1)

}


cten_to_conv_redu_eq_dyn_p_row <- make_exp_dyn_p_row(cten_to_conv_redu_overlap,"C. imbricatula vs. I.
convexus","B. reduncus-B. scotti")


cten_to_subl_redu_eq_dyn_p_row <- make_exp_dyn_p_row(cten_to_subl_redu_overlap,"C. imbricatula vs. I.
sublaevis","B. reduncus-B. scotti")


cten_to_sask_redu_eq_dyn_p_row <- make_exp_dyn_p_row(cten_to_sask_redu_overlap,"C. imbricatula vs. I.
saskatchewanensis","B. reduncus-B. scotti")


cten_to_bara_nebr_eq_dyn_p_row <- make_exp_dyn_p_row(cten_to_bara_nebr_overlap,"C. imbricatula vs. I.
barabini","D. nebrascense-E. jenneyi")


cten_to_conv_nebr_eq_dyn_p_row <- make_exp_dyn_p_row(cten_to_conv_nebr_overlap,"C. imbricatula vs. I.
convexus","D. nebrascense-E. jenneyi")


cten_to_subl_nebr_eq_dyn_p_row <- make_exp_dyn_p_row(cten_to_subl_nebr_overlap,"C. imbricatula vs. I.
sublaevis","D. nebrascense-E. jenneyi")


chla_to_bara_nebr_eq_dyn_p_row <- make_exp_dyn_p_row(chla_to_bara_nebr_overlap,"C. nebrascensis vs. I.
barabini","D. nebrascense-E. jenneyi")
```

```
chla_to_conv_nebr_eq_dyn_p_row <- make_exp_dyn_p_row(chla_to_conv_nebr_overlap,"C. nebrascensis vs. I. convexus","D. nebrascense-E. jenneyi")


chla_to_subl_nebr_eq_dyn_p_row <- make_exp_dyn_p_row(chla_to_subl_nebr_overlap,"C. nebrascensis vs. I. sublaevis","D. nebrascense-E. jenneyi")


luci_to_bara_nebr_eq_dyn_p_row <- make_exp_dyn_p_row(luci_to_bara_nebr_overlap,"L. subundata vs. I. barabini","D. nebrascense-E. jenneyi")


luci_to_conv_nebr_eq_dyn_p_row <- make_exp_dyn_p_row(luci_to_conv_nebr_overlap,"L. subundata vs. I. convexus","D. nebrascense-E. jenneyi")


luci_to_subl_nebr_eq_dyn_p_row <- make_exp_dyn_p_row(luci_to_subl_nebr_overlap,"L. subundata vs. I. sublaevis","D. nebrascense-E. jenneyi")


luci_to_sage_comp_eq_dyn_p_row <- make_exp_dyn_p_row(luci_to_sage_comp_overlap,"L. subundata vs. I. sagensis","B. compressus-B. cuneatus")


chla_to_bara_rees_eq_dyn_p_row <- make_exp_dyn_p_row(chla_to_bara_rees_overlap,"C. nebrascensis vs. I. barabini","B. reesidei-B. eliasi")


chla_to_sage_rees_eq_dyn_p_row <- make_exp_dyn_p_row(chla_to_sage_rees_overlap,"C. nebrascensis vs. I. sagensis","B. reesidei-B. eliasi")


chla_to_oblo_rees_eq_dyn_p_row <- make_exp_dyn_p_row(chla_to_oblo_rees_overlap,"C. nebrascensis vs. I. oblongus","B. reesidei-B. eliasi")


cten_to_bara_rees_eq_dyn_p_row <- make_exp_dyn_p_row(cten_to_bara_rees_overlap,"C. imbricatula vs. I. barabini","B. reesidei-B. eliasi")


cten_to_sage_rees_eq_dyn_p_row <- make_exp_dyn_p_row(cten_to_sage_rees_overlap,"C. imbricatula vs. I. sagensis","B. reesidei-B. eliasi")


cten_to_oblo_rees_eq_dyn_p_row <- make_exp_dyn_p_row(cten_to_oblo_rees_overlap,"C. imbricatula vs. I. oblongus","B. reesidei-B. eliasi")
```

```
outgroup_compare_eq_dyn_p_results <-
rbind(cten_to_conv_redu_eq_dyn_p_row,cten_to_subl_redu_eq_dyn_p_row,cten_to_sask_redu_eq_dyn_p_row
,

chla_to_bara_nebr_eq_dyn_p_row,cten_to_conv_nebr_eq_dyn_p_row,cten_to_subl_nebr_eq_dyn_p_row,

chla_to_bara_nebr_eq_dyn_p_row,chla_to_conv_nebr_eq_dyn_p_row,chla_to_subl_nebr_eq_dyn_p_row,

luci_to_bara_nebr_eq_dyn_p_row,luci_to_conv_nebr_eq_dyn_p_row,luci_to_subl_nebr_eq_dyn_p_row,
                                luci_to_sage_comp_eq_dyn_p_row,

chla_to_bara_rees_eq_dyn_p_row,chla_to_sage_rees_eq_dyn_p_row,chla_to_oblo_rees_eq_dyn_p_row,

cten_to_bara_rees_eq_dyn_p_row,cten_to_sage_rees_eq_dyn_p_row,cten_to_oblo_rees_eq_dyn_p_row)


# Combine the results of similarity tests
make_sim_p_row <- function(data1,name,interval){
  named1 <- as.data.frame(cbind(name,interval,data1[[4]]$p.D,data1[[4]]$p.I,data1[[5]]$p.D,data1[[5]]$p.I))
  colnames(named1) <-
c("species","Intervals_compared","Greater_Sim_p_D","Greater_Sim_p_I","Lower_Sim_p_D","Lower_Sim_p_I
")
  named1[,c(3:4)] <- as.numeric(as.character(named1[,c(3:4)]))
  return(named1)
}


cten_to_conv_redu_sim_row <- make_sim_p_row(cten_to_conv_redu_overlap,"C. imbricatula vs. I.
convexus","B. reduncus-B. scotti")


cten_to_subl_redu_sim_row <- make_sim_p_row(cten_to_subl_redu_overlap,"C. imbricatula vs. I.
sublaevis","B. reduncus-B. scotti")


cten_to_sask_redu_sim_row <- make_sim_p_row(cten_to_sask_redu_overlap,"C. imbricatula vs. I.
saskatchewanensis","B. reduncus-B. scotti")
```

```
cten_to_bara_nebr_sim_row <- make_sim_p_row(cten_to_bara_nebr_overlap,"C. imbricatula vs. I. barabini","D. nebrascense-E. jenneyi")

cten_to_conv_nebr_sim_row <- make_sim_p_row(cten_to_conv_nebr_overlap,"C. imbricatula vs. I. convexus","D. nebrascense-E. jenneyi")

cten_to_subl_nebr_sim_row <- make_sim_p_row(cten_to_subl_nebr_overlap,"C. imbricatula vs. I. sublaevis","D. nebrascense-E. jenneyi")

chla_to_bara_nebr_sim_row <- make_sim_p_row(chla_to_bara_nebr_overlap,"C. nebrascensis vs. I. barabini","D. nebrascense-E. jenneyi")

chla_to_conv_nebr_sim_row <- make_sim_p_row(chla_to_conv_nebr_overlap,"C. nebrascensis vs. I. convexus","D. nebrascense-E. jenneyi")

chla_to_subl_nebr_sim_row <- make_sim_p_row(chla_to_subl_nebr_overlap,"C. nebrascensis vs. I. sublaevis","D. nebrascense-E. jenneyi")

luci_to_bara_nebr_sim_row <- make_sim_p_row(luci_to_bara_nebr_overlap,"L. subundata vs. I. barabini","D. nebrascense-E. jenneyi")

luci_to_conv_nebr_sim_row <- make_sim_p_row(luci_to_conv_nebr_overlap,"L. subundata vs. I. convexus","D. nebrascense-E. jenneyi")

luci_to_subl_nebr_sim_row <- make_sim_p_row(luci_to_subl_nebr_overlap,"L. subundata vs. I. sublaevis","D. nebrascense-E. jenneyi")

luci_to_sage_comp_sim_row <- make_sim_p_row(luci_to_sage_comp_overlap,"L. subundata vs. I. sagensis","B. compressus-B. cuneatus")

chla_to_bara_rees_sim_row <- make_sim_p_row(chla_to_bara_rees_overlap,"C. nebrascensis vs. I. barabini","B. reesidei-B. eliasi")

chla_to_sage_rees_sim_row <- make_sim_p_row(chla_to_sage_rees_overlap,"C. nebrascensis vs. I. sagensis","B. reesidei-B. eliasi")
```

```
chla_to_oblo_rees_sim_row <- make_sim_p_row(chla_to_oblo_rees_overlap,"C. nebrascensis vs. I. oblongus","B. reesidei-B. eliasi")


cten_to_bara_rees_sim_row <- make_sim_p_row(cten_to_bara_rees_overlap,"C. imbricatula vs. I. barabini","B. reesidei-B. eliasi")


cten_to_sage_rees_sim_row <- make_sim_p_row(cten_to_sage_rees_overlap,"C. imbricatula vs. I. sagensis","B. reesidei-B. eliasi")


cten_to_oblo_rees_sim_row <- make_sim_p_row(cten_to_oblo_rees_overlap,"C. imbricatula vs. I. oblongus","B. reesidei-B. eliasi")




outgroup_compare_sim_results <-
rbind(cten_to_conv_redu_sim_row,cten_to_subl_redu_sim_row,cten_to_sask_redu_sim_row,
                        chla_to_bara_nebr_sim_row,cten_to_conv_nebr_sim_row,cten_to_subl_nebr_sim_row,
                        chla_to_bara_nebr_sim_row,chla_to_conv_nebr_sim_row,chla_to_subl_nebr_sim_row,
                        luci_to_bara_nebr_sim_row,luci_to_conv_nebr_sim_row,luci_to_subl_nebr_sim_row,
                        luci_to_sage_comp_sim_row,
                        chla_to_bara_rees_sim_row,chla_to_sage_rees_sim_row,chla_to_oblo_rees_sim_row,
                        cten_to_bara_rees_sim_row,cten_to_sage_rees_sim_row,cten_to_oblo_rees_sim_row)


# Combine the results of equivalency tests of Dynamics (p-vals)
make_sim_dyn_p_row <- function(data1,name,interval){
  named1 <-
as.data.frame(cbind(name,interval,data1[[4]]$p.expansion,data1[[4]]$p.stability,data1[[4]]$p.unfilling,
                      data1[[5]]$p.expansion,data1[[5]]$p.stability,data1[[5]]$p.unfilling))
  colnames(named1) <-
c("species","Intervals_compared","Greater_Sim_p_expansion","Greater_Sim_p_stability","Greater_Sim_p_unfilling",
                "Lower_Sim_p_expansion","Lower_Sim_p_stability","Lower_Sim_p_unfilling")
  named1[,c(3:4)] <- as.numeric(as.character(named1[,c(3:4)]))
  return(named1)
```

```
}

cten_to_conv_redu_sim_dyn_p_row <- make_sim_dyn_p_row(cten_to_conv_redu_overlap,"C. imbricatula vs. I. convexus","B. reduncus-B. scotti")

cten_to_subl_redu_sim_dyn_p_row <- make_sim_dyn_p_row(cten_to_subl_redu_overlap,"C. imbricatula vs. I. sublaevis","B. reduncus-B. scotti")

cten_to_sask_redu_sim_dyn_p_row <- make_sim_dyn_p_row(cten_to_sask_redu_overlap,"C. imbricatula vs. I. saskatchewanensis","B. reduncus-B. scotti")

cten_to_bara_nebr_sim_dyn_p_row <- make_sim_dyn_p_row(cten_to_bara_nebr_overlap,"C. imbricatula vs. I. barabini","D. nebrascense-E. jenneyi")

cten_to_conv_nebr_sim_dyn_p_row <- make_sim_dyn_p_row(cten_to_conv_nebr_overlap,"C. imbricatula vs. I. convexus","D. nebrascense-E. jenneyi")

cten_to_subl_nebr_sim_dyn_p_row <- make_sim_dyn_p_row(cten_to_subl_nebr_overlap,"C. imbricatula vs. I. sublaevis","D. nebrascense-E. jenneyi")

chla_to_bara_nebr_sim_dyn_p_row <- make_sim_dyn_p_row(chla_to_bara_nebr_overlap,"C. nebrascensis vs. I. barabini","D. nebrascense-E. jenneyi")

chla_to_conv_nebr_sim_dyn_p_row <- make_sim_dyn_p_row(chla_to_conv_nebr_overlap,"C. nebrascensis vs. I. convexus","D. nebrascense-E. jenneyi")

chla_to_subl_nebr_sim_dyn_p_row <- make_sim_dyn_p_row(chla_to_subl_nebr_overlap,"C. nebrascensis vs. I. sublaevis","D. nebrascense-E. jenneyi")

luci_to_bara_nebr_sim_dyn_p_row <- make_sim_dyn_p_row(luci_to_bara_nebr_overlap,"L. subundata vs. I. barabini","D. nebrascense-E. jenneyi")

luci_to_conv_nebr_sim_dyn_p_row <- make_sim_dyn_p_row(luci_to_conv_nebr_overlap,"L. subundata vs. I. convexus","D. nebrascense-E. jenneyi")
```

```
luci_to_subl_nebr_sim_dyn_p_row <- make_sim_dyn_p_row(luci_to_subl_nebr_overlap,"L. subundata vs. I. sublaevis","D. nebrascense-E. jenneyi")


luci_to_sage_comp_sim_dyn_p_row <- make_sim_dyn_p_row(luci_to_sage_comp_overlap,"L. subundata vs. I. sagensis","B. compressus-B. cuneatus")




chla_to_bara_rees_sim_dyn_p_row <- make_sim_dyn_p_row(chla_to_bara_rees_overlap,"C. nebrascensis vs. I. barabini","B. reesidei-B. eliasi")


chla_to_sage_rees_sim_dyn_p_row <- make_sim_dyn_p_row(chla_to_sage_rees_overlap,"C. nebrascensis vs. I. sagensis","B. reesidei-B. eliasi")


chla_to_oblo_rees_sim_dyn_p_row <- make_sim_dyn_p_row(chla_to_oblo_rees_overlap,"C. nebrascensis vs. I. oblongus","B. reesidei-B. eliasi")


cten_to_bara_rees_sim_dyn_p_row <- make_sim_dyn_p_row(cten_to_bara_rees_overlap,"C. imbricatula vs. I. barabini","B. reesidei-B. eliasi")


cten_to_sage_rees_sim_dyn_p_row <- make_sim_dyn_p_row(cten_to_sage_rees_overlap,"C. imbricatula vs. I. sagensis","B. reesidei-B. eliasi")


cten_to_oblo_rees_sim_dyn_p_row <- make_sim_dyn_p_row(cten_to_oblo_rees_overlap,"C. imbricatula vs. I. oblongus","B. reesidei-B. eliasi")


outgroup_compare_sim_dyn_p_results <-
rbind(cten_to_conv_redu_sim_dyn_p_row,cten_to_subl_redu_sim_dyn_p_row,cten_to_sask_redu_sim_dyn_p_
row,

chla_to_bara_nebr_sim_dyn_p_row,cten_to_conv_nebr_sim_dyn_p_row,cten_to_subl_nebr_sim_dyn_p_row,

chla_to_bara_nebr_sim_dyn_p_row,chla_to_conv_nebr_sim_dyn_p_row,chla_to_subl_nebr_sim_dyn_p_row,

luci_to_bara_nebr_sim_dyn_p_row,luci_to_conv_nebr_sim_dyn_p_row,luci_to_subl_nebr_sim_dyn_p_row,
                                luci_to_sage_comp_sim_dyn_p_row,
```

```
chla_to_bara_rees_sim_dyn_p_row,chla_to_sage_rees_sim_dyn_p_row,chla_to_oblo_rees_sim_dyn_p_row,

cten_to_bara_rees_sim_dyn_p_row,cten_to_sage_rees_sim_dyn_p_row,cten_to_oblo_rees_sim_dyn_p_row)


# Combine the results of niche dynamics
make_dynam_row <- function(data1,name,interval){
  data <- as.data.frame(data1[[6]])
  named1 <- as.data.frame(cbind(name,interval,t(data[,2])))
  colnames(named1) <- c("species","Intervals_compared","Expansion","Stability","Unfilling")
  named1[,c(3:5)] <- as.numeric(as.character(named1[,c(3:5)]))
  return(named1)
}


cten_to_conv_redu_dyn_row <- make_dynam_row(cten_to_conv_redu_overlap,"C. imbricatula vs. I.
convexus","B. reduncus-B. scotti")


cten_to_subl_redu_dyn_row <- make_dynam_row(cten_to_subl_redu_overlap,"C. imbricatula vs. I.
sublaevis","B. reduncus-B. scotti")


cten_to_sask_redu_dyn_row <- make_dynam_row(cten_to_sask_redu_overlap,"C. imbricatula vs. I.
saskatchewanensis","B. reduncus-B. scotti")


cten_to_bara_nebr_dyn_row <- make_dynam_row(cten_to_bara_nebr_overlap,"C. imbricatula vs. I.
barabini","D. nebrascense-E. jenneyi")


cten_to_conv_nebr_dyn_row <- make_dynam_row(cten_to_conv_nebr_overlap,"C. imbricatula vs. I.
convexus","D. nebrascense-E. jenneyi")


cten_to_subl_nebr_dyn_row <- make_dynam_row(cten_to_subl_nebr_overlap,"C. imbricatula vs. I.
sublaevis","D. nebrascense-E. jenneyi")


chla_to_bara_nebr_dyn_row <- make_dynam_row(chla_to_bara_nebr_overlap,"C. nebrascensis vs. I.
barabini","D. nebrascense-E. jenneyi")
```

```
chla_to_conv_nebr_dyn_row <- make_dynam_row(chla_to_conv_nebr_overlap,"C. nebrascensis vs. I. convexus","D. nebrascense-E. jenneyi")


chla_to_subl_nebr_dyn_row <- make_dynam_row(chla_to_subl_nebr_overlap,"C. nebrascensis vs. I. sublaevis","D. nebrascense-E. jenneyi")


luci_to_bara_nebr_dyn_row <- make_dynam_row(luci_to_bara_nebr_overlap,"L. subundata vs. I. barabini","D. nebrascense-E. jenneyi")


luci_to_conv_nebr_dyn_row <- make_dynam_row(luci_to_conv_nebr_overlap,"L. subundata vs. I. convexus","D. nebrascense-E. jenneyi")


luci_to_subl_nebr_dyn_row <- make_dynam_row(luci_to_subl_nebr_overlap,"L. subundata vs. I. sublaevis","D. nebrascense-E. jenneyi")


luci_to_sage_comp_dyn_row <- make_dynam_row(luci_to_sage_comp_overlap,"L. subundata vs. I. sagensis","B. compressus-B. cuneatus")


chla_to_bara_rees_dyn_row <- make_dynam_row(chla_to_bara_rees_overlap,"C. nebrascensis vs. I. barabini","B. reesidei-B. eliasi")


chla_to_sage_rees_dyn_row <- make_dynam_row(chla_to_sage_rees_overlap,"C. nebrascensis vs. I. sagensis","B. reesidei-B. eliasi")


chla_to_oblo_rees_dyn_row <- make_dynam_row(chla_to_oblo_rees_overlap,"C. nebrascensis vs. I. oblongus","B. reesidei-B. eliasi")




cten_to_bara_rees_dyn_row <- make_dynam_row(cten_to_bara_rees_overlap,"C. imbricatula vs. I. barabini","B. reesidei-B. eliasi")


cten_to_sage_rees_dyn_row <- make_dynam_row(cten_to_sage_rees_overlap,"C. imbricatula vs. I. sagensis","B. reesidei-B. eliasi")
```

```
cten_to_oblo_rees_dyn_row <- make_dynam_row(cten_to_oblo_rees_overlap,"C. imbricatula vs. I.
oblongus","B. reesidei-B. eliasi")


outgroup_compare_dyn_results <-
rbind(cten_to_conv_redu_dyn_row,cten_to_subl_redu_dyn_row,cten_to_sask_redu_dyn_row,
                    chla_to_bara_nebr_dyn_row,cten_to_conv_nebr_dyn_row,cten_to_subl_nebr_dyn_row,
                    chla_to_bara_nebr_dyn_row,chla_to_conv_nebr_dyn_row,chla_to_subl_nebr_dyn_row,
                    luci_to_bara_nebr_dyn_row,luci_to_conv_nebr_dyn_row,luci_to_subl_nebr_dyn_row,
                    luci_to_sage_comp_dyn_row,
                    chla_to_bara_rees_dyn_row,chla_to_sage_rees_dyn_row,chla_to_oblo_rees_dyn_row,
                    cten_to_bara_rees_dyn_row,cten_to_sage_rees_dyn_row,cten_to_oblo_rees_dyn_row)


## Write csv files of results
write.csv(outgroup_compare_I_D_results,file="tables/outgroup_compare_I_D_results.csv")
write.csv(outgroup_compare_eq_results,file="tables/outgroup_compare_eq_results.csv")
write.csv(outgroup_compare_sim_results,file="tables/outgroup_compare_sim_results.csv")
write.csv(outgroup_compare_dyn_results,file="tables/outgroup_compare_dyn_results.csv")
write.csv(outgroup_compare_eq_dyn_p_results,file="tables/outgroup_compare_eq_dyn_p_results.csv")
write.csv(outgroup_compare_sim_dyn_p_results,file="tables/outgroup_compare_sim_dyn_p_results.csv")


#### SPECIES TO GENUS COMPARISONS ####


conv_to_inoceram_redu_overlap <- run_ecospat_dyn(conv_redu_prep,Inoceramus_redu_prep,
                        "I. convextus to Inoceramus\nB. reduncus-B. scotti","Iconv_to_Inocer_redu_")



subl_to_inoceram_redu_overlap <- run_ecospat_dyn(subl_redu_prep,Inoceramus_redu_prep,
                        "I. sublaevis to Inoceramus\nB. reduncus-B. scotti","Isubl_to_Inocer_redu_")


sask_to_inoceram_redu_overlap <- run_ecospat_dyn(sask_redu_prep,Inoceramus_redu_prep,
                        "I. saskatchewanensis to Inoceramus\nB. reduncus-B.
scotti","Isask_to_Inocer_redu_")
```

```
conv_to_inoceram_nebr_overlap <- run_ecospat_dyn(conv_nebr_prep,Inoceramus_nebr_prep,
                                "I. convextus to Inoceramus\nD. nebrascense-E. jenneyi","Iconv_to_Inocer_nebr")

subl_to_inoceram_nebr_overlap <- run_ecospat_dyn(subl_nebr_prep,Inoceramus_nebr_prep,
                                "I. sublaevis to Inoceramus\nD. nebrascense-E. jenneyi","Isubl_to_Inocer_nebr")

bara_to_inoceram_nebr_overlap <- run_ecospat_dyn(bara_nebr_prep,Inoceramus_nebr_prep,
                                "I. barabini to Inoceramus\nD. nebrascense-E. jenneyi","Ibara_to_Inocer_nebr")

bara_to_inoceram_rees_overlap <- run_ecospat_dyn(bara_rees_prep,Inoceramus_rees_prep,
                                "I. barabini to Inoceramus\nB. reesidei-B. eliasi","Ibara_to_Inocer_rees")

sage_to_inoceram_rees_overlap <- run_ecospat_dyn(sage_rees_prep,Inoceramus_rees_prep,
                                "I. sagensis to Inoceramus\nB. reesidei-B. eliasi","Isage_to_Inocer_rees")

oblo_to_inoceram_rees_overlap <- run_ecospat_dyn(oblo_rees_prep,Inoceramus_rees_prep,
                                  "I. oblongus to Inoceramus\nB. reesidei-B. eliasi","Ioblo_to_Inocer_rees")



#### Summarize Specise to Genus Comparisons and Export Tables ####
# Combine the results of observed overlap
make_D_I_row <- function(data1,name,interval){
  named1 <- as.data.frame(cbind(name,interval,data1[[2]]$obs$D,data1[[2]]$obs$I))
  colnames(named1) <- c("species","Intervals_compared","D_val","I_val")
  named1[,c(3:4)] <- as.numeric(as.character(named1[,c(3:4)]))
  return(named1)
}

conv_to_inoceram_redu_D_I_row <- make_D_I_row(conv_to_inoceram_redu_overlap,"I. convexus","B.
reduncus-B. scotti")
```

```
subl_to_inoceram_redu_D_I_row <- make_D_I_row(subl_to_inoceram_redu_overlap,"I. sublaevis","B. reduncus-B. scotti")


sask_to_inoceram_redu_D_I_row <- make_D_I_row(sask_to_inoceram_redu_overlap,"I. saskatchewanensis","B. reduncus-B. scotti")

conv_to_inoceram_nebr_D_I_row <- make_D_I_row(conv_to_inoceram_nebr_overlap,"I. convexus","D. nebrascense-E. jenneyi")


subl_to_inoceram_nebr_D_I_row <- make_D_I_row(subl_to_inoceram_nebr_overlap,"I. sublaevis","D. nebrascense-E. jenneyi")

bara_to_inoceram_nebr_D_I_row <- make_D_I_row(bara_to_inoceram_nebr_overlap,"I. barabini","D. nebrascense-E. jenneyi")


bara_to_inoceram_rees_D_I_row <- make_D_I_row(bara_to_inoceram_rees_overlap,"I. barabini","B. reesidei-B. eliasi")

sage_to_inoceram_rees_D_I_row <- make_D_I_row(sage_to_inoceram_rees_overlap,"I. sagensis","B. reesidei-B. eliasi")


oblo_to_inoceram_rees_D_I_row <- make_D_I_row(oblo_to_inoceram_rees_overlap,"I. oblongus","B. reesidei-B. eliasi")



gen_to_sp_I_D_results <- rbind(conv_to_inoceram_redu_D_I_row,subl_to_inoceram_redu_D_I_row,
                        sask_to_inoceram_redu_D_I_row,conv_to_inoceram_nebr_D_I_row,
                        subl_to_inoceram_nebr_D_I_row,bara_to_inoceram_nebr_D_I_row,
                        bara_to_inoceram_rees_D_I_row,sage_to_inoceram_rees_D_I_row,
                        oblo_to_inoceram_rees_D_I_row)


# Combine the results of equivalency tests

make_exp_p_row <- function(data1,name,interval){

  named1 <- as.data.frame(cbind(name,interval,data1[[2]]$p.D,data1[[2]]$p.I,data1[[3]]$p.D,data1[[3]]$p.I))

  colnames(named1) <-
c("species","Intervals_compared","Greater_Eq_p_D","Greater_Eq_p_I","Lower_Eq_p_D","Lower_Eq_p_I")

  named1[,c(3:4)] <- as.numeric(as.character(named1[,c(3:4)]))

  return(named1)
```

```
}


conv_to_inoceram_redu_eq_row <- make_exp_p_row(conv_to_inoceram_redu_overlap,"I. convexus","B.
reduncus-B. scotti")

subl_to_inoceram_redu_eq_row <- make_exp_p_row(subl_to_inoceram_redu_overlap,"I. sublaevis","B.
reduncus-B. scotti")


sask_to_inoceram_redu_eq_row <- make_exp_p_row(sask_to_inoceram_redu_overlap,"I.
saskatchewanensis","B. reduncus-B. scotti")

conv_to_inoceram_nebr_eq_row <- make_exp_p_row(conv_to_inoceram_nebr_overlap,"I. convexus","D.
nebrascense-E. jenneyi")


subl_to_inoceram_nebr_eq_row <- make_exp_p_row(subl_to_inoceram_nebr_overlap,"I. sublaevis","D.
nebrascense-E. jenneyi")

bara_to_inoceram_nebr_eq_row <- make_exp_p_row(bara_to_inoceram_nebr_overlap,"I. barabini","D.
nebrascense-E. jenneyi")


bara_to_inoceram_rees_eq_row <- make_exp_p_row(bara_to_inoceram_rees_overlap,"I. barabini","B. reesidei-
B. eliasi")

sage_to_inoceram_rees_eq_row <- make_exp_p_row(sage_to_inoceram_rees_overlap,"I. sagensis","B. reesidei-
B. eliasi")


oblo_to_inoceram_rees_eq_row <- make_exp_p_row(oblo_to_inoceram_rees_overlap,"I. oblongus","B.
reesidei-B. eliasi")


gen_to_sp_eq_results <- rbind(conv_to_inoceram_redu_eq_row,subl_to_inoceram_redu_eq_row,
                              sask_to_inoceram_redu_eq_row,conv_to_inoceram_nebr_eq_row,
                              subl_to_inoceram_nebr_eq_row,bara_to_inoceram_nebr_eq_row,
                              bara_to_inoceram_rees_eq_row,sage_to_inoceram_rees_eq_row,
                              oblo_to_inoceram_rees_eq_row)


# Combine the results of equivalency tests of Dynamics (p-vals)

make_exp_dyn_p_row <- function(data1,name,interval){

  named1 <-
as.data.frame(cbind(name,interval,data1[[2]]$p.expansion,data1[[2]]$p.stability,data1[[2]]$p.unfilling,
```

```
                data1[[3]]$p.expansion,data1[[3]]$p.stability,data1[[3]]$p.unfilling))
  colnames(named1) <-
c("species","Intervals_compared","Greater_Eq_p_expansion","Greater_Eq_p_stability","Greater_Eq_p_unfillin
g",
                "Lower_Eq_p_expansion","Lower_Eq_p_stability","Lower_Eq_p_unfilling")
  named1[,c(3:4)] <- as.numeric(as.character(named1[,c(3:4)]))
  return(named1)
}


conv_to_inoceram_redu_eq_dyn_p_row <- make_exp_dyn_p_row(conv_to_inoceram_redu_overlap,"I.
convexus","B. reduncus-B. scotti")
subl_to_inoceram_redu_eq_dyn_p_row <- make_exp_dyn_p_row(subl_to_inoceram_redu_overlap,"I.
sublaevis","B. reduncus-B. scotti")


sask_to_inoceram_redu_eq_dyn_p_row <- make_exp_dyn_p_row(sask_to_inoceram_redu_overlap,"I.
saskatchewanensis","B. reduncus-B. scotti")
conv_to_inoceram_nebr_eq_dyn_p_row <- make_exp_dyn_p_row(conv_to_inoceram_nebr_overlap,"I.
convexus","D. nebrascense-E. jenneyi")


subl_to_inoceram_nebr_eq_dyn_p_row <- make_exp_dyn_p_row(subl_to_inoceram_nebr_overlap,"I.
sublaevis","D. nebrascense-E. jenneyi")
bara_to_inoceram_nebr_eq_dyn_p_row <- make_exp_dyn_p_row(bara_to_inoceram_nebr_overlap,"I.
barabini","D. nebrascense-E. jenneyi")


bara_to_inoceram_rees_eq_dyn_p_row <- make_exp_dyn_p_row(bara_to_inoceram_rees_overlap,"I.
barabini","B. reesidei-B. eliasi")
sage_to_inoceram_rees_eq_dyn_p_row <- make_exp_dyn_p_row(sage_to_inoceram_rees_overlap,"I.
sagensis","B. reesidei-B. eliasi")


oblo_to_inoceram_rees_eq_dyn_p_row <- make_exp_dyn_p_row(oblo_to_inoceram_rees_overlap,"I.
oblongus","B. reesidei-B. eliasi")



gen_to_sp_eq_dyn_p_results <-
rbind(conv_to_inoceram_redu_eq_dyn_p_row,subl_to_inoceram_redu_eq_dyn_p_row,
                sask_to_inoceram_redu_eq_dyn_p_row,conv_to_inoceram_nebr_eq_dyn_p_row,
```

```
subl_to_inoceram_nebr_eq_dyn_p_row,bara_to_inoceram_nebr_eq_dyn_p_row,
bara_to_inoceram_rees_eq_dyn_p_row,sage_to_inoceram_rees_eq_dyn_p_row,
oblo_to_inoceram_rees_eq_dyn_p_row)



# Combine the results of similarity tests
make_sim_p_row <- function(data1,name,interval){
  named1 <- as.data.frame(cbind(name,interval,data1[[4]]$p.D,data1[[4]]$p.I,data1[[5]]$p.D,data1[[5]]$p.I))
  colnames(named1) <-
c("species","Intervals_compared","Greater_Sim_p_D","Greater_Sim_p_I","Lower_Sim_p_D","Lower_Sim_p_I
")
  named1[,c(3:4)] <- as.numeric(as.character(named1[,c(3:4)]))
  return(named1)
}


conv_to_inoceram_redu_sim_row <- make_sim_p_row(conv_to_inoceram_redu_overlap,"I. convexus","B.
reduncus-B. scotti")
subl_to_inoceram_redu_sim_row <- make_sim_p_row(subl_to_inoceram_redu_overlap,"I. sublaevis","B.
reduncus-B. scotti")


sask_to_inoceram_redu_sim_row <- make_sim_p_row(sask_to_inoceram_redu_overlap,"I.
saskatchewanensis","B. reduncus-B. scotti")
conv_to_inoceram_nebr_sim_row <- make_sim_p_row(conv_to_inoceram_nebr_overlap,"I. convexus","D.
nebrascense-E. jenneyi")


subl_to_inoceram_nebr_sim_row <- make_sim_p_row(subl_to_inoceram_nebr_overlap,"I. sublaevis","D.
nebrascense-E. jenneyi")
bara_to_inoceram_nebr_sim_row <- make_sim_p_row(bara_to_inoceram_nebr_overlap,"I. barabini","D.
nebrascense-E. jenneyi")


bara_to_inoceram_rees_sim_row <- make_sim_p_row(bara_to_inoceram_rees_overlap,"I. barabini","B.
reesidei-B. eliasi")
sage_to_inoceram_rees_sim_row <- make_sim_p_row(sage_to_inoceram_rees_overlap,"I. sagensis","B.
reesidei-B. eliasi")
```

```
oblo_to_inoceram_rees_sim_row <- make_sim_p_row(oblo_to_inoceram_rees_overlap,"I. oblongus","B.
reesidei-B. eliasi")



gen_to_sp_sim_results <- rbind(conv_to_inoceram_redu_sim_row,subl_to_inoceram_redu_sim_row,
                        sask_to_inoceram_redu_sim_row,conv_to_inoceram_nebr_sim_row,
                        subl_to_inoceram_nebr_sim_row,bara_to_inoceram_nebr_sim_row,
                        bara_to_inoceram_rees_sim_row,sage_to_inoceram_rees_sim_row,
                        oblo_to_inoceram_rees_sim_row)


# Combine the results of equivalency tests of Dynamics (p-vals)
make_sim_dyn_p_row <- function(data1,name,interval){
  named1 <-
as.data.frame(cbind(name,interval,data1[[4]]$p.expansion,data1[[4]]$p.stability,data1[[4]]$p.unfilling,
                      data1[[5]]$p.expansion,data1[[5]]$p.stability,data1[[5]]$p.unfilling))
  colnames(named1) <-
c("species","Intervals_compared","Greater_Sim_p_expansion","Greater_Sim_p_stability","Greater_Sim_p_unfi
lling",
                "Lower_Sim_p_expansion","Lower_Sim_p_stability","Lower_Sim_p_unfilling")
  named1[,c(3:4)] <- as.numeric(as.character(named1[,c(3:4)]))
  return(named1)
}


conv_to_inoceram_redu_sim_dyn_p_row <- make_sim_dyn_p_row(conv_to_inoceram_redu_overlap,"I.
convexus","B. reduncus-B. scotti")
subl_to_inoceram_redu_sim_dyn_p_row <- make_sim_dyn_p_row(subl_to_inoceram_redu_overlap,"I.
sublaevis","B. reduncus-B. scotti")


sask_to_inoceram_redu_sim_dyn_p_row <- make_sim_dyn_p_row(sask_to_inoceram_redu_overlap,"I.
saskatchewanensis","B. reduncus-B. scotti")
conv_to_inoceram_nebr_sim_dyn_p_row <- make_sim_dyn_p_row(conv_to_inoceram_nebr_overlap,"I.
convexus","D. nebrascense-E. jenneyi")
```

```
subl_to_inoceram_nebr_sim_dyn_p_row <- make_sim_dyn_p_row(subl_to_inoceram_nebr_overlap,"I. sublaevis","D. nebrascense-E. jenneyi")

bara_to_inoceram_nebr_sim_dyn_p_row <- make_sim_dyn_p_row(bara_to_inoceram_nebr_overlap,"I. barabini","D. nebrascense-E. jenneyi")


bara_to_inoceram_rees_sim_dyn_p_row <- make_sim_dyn_p_row(bara_to_inoceram_rees_overlap,"I. barabini","B. reesidei-B. eliasi")

sage_to_inoceram_rees_sim_dyn_p_row <- make_sim_dyn_p_row(sage_to_inoceram_rees_overlap,"I. sagensis","B. reesidei-B. eliasi")


oblo_to_inoceram_rees_sim_dyn_p_row <- make_sim_dyn_p_row(oblo_to_inoceram_rees_overlap,"I. oblongus","B. reesidei-B. eliasi")


gen_to_sp_sim_dyn_p_results <-
rbind(conv_to_inoceram_redu_sim_dyn_p_row,subl_to_inoceram_redu_sim_dyn_p_row,
                        sask_to_inoceram_redu_sim_dyn_p_row,conv_to_inoceram_nebr_sim_dyn_p_row,
                        subl_to_inoceram_nebr_sim_dyn_p_row,bara_to_inoceram_nebr_sim_dyn_p_row,
                        bara_to_inoceram_rees_sim_dyn_p_row,sage_to_inoceram_rees_sim_dyn_p_row,
                        oblo_to_inoceram_rees_sim_dyn_p_row)


# Combine the results of niche dynamics
make_dynam_row <- function(data1,name,interval){
  data <- as.data.frame(data1[[6]])
  named1 <- as.data.frame(cbind(name,interval,t(data[,2])))
  colnames(named1) <- c("species","Intervals_compared","Expansion","Stability","Unfilling")
  named1[,c(3:5)] <- as.numeric(as.character(named1[,c(3:5)]))
  return(named1)
}


conv_to_inoceram_redu_dyn_row <- make_dynam_row(conv_to_inoceram_redu_overlap,"I. convexus","B. reduncus-B. scotti")

subl_to_inoceram_redu_dyn_row <- make_dynam_row(subl_to_inoceram_redu_overlap,"I. sublaevis","B. reduncus-B. scotti")
```

```
sask_to_inoceram_redu_dyn_row <- make_dynam_row(sask_to_inoceram_redu_overlap,"I. saskatchewanensis","B. reduncus-B. scotti")

conv_to_inoceram_nebr_dyn_row <- make_dynam_row(conv_to_inoceram_nebr_overlap,"I. convexus","D. nebrascense-E. jenneyi")


subl_to_inoceram_nebr_dyn_row <- make_dynam_row(subl_to_inoceram_nebr_overlap,"I. sublaevis","D. nebrascense-E. jenneyi")

bara_to_inoceram_nebr_dyn_row <- make_dynam_row(bara_to_inoceram_nebr_overlap,"I. barabini","D. nebrascense-E. jenneyi")


bara_to_inoceram_rees_dyn_row <- make_dynam_row(bara_to_inoceram_rees_overlap,"I. barabini","B. reesidei-B. eliasi")

sage_to_inoceram_rees_dyn_row <- make_dynam_row(sage_to_inoceram_rees_overlap,"I. sagensis","B. reesidei-B. eliasi")


oblo_to_inoceram_rees_dyn_row <- make_dynam_row(oblo_to_inoceram_rees_overlap,"I. oblongus","B. reesidei-B. eliasi")



gen_to_sp_dyn_results <- rbind(conv_to_inoceram_redu_dyn_row,subl_to_inoceram_redu_dyn_row,
                    sask_to_inoceram_redu_dyn_row,conv_to_inoceram_nebr_dyn_row,
                    subl_to_inoceram_nebr_dyn_row,bara_to_inoceram_nebr_dyn_row,
                    bara_to_inoceram_rees_dyn_row,sage_to_inoceram_rees_dyn_row,
                    oblo_to_inoceram_rees_dyn_row)



## Write csv files of results

write.csv(gen_to_sp_I_D_results,file="tables/gen_to_sp_I_D_results.csv")

write.csv(gen_to_sp_eq_results,file="tables/gen_to_sp_eq_results.csv")

write.csv(gen_to_sp_sim_results,file="tables/gen_to_sp_sim_results.csv")

write.csv(gen_to_sp_dyn_results,file="tables/gen_to_sp_dyn_results.csv")

write.csv(gen_to_sp_eq_dyn_p_results,file="tables/gen_to_sp_eq_dyn_p_results.csv")

write.csv(gen_to_sp_sim_dyn_p_results,file="tables/gen_to_sp_sim_dyn_p_results.csv")
```

```
##### RUN PCA OF MORE THAN 2 SPP TO GENUS ####
## Extract values to matrix for each raster stack of env proxy variables:

redu_taxa_list <- list(conv_redu_prep,subl_redu_prep,sask_redu_prep,Inoceramus_redu_prep)

nebr_taxa_list <- list(conv_nebr_prep,subl_nebr_prep,bara_nebr_prep,Inoceramus_nebr_prep)

rees_taxa_list <- list(bara_rees_prep,sage_rees_prep,oblo_rees_prep,Inoceramus_rees_prep)



name_list_redu <- c("I. convexus","I. sublaevis", "I. saskatchewanensis","Inoceramus")
name_list_nebr <- c("I. convexus","I. sublaevis","I. barabini","Inoceramus")
name_list_rees <- c("I. barabini","I. sagensis", "I. oblongus","Inoceramus")

# Create function to tabulate PCA scores for plotting spp and genus information
  # Function will only take a list of prepped occurrence/raster stack data
  # and output a list of 1-Joined (global) environment pts, 2-pts for each spp/gen,
  # and 3-the convex hull index values for creating polygons around spp/gen
run_gen_spp_pca <- function(occ_raster_list,name_list){

  bg_score_list <- list() # make list to put bg scores into

  for (i in 1:length(occ_raster_list)){

    # Get scores for each environment's backgrounds
    first_bg<- getValues(occ_raster_list[[i]][[2]])

    ## Clean out missing values:
    first_bg <- first_bg[complete.cases(first_bg), ]

    bg_score_list[[i]] <- first_bg
```

```
}

## Combined global environment:
joined_bg <- list.rbind(bg_score_list)

# Run PCA analysis on combine BG pts (all env data)
pca_joined_bg <- dudi.pca(joined_bg, center = TRUE,
                          scale = TRUE, scannf = FALSE, nf = 2)
joined_bg_scores <- pca_joined_bg$li # get the pca scores

occ_loc_list <- list() # Make list to put occ locations in PCA into
for (i in 1:length(occ_raster_list)){
  # Map occurrence data into the 2d ordination (have to coerce spatialPointsDataFrame into data.frame)
  # Explicitly match the colnames so only using the right ones
  # Only selected the li element (doesn't include others)
  first_occ_scores <-
    suprow(pca_joined_bg,
           data.frame(occ_raster_list[[i]][[1]])[, colnames(joined_bg)])$li

  occ_loc_list[[i]] <- first_occ_scores

}

# Add column with spp/gen names to occ location list:
for (i in 1:length(occ_raster_list)) {

  occ_loc_list[[i]]$tax <- name_list[i]

}

# Bind all the taxa localities together
```

```
  occ_loc_df <- list.rbind(occ_loc_list)

  bg_loc_list <- list() # make list to put bg locations in PCA into

  for (i in 1:length(occ_raster_list)){

    # Map BG data to the 2d ordination
    first_bg_scores <- suprow(pca_joined_bg, bg_score_list[[i]])$li

    bg_loc_list[[i]] <- first_bg_scores

  }
  #
  # NOT USED ANYMORE, BUT MAYBE STILL WANT TO KEEP IN CASE
  # hull_list <- list() # make list to hold convex hulls
  #
  # for (i in 1:length(occ_raster_list)){
  #
  #   # Make convex hulls to put around points for each spp/genus
  #   first_hull <- chull(occ_loc_list[[i]])
  #   first_hull <- c(first_hull,first_hull[1]) # Get list of index values for making a polygon
  #
  #   hull_list[[i]] <- first_hull
  #
  # }

  final_list <- list(joined_bg_scores,occ_loc_df)

  return(final_list)
}
```

```
# Use above function to make list of PCA information for plotting in each interval
redu_taxa_pca_list <- run_gen_spp_pca(redu_taxa_list,name_list_redu)
nebr_taxa_pca_list <- run_gen_spp_pca(nebr_taxa_list,name_list_nebr)
rees_taxa_pca_list <- run_gen_spp_pca(rees_taxa_list,name_list_rees)

# Use listed outputs from above function run on list of spp/gen to make plots:

library(devtools)
library(ggConvexHull)

pdf("Figures/pca_genus_to_spp_figures.pdf")
# REDU
ggplot(redu_taxa_pca_list[[2]], aes(x=Axis1, y=Axis2)) + # first spp/genus pts
  geom_point(data=redu_taxa_pca_list[[1]],aes(x=Axis1, y=Axis2),color='grey25') + # global env pts
  geom_convexhull(alpha = 0.3, aes(fill = tax,color = tax), linewidth = 1) +
  scale_fill_manual(values = c("#56B4E9", "#E69F00",'#F0E442', '#999999')) +
  scale_color_manual(values = c("#56B4E9", "#E69F00",'#F0E442', '#999999')) +
  ggtitle("PCA of Inoceramus genus and Species in B. reduncus-B. scotti") +
  theme_light() +
  theme(legend.title = element_blank())

# NEBR
ggplot(nebr_taxa_pca_list[[2]], aes(x=Axis1, y=Axis2)) + # first spp/genus pts
  geom_point(data=nebr_taxa_pca_list[[1]],aes(x=Axis1, y=Axis2),color='grey25') + # global env pts
  geom_convexhull(alpha = 0.3, aes(fill = tax, color = tax), linewidth = 1) +
  scale_fill_manual(values = c("#009E73", "#56B4E9", "#F0E442","#999999")) +
  scale_color_manual(values = c("#009E73", "#56B4E9", "#F0E442","#999999")) +
  ggtitle("PCA of Inoceramus genus and Species in D. nebrascense-E. jenneyi") +
  theme_light() +
  theme(legend.title = element_blank())
```

```
# REES
ggplot(rees_taxa_pca_list[[2]], aes(x=Axis1, y=Axis2)) + # first spp/genus pts
  geom_point(data=rees_taxa_pca_list[[1]],aes(x=Axis1, y=Axis2),color='grey25') + # global env pts
  geom_convexhull(alpha = 0.3, aes(fill = tax, color = tax), linewidth = 1) +
  scale_fill_manual(values = c("#009E73", "#CC79A7", '#0072B2', "#999999")) +
  scale_color_manual(values = c("#009E73", "#CC79A7", '#0072B2', "#999999")) +
  ggtitle("PCA of Inoceramus genus and Species in B. reesidei-B. eliasi") +
  theme_light() +
  theme(legend.title = element_blank())
dev.off()



#### SPECIES TO GENUS (ALL OCC, EVEN NO SPP ID) COMPARISONS ####


conv_to_inoceram_all_redu_overlap <- run_ecospat_dyn(conv_redu_prep,Inoceramus_all_redu_prep,
                                    "I. convextus to Inoceramus\nB. reduncus-B.
scotti","Iconv_to_Inocer_all_redu_")


subl_to_inoceram_all_redu_overlap <- run_ecospat_dyn(subl_redu_prep,Inoceramus_all_redu_prep,
                                    "I. sublaevis to Inoceramus\nB. reduncus-B. scotti","Isubl_to_Inocer_all_redu_")


sask_to_inoceram_all_redu_overlap <- run_ecospat_dyn(sask_redu_prep,Inoceramus_all_redu_prep,
                                    "I. saskatchewanensis to Inoceramus\nB. reduncus-B.
scotti","Isask_to_Inocer_all_redu_")


conv_to_inoceram_all_nebr_overlap <- run_ecospat_dyn(conv_nebr_prep,Inoceramus_all_nebr_prep,
                                    "I. convextus to Inoceramus\nD. nebrascense-E.
jenneyi","Iconv_to_Inocer_all_nebr")


subl_to_inoceram_all_nebr_overlap <- run_ecospat_dyn(subl_nebr_prep,Inoceramus_all_nebr_prep,
                                    "I. sublaevis to Inoceramus\nD. nebrascense-E.
jenneyi","Isubl_to_Inocer_all_nebr")
```

```
bara_to_inoceram_all_nebr_overlap <- run_ecospat_dyn(bara_nebr_prep,Inoceramus_all_nebr_prep,
                                    "I. barabini to Inoceramus\nD. nebrascense-E.
jenneyi","Ibara_to_Inocer_all_nebr")



bara_to_inoceram_all_rees_overlap <- run_ecospat_dyn(bara_rees_prep,Inoceramus_all_rees_prep,
                                    "I. barabini to Inoceramus\nB. reesidei-B. eliasi","Ibara_to_Inocer_all_rees")

sage_to_inoceram_all_rees_overlap <- run_ecospat_dyn(sage_rees_prep,Inoceramus_all_rees_prep,
                                    "I. sagensis to Inoceramus\nB. reesidei-B. eliasi","Isage_to_Inocer_all_rees")

oblo_to_inoceram_all_rees_overlap <- run_ecospat_dyn(oblo_rees_prep,Inoceramus_all_rees_prep,
                                    "I. oblongus to Inoceramus\nB. reesidei-B. eliasi","Ioblo_to_Inocer_all_rees")



#### Summarize Species to Genus (all occ) Comparisons and Export Tables ####
# Combine the results of observed overlap
make_D_I_row <- function(data1,name,interval){
  named1 <- as.data.frame(cbind(name,interval,data1[[2]]$obs$D,data1[[2]]$obs$I))
  colnames(named1) <- c("species","Intervals_compared","D_val","I_val")
  named1[,c(3:4)] <- as.numeric(as.character(named1[,c(3:4)]))
  return(named1)
}

conv_to_inoceram_all_redu_D_I_row <- make_D_I_row(conv_to_inoceram_all_redu_overlap,"I.
convexus","B. reduncus-B. scotti")
subl_to_inoceram_all_redu_D_I_row <- make_D_I_row(subl_to_inoceram_all_redu_overlap,"I. sublaevis","B.
reduncus-B. scotti")

sask_to_inoceram_all_redu_D_I_row <- make_D_I_row(sask_to_inoceram_all_redu_overlap,"I.
saskatchewanensis","B. reduncus-B. scotti")
```

```
conv_to_inoceram_all_nebr_D_I_row <- make_D_I_row(conv_to_inoceram_all_nebr_overlap,"I. convexus","D. nebrascense-E. jenneyi")


subl_to_inoceram_all_nebr_D_I_row <- make_D_I_row(subl_to_inoceram_all_nebr_overlap,"I. sublaevis","D. nebrascense-E. jenneyi")

bara_to_inoceram_all_nebr_D_I_row <- make_D_I_row(bara_to_inoceram_all_nebr_overlap,"I. barabini","D. nebrascense-E. jenneyi")


bara_to_inoceram_all_rees_D_I_row <- make_D_I_row(bara_to_inoceram_all_rees_overlap,"I. barabini","B. reesidei-B. eliasi")

sage_to_inoceram_all_rees_D_I_row <- make_D_I_row(sage_to_inoceram_all_rees_overlap,"I. sagensis","B. reesidei-B. eliasi")


oblo_to_inoceram_all_rees_D_I_row <- make_D_I_row(oblo_to_inoceram_all_rees_overlap,"I. oblongus","B. reesidei-B. eliasi")


gen_to_sp_I_D_results <- rbind(conv_to_inoceram_all_redu_D_I_row,subl_to_inoceram_all_redu_D_I_row,
                    sask_to_inoceram_all_redu_D_I_row,conv_to_inoceram_all_nebr_D_I_row,
                    subl_to_inoceram_all_nebr_D_I_row,bara_to_inoceram_all_nebr_D_I_row,
                    bara_to_inoceram_all_rees_D_I_row,sage_to_inoceram_all_rees_D_I_row,
                    oblo_to_inoceram_all_rees_D_I_row)



# Combine the results of equivalency tests

make_exp_p_row <- function(data1,name,interval){
  named1 <- as.data.frame(cbind(name,interval,data1[[2]]$p.D,data1[[2]]$p.I,data1[[3]]$p.D,data1[[3]]$p.I))
  colnames(named1) <-
c("species","Intervals_compared","Greater_Eq_p_D","Greater_Eq_p_I","Lower_Eq_p_D","Lower_Eq_p_I")
  named1[,c(3:4)] <- as.numeric(as.character(named1[,c(3:4)]))
  return(named1)
}


conv_to_inoceram_all_redu_eq_row <- make_exp_p_row(conv_to_inoceram_all_redu_overlap,"I. convexus","B. reduncus-B. scotti")
```

```
subl_to_inoceram_all_redu_eq_row <- make_exp_p_row(subl_to_inoceram_all_redu_overlap,"I. sublaevis","B. reduncus-B. scotti")


sask_to_inoceram_all_redu_eq_row <- make_exp_p_row(sask_to_inoceram_all_redu_overlap,"I. saskatchewanensis","B. reduncus-B. scotti")

conv_to_inoceram_all_nebr_eq_row <- make_exp_p_row(conv_to_inoceram_all_nebr_overlap,"I. convexus","D. nebrascense-E. jenneyi")


subl_to_inoceram_all_nebr_eq_row <- make_exp_p_row(subl_to_inoceram_all_nebr_overlap,"I. sublaevis","D. nebrascense-E. jenneyi")

bara_to_inoceram_all_nebr_eq_row <- make_exp_p_row(bara_to_inoceram_all_nebr_overlap,"I. barabini","D. nebrascense-E. jenneyi")


bara_to_inoceram_all_rees_eq_row <- make_exp_p_row(bara_to_inoceram_all_rees_overlap,"I. barabini","B. reesidei-B. eliasi")

sage_to_inoceram_all_rees_eq_row <- make_exp_p_row(sage_to_inoceram_all_rees_overlap,"I. sagensis","B. reesidei-B. eliasi")


oblo_to_inoceram_all_rees_eq_row <- make_exp_p_row(oblo_to_inoceram_all_rees_overlap,"I. oblongus","B. reesidei-B. eliasi")


gen_to_sp_eq_results <- rbind(conv_to_inoceram_all_redu_eq_row,subl_to_inoceram_all_redu_eq_row,
                              sask_to_inoceram_all_redu_eq_row,conv_to_inoceram_all_nebr_eq_row,
                              subl_to_inoceram_all_nebr_eq_row,bara_to_inoceram_all_nebr_eq_row,
                              bara_to_inoceram_all_rees_eq_row,sage_to_inoceram_all_rees_eq_row,
                              oblo_to_inoceram_all_rees_eq_row)


# Combine the results of equivalency tests of Dynamics (p-vals)

make_exp_dyn_p_row <- function(data1,name,interval){

  named1 <- as.data.frame(cbind(name,interval,data1[[2]]$p.expansion,data1[[2]]$p.stability,data1[[2]]$p.unfilling,
                                data1[[3]]$p.expansion,data1[[3]]$p.stability,data1[[3]]$p.unfilling))

  colnames(named1) <- c("species","Intervals_compared","Greater_Eq_p_expansion","Greater_Eq_p_stability","Greater_Eq_p_unfilling",
                        "Lower_Eq_p_expansion","Lower_Eq_p_stability","Lower_Eq_p_unfilling")
```

```
  named1[,c(3:4)] <- as.numeric(as.character(named1[,c(3:4)]))
  return(named1)
}



conv_to_inoceram_all_redu_eq_dyn_p_row <- make_exp_dyn_p_row(conv_to_inoceram_all_redu_overlap,"I.
convexus","B. reduncus-B. scotti")

subl_to_inoceram_all_redu_eq_dyn_p_row <- make_exp_dyn_p_row(subl_to_inoceram_all_redu_overlap,"I.
sublaevis","B. reduncus-B. scotti")


sask_to_inoceram_all_redu_eq_dyn_p_row <- make_exp_dyn_p_row(sask_to_inoceram_all_redu_overlap,"I.
saskatchewanensis","B. reduncus-B. scotti")

conv_to_inoceram_all_nebr_eq_dyn_p_row <- make_exp_dyn_p_row(conv_to_inoceram_all_nebr_overlap,"I.
convexus","D. nebrascense-E. jenneyi")


subl_to_inoceram_all_nebr_eq_dyn_p_row <- make_exp_dyn_p_row(subl_to_inoceram_all_nebr_overlap,"I.
sublaevis","D. nebrascense-E. jenneyi")

bara_to_inoceram_all_nebr_eq_dyn_p_row <- make_exp_dyn_p_row(bara_to_inoceram_all_nebr_overlap,"I.
barabini","D. nebrascense-E. jenneyi")


bara_to_inoceram_all_rees_eq_dyn_p_row <- make_exp_dyn_p_row(bara_to_inoceram_all_rees_overlap,"I.
barabini","B. reesidei-B. eliasi")

sage_to_inoceram_all_rees_eq_dyn_p_row <- make_exp_dyn_p_row(sage_to_inoceram_all_rees_overlap,"I.
sagensis","B. reesidei-B. eliasi")


oblo_to_inoceram_all_rees_eq_dyn_p_row <- make_exp_dyn_p_row(oblo_to_inoceram_all_rees_overlap,"I.
oblongus","B. reesidei-B. eliasi")



gen_to_sp_eq_dyn_p_results <-
rbind(conv_to_inoceram_all_redu_eq_dyn_p_row,subl_to_inoceram_all_redu_eq_dyn_p_row,
                    sask_to_inoceram_all_redu_eq_dyn_p_row,conv_to_inoceram_all_nebr_eq_dyn_p_row,
                    subl_to_inoceram_all_nebr_eq_dyn_p_row,bara_to_inoceram_all_nebr_eq_dyn_p_row,
                    bara_to_inoceram_all_rees_eq_dyn_p_row,sage_to_inoceram_all_rees_eq_dyn_p_row,
                    oblo_to_inoceram_all_rees_eq_dyn_p_row)
```

```
# Combine the results of similarity tests

make_sim_p_row <- function(data1,name,interval){

  named1 <- as.data.frame(cbind(name,interval,data1[[4]]$p.D,data1[[4]]$p.I,data1[[5]]$p.D,data1[[5]]$p.I))

  colnames(named1) <-
c("species","Intervals_compared","Greater_Sim_p_D","Greater_Sim_p_I","Lower_Sim_p_D","Lower_Sim_p_I
")

  named1[,c(3:4)] <- as.numeric(as.character(named1[,c(3:4)]))

  return(named1)

}


conv_to_inoceram_all_redu_sim_row <- make_sim_p_row(conv_to_inoceram_all_redu_overlap,"I.
convexus","B. reduncus-B. scotti")

subl_to_inoceram_all_redu_sim_row <- make_sim_p_row(subl_to_inoceram_all_redu_overlap,"I.
sublaevis","B. reduncus-B. scotti")


sask_to_inoceram_all_redu_sim_row <- make_sim_p_row(sask_to_inoceram_all_redu_overlap,"I.
saskatchewanensis","B. reduncus-B. scotti")

conv_to_inoceram_all_nebr_sim_row <- make_sim_p_row(conv_to_inoceram_all_nebr_overlap,"I.
convexus","D. nebrascense-E. jenneyi")


subl_to_inoceram_all_nebr_sim_row <- make_sim_p_row(subl_to_inoceram_all_nebr_overlap,"I.
sublaevis","D. nebrascense-E. jenneyi")

bara_to_inoceram_all_nebr_sim_row <- make_sim_p_row(bara_to_inoceram_all_nebr_overlap,"I.
barabini","D. nebrascense-E. jenneyi")


bara_to_inoceram_all_rees_sim_row <- make_sim_p_row(bara_to_inoceram_all_rees_overlap,"I. barabini","B.
reesidei-B. eliasi")

sage_to_inoceram_all_rees_sim_row <- make_sim_p_row(sage_to_inoceram_all_rees_overlap,"I. sagensis","B.
reesidei-B. eliasi")


oblo_to_inoceram_all_rees_sim_row <- make_sim_p_row(oblo_to_inoceram_all_rees_overlap,"I.
oblongus","B. reesidei-B. eliasi")



gen_to_sp_sim_results <- rbind(conv_to_inoceram_all_redu_sim_row,subl_to_inoceram_all_redu_sim_row,
```

```
                sask_to_inoceram_all_redu_sim_row,conv_to_inoceram_all_nebr_sim_row,
                subl_to_inoceram_all_nebr_sim_row,bara_to_inoceram_all_nebr_sim_row,
                bara_to_inoceram_all_rees_sim_row,sage_to_inoceram_all_rees_sim_row,
                oblo_to_inoceram_all_rees_sim_row)


# Combine the results of equivalency tests of Dynamics (p-vals)
make_sim_dyn_p_row <- function(data1,name,interval){
  named1 <-
as.data.frame(cbind(name,interval,data1[[4]]$p.expansion,data1[[4]]$p.stability,data1[[4]]$p.unfilling,
                    data1[[5]]$p.expansion,data1[[5]]$p.stability,data1[[5]]$p.unfilling))
  colnames(named1) <-
c("species","Intervals_compared","Greater_Sim_p_expansion","Greater_Sim_p_stability","Greater_Sim_p_unfi
lling",
          "Lower_Sim_p_expansion","Lower_Sim_p_stability","Lower_Sim_p_unfilling")
  named1[,c(3:4)] <- as.numeric(as.character(named1[,c(3:4)]))
  return(named1)
}


conv_to_inoceram_all_redu_sim_dyn_p_row <- make_sim_dyn_p_row(conv_to_inoceram_all_redu_overlap,"I.
convexus","B. reduncus-B. scotti")
subl_to_inoceram_all_redu_sim_dyn_p_row <- make_sim_dyn_p_row(subl_to_inoceram_all_redu_overlap,"I.
sublaevis","B. reduncus-B. scotti")


sask_to_inoceram_all_redu_sim_dyn_p_row <- make_sim_dyn_p_row(sask_to_inoceram_all_redu_overlap,"I.
saskatchewanensis","B. reduncus-B. scotti")
conv_to_inoceram_all_nebr_sim_dyn_p_row <- make_sim_dyn_p_row(conv_to_inoceram_all_nebr_overlap,"I.
convexus","D. nebrascense-E. jenneyi")


subl_to_inoceram_all_nebr_sim_dyn_p_row <- make_sim_dyn_p_row(subl_to_inoceram_all_nebr_overlap,"I.
sublaevis","D. nebrascense-E. jenneyi")
bara_to_inoceram_all_nebr_sim_dyn_p_row <- make_sim_dyn_p_row(bara_to_inoceram_all_nebr_overlap,"I.
barabini","D. nebrascense-E. jenneyi")


bara_to_inoceram_all_rees_sim_dyn_p_row <- make_sim_dyn_p_row(bara_to_inoceram_all_rees_overlap,"I.
barabini","B. reesidei-B. eliasi")
```

```
sage_to_inoceram_all_rees_sim_dyn_p_row <- make_sim_dyn_p_row(sage_to_inoceram_all_rees_overlap,"I.
sagensis","B. reesidei-B. eliasi")


oblo_to_inoceram_all_rees_sim_dyn_p_row <- make_sim_dyn_p_row(oblo_to_inoceram_all_rees_overlap,"I.
oblongus","B. reesidei-B. eliasi")



gen_to_sp_sim_dyn_p_results <-
rbind(conv_to_inoceram_all_redu_sim_dyn_p_row,subl_to_inoceram_all_redu_sim_dyn_p_row,

sask_to_inoceram_all_redu_sim_dyn_p_row,conv_to_inoceram_all_nebr_sim_dyn_p_row,

subl_to_inoceram_all_nebr_sim_dyn_p_row,bara_to_inoceram_all_nebr_sim_dyn_p_row,

bara_to_inoceram_all_rees_sim_dyn_p_row,sage_to_inoceram_all_rees_sim_dyn_p_row,
                        oblo_to_inoceram_all_rees_sim_dyn_p_row)


# Combine the results of niche dynamics
make_dynam_row <- function(data1,name,interval){
  data <- as.data.frame(data1[[6]])
  named1 <- as.data.frame(cbind(name,interval,t(data[,2])))
  colnames(named1) <- c("species","Intervals_compared","Expansion","Stability","Unfilling")
  named1[,c(3:5)] <- as.numeric(as.character(named1[,c(3:5)]))
  return(named1)
}


conv_to_inoceram_all_redu_dyn_row <- make_dynam_row(conv_to_inoceram_all_redu_overlap,"I.
convexus","B. reduncus-B. scotti")
subl_to_inoceram_all_redu_dyn_row <- make_dynam_row(subl_to_inoceram_all_redu_overlap,"I.
sublaevis","B. reduncus-B. scotti")


sask_to_inoceram_all_redu_dyn_row <- make_dynam_row(sask_to_inoceram_all_redu_overlap,"I.
saskatchewanensis","B. reduncus-B. scotti")
conv_to_inoceram_all_nebr_dyn_row <- make_dynam_row(conv_to_inoceram_all_nebr_overlap,"I.
convexus","D. nebrascense-E. jenneyi")
```

```
subl_to_inoceram_all_nebr_dyn_row <- make_dynam_row(subl_to_inoceram_all_nebr_overlap,"I. sublaevis","D. nebrascense-E. jenneyi")

bara_to_inoceram_all_nebr_dyn_row <- make_dynam_row(bara_to_inoceram_all_nebr_overlap,"I. barabini","D. nebrascense-E. jenneyi")


bara_to_inoceram_all_rees_dyn_row <- make_dynam_row(bara_to_inoceram_all_rees_overlap,"I. barabini","B. reesidei-B. eliasi")

sage_to_inoceram_all_rees_dyn_row <- make_dynam_row(sage_to_inoceram_all_rees_overlap,"I. sagensis","B. reesidei-B. eliasi")


oblo_to_inoceram_all_rees_dyn_row <- make_dynam_row(oblo_to_inoceram_all_rees_overlap,"I. oblongus","B. reesidei-B. eliasi")



gen_to_sp_dyn_results <- rbind(conv_to_inoceram_all_redu_dyn_row,subl_to_inoceram_all_redu_dyn_row,
                    sask_to_inoceram_all_redu_dyn_row,conv_to_inoceram_all_nebr_dyn_row,
                    subl_to_inoceram_all_nebr_dyn_row,bara_to_inoceram_all_nebr_dyn_row,
                    bara_to_inoceram_all_rees_dyn_row,sage_to_inoceram_all_rees_dyn_row,
                    oblo_to_inoceram_all_rees_dyn_row)


## Write csv files of results

write.csv(gen_to_sp_I_D_results,file="tables/gen_all_to_sp_I_D_results.csv")

write.csv(gen_to_sp_eq_results,file="tables/gen_all_to_sp_eq_results.csv")

write.csv(gen_to_sp_sim_results,file="tables/gen_all_to_sp_sim_results.csv")

write.csv(gen_to_sp_dyn_results,file="tables/gen_all_to_sp_dyn_results.csv")

write.csv(gen_to_sp_eq_dyn_p_results,file="tables/gen_all_to_sp_eq_dyn_p_results.csv")

write.csv(gen_to_sp_sim_dyn_p_results,file="tables/gen_all_to_sp_sim_dyn_p_results.csv")


##### RUN PCA OF MORE THAN 2 SPP TO GENUS ####

## Extract values to matrix for each raster stack of env proxy variables:


redu_taxa_list <- list(conv_redu_prep,subl_redu_prep,sask_redu_prep,Inoceramus_all_redu_prep)
```

```
nebr_taxa_list <- list(conv_nebr_prep,subl_nebr_prep,bara_nebr_prep,Inoceramus_all_nebr_prep)

rees_taxa_list <- list(bara_rees_prep,sage_rees_prep,oblo_rees_prep,Inoceramus_all_rees_prep)


name_list_redu <- c("I. convexus","I. sublaevis", "I. saskatchewanensis","Inoceramus")
name_list_nebr <- c("I. convexus","I. sublaevis","I. barabini","Inoceramus")
name_list_rees <- c("I. barabini","I. sagensis", "I. oblongus","Inoceramus")

# Create function to tabulate PCA scores for plotting spp and genus information
# Function will only take a list of prepped occurrence/raster stack data
# and output a list of 1-Joined (global) environment pts, 2-pts for each spp/gen,
# and 3-the convex hull index values for creating polygons around spp/gen
run_gen_spp_pca <- function(occ_raster_list,name_list){

  bg_score_list <- list() # make list to put bg scores into

  for (i in 1:length(occ_raster_list)){

    # Get scores for each environment's backgrounds
    first_bg<- getValues(occ_raster_list[[i]][[2]])

    ## Clean out missing values:
    first_bg <- first_bg[complete.cases(first_bg), ]

    bg_score_list[[i]] <- first_bg
  }

  ## Combined global environment:
  joined_bg <- list.rbind(bg_score_list)
```

```
# Run PCA analysis on combine BG pts (all env data)
pca_joined_bg <- dudi.pca(joined_bg, center = TRUE,
                          scale = TRUE, scannf = FALSE, nf = 2)
joined_bg_scores <- pca_joined_bg$li # get the pca scores



occ_loc_list <- list() # Make list to put occ locations in PCA into
for (i in 1:length(occ_raster_list)){
  # Map occurrence data into the 2d ordination (have to coerce spatialPointsDataFrame into data.frame)
  # Explicitly match the colnames so only using the right ones
  # Only selected the li element (doesn't include others)
  first_occ_scores <-
    suprow(pca_joined_bg,
           data.frame(occ_raster_list[[i]][[1]])[, colnames(joined_bg)])$li

  occ_loc_list[[i]] <- first_occ_scores

}

# Add column with spp/gen names to occ location list:
for (i in 1:length(occ_raster_list)) {

  occ_loc_list[[i]]$tax <- name_list[i]

}

# Bind all the taxa localities together
occ_loc_df <- list.rbind(occ_loc_list)

bg_loc_list <- list() # make list to put bg locations in PCA into
```

```
  for (i in 1:length(occ_raster_list)){

    # Map BG data to the 2d ordination
    first_bg_scores <- suprow(pca_joined_bg, bg_score_list[[i]])$li

    bg_loc_list[[i]] <- first_bg_scores

  }
  #
  # NOT USED ANYMORE, BUT MAYBE STILL WANT TO KEEP IN CASE
  # hull_list <- list() # make list to hold convex hulls
  #
  # for (i in 1:length(occ_raster_list)){
  #
  #   # Make convex hulls to put around points for each spp/genus
  #   first_hull <- chull(occ_loc_list[[i]])
  #   first_hull <- c(first_hull,first_hull[1]) # Get list of index values for making a polygon
  #
  #   hull_list[[i]] <- first_hull
  #
  # }

  final_list <- list(joined_bg_scores,occ_loc_df)

  return(final_list)
}

# Use above function to make list of PCA information for plotting in each interval
redu_taxa_pca_list <- run_gen_spp_pca(redu_taxa_list,name_list_redu)
nebr_taxa_pca_list <- run_gen_spp_pca(nebr_taxa_list,name_list_nebr)
```

```
rees_taxa_pca_list <- run_gen_spp_pca(rees_taxa_list,name_list_rees)



# Use listed outputs from above function run on list of spp/gen to make plots:

library(devtools)
library(ggConvexHull)

pdf("Figures/pca_genus_all_to_spp_figures.pdf")
# REDU
ggplot(redu_taxa_pca_list[[2]], aes(x=Axis1, y=Axis2)) + # first spp/genus pts
  geom_point(data=redu_taxa_pca_list[[1]],aes(x=Axis1, y=Axis2),color='grey25') + # global env pts
  geom_convexhull(alpha = 0.3, aes(fill = tax,color = tax), linewidth = 1) +
  scale_fill_manual(values = c("#56B4E9", "#E69F00",'#F0E442', '#999999')) +
  scale_color_manual(values = c("#56B4E9", "#E69F00",'#F0E442', '#999999')) +
  ggtitle("PCA of Inoceramus genus and Species in B. reduncus-B. scotti") +
  theme_light() +
  theme(legend.title = element_blank())

# NEBR
ggplot(nebr_taxa_pca_list[[2]], aes(x=Axis1, y=Axis2)) + # first spp/genus pts
  geom_point(data=nebr_taxa_pca_list[[1]],aes(x=Axis1, y=Axis2),color='grey25') + # global env pts
  geom_convexhull(alpha = 0.3, aes(fill = tax, color = tax), linewidth = 1) +
  scale_fill_manual(values = c("#56B4E9", "#E69F00", "#009E73","#999999")) +
  scale_color_manual(values = c("#56B4E9", "#E69F00", "#009E73","#999999")) +
  ggtitle("PCA of Inoceramus genus and Species in D. nebrascense-E. jenneyi") +
  theme_light() +
  theme(legend.title = element_blank())

# REES
ggplot(rees_taxa_pca_list[[2]], aes(x=Axis1, y=Axis2)) + # first spp/genus pts
```

```
  geom_point(data=rees_taxa_pca_list[[1]],aes(x=Axis1, y=Axis2),color='grey25') + # global env pts
  geom_convexhull(alpha = 0.3, aes(fill = tax, color = tax), linewidth = 1) +
  scale_fill_manual(values = c("#009E73", "#CC79A7", '#0072B2', "#999999")) +
  scale_color_manual(values = c("#009E73", "#CC79A7", '#0072B2', "#999999")) +
  ggtitle("PCA of Inoceramus genus and Species in B. reesidei-B. eliasi") +
  theme_light() +
  theme(legend.title = element_blank())
dev.off()

#### Data review (correlation charts) ####

library(ggplot2)

setwd("C:/Users/ceara/Documents/Province Project/ENM_Analysis_Code-and-Files")

## Duration vs D in interval comparisons

dat <- read.csv("csv_for_dat_analysis/duration_d_stability_vals.csv")

#fit linear regression model to dataset and view model summary
dat_model <- lm(d~ratio, data=dat)
summary(dat_model)
names(summary(dat_model))

r <- summary(dat_model)$adj.r.squared

## Mean Duration vs mean D in interval



dat_mean <- read.csv("csv_for_dat_analysis/mean_duration_d_stability_vals.csv")
```

```
#fit linear regression model to dataset and view model summary
mean_dat_model <- lm(d_mean~ratio, data=dat_mean)
r_mean_dat <- summary(mean_dat_model)$adj.r.squared


# Export corr plots


cbPalette <- c("#999999", "#E69F00", "#56B4E9", "#009E73", "#F0E442", "#0072B2", "#D55E00",
"#CC79A7")


dat$intervals_compared<-sub("to", "to \n ", dat$intervals_compared)
dat_mean$t_comparison<-sub("to", "to \n ", dat_mean$t_comparison)


pdf(file = "Figures/corr_plots_duration_overlap.pdf",width = 2.75591, height = 2.75591 )
ggplot(data = dat) +
  geom_point(aes(x = ratio, y = d, color = intervals_compared), shape = 19, size = 0.3) +
  geom_smooth(method='lm', mapping = aes(x = ratio, y = d), size=0.3) +
  theme_light() +
  labs(x ="Duration Ratio (biozone : substage/stage)", y = "D",color = "Intervals Compared") +
  scale_colour_manual(values=cbPalette) +
  theme(plot.title = element_text(size=9, hjust=0.5),
        axis.title.x = element_text(size=9, vjust=2.5),
        axis.title.y = element_text(size=9, vjust=-1.5),
        axis.text.x = element_text(size=6, vjust=2),
        axis.text.y = element_text(size=6, angle=90, hjust=0.5, vjust=-0.5),
        legend.position="bottom",
        legend.text = element_text(size=6),
        legend.title= element_blank(),
        legend.spacing.x = unit(-0.15, 'cm'),
        legend.spacing.y = unit(-0.15, 'cm'),
        legend.box.margin=margin(-12,15,-10,0)) +
  annotate(geom="text", x=0.4, y=0.9, size = 3, label=paste("R^2 == ", round(r, digits = 3)), parse=TRUE) +
  guides(color = guide_legend(byrow=TRUE, nrow = 4))
```

```
ggplot(data = dat_mean) +
  geom_point(aes(x = ratio, y = d_mean, color = t_comparison), shape = 19, size = 0.3) +
  geom_smooth(method='lm', mapping = aes(x = ratio, y = d_mean), size=0.3) +
  theme_light() +
  labs(x ="Duration Ratio (biozone : substage/stage)", y = "Mean D",color = "Intervals Compared") +
  scale_colour_manual(values=cbPalette) +
  theme(plot.title = element_text(size=9, hjust=0.5),
        axis.title.x = element_text(size=9, vjust=2.5),
        axis.title.y = element_text(size=9, vjust=-1.5),
        axis.text.x = element_text(size=6, vjust=2),
        axis.text.y = element_text(size=6, angle=90, hjust=0.5, vjust=-0.5),
        legend.position="bottom",
        legend.text = element_text(size=6),
        legend.title= element_blank(),
        legend.spacing.x = unit(-0.15, 'cm'),
        legend.spacing.y = unit(-0.15, 'cm'),
        legend.box.margin=margin(-12,15,-10,0)) +
  annotate(geom="text", x=0.3, y=1, size = 3, label=paste("R^2 == ", round(r_mean_dat, digits = 3)),
parse=TRUE) +
  guides(color = guide_legend(byrow=TRUE,nrow = 4))

dev.off()



## Duration vs mean D in interval (spp comparisons)

dat_spp <- read.csv("spp_d_stability_vals.csv")

ggplot(data = dat_spp) +
  geom_point(aes(x = short_dur, y = d, color = species)) +
```

```
  geom_smooth(method='lm', mapping = aes(x = short_dur, y = d))



#fit linear regression model to dataset and view model summary
model <- lm(d~short_dur, data=dat_spp)
summary(model)

## Duration ratio vs mean D in interval

dat <- read.csv("duration_d_stability_vals.csv")

# plot the data
plot(dat$ratio, dat$d, pch = 19, col = "lightblue")

# Add regression line
abline(lm(dat$d ~ dat$ratio), col = "red", lwd = 3)

# Pearson correlation
text(paste("Correlation:", round(cor(dat$ratio, dat$d), digits = 2)), x = 0.25, y = 0.78)

## Duration ratio vs mean D in interval

dat_mean <- read.csv("mean_duration_d_stability_vals.csv")

# plot the data
plot(dat_mean$ratio, dat_mean$d_mean, pch = 19, col = "lightblue")

# Add regression line
abline(lm(dat_mean$d_mean ~ dat_mean$ratio), col = "red", lwd = 3)

# Pearson correlation
```

```
text(paste("Correlation:", round(cor(dat_mean$ratio, dat_mean$d_mean), digits = 2)), x = 0.25, y = 0.78)

## Duration vs mean D in interval (spp comparisons)

dat_spp <- read.csv("spp_d_stability_vals.csv")

# plot the data
plot(dat_spp$short_dur, dat_spp$d, pch = 19, col = "lightblue")

# Add regression line
abline(lm(dat_spp$d ~ dat_spp$short_dur), col = "red", lwd = 3)

# Pearson correlation
text(paste("Correlation:", round(cor(dat_spp$short_dur, dat_spp$d), digits = 2)), x = 1.3, y = 0.7)
```