Diversity dynamics of New Zealand's temperate shallow-marine ecosystems through the Cenozoic

by

Thomas Michael Womack

A thesis submitted to the Victoria University of Wellington in fulfilment of the requirements for the degree of Doctor of Philosophy.

Victoria University of Wellington

(2022)

Abstract

The fossil record provides the only evidence of how ecosystems have evolved over long timescales (multi-centennial to millennial) and how they have responded to changes in geological and environmental processes. Quantifying the impact of these geological and environmental processes on biodiversity is fundamental to understanding the paradigm of modern biodiversity loss and the ecological reverberations of anthropogenic climate change. Crucially, biodiversity is spatially dependent, and our perception of biodiversity may differ depending on the spatial scale of observation. As a result, understanding the diversity dynamics of the fossil record is both a question of time and space. Recent evidence suggests that studies at the regional spatial scale are optimal for understanding macroevolutionary and macroecological processes. Despite this, regional-scale studies using data from the fossil record are limited, particularly in the southern hemisphere.

This thesis focuses on the temporal and spatial impacts of geological, climatic, and oceanographic processes on New Zealand's regional shallow-marine molluscan ecosystems through the Cenozoic Era (the last 66 Myrs). New Zealand hosts remarkably complete Cenozoic stratigraphic and fossil records; the latter is regarded as the most complete shallow-marine fossil record in the southern hemisphere. As a result, this record provides an ideal case study to examine macroevolutionary and macroecological processes in the shallow marine realm at the regional scale. In addition, New Zealand's modern and Cenozoic shallow-marine ecosystems are temperate. Temperate marine ecosystems are vastly understudied compared to their tropical counterparts in both modern ecosystems and the fossil record, and their response to climate change is poorly understood, particularly in the southern hemisphere.

This thesis comprises four related research papers that form the main chapters. These are introduced by detailed background, data analysis, and methods chapters. The overarching aim of these papers is to: (1) quantify the impact of known confounding effects in the fossil record, particularly the Pull of the Recent; (2) evaluate the spatial dependence of biodiversity in the fossil record, particularly on temporal patterns of beta diversity; (3) understand the impact of climate change on taxonomic and functional biogeographic patterns of shallow-marine biodiversity; (4) employ the methods developed in the previous chapters to test the importance and role of several important geological and climatic factors on the spatial structuring New Zealand's shallow-marine biodiversity.

First, I find that the Pull of the Recent effects <2 % of genera and <4 % of species when considering taxa missing from the youngest ~5 Myrs of the New Zealand shallow-marine molluscan fossil record. Furthermore, the taxonomic composition of molluscs missing from the youngest part of the fossil record cannot easily be explained by effects related to shell mineralogical composition, body size, habitat, or taxonomic class. Lithification has minimal effect on the impact of the Pull of the Recent but does have a notable effect on apparent range-through diversity in the Pleistocene. Secondly, I show that beta diversity is spatially dependent at local to regional spatial scales and that uneven spatial sampling can influence recovered temporal trends in beta diversity. We argue that newly developed methods that allow for multisite beta diversity, partitioned into total dissimilarity, spatial turnover and nestedness, are the most suitable for elucidating patterns of beta diversity in the fossil record. Thirdly, I find that the interrelationship between two biogeographic patterns, the species-area relationship and functional diversity-area relationship, shows a strong, positive correlation to regional oceanic temperature, suggesting a long-lived and persistent association in the structuring of biodiversity, and temperature-dependence of functional redundancy, over the last ~45 Ma. Lastly, I find that oceanic warming seemingly has both a

long-term positive effect on regional taxonomic diversity and rates of origination, but a negative correlation with regional functional diversity and its spatial distribution. This negative long-term relationship between regional functional diversity and oceanic temperature is likely driven by functional resilience to extinction during cool intervals rather than a negative response to warming per se.

In summary, this thesis confirms the high quality of the New Zealand shallow-marine fossil record and identifies oceanic temperature as a major positive correlate of regional taxonomic diversity, origination, and functional redundancy. This correlation is not consistent across spatial scales, a reflection of other abiotic and biotic processes that are likely important. These findings are relevant to our understanding of macroevolutionary and macroecological processes, particularly how temperate shallow-marine ecosystems may respond to future warming. For New Zealand, this research suggests, as a baseline, that natural warming may increase equilibrium taxonomic diversity and functional redundancy in New Zealand. However, anthropogenic climate change is associated with increased loss of biodiversity and habitat destruction in the geological short term. This is likely to alter the future trajectory of New Zealand's biodiversity anticipated from natural warming, which may have far-reaching impacts for the short- and long-term future of New Zealand's unique and highly endemic marine fauna. Furthermore, the results discussed of this thesis also suggest the unfolding sixth mass extinction may be qualitatively different (i.e., higher levels of functional extinction) from previous mass extinction events and have correspondingly unforeseeable consequences. Lastly, and importantly, this thesis highlights the need for further study in shallow-marine temperate ecosystems, particularly at finer time scales, to better our understanding of diversity dynamics, and extinction risk.

Acknowledgments

This thesis would have not been possible without the help and support of many people. Firstly, I would like to thank my supervisors James Crampton and Michael Hannah. Your continued enthusiasm, encouragement, and positive feedback – in the face of numerous graphs with limited keys – has been fundamental in shaping and completing this PhD. I would particularly like to thank James for supporting my application for funding for this project (not only once but twice), always being positive and excited about my research, and ultimately inspiring much of the work I have produced. I would also be remised if did not take this opportunity to acknowledge your seemingly endless wardrobe of vibrant (dare I say stylish?) shirts.

Second, I would like to thank several other key contributors to this thesis. Alan Beu and Katie Collins thank you for all your taxonomic molluscan knowledge, without your help identifying and categorising molluscs I would surely still be agonising over *Stiracolpus* and *Zeacolpus*. I would like to further acknowledge Katie for helping me create a molluscan ecospace, a task that would have taken a lot longer without your help. In addition, I would like to thank Ben Hines, who has been a compendium of New Zealand geology and a worthy squash foe. I would also like to thank the contributors to the Fossil Record Electronic Database (FRED) and particularly Chris Clowes (but also James Crampton and Joe Prebble) for helping me access this exemplary (although computationally archaic) resource. I am also grateful to Hannu Seebeck for showing me how to use GPlates and Dominic Strogen for his assistance and provision of the base maps for the paleobathymetry of New Zealand. Furthermore, I would like to thank the organizers of the Analytical Paleobiology Workshop and particularly the other attendees who have been a shared source of knowledge (and dog photos) – particularly on statistics. Lastly within this paragraph, I would like to acknowledge all the named and

anonymous reviewers of the papers that form the chapters of this thesis, and John Alroy for his advice on sampling and spatial standardization.

During this project I completed several weeks of fieldwork in Hawke's Bay and along the Whanganui River. Although this fieldwork did not make it into the final thesis (in part thanks to the global pandemic), I would like to extend particular thanks to Tim Naish for his assistance and expert knowledge on the sequence stratigraphy on the Whanganui River and Kyle Bland for his advice on East Coast. The sieving of these fossil is still underway and on this note I would like to thank Jane Chewings for accommodating my many bags of fossil laden sand. Furthermore, I would also like to extend my thanks to Dennis Gordon for help in identifying (and photographing) the diverse array of bryozoans I have collected – hopefully, this alongside the molluscan communities, will be published at some point.

I would also like to thank several other people at Victoria University who may not have directly contributed to my PhD thesis but have shown me the geology of New Zealand and kept me gainfully employed over the last several years. Thanks particularly to Dene Carroll and Cliff Atkins (but also James Crampton and Michael Hannah), I do not think this PhD would have been the same without the Te Muna fieldtrip – every year I learn something new (and probably knock a few old things out at the same time). There are also many other fundamental staff at Victoria University that have facilitated my fieldwork and general logistics of tutoring, whilst there are too many to name your help has not gone unnoticed, although I will single out Dez Tessler who has always helped and solved all my logistical woes.

It has been a pleasure to spend the last few years in the School of Geography, Environmental and Earth Sciences, and particularly in CO418. Although the blinding glare of the sun, the astronomical temperatures within the "Austrian sauna" during the winter and summer, and the still unwashed tea mug were sometimes a hinderance – many friendships have

Table of Contents

List of Figures

Chapter 5

Chapter 6

Chapter 7

Chapter 8

List of Tables

Chapter 7

Chapter 8

Table of Appendices

Digital Appendices (code and datafiles, see Chapter 3).

Chapter 6_150721.zip

Chapter 7_150721.zip

Chapter 8_150721.zip

Chapter 1

Introduction

"It's not possible! No, it's necessary"

(Cooper and Case, Interstellar, 2014*)*

1.1. Introduction

Ecological observations are crucial to predicting and understanding short-term anthropogenic impacts on diversity. However, the transient nature of ecological observations confounds our understanding of biological processes over longer time spans (centennial to millennial). The fossil record provides the only direct evidence of the long-term impacts of climatic and environmental perturbations on biodiversity; fundamental to our understanding of macroecological and macroevolutionary processes. Reconciling these long-term processes with short-term observations and predictions of subrecent and modern changes in biodiversity and biogeography is paramount to protecting the future of our biosphere.

The development of new quantitative techniques and the creation and modification of new and existing datasets over the last few decades has progressed the field of analytical paleobiology, fuelled by advances in computing power and technology. These advances have provided new opportunities to study the quantitative aspects of macroecology and macroevolution, and revisit and postulate new questions on the nature of biological processes at different temporal and spatial scales. Within this thesis I focus on the application and extension of these quantitative techniques to examine and quantify the long-term driving mechanisms of New Zealand's temperate shallow-marine biodiversity change through the Cenozoic.

I focus on New Zealand's Cenozoic shallow-marine fossil record for four reasons. Firstly, New Zealand's Cenozoic molluscan fossil record is regarded as the most complete in the southern hemisphere (Cooper et al. 2006, Crampton et al. 2006). Second, New Zealand's current and past shallow-marine ecosystems are temperate, which are understudied in the context of their ecology and vulnerability to climatic shifts (Vergés et al. 2014, Wernberg et al. 2016, Antão et al. 2020). Third, New Zealand's past and present marine biota have a high

degree of endemism at the species level (> 95 % endemism in the recent molluscan fauna and ~85 % during the Cenozoic) (Beu and Maxwell 1990, Crampton et al. 2006, Gordon et al. 2010, Mittermeier et al. 2011). Furthermore, New Zealand's modern marine ecosystems are currently identified as an international biodiversity hotspot (Myers et al. 2000), but are also thought to be at significant risk from future anthropogenic climate change (Finnegan et al. 2015, Urban 2015, Collins et al. 2018). Lastly, recent studies suggest that the global fossil record may not represent true patterns of global biodiversity and are confounded by the uneven spatial distribution of fossil collection through time (Close et al. 2020). This implies that focus should be placed on studies at the regional spatial scale, which is regarded as an important scale for macroevolutionary processes (Vermeij and Leighton 2003). Thus, New Zealand provides an ideal case-study to examine the long-term regional drivers of temperate marine biodiversity change, but also provides an important historic perspective relevant to the long-term conservation strategy of New Zealand's vulnerable marine biota. Despite the exemplary nature of New Zealand's Cenozoic shallow-marine fossil record, it is imperfect. Temporal patterns of apparent diversity (i.e., raw data) are likely to be confounded by effects related to uneven sampling of fossil collections across space and time. Thus, quantifying the long-term drivers of shallow-marine biodiversity change requires an understanding of the spatiotemporal structure of the data – a consideration that underpins many analyses throughout this thesis.

1.2. Research Questions

Although this thesis is a compilation of four research papers, there are three research questions that underpin the study. These questions are outlined below and discussed in more detail in section 1.3.

1. What is the structure of apparent diversity patterns in New Zealand's shallow-marine fossil record?

    a. Is this consistent with the nature of the global diversity history that records a putative rise in taxonomic diversity towards the recent?

    b. Are the resultant patterns impacted by biases that are known to affect the global diversity trajectory, in particular the Pull of the Recent (see section 1.3.4)?

2. Are patterns of biodiversity at the regional scale confounded by the uneven distribution of fossil collections?

    a. Is there a pattern of spatial scaling in beta diversity and are temporal patterns of beta diversity in the fossil record dependent on the spatial distribution of data?

    b. Does the spatial scaling of both taxonomic and functional richness conform to similar patterns observed in modern ecology?

3. What are the key drivers of long-term biodiversity change and macroevolution in New Zealand's shallow-marine ecosystems through the Cenozoic?

   a. Are these drivers consistent across spatial partitions of biodiversity?

   b. Do these drivers affect taxonomic and functional components of biodiversity in the same way through space and time?

   c. Are the long-term drivers of biodiversity consistent with predictions and observations from modern ecology?

   d. What are the implications for the future of New Zealand's temperate marine ecosystems?

1.3. Thesis Outline

This thesis comprises four published, or soon to be submitted, research papers, which form the four main results Chapters (Chapters 5 - 8) (Table 1.1). These papers have been reformatted here, but the content remains largely unchanged from the published versions. I also include four supplementary Chapters which expand on the background and methods used within the papers. A preface is provided for each of the main research Chapters to: (1) provide a brief background on the focus of the paper, (2) contextualise the research in relation to the other Chapters and overall narrative of the thesis, and (3) outline the role of co-authors and other contributors. Outlines of all the Chapters are given below. In addition, during my doctoral studies I have co-authored two papers. Although these co-authored papers are not included within this thesis, they exploit elements of the knowledge and methods developed throughout this doctorate (Table 1.1).

Thesis Chapter	Paper title	Journal	Status
Lead author			
5	*The Pull of the Recent revisited – negligible species-level effect in a regional marine fossil record.*	Paleobiology	Published (vol. *46*, pp. 470-477)
6	*Spatial scaling of beta diversity in the shallow-marine fossil record*	Paleobiology	Published (vol. *47,* pp. 39-53)
7	*A positive relationship between functional redundancy and temperature in Cenozoic marine ecosystems*	Science	Second revision
8	*Biogeographic trends and climatic change: comparisons through space and time*	Proceedings of the Royal Society B	Unsubmitted
Co-authored			
-	*The New Zealand Fossil Record File: a unique database of biological history*	New Zealand Journal of Geology and Geophysics	Published (vol. *64,* pp. 62-71)
-	*A 100 million year composite pollen record from New Zealand shows maximum angiosperm abundance delayed until Eocene*	Paleogeography, Paleoclimatology, Paleoecology	Published online (vol. *566*)

TABLE 1.1. Summary of the publication status of the Chapters included within this thesis and additional co-authored papers.

1.3.1. Chapter 2 – *Background*

This Chapter provides a broad review of the literature relevant to this thesis. The aim of this Chapter is to provide a primer on New Zealand's geological, climatic, and oceanographic history through the Cenozoic, and a summary of the previous research on New Zealand's Cenozoic shallow-marine molluscan fossil record. This Chapter can be read in isolation or as a supplement to the main results Chapters and is cross-referenced throughout.

1.3.2. Chapter 3 – *Data and Analysis*

The purpose of this Chapter is to discuss the workflows used to generate the main datasets employed within these Chapters and provide details on the computation methods of data analysis. This Chapter can be read in isolation or as a supplement to the main results Chapters and is cross-referenced throughout.

1.3.3. Chapter 4 – *Methods Review*

This Chapter provides a review of the methods used within this thesis. The purpose of this Chapter is to contextualise, explain and justify methodological decisions. This Chapter can be read in isolation or as a supplement to the main results Chapters and is cross-referenced throughout.

1.3.4. Chapter 5 – *The Pull of the Recent revisited – Negligible Species-Level Effect in a Regional Marine Fossil Record.*

This Chapter addresses research question 1 and quantifies elements of the raw structure and the impact of the Pull of the Recent on New Zealand's shallow-marine molluscan fossil record through the Cenozoic. The Pull of the Recent describes distortion of apparent biodiversity patterns over time. This distortion results from comparatively complete knowledge and sampling of the living biota in the recent compared to fossil biotas of the past. In effect, the Pull of the Recent extends the biostratigraphic ranges of fossil taxa, and therefore increases range-through diversity estimates toward the recent through the inclusion of living "Lazarus" taxa (i.e., those with gaps in their stratigraphic records). The Chapter takes several novel approaches to assess the impact of the Pull of the Recent, including: (1) quantifying the Pull of the Recent at species level; (2) focusing on the Cenozoic Era rather than the Phanerozoic Eon (66 Ma to present rather than 541 Ma to present); (3) considering

smaller time intervals than previous studies; and (4) examining whether particular attributes of taxa control their likely occurrence in the youngest fossil record.

1.3.5. Chapter 6 – *Spatial Scaling of Beta Diversity in the Shallow-Marine Fossil Record*

This Chapter addresses research question 2 and quantifies the spatial scaling of beta diversity in the shallow-marine fossil record, providing the basis for several of the methodological approaches utilised in Chapters 7 and 8. The aim of this Chapter is to assess various measures of beta diversity in order to identify the best methods to use in the New Zealand setting and to quantify the spatial scaling of beta diversity. Beta diversity describes the spatial (geographic) structuring of ecological communities and is central to understanding many macroecological and macroevolutionary phenomena. However, studies of beta diversity at the regional scale are relatively limited, and even scarcer are studies on the spatial dependence of beta diversity – in other words, how are measures of beta diversity influenced by the geographic extent of observation? This is particularly important in the fossil record, where temporal variation in spatial distribution of fossil localities is a known and pervasive bias.

1.3.6. Chapter 7 – *A Positive Relationship Between Functional Redundancy and Temperature in Cenozoic Marine Ecosystems*

This Chapter addresses components of research questions 2 and 3 and quantifies the spatial scaling of taxonomic and functional diversities (i.e., the species-area and functional diversity-area relationships, respectively). This Chapter builds on observations from the modern latitudinal diversity gradient, which suggest that the interrelationship between the species-area relationship and functional-diversity area relationship - i.e., the relationship between species and functional richness at increasing geographic scales - may be temperature dependent (Schumm et al. 2019). The aims of this Chapter are to (1) quantify the species-area

relationship and functional diversity-area relationhship in New Zealand's shallow-marine fossil record to see if they conform to patterns observed in modern ecology, and (2) determine whether their interrelationship – termed the functional redundancy relationship within this thesis, which represents an aspect of functional redundancy – is correlated to proxies of oceanic temperature during the Cenozoic.

1.3.7.   Chapter 8 – *Biogeographic Trends with Climatic Change: Comparisons Through Space and Time*

This Chapter addresses research question 3 and focuses on quantifying the role of oceanic temperature, climatic variability, environmental heterogeneity, and habitable area in controlling the spatial structuring of New Zealand's shallow-marine molluscan biodiversity and macroevolutionary rates (i.e., origination and extinction) through the Cenozoic. This Chapter builds on two observations/predictions from modern ecology. First, observations from the modern latitudinal diversity gradient in marine molluscs reveal a poleward decline in both species and functional richness (Edie et al. 2018a, b, Schumm et al. 2019). This is attributed, in part, to latitudinal gradients of climate and temperature. Secondly, short-term observations and predictions of anthropogenic climate change suggest an asymmetric response in tropical versus temperate marine ecosystems, whereby temperate ecosystems may see increases in regional species richness in response to oceanic warming while tropical richness declines (Antão et al. 2020). Within this Chapter I focus on both the taxonomic and functional aspects of biodiversity and examine: (1) whether these two components exhibit congruent temporal and spatial patterns, and (2) whether they respond in the same way to the explanatory factors considered here.

1.3.8. Chapter 9 – *Conclusion*

This Chapter summarises and concludes the thesis and outlines potential future work.

Chapter 2

Background

2.1. Introduction

This thesis comprises several published, or soon to be submitted, research papers. As a result, the background information provided within each results Chapter (i.e., paper) (Chapters 5 to 8) is limited by the specific journal formats. The purpose of this Chapter is to elaborate on the background information, particularly the geological, climatic, and oceanographic history of New Zealand, relevant to the subsequent Chapters and the research questions outlined in Chapter 1.

This Chapter is split into three sections. The first and second provide a review of New Zealand's geological, climatic, and oceanographic history throughout the Cenozoic. The third section provides a summary of previous research on New Zealand's Cenozoic shallow-marine fossil record. Lastly, a synthesis of the above information is provided in the context of the narrative of this thesis, including a summary of the key events discussed in sections 1 and 2 (Table 2.1 – included at the end of this Chapter).

2.2. Cenozoic Geological History of New Zealand

New Zealand hosts a remarkably complete Cenozoic stratigraphic and fossil record, regarded as the most complete shallow-marine fossil record in the southern hemisphere (Cooper et al. 2006, Crampton et al. 2006a). The quality of these records can be attributed to New Zealand's unique tectonic history and setting. However, New Zealand's geological history is also intertwined with the evolution of its highly endemic terrestrial and marine faunas (Cooper and Cooper 1995). As a result, understanding the geological history of New Zealand is important in quantifying the drivers of biodiversity patterns through time. This section focuses on the geological evolution of New Zealand through the Cenozoic and examines the geological events and processes that contributed to both its exemplary stratigraphic and fossil records, and unique marine biota.

Importantly, New Zealand's geological timescale is not the same as the international geological timescale (Fig. 2.1). For clarity, in this Chapter I predominantly use the international timescale at the resolution of epochs to discuss events of a global nature (see Fig. 2.1, "International units"). For specific timings and regional events I also refer to the New Zealand units at the resolution of stages (see Fig. 2.1, "NZ units") (after Raine et al. 2015).
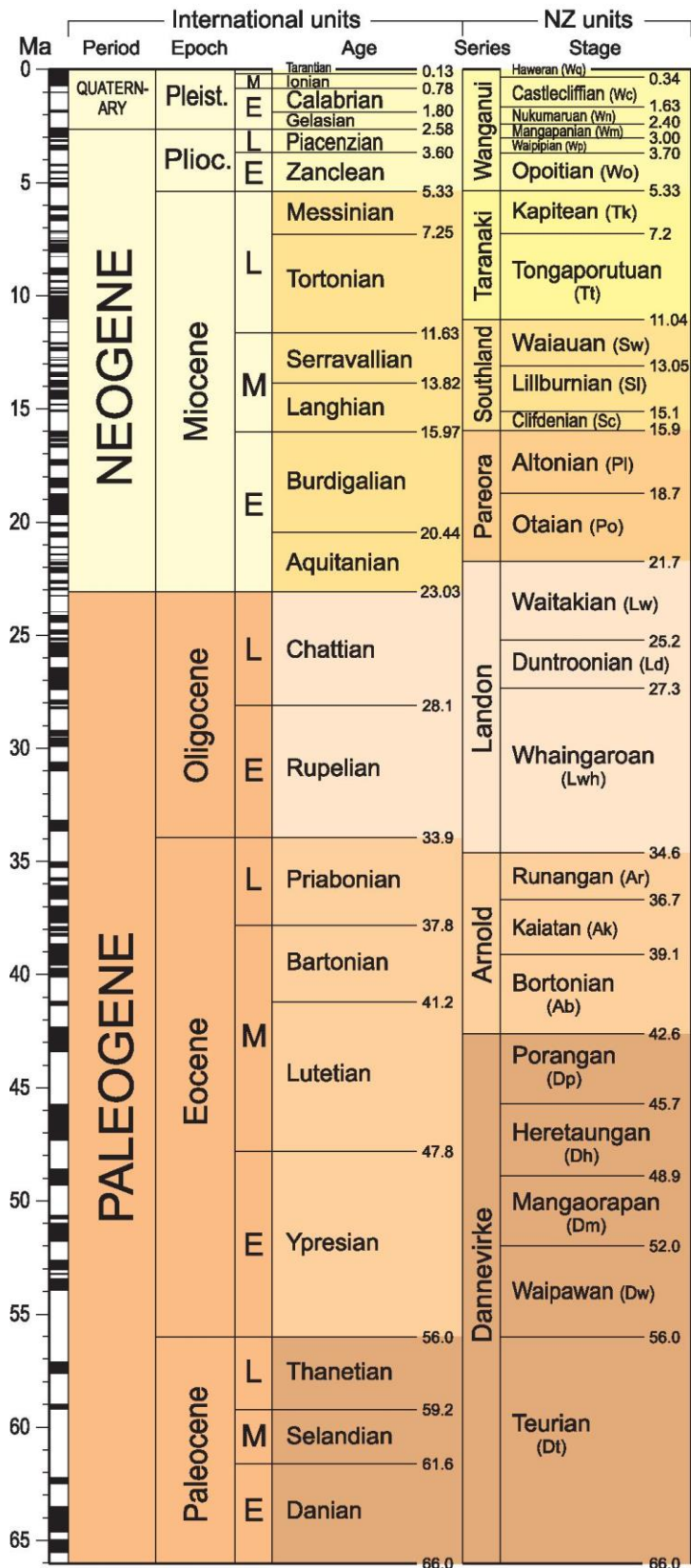
FIGURE 2.1. New Zealand geological timescale after Raine et al. (2015). Ages provided are in millions of years ago (Ma).

2.2.1. Geological Overview of Zealandia

The landmass of New Zealand is situated on the continent of Zealandia, now straddling the Pacific and Australian tectonic plates, separated by the modern plate boundary that is marked by the Alpine Fault (Mortimer et al. 2017) (Fig. 2.2). Zealandia was originally part of Gondwana, separating in the Late Cretaceous during the opening of the Tasman Sea. Presently, Zealandia's continental mass is approximately 94 % submerged, and although this proportion has fluctuated during the Cenozoic, Zealandia has remained largely submerged throughout its Cenozoic history (Mortimer et al. 2017). Following Zealandia's separation from Gondwana, the continent drifted (predominately northwards) at relatively static southern mid-latitudes until arriving at its modern-day position (with the majority of New Zealand's main landmass between ~35°S and 50°S). As a result, New Zealand has been relatively geographically isolated for the duration of the Cenozoic. This is considered to have limited migration to and from New Zealand, allowing the evolution of its highly endemic marine and terrestrial biota (Cooper and Cooper 1995, Spencer et al. 2009, Spencer et al. 2017). Whereas Zealandia was tectonically relatively quiescent during the early Cenozoic, tectonic activity intensified after the Eocene, marked by several pulses of increased tectonic activity (King 2000, Crampton et al. 2006b), discussed in detail in the subsequent sections. Perhaps the most notable phases of increased tectonic activity are the development of the Alpine Fault in the Late Oligocene/Early Miocene and the sharp change in the Australian-Pacific plate relative motion at the beginning of the Pliocene, leading to increased convergence and uplift along the Alpine Fault (King 2000). Although several processes are responsible for the exposure of the New Zealand's Cenozoic stratigraphy, the recent increase in uplift (5.33 Ma to recent) has exposed young (including < 0.34 Ma) Plio-Pleistocene marine sediments above sea-level. Consequently, New Zealand's marine stratigraphic and

fossil records extend to the recent, an attribute that is uncommon in the global marine fossil record.
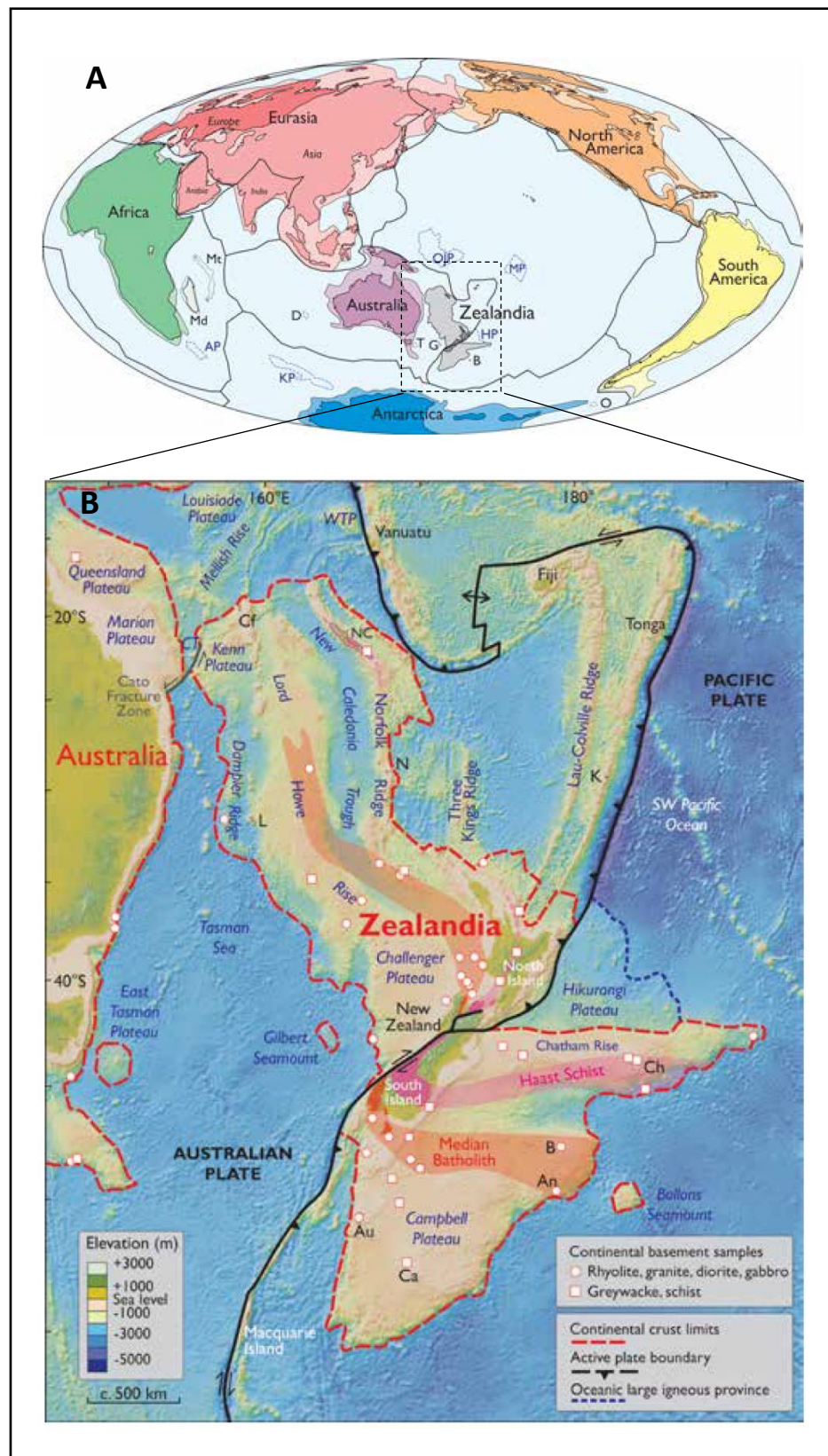
FIGURE 2.2. Modern global (A) and regional (B) maps of Zealandia. (A) Global map, showing New Zealand's position on the largely submerged continent of Zealandia, straddling the Australian and Pacific plates. (B) Regional map of New Zealand highlighting several structural and bathymetric features. The map highlights the transform fault boundary, known as the Alpine Fault, which joins the active plate boundaries between the subducting Australian Plate (beneath the Pacific Plate) in the south west and the subducting Pacific Plate (beneath the Australian Plate) in the north east. Figures adapted from Mortimer et al. (2017).

2.2.2. Cenozoic Stratigraphic Architecture

This thesis focuses on the shallow-marine fossil record of mainland New Zealand (North and South Island, and smaller islands in proximity, e.g., the Chatham Islands). Throughout the Cenozoic, sedimentation in New Zealand has been focused within several major sedimentary basins (Fig. 2.3): Northland basin, Taranaki Basin, East Coast Basin, Wanganui Basin, West Coast Basin, Canterbury Basin and Westland Southland Basin. The composite Cretaceous-Cenozoic stratigraphic succession compiled from these basins comprises a first-order transgressive-regressive 'megasequence' (King et al. 1999, King 2000). Within this megasequence there are several second-order unconformities that can be correlated across the basins, forming seven distinct unconformity-bound second-order cycles (King 2000) (Fig. 2.4). Whereas several of the second-order sequence boundaries line-up with base-level shifts in the global sea-level curve, these boundaries are considered to largely reflect local tectonic phases (discussed below), rather than eustasy (King et al. 1999). Therefore, large-scale sedimentation patterns in New Zealand are, in part, independent from the global sea-level curve. These first- and second-order tectono-stratigraphic cycles provide an ideal framework to sequentially examine New Zealand's geological history. Importantly,

17

these cycles represent a composite sequence, i.e., it is not vertically continuous, and many of the second-order cycles directly overlie metamorphosed Jurassic/Triassic age basement rocks.
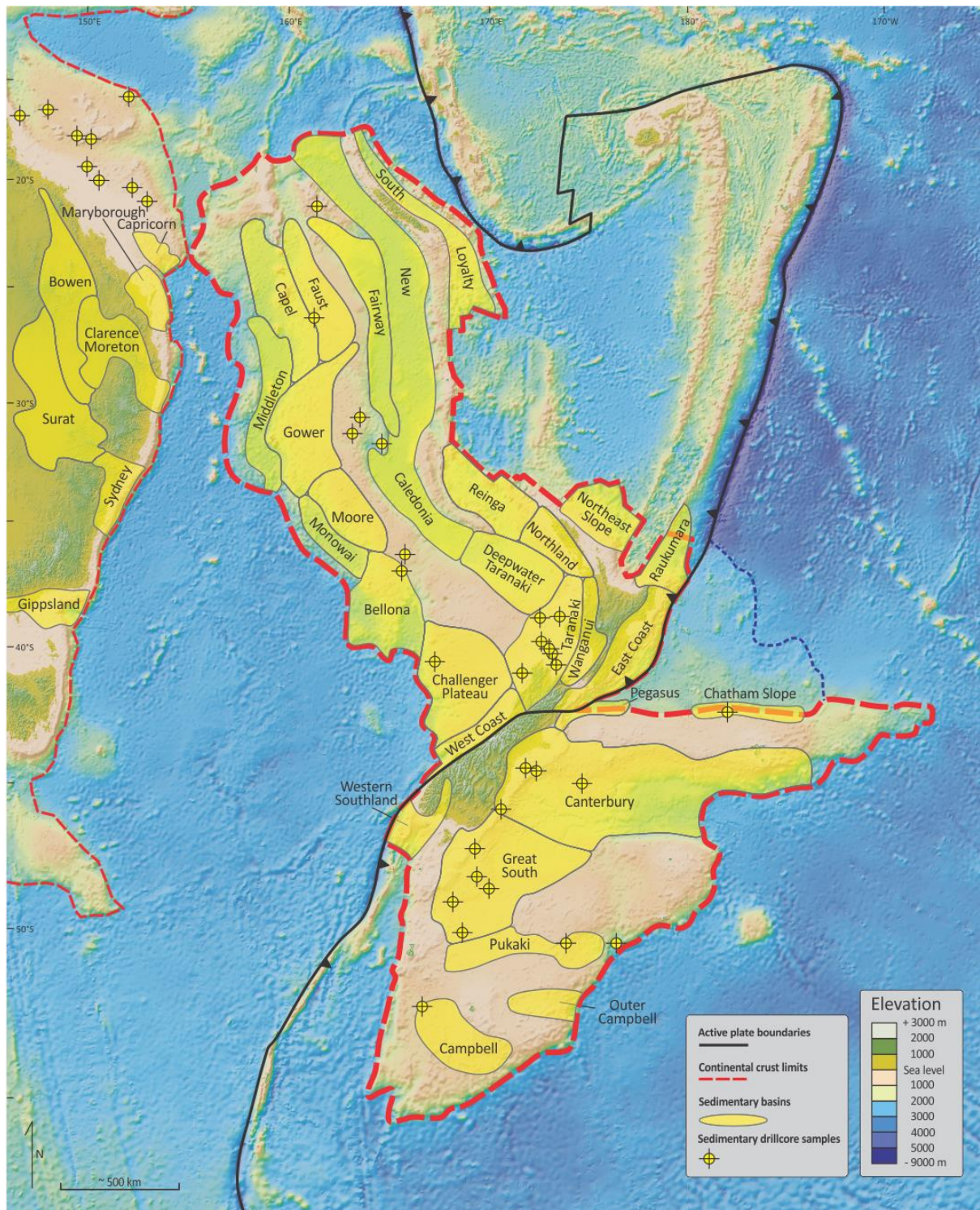
FIGURE 2.3. Modern regional map of Zealandia highlighting major sedimentary basins.

Figure from Mortimer and Campbell (2014).

Timescale | Global eustatic sea-level curve | New Zealand-wide Cretaceous-Cenozoic megasequence (1st order, 100 Ma)

Timescale

Global eustatic sea-level curve

New Zealand-wide Cretaceous-Cenozoic megasequence (1st order, 100 Ma)

Tectono-stratigraphic cycles

Relative base-level

1st-order sequence stratigraphy

2nd-order cycles

(Available marine accommodation)

Fall ... Rise

Ma | NZ stages

0

Wc — Pleistocene
Wn
Wm / Wo — Plio-cene
Wo
Tk
Tt
MIOCENE
Sw
Sl
Sc
Pl
Po
Lw
Ld
OLIGOCENE
Lwh
Ar
Ak
Ab
EOCENE
Dp
Dh
Dm
Dw
Dt
PALEOCENE
Mh
CRETACEOUS
Mp
Rt
Rm
Ra
Cn

0 100 200 300 400 (m)

1st and 2nd order curve

⑦
⑥
⑤

Highstand / regressive systems tract

Downlap surface

--- DLS ---
Maximum marine flooding of NZ subcontinent

Condensed section

④

Composite unconformity: possible erosion (currents, uplift, eustatic fall), or sediment starvation and stratigraphic condensation

Marshall paraconformity

--- RFS ---
Regional marine flooding surface

③

Regional onlap of basement

Erosion in Taranaki and Great South Basins

Transgressive systems tract

②

Passive margin subsidence in east and extension and basin formation in the west

Regional onlap of basement begins

West ... Coast ... Basins.

▲ Raukumara

①

East Coast basins

Distal hiatus beyond western foredeep

MODERN PLATE BOUNDARY — Dextral transpression and oblique subduction

PROTO-PLATE BOUNDARY — Emerald Basin spreading / continental rifting

Transition

SEA FLOOR SPREADING BETWEEN NZ AND GONDWANA — Continental drifting and rifting

GONDWANA BREAK-UP — Intracontinental rifting

20

FIGURE 2.4. Summary of the Cenozoic (and early Cretaceous) stratigraphic succession of New Zealand. The figure depicts a comparison between the global eustatic sea-level curve (after Haq et al. 1987) with the 1st-order and 2nd-order stratigraphic cycles of New Zealand (after King 2000). Timescale shown on the left is in Ma. Figure from King (2000) . *Pleistocene boundary was ratified after this publication and is now located at 2.58 Ma (Gibbard et al. 2010).

Cycle 2: Separation from Gondwana (Late Cretaceous – Paleocene).— The basal sequence boundary of the oldest first-order sequence (which contains Cycle 1 of the second-order sequence) (~100 Ma) is regarded as representing the breakup of Gondwana and intracontinental rifting. This marks the beginning of the transgressive sequence tract (TST) within the first-order megasequence. Here we skip to the second second-order sequence boundary (~80 Ma) that marks the beginning of the separation of Zealandia from Gondwana and the initiation of spreading of the Tasman Sea (King 2000). In general, deposition in the west is associated with syn-rift extensional basins, dominated by terrestrial and shallow-marine facies. Deposition in the east is associated with passive margin sediments, dominated by transgressive shallow-marine to bathyal sedimentary environments (King 2000). The terrestrial pollen record is relatively rich within this sequence (Prebble et al. 2021), however, the shallow-marine fossil record is comparatively sparse and concentrated along the eastern margin of New Zealand (Beu and Maxwell 1990), particularly in the Canterbury Basin. Similarly, the fossil records of other marine groups are also relatively poor (Hayward et al. 2009, Schwarzhans 2019). Despite this, the molluscan faunas are globally important as they represent just a small number of global molluscan faunas known from the circum-Pacific during the Paleocene (Beu and Maxwell 1990).

Cycle 3: Transition and Formation of the Proto-Plate Boundary (Latest Paleocene –
Eocene).— The early Eocene marks the end of the Tasman Sea spreading (~52 Ma) and the
cycle is characterized by a period of relative tectonic quiescence as New Zealand drifted
northwards away from Antarctica. Notwithstanding this, the cycle includes the formation of
the proto-plate boundary in the north (subduction zone) and a rift to the south, creating the
Emerald Basin (King 2000). In general, the stratigraphy is characterized by a progressive
onlap of marginal-marine and shelf facies across coastal plain regions (King 2000). Similar to
cycle 2, the Early and Late Eocene shallow-marine faunas are generally sparse and restricted
to the south east of New Zealand (Canterbury Basin). However, the Middle Eocene shallow-
marine faunas are much better represented in the fossil record, and are more extensive
spatially, although still predominately focused within the Canterbury Basin (Beu and
Maxwell 1990). The increase in fidelity of the Middle Eocene fossil record is likely related to
a higher order regression-transgression sequence, depositing fossiliferous shallow-marine
siliciclastic (and limited carbonates) sediments (Beu and Maxwell 1990).

Cycle 4: Maximum Flooding of Zealandia (Oligocene – earliest Miocene).— Cycle 4
represents the condensed section of the first-order megasequence and the maximum flooding
of the proto-New Zealand landmass (King 2000). The cycle records a transition from tectonic
quiescence characterized by slow regional subsidence, to the start of local tectonism
associated with continued development of the proto-plate boundary (King et al. 1999). The
early Oligocene section (Whaingaroan Stage) of the cycle contains numerous unconformities
(e.g. Marshall paraunconformity), many of which still have unknown origins (King 2000). In
general, sedimentation was carbonate-dominated across most of New Zealand, with minimal
clastic supply. The early Oligocene shallow-marine fossil record is relatively poor,
particularly considering the length of the Whaingaroan Stage (7.3 Myrs). This is likely a
result of the erosive nature of the cycle during this stage. Collections from the dominant

22

carbonate facies are generally sparse and represented by few genera, and collections from areas with higher terrigenous input are limited. In contrast, the Late Oligocene (Duntroonian and Waitakian Stages) shallow-marine faunas are both diverse and laterally extensive (Beu and Maxwell 1990).

Cycle 5: Re-emergence of New Zealand and the Modern Plate Boundary (Early to Late Miocene).— The basal sequence boundary of Cycle 5 marks the downlap surface (a marine flooding surface, prior to an increase in the rate of progradation) of the first-order megasequence and the beginning of the regressive portion of the megasequence (highstand/regressive systems tract) (King 2000). This is largely due to a shift in the regional tectonic regime associated with the development of the Alpine Fault, leading to increased uplift and erosion. This in turn led to a shift from carbonate dominated sedimentation to primarily terrigenous, yielding much thicker sedimentary sequences. The shallow-marine fossil record is well-preserved during this cycle and the record is laterally extensive, including sites from far north of New Zealand (Northland). This is mirrored in the fossil record of other marine groups (Hayward et al. 2009, Schwarzhans 2019) and the terrestrial pollen record (Prebble et al. 2021).

Cycle 6: Southern Subduction and Continued Uplift (Late Miocene).— The sequence boundary between cycles 5 and 6 marks a distinct tectonic pulse that caused a shift in tectono-sedimentary patterns (King et al. 1999). This shift is attributed to a change in the plate boundary configuration (King et al. 1999), indicated by the initiation of the Puysegur Trench in the south, culminating in the uplift and erosion in the proximal margins of many of the depositional basins (King 2000). In general, sedimentation was relatively similar to cycle 5, i.e., dominated by terrigenous sediments, but was generally coarser grained (King et al. 1999), and similarly, the shallow-marine fossil record is generally well-preserved and laterally extensive (Beu and Maxwell 1990).

Cycle 7: Uplift and Glacio-Eustacy (Pliocene to Recent).— The sequence boundary between cycles 6 and 7 marks another distinct tectonic pulse, caused by a sharp change in the relative motion of the Australian and Pacific plates (King 2000). This culminated in an increased tempo of uplift, erosion, and clastic sedimentation (King 2000). The shallow-marine fossil record is generally very well-preserved during this cycle, and laterally extensive, although collections from the Pleistocene shallow-marine fossil record are generally focused within the East Coast and Wanganui basins.

Cenozoic Paleogeography. — Whilst New Zealand has remained at relatively static southern mid-latitudes for much of the Cenozoic, there are some general paleogeographic trends that are not discussed above. Firstly, Zealandia has drifted slowly northwards since the breakup of Gondwana ~80 Ma, apart from a phase of primarily westward drift during the Eocene (cycle 3). Secondly, Zealandia has rotated broadly clockwise throughout the Cenozoic (although the Pacific and Australian plates have rotated independently). As Zealandia rotated, the latitudinal range of the continent and the emergent landmass of New Zealand has increased (see Fig. 2.5 for paleographic maps throughout the Cenozoic).
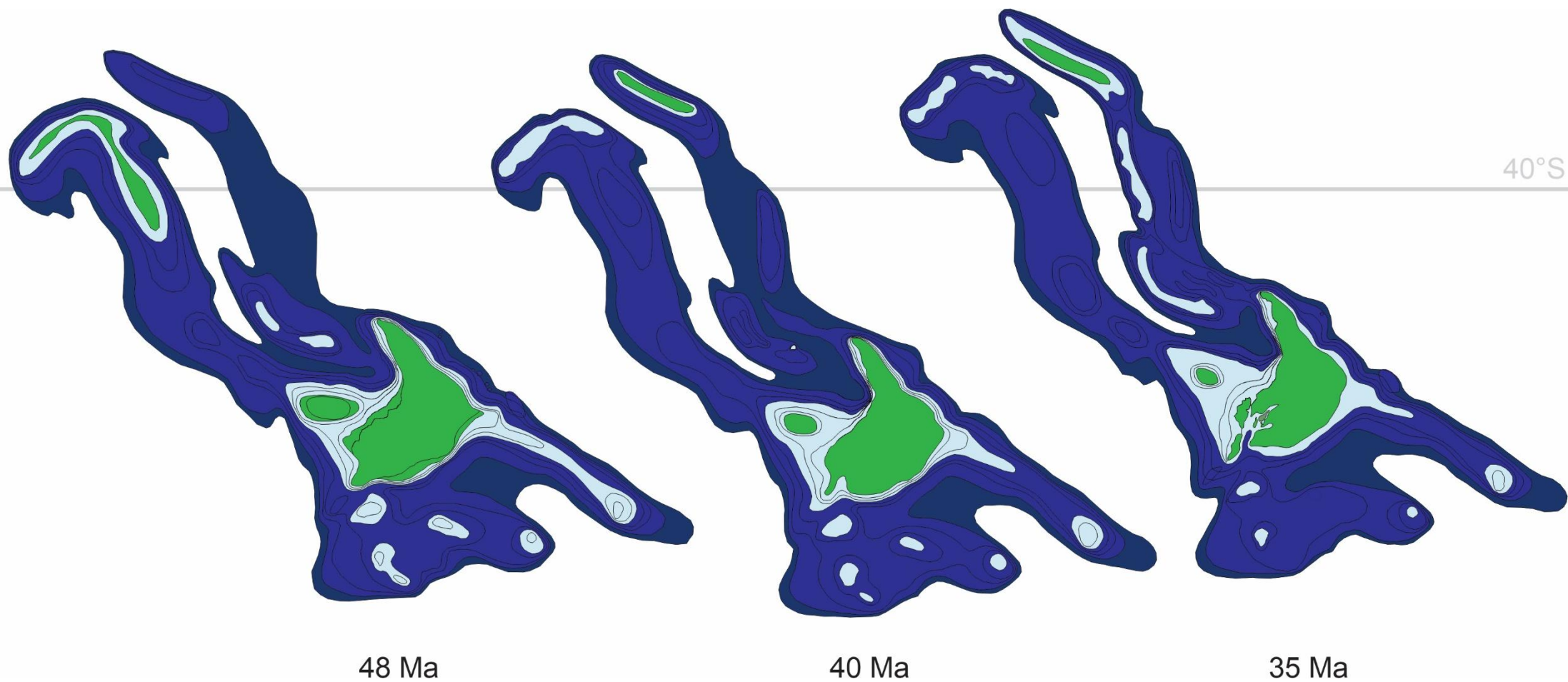
	48 Ma	40 Ma	35 Ma

40°S

KEY

= Landmass

= Inner-outer shelf (0-200 m)

= Upper-mid bathyal (200-1000 m)

= Lower bathyal (1000-1500 m)

40°S

30 Ma

19 Ma

KEY

= Landmass

= Inner-outer shelf (0-200 m)

= Upper-mid bathyal (200-1000 m)

= Lower bathyal (1000-1500 m)

14 Ma

10 Ma

Present

Chatham Rise

Campbell Plateau

40°S

KEY
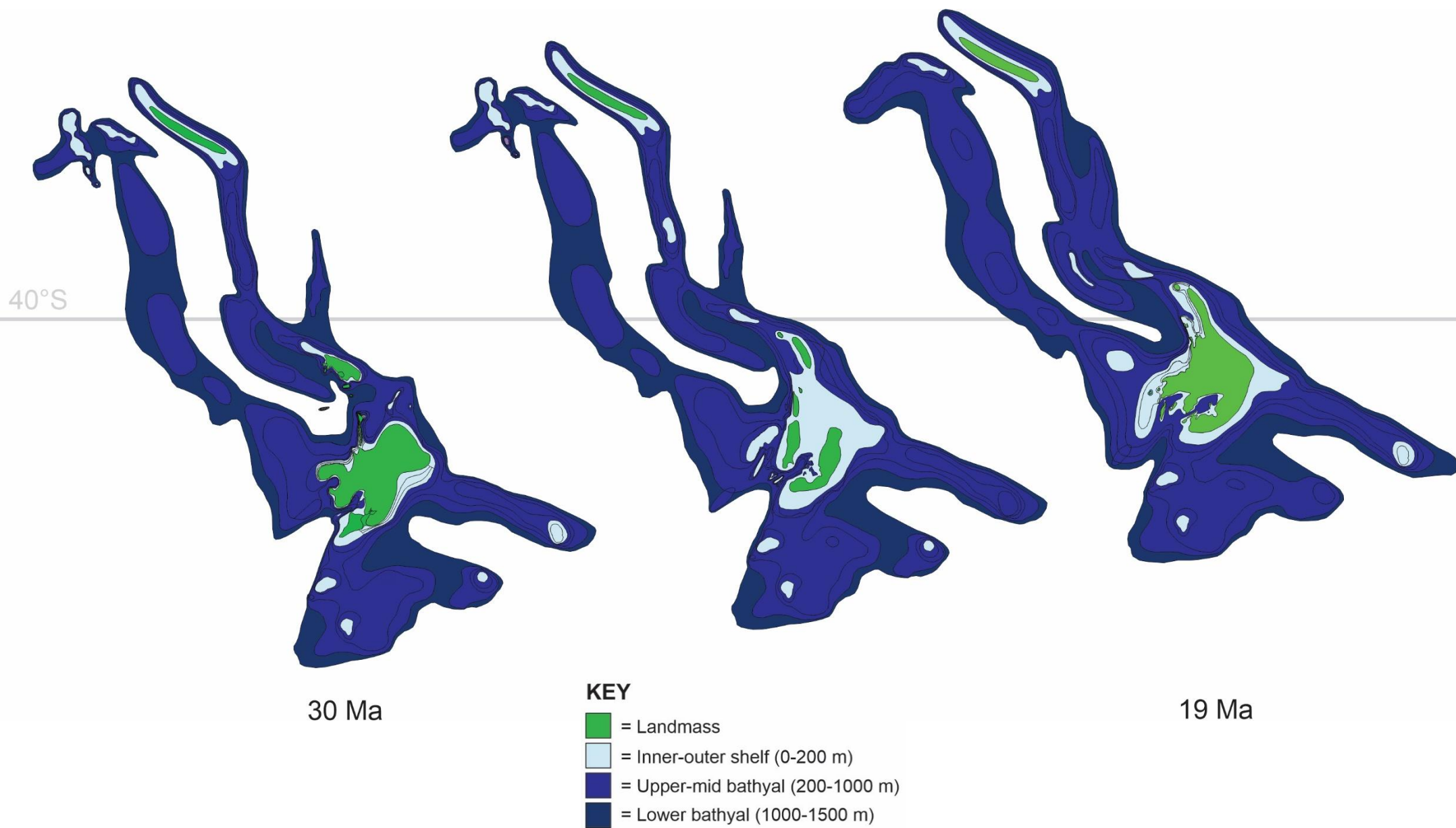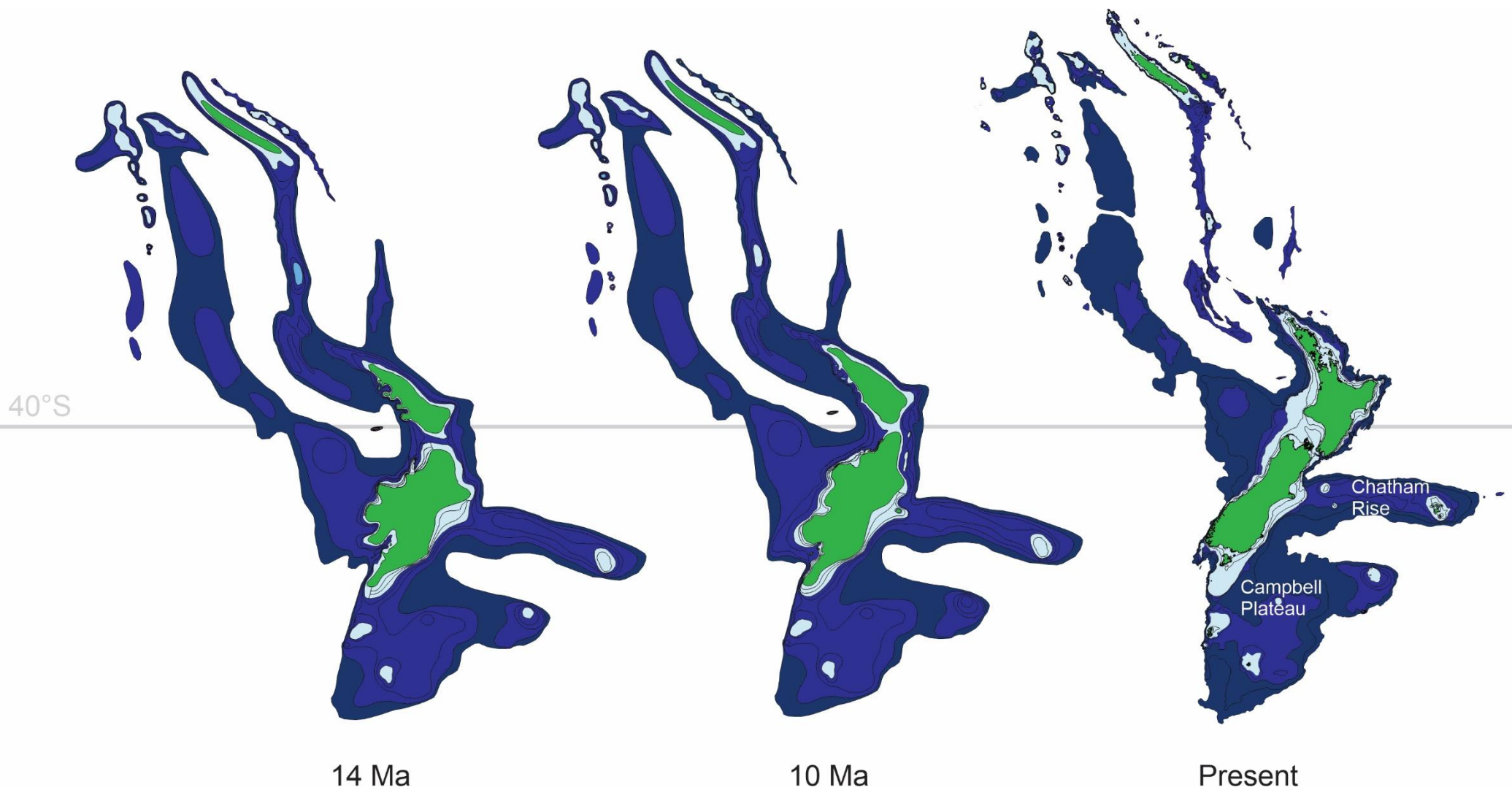
= Landmass

= Inner-outer shelf (0-200 m)

= Upper-mid bathyal (200-1000 m)

= Lower bathyal (1000-1500 m)

FIGURE 2.5. Paleogeographic maps of New Zealand through the Cenozoic (48 – 0 Ma). Maps are modified from paleogeographic maps provided by D.Strogen (GNS Science). For each series of three maps (i.e., each of the last three pages) the relative position of the 40°S latitude line is shown (grey line). Two important bathymetric features, the Chatham Rise and Campbell Plateau, are indicated in the present map of New Zealand for reference. The maps depict the clockwise rotation of Zealandia (although the Pacific and Australian plates have rotated independently), which increased the latitudinal range of the continent throughout the Cenozoic, and the general northward drift of Zealandia.

2.3. Cenozoic Climatic and Oceanographic History of New Zealand

The Cenozoic Era records major changes in global climate that included warm climatic optima and cooling events superimposed on an overall cooling trend, and large-scale changes in the distribution of the Earth's continents and oceans. These changes have clearly affected the history of life on Earth and shaped the spatial structuring of modern biodiversity. New Zealand currently spans several major climatic, oceanographic (Nelson and Cooke 2001), and biogeochemical boundaries (Reygondeau et al. 2013). The interplay between the development of these boundaries, shift in the global climatic state, distribution of global continents and oceans, and geological evolution of Zealandia through the Cenozoic has had a significant impact on New Zealand's regional climate and oceanography. This in turn has impacted New Zealand's shallow-marine ecosystems, particularly through the origination, extinction and mixing of cool and warm shallow-marine faunas throughout the Cenozoic (Beu and Maxwell 1990). Importantly, the evolution of New Zealand's regional Cenozoic climate and oceanography is related to both regional events and processes, superimposed on top of large scale global and ocean-basin-wide trends. As a result, this subsection is subdivided into two parts: (1) a broad overview of the global Cenozoic climate, and (2) a discussion of the regional climatic and oceanographic events and processes relevant to New Zealand.

2.3.1. Global Cenozoic Climate

Broadly, the global Cenozoic climate can be subdivided into four climatic states: Hothouse, Warmhouse, Coolhouse and Icehouse (Westerhold et al. 2020) (Fig. 2.6). These states are identified from reconstructions using carbon and oxygen isotope records from deep-sea benthic foraminifera and are described individually below.
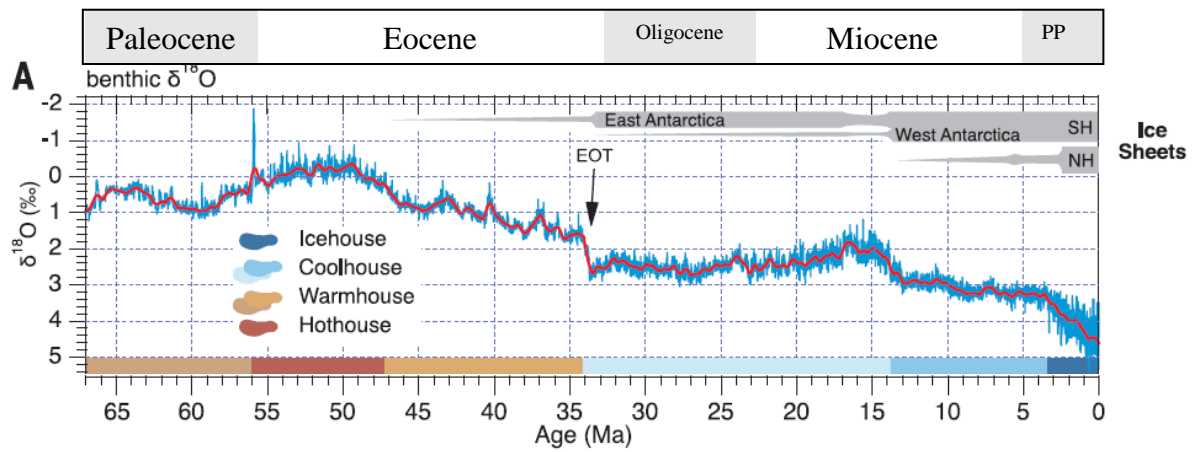
FIGURE 2.6. Global climatic states through the Cenozoic based on analysis of deep-sea foraminiferal δ^{18}O values. Climatic states include early to mid-Cenozoic Warmhouse state, punctuated by a Hothouse interval, before transitioning to a Coolhouse state in the Oligocene and Icehouse state in the late Pliocene. Development of southern hemisphere (SH) and northern hemisphere (NH) ice sheets shown along the top margin of the graph (grey). Figure modified from Westerhold et al. (2020). (PP = Plio-Pleistocene).

Warmhouse and Hothouse States (66 – 34 Ma).— The Cretaceous/Paleogene boundary marks the beginning of the Warmhouse and Hothouse climate states, which ended during the Eocene/Oligocene transition (~34 Ma). The Hothouse climate state is marked by the beginning of the Paleocene-Eocene thermal maximum (PETM) and persisted to the end of the Early Eocene climatic optimum (EECO) (56 – 47 Ma) (Westerhold et al. 2020). The Hothouse climate state is separated by two Warmhouse states, which were broadly similar and characterized by global temperatures > 5°C warmer than today (Westerhold et al. 2020). The Hothouse climate state marked a period of global temperature > 10°C warmer than today, characterized by greater amplitude variability in oxygen and carbon isotope values compared to the Warmhouse State, and numerous transient warming events (hyperthermals) (Westerhold et al. 2020). Comparison between ocean basins oxygen and carbon isotopes suggest that during the Hothouse and second Warmhouse state (Early Eocene), the ocean interiors switched from a homogenous to a heterogenous state, particularly between the

30

Southern Ocean and neighboring basins (Cramer et al. 2009). This is attributed to the separation of South American and Australia from Antarctica, leading to the opening of a partial to complete (there is equivocal evidence Australia had not fully separated from Antarctica - see Cramer et al. 2009 for further discussion) circumpolar passage in the Late Eocene.

Coolhouse State (34 – 3.3 Ma).— The Coolhouse climatic state marked a significant, stepwise decrease in global mean temperature, an increase in continental ice volume, and the appearance of major ice-sheets in Antarctica (Westerhold et al. 2020). The formation of major ice-sheets in Antarctica was particularly important as they have modulated the global climate since their inception (Westerhold et al. 2020). The Coolhouse state is broadly split into two phases, separated by the large increase in $\delta^{18}O$ at ~13.9 Ma during the Miocene climate transition, which was associated with the expansion of the Antarctic ice-sheets (Westerhold et al. 2020). During the first phase, prior to the Miocene climate transition, temperatures were generally warmer, particularly during the Miocene Climatic Optimum (~17 – 14 Ma). The second phase was cooler and includes a major transition in the carbon cycle ~7 Ma (Westerhold et al. 2020). Additionally, increased gradients between ocean basins of carbon and oxygen isotopes developed during the Oligocene. This is attributed to the development of the Antarctic Circumpolar Current (ACC) (discussed in detail in subsequent sections).

Icehouse State (~3.3 Ma – recent).— The Icehouse climatic state was driven by the appearance of fluctuating ice sheets in the northern hemisphere (Westerhold et al. 2020). Whilst evidence of ice sheets in the northern hemisphere precedes the Pliocene boundary, the Milankovich paced waxing and waning that characterizes the Icehouse state is thought to have been fully established globally in the region of Marine Isotope Stage M2 (3.3 Ma). Thus, Marine Isotope Stage M2 marks the approximate boundary between the preceding

Coolhouse state and the Icehouse state (Westerhold et al. 2020). Statistical analyses (recurrence plots based on high resolution isotopic data – see Westerhold et al. [2020] for further details) suggest that the Icehouse climate state is not comparable to any other climatic state during the Cenozoic (Westerhold et al. 2020).

2.3.2. Cenozoic Climatic and Oceanographic History in New Zealand

During the first Warmhouse and Hothouse climatic states (66 - ~45 Ma), Zealandia was bathed in warm and cool subtropical surface waters transported from the north west to east as part of a South Pacific warm-water gyre (Nelson and Cooke 2001). There is no evidence for the presence of major oceanic fronts across Zealandia, but a gradient between warm subtropical waters in the north and cool subtropical waters in the south was likely present (Nelson and Cooke 2001). The observed warm temperatures from deep-ocean oxygen isotope records are also evident in the fossil record of New Zealand and shown by the presence of tropical shallow-marine molluscan faunas and hermatypic (reef building) corals (Squires 1958, Beu 1990, Beu and Maxwell 1990, Hornibrook 1992). During the Paleocene, Antarctica was likely ice free with no major oceanic flow between the continental masses of Australia and Antarctica (Nelson and Cooke 2001, Zachos et al. 2001).

By the Middle Eocene there is some evidence of oceanic flow between Australia and Antarctica (Tasman Passage) and the formation of ephemeral ice-sheets in eastern Antarctica (Nelson and Cooke 2001). However, evidence of a cool water flow (through the Tasman Passage) is restricted to the latest Eocene only (Nelson and Cooke 2001, Lyle et al. 2007, Cramer et al. 2009). There is also evidence of a proto-Subtropical Front (STF) forming south of Zealandia in the Late Eocene, separating cool-water and warm-water masses offshore from Antarctica (Nelson and Cooke 2001). The STF was likely fully established by the Oligocene/Eocene boundary, and was eventually displaced by the Antarctic Polar Front at a latitude of ~65°S sometime between the Oligocene/Eocene boundary and earliest Oligocene,

as the glaciation in eastern Antarctica reached sea-level (Nelson and Cooke 2001). Modern and subrecent (i.e., latest Pleistocene and Holocene) evidence indicates that the convergence of the STF and Subantarctic front (SAF) occurs over a wide zone (up to 500 km in latitudinal width), termed the Subtropical Frontal Zone (Bostock et al. 2015). For clarity, we follow the definitions within the paleoceanographic literature and discuss the STF and SAF as singular fronts. However, it is likely that the two converged in a similar frontal zone for much of the Cenozoic.

Perhaps the most long-lived and significant change in New Zealand's oceanography and climate can be related to the formation of the ACC. The ACC isolated Antarctica thermally and contributed to the development of distinct oceanic fronts, currents, and winds in the Pacific/Southern Ocean. The development of the ACC was restricted by two passages: (1) the Tasman Passage between Australia and Antarctica, and (2) the Drake Passage between South America and Antarctica. The exact timing of complete opening of the Tasman and Drake passages, and the subsequent development of the ACC is debated (Zachos et al. 2001, Livermore et al. 2005, Scher and Martin 2006, Toumoulin et al. 2020). Notwithstanding this, it is likely that all three events were gradational processes with a proto-ACC forming in the latest Eocene and the Tasman and Drake Passages becoming fully open, with consequent establishment of an unrestricted ACC, sometime in the Early Oligocene (Nelson and Cooke 2001). Development of the ACC likely had a significant ecological impact on the dispersal of Southern Ocean marine biota, particularly those with a planktonic larval stage. Prior to the opening of the ACC it is unlikely that widespread planktonic dispersal in the Southern Ocean was possible. Following the inception of the ACC there is evidence in New Zealand for the arrival of several molluscan genera from South America (Beu et al. 1997). However, despite strong evidence of global cooling during the Eocene/Oligocene transition (i.e., the transition to the Coolhouse climatic state) and a short-lived temperature minimum (Oi-1 Glaciation)

(Westerhold et al. 2020), there is no clear-cut evidence in New Zealand's molluscan fauna that support a dramatic cooling (Beu and Maxwell 1990). This could be a result of bias in the fossil record, particularly considering the erosive nature of stratigraphic successions during this stage (King et al. 1999), but could also be related to the southern position of the STF.

During the Early and Middle Miocene the STF migrated southwards, coinciding with the southern expansion of warm subtropical waters at the expense of cool subtropical and subantarctic surface flows that had been established during the Oligocene (Nelson and Cooke 2001). The southward expansion of warm subtropical water and increase in oceanic temperature is widely evidenced in the Early and Middle Miocene marine fossil record in New Zealand. Hermatypic corals are recorded from the Early Miocene, suggestive of subtropical to marginally tropical oceanic temperatures (Squires 1958, Hayward 1977). This is mirrored in the molluscan fossil record by the prevalence of warmer water genera and faunas (Beu and Maxwell 1990), and in the foram record by larger species that are typical of warmer waters (Nelson and Cooke 2001). In addition, there is evidence of increased homogeneity of the planktic foraminiferal faunas compared to previous periods in the Cenozoic, attributed to the continued widening of the Tasman Passage and the ACC. The exact timing of peak warmth during the Miocene Climatic Optimum in New Zealand is debatable, but it is estimated to be close to the boundary between the Early and Middle Miocene (centering on the Altonian Stage). This is based on evidence from the fossil record (nannofossils, foraminifera, and macrofossils) and the timing of the tectonic events that caused restricted flow through the Indonesian gateway (Nelson and Cooke 2001). The restricted flow through the Indonesian gateway, which forms a connection between the Pacific and Indian Oceans, is thought to have redirected warm equatorial water flows southwards into the Tasman Sea, perhaps controlling the timing of the Miocene thermal maximum in the southern Pacific Ocean (Nelson and Cooke 2001, Hall et al. 2011).

The end of the Middle Miocene marks the transition into the second phase of the Coolhouse climate state and the beginning of the Miocene climatic transition and global cooling. This transition is associated with the major development of ice-sheets in Antarctica, particularly in the west, and expansion and intensification of the ACC, attributed to widening of the Southern Ocean (Nelson and Cooke 2001). Whereas there is some evidence of a proto-SAF forming during the early and middle Miocene to the south of New Zealand, there is clear evidence of the SAF and STF in the Late Miocene (Nelson and Cooke 2001). The relative position of both fronts is thought to have been similar to their modern day position, with the STF tied to the Campbell Rise on the eastern margin of New Zealand, and the SAF across the southern edge of the submerged continental margin of Zealandia (Nelson and Cooke 2001). New Zealand's shallow-marine molluscan fossil record provides a similar narrative, with the progressive extinction of warm-water genera that appeared during the Early and Middle Miocene, culminating in a low diversity terminal Miocene fauna (Beu and Maxwell 1990).

At the beginning of the Pliocene (Opoitian Stage) there is evidence of a phase of transient warming, particularly from shallow-marine molluscan faunas, marked by the appearance of several warm-water taxa absent from the Late Miocene (Beu and Maxwell 1990). This is not represented in the deep-sea benthic foraminiferan oxygen isotope record, but is clear in the marine sea surface temperature (SST) estimates based on the organic U^k_{37} proxy (a proxy for temperature based on the ratio between alkenones created by haptophyte algae of different chain lengths) in the southern hemisphere (Herbert et al. 2016) (see Chapter 4, section 4.5.2). During the middle-late Pliocene (late Opoitian to Mangapanian stages) there is evidence of a strong regional cooling of subtropical water, with several strong cooler excursions (3.35 Ma, 3 Ma and 2.8 Ma) (Sabaa et al. 2004). The overall trend of this regional cooling is attributed to increased upwelling, although northwards migration of the STF is not

ruled out (Sabaa et al. 2004). This phase of cooling is also recorded in the shallow-marine molluscan fossil record by the progressive extinction of warm-water taxa at the end of each New Zealand stage toward the Pleistocene boundary (Beu and Maxwell 1990).

Dramatic cooling in New Zealand during the Pliocene coincides with the arrival of the cold-water scallop *Zygochlamys delicatula* at the beginning of the Nukumaruan Stage (Beu and Maxwell 1990). Whereas there is evidence that the Antarctic Polar Front, SAF and the STF migrated northwards during glacial intervals (Nelson and Cooke 2001, Hayward et al. 2012, Bostock et al. 2015), they likely maintained positions similar to their modern day locations (discussed below) during interglacial intervals (Nelson and Cooke 2001). As a result, New Zealand was exposed to periodic incursions of cool subantarctic waters throughout the Pleistocene (Hayward et al. 2012). This is reflected in the shallow-marine molluscan fossil record by the episodic latitudinal migration of cool and warm water taxa (Beu and Maxwell 1990, Beu et al. 2004, Beu 2012). There is strong evidence that this process is continuing today, as part of a longer term expansion as oceanic temperature continues to rise in the present interglacial period, with recorded occurrences of several molluscan tropical species (with planktotrophic larval stages) in northern Northland (Beu et al. 2004).

The current oceanographic setting of New Zealand is not too dissimilar from the interglacial intervals during the Pleistocene. The STF (or northern extent of the Subtropical Frontal Zone) is topographically constrained by the Campbell Plateau and Chatham Rise in the east (see Fig. 2.2, panel B) and follows the New Zealand coastline, transecting the Solander Trough and extending eastward to the south of Tasmania (Nelson and Cooke 2001, Bostock et al. 2015). The STF separates cool subtropical water to the north from subantarctic surface waters to the south. The SAF is topographically constrained to the southern margin of Zealandia, separating the circumpolar surface water to the south from the subantarctic surface

waters in the north (Nelson and Cooke 2001). Although, as discussed earlier the STF is not a singular front, but is marked by the upper and lower limits of the Subtropical Frontal Zone (Bostock et al. 2015) (Fig. 2.7). As a result, most of New Zealand's shallow-marine shelfal ecosystems are bathed in cool subtropical waters except for southern New Zealand. This is reflected in the modern day molluscan faunas in Southland and islands south of Southland within New Zealand's oceanic territory, which contain cool-water elements from subantarctic faunas (Beu 2012)
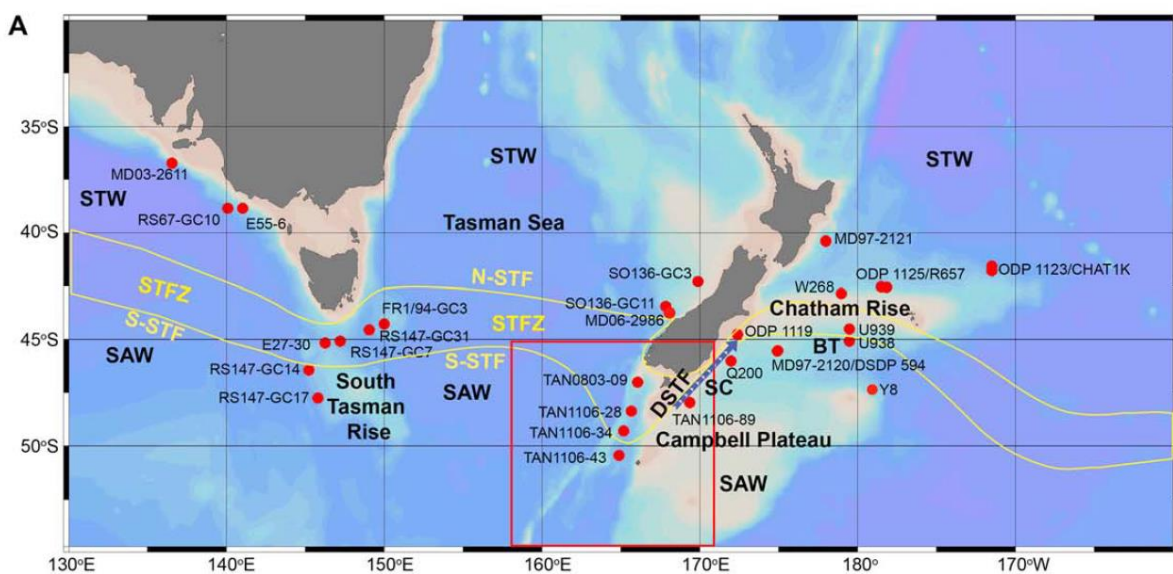
FIGURE 2.7. Oceanographic map of the modern southwest Pacific Ocean. The map shows the position of the Subtropical Frontal Zone (STFZ) which marks the convergence of the subtropical surface waters from subantarctic surface waters. Within the main text the paleoceanographic position of the Subtropical Front (STF) is simplified to represent an individual front. In the modern and subrecent history of New Zealand the Subtropical Front is marked by two boundaries the northern Subtropical Front (N-STF) and the southern Subtropical Front (S-STF). Figure from Bostock et al. (2015).

While the global climate reveals a story of progressive stepwise cooling from the Late Eocene onwards, New Zealand's climatic and oceanographic history is more complex and dynamic. This is largely due to its relatively static southern mid-latitude position throughout the Cenozoic, placing New Zealand at the interface between the Southern and Pacific oceans, and the evolution of oceanic fronts (predominantly the STF and SAF) and currents. This complexity, along with the lack of continuous regional oceanic temperature records, makes quantifying relative changes in New Zealand's oceanic temperature difficult. Despite this, early attempts at semi-quantitatively interpreting relative change in oceanic temperature around New Zealand based on fossil faunas (including molluscs) (Hornibrook 1992), shares striking similarities to the available geochemical proxies from New Zealand, the Pacific Ocean and global records (Nelson and Cooke 2001, Cramer et al. 2009, Herbert et al. 2016, Westerhold et al. 2020). Nevertheless, despite uncertainties, it is clear that variation in oceanic temperature has had a major impact on the evolution of New Zealand's shallow-marine ecosystems.

2.4. Cenozoic Shallow-Marine Biodiversity

Quantifying the relative importance of different driving mechanisms of biodiversity during the Phanerozoic has been the source of significant debate since the genesis of quantitative paleobiology. Recognition of the role of the fossil record in understanding the long-term impacts of anthropogenic climate change has bought this problem to the forefront of paleobiology. Although the fossil record provides our only means of understanding patterns of biodiversity change over geological timescales, the modern global distribution of species provides a wealth of information about the current spatial driving mechanism of modern diversity. This serves as an important analogue for how diversity may have been distributed in the past and how modern spatial patterns may manifest temporally. The Cenozoic Era is of particular interest for two reasons. First, most evidence from the fossil record points to a putative rise in global diversity during the Cenozoic relative to the rest of the Phanerozoic (Sepkoski et al. 1981, Alroy 2008, Close et al. 2020). Secondly, as previously discussed, the Cenozoic Era records major changes in global climate and large-scale changes in the distribution of the Earth's continents and oceans. These processes shaped the modern global distribution of species, and thus provide the best line of evidence of how modern diversity may respond in the future to large-scale environmental change. This may appear to be a question of global scale, but changes in diversity are hierarchical (Ricklefs 1987, 2004) and manifest at local and regional scale. Any truly global processes should be evident across numerous smaller spatial studies, whereas regional and local signals will be masked within global scale studies (Vermeij and Leighton 2003). Furthermore, the quality of data is much higher at regional spatial scales, particularly in the fossil record, and future conservation efforts are generally focused on regional spatial scales or smaller. As a result,

paleodiversity studies at the regional scale are critical for the prediction of the impact of large-scale environmental change in the future.

Within this section I focus predominately on the previous studies of New Zealand's Cenozoic shallow-marine molluscan record. These studies have informed this thesis and provide the basis for the subsequent results Chapters. To provide context, I also examine our understanding of the modern latitudinal diversity gradient (LDG), which serves as an important analogue for understanding temporal patterns of diversity change. Lastly, I provide a brief background on global scale paleodiversity studies. The aim of this section is not to provide a complete literature review of paleodiversity studies, which are simply too numerous, but instead it aims to provide a general background and supply suitable context for the results Chapters that follow.

2.4.1. Modern Latitudinal Diversity Gradient

The modern latitudinal diversity gradient is characterized by the distinct global gradient of high to low biological diversity from tropical to polar latitudes in both marine and terrestrial realms. Although recent evidence suggests that the modern LDG is largely a proxy for other explanatory variables (Gagné et al. 2020), i.e., omitting latitude as a factor per se, this is based on species richness only. For clarity, I discuss the LDG in relation to both the patterns observed in biodiversity and associated explanatory variables. The original definition of the LDG was based on species richness, however, recent research suggests that a similar gradient exists for functional richness (measured as the number of unique ecological functional groups – defined based on a common set of functional traits between species – see Chapter 7), and an inverse pattern for functional evenness (i.e., species become more unevenly distributed in functional groups from the tropics to the poles) (Edie et al. 2018b, Schumm et al. 2019) (Fig. 2.8). Whereas the exact timing of the development of the LDG is

still debated, it is largely accepted that it has existed since at least the Eocene/Oligocene boundary (~34 Ma) (Mannion et al. 2014, Crame 2020), and has prevailed intermittently from the mid-Paleozoic (Alroy et al. 2008, Erwin 2009). Furthermore, there is evidence in the marine realm that the steepness of this gradient and its latitudinal position has shifted through geological time as global climate has warmed and cooled (Crame 2000, Powell et al. 2015, Edie et al. 2018b, Song et al. 2020). This suggests the LDG is a long-lived ecologically, and biogeographically important phenomenon. As a result, the driving mechanisms associated with the spatial patterns observed across the modern LDG provide a potential analogue to understand how these variables may impact or explain temporal patterns in biodiversity at the regional scale. This is particularly important for New Zealand as it currently spans the latitudinal range where the largest changes in modern biodiversity are observed (Edie et al. 2018b, Schumm et al. 2019). On this basis, fluctuations in the position and steepness of the LDG/associated variables through time are likely to be an important factor in New Zealand's biodiversity history. Despite this, the mechanisms responsible for the LDG are still widely debated and unresolved (Hillebrand 2004, Mittelbach et al. 2007, Fine 2015, Jablonski et al. 2017, Edie et al. 2018b, Schumm et al. 2019).
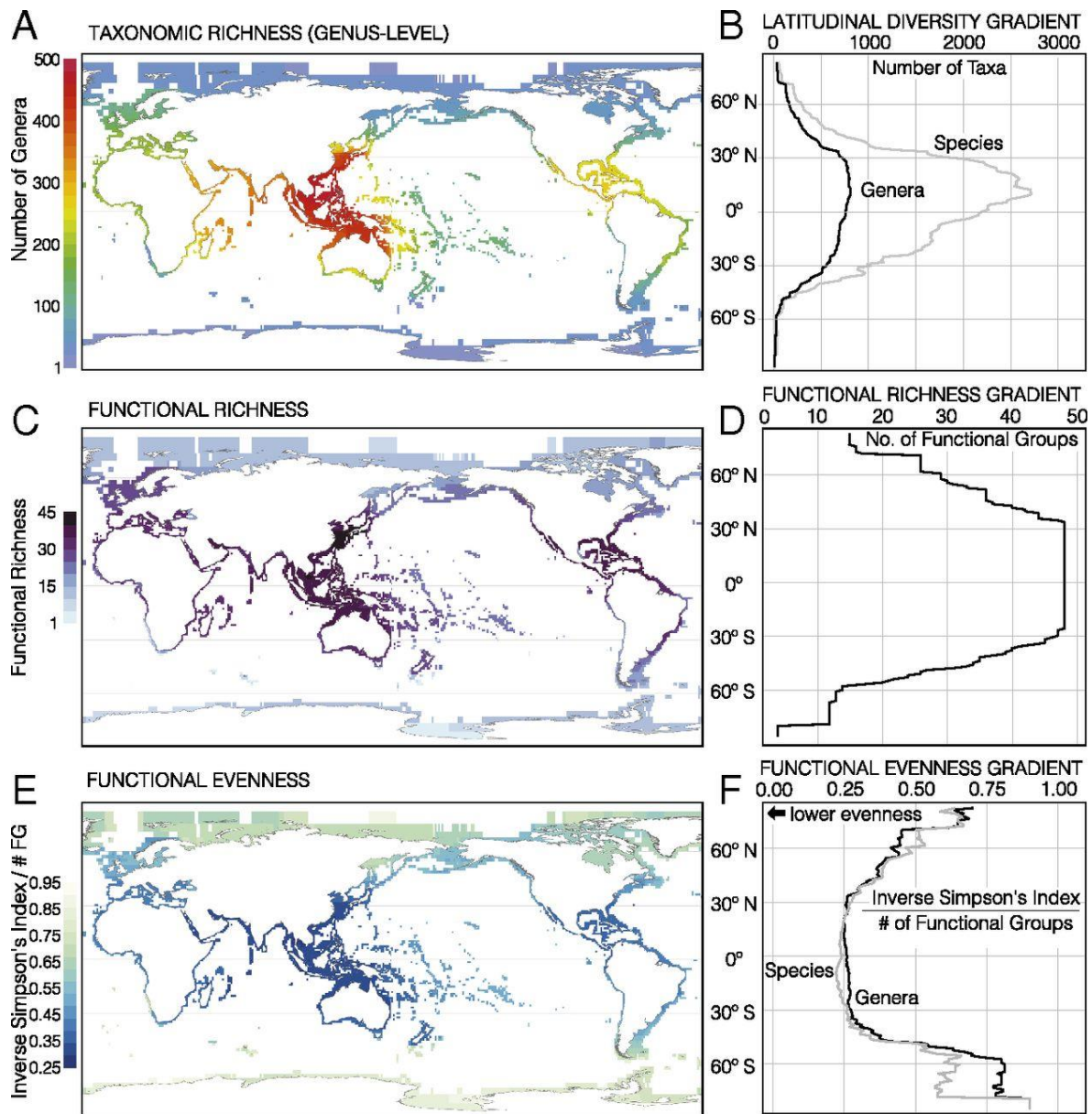
FIGURE 2.8. Global taxonomic richness, functional richness and evenness of modern bivalves. Both taxonomic and functional diversity (B and D) show a similar trend of high diversity in the tropics decreasing poleward. Functional evenness (F) shows an inverse pattern where evenness (the evenness of the number of species within unique functional groups) is lower in the tropics and increases poleward. An increase in evenness suggests that taxa (depicted as species and genera) are more evenly spread amongst functional groups at higher latitudes. Figure from Edie et al. (2018b).

Commonly invoked factors in the spatial distribution of diversity across modern marine ecosystems range from in-situ evolution, environmental heterogeneity (at different spatial scales), thermal and trophic settings, environmental and climatic stability, habitable area (extending to the size and structures of the ecosystems) and primary productivity (Edie et al. 2018b). However, perhaps the most convincing and empirically supported factors in the marine realm are the latitudinal gradients in climate and temperature, and the latitudinal asymmetry in the strength of seasonal climatic variability (Hawkins et al. 2003, Willig et al. 2003, Hillebrand 2004, Erwin 2009, Field et al. 2009, Archibald et al. 2010). In addition, recent research suggests that the most important environmental determinants on living marine biodiversity are water depth, average water temperature and sunlight intensity (Gagné et al. 2020). These climatic factors could facilitate higher taxonomic richness in the tropics through a number of mechanisms, e.g., increased energy availability, allowing greater populations sizes and therefore decreased extinction risk, increased exploitation of specialized niches, increased metabolic scope and faster origination rates and/or diminished extinction risk (Erwin 2009). These mechanisms have been used to support two prominent models in the origin of the LDG: (1) the LDG originated from in-situ local environmental factors, i.e., the tropics can essentially support a higher number of niches and species, or (2) is a result of spatial dynamics, e.g. preferential origination in the tropics and poleward expansion (e.g. the "out of the tropics" hypothesis and the idea that the tropics are an evolutionary cradle) (Jablonski et al. 2017). However, in the case of bivalves, both models are problematic based on modern observations (Jablonski et al. 2017). Firstly, in-situ models imply that standing richness in modern tropical regions should exceed that of temperate regions. However, Caribbean bivalve species richness is surpassed individually by the temperate faunas of the West Pacific (south east Australia) and western Pacific (south east Japan), despite the vastly larger shelf area of the Caribbean (Jablonski et al. 2017). Second, although there is evidence

that bivalves have a higher origination rate in the tropics (see Mittelbach et al. 2014 and

Jablonski et al. 2017), the majority of the clades that expand out of the tropics still maintain

their tropical presence (Jablonski 2005, Jablonski et al. 2013, Jablonski et al. 2017). As a

result, the tropics likely act as both a cradle and museum for diversity (Jablonski et al. 2017).

Jablonski et al. (2017) suggested that both models are implicated in the formation of the

modern LDG for bivalves, and this likely extends to other marine groups. New Zealand's

geographic isolation throughout the Cenozoic likely precludes the role of wholesale

immigration and migration processes (Crampton et al. 2006a). Notwithstanding this, it is

possible that both the in-situ and spatial dynamics models may act at a regional scale and

increased origination rates during warmer periods, and at lower paleolatitudes are plausible.

The distribution of molluscan marine biodiversity across the modern LDG provides

several *a priori* predictions for how molluscan biodiversity may have evolved through the

Cenozoic in New Zealand. Firstly, we may expect species richness to increase in New

Zealand in intervals where global and Pacific Ocean temperatures increased. This could be

facilitated by a poleward expansion of warmer oceanic waters associated with oceanic fronts

(particularly the STF) and associated gradients. Ecologically, this could be facilitated by: (1)

southward expansion in the geographic ranges of tropical species to New Zealand (2) and/or

existing warm-water species, and/or (3) increased origination rates in situ. Conversely, the

northward expansion of cool-water faunas and/or subantarctic migrants could also happen in

cooler periods with equatorial migration of cool-water species. However, there is no major

landmass between New Zealand and Antarctica, and based on evidence from the modern

LDG, we would anticipate these faunas to be low in diversity.

2.4.2. Global Cenozoic Diversity Patterns

The integration of biology and ecology in order to rigorously quantify and interpret global diversity patterns in the fossil record started in the late 19th century and provided a quantitative basis for future research. These studies largely focused on global patterns of marine diversity and are historically and scientifically well-documented and synthesized (for example see Sepkoski 2005, 2012). As a result, this thesis does not dwell on these early works and focuses instead on more recent studies (although see Chapter 5 preface). Notwithstanding this, these early works identified two distinct problems which are still currently contested: (1) is the observed global rise in Cenozoic diversity (relative to the earlier Phanerozoic) a true biological pattern, and (2) how does the spatial structuring of diversity accommodate changes in global taxonomic diversity, i.e. "where does all the diversity go" (Sepkoski Jr 1988).

Perhaps the most influential development in the 20th century has been the development of new methods of sampling standardization, particularly those that standardize to equal completeness ("coverage") rather than equal sampling size, herein referred to as coverage-based subsampling (also termed shareholder quorum subsampling and coverage-based rarefaction – see Chapter 4, section 4.3). This new method was developed and applied to the global marine fossil record by Alroy (2010a, b, c). Alroy (2010b) observed two prominent differences in the global Phanerozoic marine invertebrate fossil record once standardization to equal coverage was implemented. Firstly, there was no evidence of an interval of unregulated and sustained exponential growth through the Phanerozoic, as suggested in earlier research (Raup 1972, Lane and Benton 2003), particularly during the Cenozoic. Previous analyses had proposed a two- to three-fold increase in diversity in the latest Cenozoic (see Miller 2000). Second, that tropical genus richness through the Phanerozoic was relatively static, suggesting that most large shifts in the global diversity

45

curve reflect trends at non-tropical latitudes. Subsequently, Alroy (2010c) employed coverage-based subsampling to individual major marine groups rather than their combined diversity. Alroy observed that each major taxonomic group of marine invertebrates had characteristically different diversity trajectories through the Phanerozoic suggesting that each major taxonomic group had a consistent net rate of diversification and a limit to its species richness (Alroy 2010c). A similar result for the Phanerozoic global diversity curve for marine invertebrates was observed, consistent with the theory of a global limit to biodiversity (Alroy et al. 2008, Alroy 2010b, c). This is turn suggests that for the majority of the Phanerozoic, excluding major extinction episodes, global diversity was saturated (Alroy et al. 2008, Alroy 2010b, c). Notwithstanding this, the idea of bounded (i.e., a carrying capacity) versus unbounded diversity is still widely debated and contentious (Cornell 2013, Harmon and Harrison 2015, Rabosky and Hurlbert 2015, Storch and Okie 2019).

Importantly, the development of new methods of sampling standardization, along with refined methods of quantifying taxonomic rates (for example see Foote 2000, Alroy 2010a, 2014, Liow and Nichols 2010, Silvestro et al. 2014 – and see Chapter 4, section 4.8), have allowed the focus to shift from identifying patterns within the fossil record to interpreting them. As a result, global scale paleodiversity studies over the last decade have focused on quantifying the driving mechanisms of diversity, particularly the role of climate change and other environmental factors. Most of these studies have been rooted on the observations of environmental gradients along the modern LDG.

Based on empirical evidence that the modern LDG is driven by temperature and climate, Mayhew et al. (2012) examined the role of global temperature and marine macroinvertebrate macroevolution over the Phanerozoic, utilizing coverage-based subsampling and newly developed rates of extinction and origination (Alroy 2008). To measure variation in global climate, Mayhew et al. (2012) used proxies for atmospheric CO_2

and oceanic temperature using $\delta^{18}O$. Their key results included: (1) a broad pattern of diversity increase during warmer periods, (2) temperature is a strong predictor of taxonomic rates, particularly origination, alongside other abiotic and biotic variables, and (3) during mass extinction intervals, the temperature-diversity relationship can break down.

Notwithstanding the above, recent evidence questions the nature of the global fossil record and suggests it may be confounded by uneven spatial distribution of fossil collections through time, and thus, is not truly representative of global patterns (Close et al. 2020). Close et al. (2020) suggested that focus should be placed on regional and among-region diversity patterns rather compiling global data, and concluded that paleodiversity studies, particularly of global nature, need to be refocused. Nevertheless, after employing a spatially explicit approach to quantifying variation in global diversity through the Phanerozoic, they concluded that Cenozoic diversity of marine invertebrates was still approximately twice as high relative to the earlier Phanerozoic.

2.4.3. Previous Studies in New Zealand

The Cenozoic New Zealand mollsucan fossil record has been studied in detail, particularly in terms of taxonomy and biostratigraphy (Powell 1979, Beu 1990, Beu et al. 2004, Beu 2006, Beu and Raine 2009, Beu 2011, 2012, Spencer et al. 2017). Perhaps the earliest study to examine patterns of biodiversity was by Beu (1990), who quantified raw generic diversity through the Neogene, observing three important features. Firstly, Beu (1990) observed that the generic diversity curve correlates strongly to the trend of published New Zealand paleotemperature curves, with the exception of the Plio-Pleistocene, observing peak regional diversity during the Altonian Stage (i.e., during the Miocene Climatic Optimum) (Fig. 2.9). Notwithstanding this, the paleotemperature curve was derived from the presence of certain fossil faunas in New Zealand, including molluscs themselves, and thus the

study may have been influenced to some degree by circular reasoning. Second, it was observed that the most marked Neogene extinction was during the Late Miocene (Tongaporutuan Stage), attributed to a sustained period of cooling. Lastly, Beu (1990) derived estimates for the preservation of shallow water facies, recording a marked increase during the Plio-Pleistocene, attributing this as a causal factor to the observed increase in generic molluscan diversity, particularly during the Pleistocene.
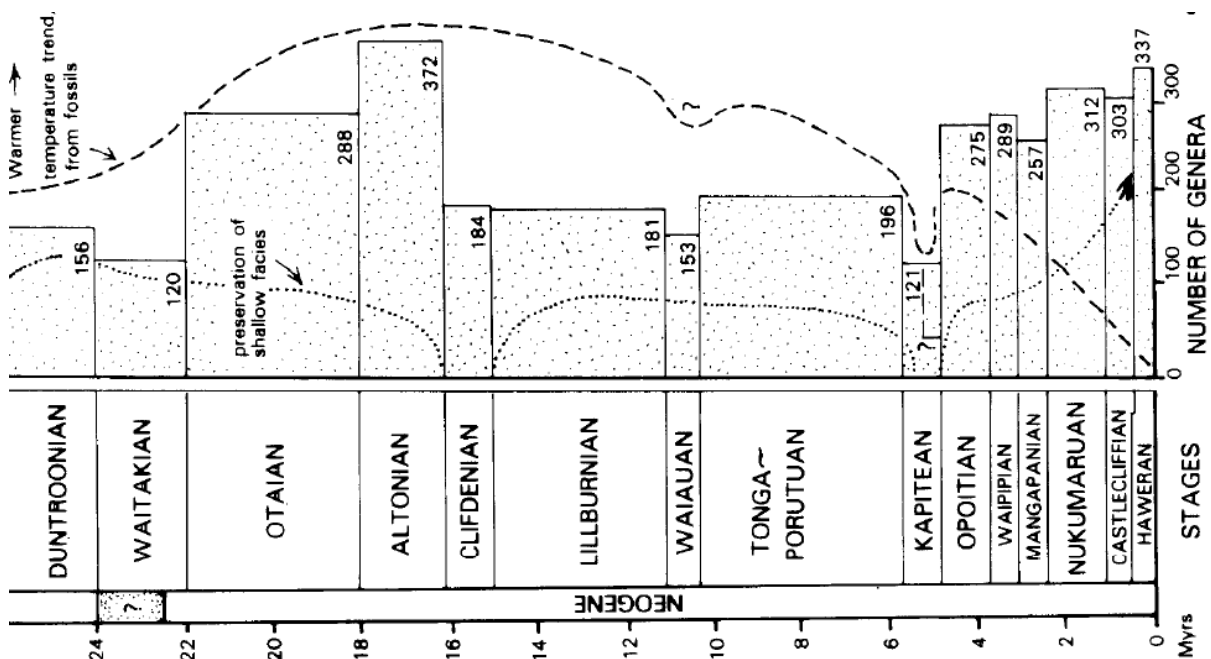
FIGURE 2.9. Generic-level marine molluscan diversity through the Neogene in New Zealand. Histograms depict the number of occurrences of individual genera recorded in each time bin (at the resolution of New Zealand stages), recording peak taxonomic diversity in the Altonian, during the global Miocene climatic optimum. Dashed line indicates paleotemperature in New Zealand estimated from fossils and dotted line represents the preservation of shallow-water facies. Figure from Beu (1990).

The early work quantifying patterns of biodiversity in New Zealand's shallow-marine fossil record was furthered by several papers in the early 20th century. These papers adopted quantitative methods from ecology and started to focus on the controls on the quality of the

48

fossil record. The earliest research focused on estimating the rock volume and effects of lithification in paleobiodiversity studies, using the Cenozoic shallow-marine mollsucan fossil record in New Zealand as a case study. Crampton et al. (2003) observed that outcrop area was likely to be a reliable proxy of rock volume in both stable and cratonic regions, and that outcrop volume was moderately correlated to the number of fossil collections, and the number of collections was moderately correlated to raw species richness. Hendy (2005, 2009) observed that lithification negatively affected raw generic richness, and as a result, the relative increase in unlithified fossil collections toward the recent may inflate apparent diversity trends, particularly within community (i.e., alpha) diversity. Furthermore, Hendy (2005, 2009) observed that the unlithified and lithified sediments preserved different pools of taxa that are compositionally distinct. Notwithstanding this, the comparisons and interpretations between lithified or unlithified faunas at similar time intervals were based on classical rarefaction, i.e., rarefied by sample size. As discussed in Chapter 4 (see section 4.3), this method of sampling standardization is itself confounded by sampling size and has been superseded by other methods based on sampling to equal completeness or extrapolation. Regardless, both the studies of Crampton and Hendy showed that raw uncorrected patterns of species and generic richness are likely confounded by both the quality of the rock record and lithification, and these factors need to be considered when quantifying measures of diversity.

Following this, two studies focused on quantifying the completeness (the proportion of species that have been sampled and recorded at least once in the fossil record) of the fossil record. Both studies came to similar conclusions, but using contrasting methodologies. Cooper et al. (2006) estimated completeness using the method of extant forms (Paul and Donovan 1998), whereby completeness was estimated by the proportion of the living fauna that has been collected as fossils. Conversely, Crampton et al. (2006b) estimated completeness based on the ratio between per-stage sampling and extinction rates. Both

studies recorded similar levels of completeness, quantified by Crampton et al. (2006b) as ~32 % for the Cenozoic and 41 % for the Neogene. These two studies quantitatively highlighted the exceptional nature of New Zealand's Cenozoic shallow-marine fossil record, establishing it as the most complete Cenozoic molluscan fossil record in the southern hemisphere. Notwithstanding this, the quality of the Cenozoic New Zealand fossil record was known qualitatively prior to this (and quantitatively in the Wanganui Basin, estimated to be up to 88 % complete – see Johnson and Curry [2001]). Crampton et al. (2006b) also focused on the impact of second-order sequence stratigraphic controls on the nature of the fossil record (discussed earlier in this Chapter). They observed that short-term patterns of variation in per-stage sampling probability reflect enhanced preservation in mid-cycle positions (Fig. 2.10). Whereas this is also observed within stable cratonic settings, Crampton et al. (2006b) suggested this may be more prominent in New Zealand due to its location on an active margin.
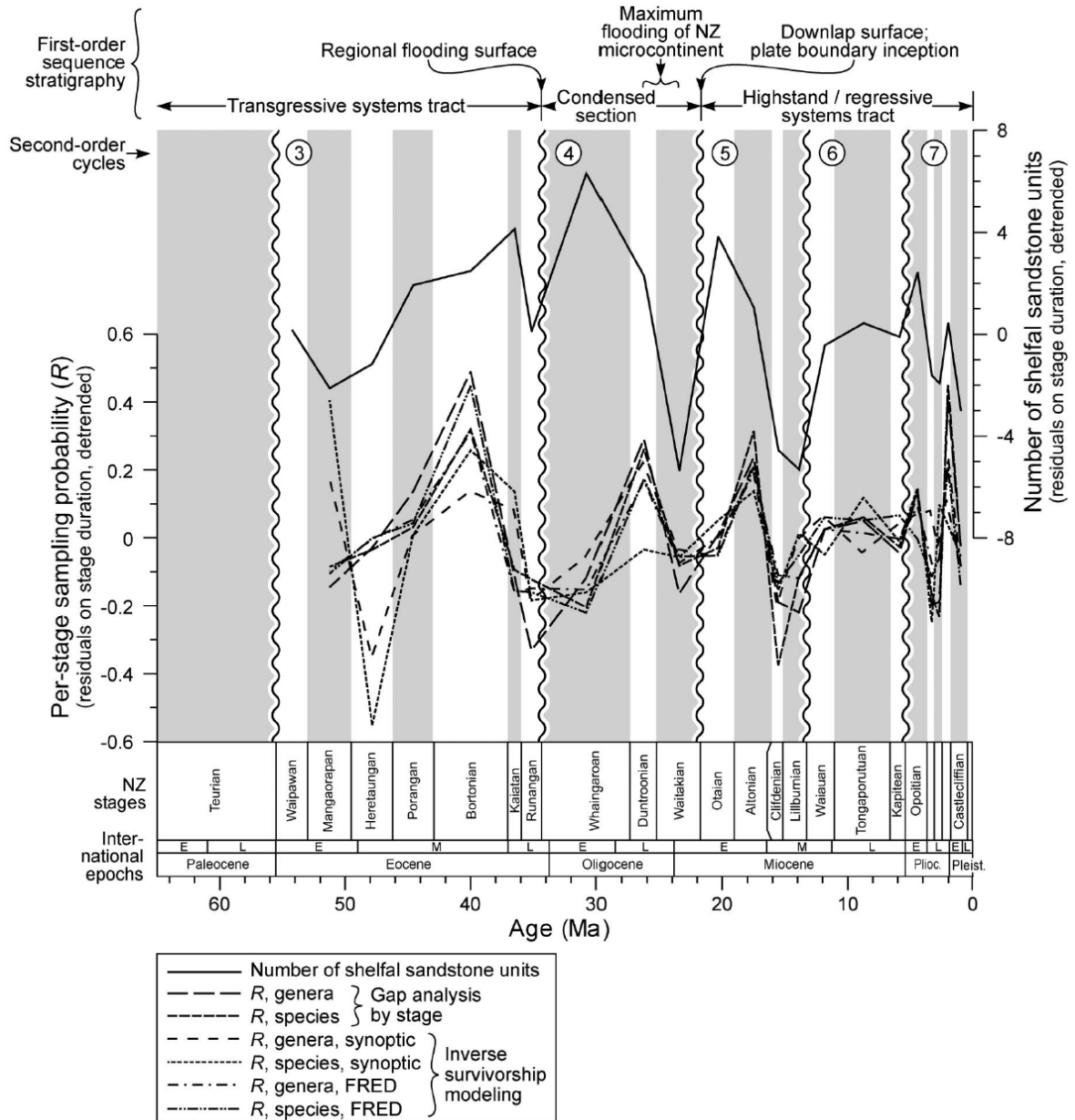
FIGURE 2.10. Plots of detrended per-stage sampling probability of shallow-marine molluscs and the number of shelfal sandstone units (detrended) through the Cenozoic. Vertical undulating lines show the second-order cycles identified by King et al. (1999). The figure shows clearly that per-stage sampling probability is enhanced in mid-cycle positions. Figure from Crampton et al. (2006b).

The first application of sampling standardization to the New Zealand Cenozoic shallow-marine fossil record was by Crampton et al. (2006a), and recorded contrasting results

to earlier works. This study synthesized the previous work on the quality of the New Zealand fossil record and evaluated several other taphonomic and systematic biases. These biases included the effect of non-uniform loss of aragonitic faunas, biostratigraphic range errors, taxonomic errors, and taxonomic rank. Crampton et al. (2006a) argued that once sampling standardization was implemented the resultant patterns were largely robust, to a first order, to these biases. The resultant sampling standardized regional scale molluscan diversity curve suggested that there was little increase in diversity through the middle and late Cenozoic and, in contrast, a decrease in diversity through the Plio-Pleistocene, particularly at the generic-level (Crampton et al. 2006a) (Fig. 2.11). Furthermore, Crampton et al. (2006a) observed no evidence for a species-area effect on diversity, or a relationship between temperature and diversity, but suggested that both should be re-examined once refined estimates of shelf area and temperature become available.
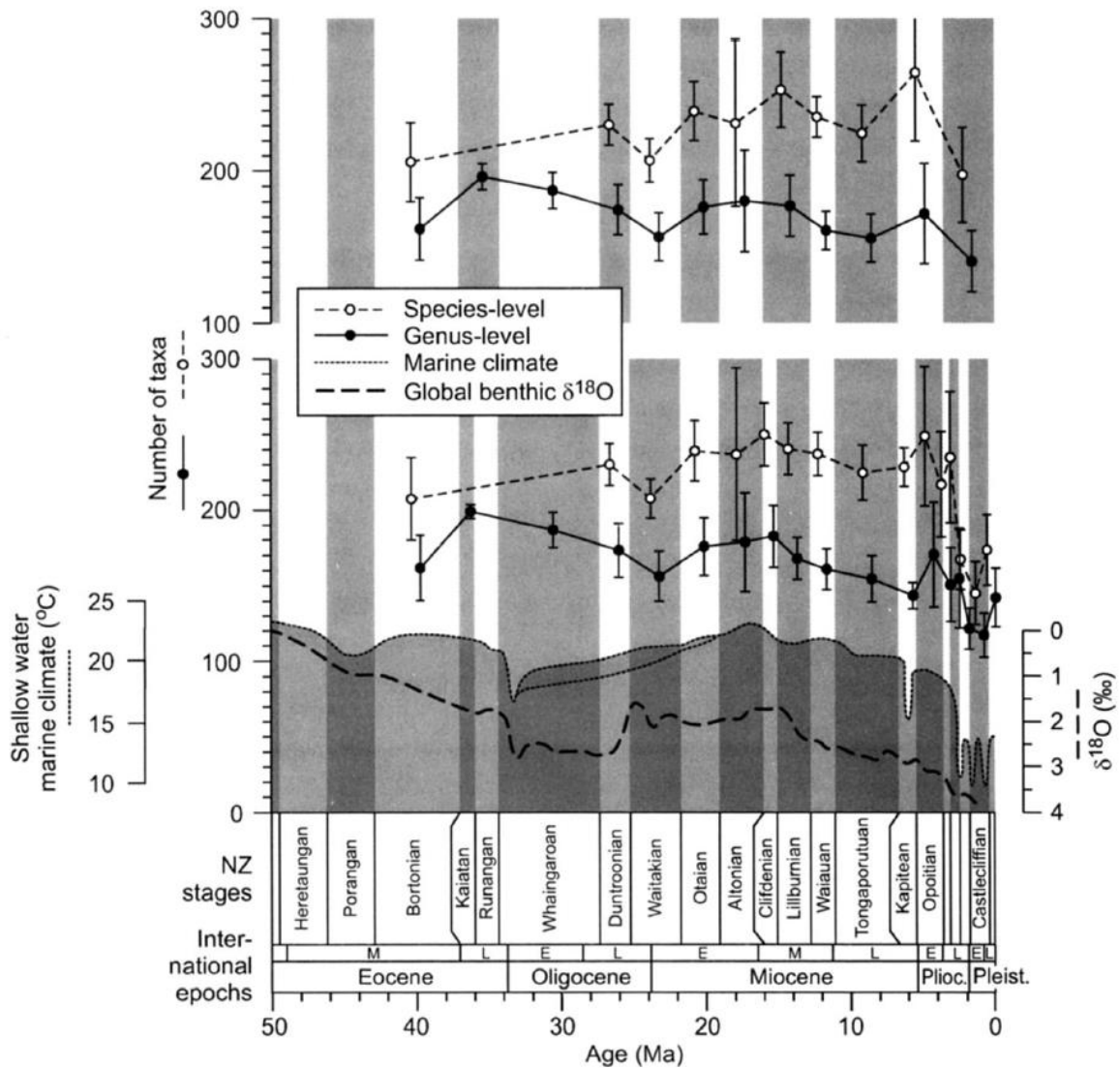
FIGURE 2.11. Plots of sampling standardized species and genus level diversity of shallow-marine molluscs in New Zealand during the Cenozoic. Diversity curves from the upper and lower graphs are based on the same data but show analysis using different configurations of time bins. Dashed and dotted lines depict estimates of marine temperature (climate). Dashed line (bottom) is a normalized oceanic temperature from Hornibrook (1992) normalized to a latitude of 42°S. Dashed line represents the global deep-sea benthic foraminifera $\delta^{18}O$ record from Zachos et al. (2001). Figure from Crampton et al. (2006a).

The research discussed above predominantly focused on quantifying patterns of regional diversity (gamma diversity), although Hendy (2009) did examine alpha diversity

during the Late Miocene to recent fossil record. Crampton et al. (2011) extended their previous analyses and examined patterns of gamma and beta diversity. The focus of this study was to examine the role of spatial structuring of both environments and biodiversity (i.e., beta diversity) in driving macroevolutionary dynamics and biases in the fossil record (Crampton et al. 2011). This was founded on the hypothesis that the drivers of bias in the fossil record have also impacted actual biodiversity patterns – forming the basis of the common cause hypothesis (see Peters and Heim 2011). In this case, Crampton et al. (2011) speculated that the factors that may have affected the quality of the marine fossil record, e.g., areal extent and distribution of depositional versus non-depositional environments, may also have affected habitat fragmentation and therefore biodiversity in New Zealand's marine ecosystems. Whereas ecological spatial structuring can be measured by beta diversity, a new metric was required to measure the spatial structuring of environments. Crampton et al. (2011) utilised the principles of macrostratigraphy (Peters 2008), compiling lithological section data from King et al. (1999) into the macrostrat database (Peters et al. 2018), to develop several proxies of environmental partitioning. Crampton et al. (2011) found no evidence to suggest that the spatial structuring of molluscan faunas, or their derived taxonomic rates, were correlated with the spatial structuring of environments. However, they did infer relationships between origination rate and net diversity, and extinction rate and the quality of the fossil record.

Following the work of Crampton et al. (2006a), Foote et al. (2015) revisited the impact of aragonitic bias within the shallow-marine molluscan fossil record of New Zealand (i.e., preferential loss of fossil species based on their predominant mineralogical composition). Considering that mineralogical composition is typically consistent at the family level in molluscs (Carter 1990), any preferential preservation of calcitic or aragonitic taxa could lead to taphonomic distortion and confound estimates of biodiversity (Foote et al.

2015). As noted earlier, Crampton et al. (2006b) concluded that substantial loss of aragonitic taxa was unlikely during the early to late Cenozoic (~45 Myrs to recent). Foote et al. (2015) extended the previous estimates by Crampton et al. (2006b) by comparing: (1) the temporal variation in sampling probability of aragonitic versus calcitic mollusc species, and (2) comparing and testing regional versus local (defined within this study as an individual collection) spatial scales. They concluded that aragonitic bias was effectively absent in terrigenous clastic sediment, but there was preferential preservation of calcitic taxa in limestones, particularly during intervals of widespread carbonate deposition (e.g., Oligocene and Plio-Pleistocene) (Foote et al. 2015). However, they observed that the absence of aragonitic forms in carbonates was due to taphonomic loss but, equally, also reflects a true ecological signal, i.e., an ecological preference of calcitic taxa during pervasive carbonate deposition (Foote et al. 2015).

The work by Crampton et al. (2011) was the last to examine the diversity dynamics of the shallow-marine fossil record of New Zealand through the Cenozoic. Over the last decade, several new methods and higher resolution datasets have become available. The development of these new methods coincides with an increased understanding and recognition of the role of the fossil record in understanding the impact of anthropogenic climate change and the paradigm of modern diversity loss. As outlined in Chapter 1, one of the primary aims of this thesis are to build upon previous research on global and regional diversity dynamics, incorporate and synthesize new datasets and methods, and develop new methods to examine the drivers of macroecological and macroevolutionary processes in the shallow-marine Cenozoic fossil-record of New Zealand.

2.5. Summary

The major geological, climatic, and oceanographic events during the Cenozoic Era are summarized in Table 2.1. The events considered most relevant to this thesis, and the subsequent Chapters, are summarised below:

- Intensification of tectonic activity within New Zealand post-Eocene and an overall trend of climatic cooling with superimposed climatic optima (end Eocene thermal maximum, Miocene climatic optimum, and transient warming in the early Pliocene).

- Formation of the Antarctic circumpolar current during the latest Eocene to Late Oligocene. This thermally isolated Antarctica and contributed to the development of distinct oceanic fronts, currents, and winds in the Pacific/Southern Ocean.

- Full establishment of the subtropical front and the subantarctic front (and subtropical frontal zone) post formation of the Antarctic circumpolar current, which separated cool subantarctic surface waters in the south from subtropical waters in the north. Although the latitudinal position of these two fronts have shifted since their inception, the majority of New Zealand continental shelf has been bathed in warm-cool subtropical surface waters derived from the Pacific gyre (via the East Australian Current) through the Cenozoic.

In Chapter 1 (section 1.2) I highlight three research questions. Within this Chapter I provide background to contextualize these questions. Firstly, I show that the nature and magnitude of a putative rise in global marine diversity is still unknown. Although recent evidence questions the nature of the global fossil record, estimates of global marine diversity during the Phanerozoic, employing a spatially explicit approach, record a two-fold increase in diversity during the Cenozoic relative to the earlier Phanerozoic (Close et al. 2020).

However, previous paleodiversity studies in New Zealand suggest different and contrasting diversity trajectories of the Cenozoic shallow-marine molluscan fossil record and it is not clear how New Zealand's record relates to this global pattern. Second, in addition to the contrasting diversity trajectories recorded in previous studies, variation in the spatial distribution of the data has not been previously quantified in New Zealand and it is possible that prior estimates of regional diversity are confounded by the uneven geographic distribution of fossil collections through the Cenozoic. Furthermore, there has been limited study on the spatial partitions of biodiversity in the Cenozoic New Zealand fossil record, and no studies on patterns of functional diversity. Lastly, global studies of taxonomic diversity in the fossil record, and the geographic distribution of species richness along the modern latitudinal gradient, implicate climate, particularly oceanic temperature, as a major positive correlate of marine biodiversity (Mayhew et al. 2008, Mayhew et al. 2012, Edie et al. 2018b, Schumm et al. 2019, Gagné et al. 2020). On this basis, we may expect species richness to increase in New Zealand in intervals where global and Pacific Ocean temperatures increase. Whilst earlier works on the Cenozoic shallow-marine molluscan fossil record suggested that oceanic temperature was an important positive explanatory variable (Beu 1990), more recent studies have been unable to demonstrate a relationship between temperature and biodiversity. Several new methods and higher resolution datasets have now become available and permit the re-evaluation of the Cenozoic shallow-marine fossil record of New Zealand.

International Epoch (Age Ma)	New Zealand Stage	Age (Ma)	1st-order cycle	2nd-order cycle	Global climatic state	Geological events	Climatic and Oceanographic events	Quality of the shallow-marine molluscan fossil record
Pleistocene (2.58 - 0)	Castlecliffian - recent	1.63 - 0			Icehouse		Glacial cyclicity and fluctuating Subtropical front position, with incursions of cool subantarctic waters	Well-preserved
	Nukumaruan	2.4 - 1.63						Well-preserved
	Mangapanian	3 - 2.4		7			Progressive cooling	Well-preserved
Pliocene (5.33 - 2.58)	Waipipian	3.7 - 3						Well-preserved
	Opoitian	5.33 - 3.7				Tectonic pulse related to sharp change in relative motion of plates	Period of transient warming	Well-preserved
	Kapitean	7.2 - 5.33						Relatively well-preserved
	Tongaporutuan	11.04 - 7.2		6		Beginning of Puysegur subduction	Progressive cooling	Moderately preserved
	Waiauan	13.05 - 11.04						Relatively well-preserved
	Lillburnian	15.1 - 13.05	HST /RST		Coolhouse		Miocene Climatic Transition, major development of Antarctic icesheets and expansion and intensification of the Antarctic Circumpolar Current. Clear evidence of the Subantarctic front south of New Zealand, with the Subantarctic front and Subtropical front occupying similar positions relative to today	Well-preserved
Miocene (23.03 - 5.33)	Clifdenian	15.9 - 15.1		5			Miocene Climatic Optimum and restricted flow through the Indonesian gateway (~16 Ma)	Well-preserved
	Altonian	18.7 - 15.9						Well-preserved
	Otaian	21.7 - 18.7				Tectonic pulse related to the development of the Alpine Fault (downlap surface) and shift from carbonate dominated sedimentation to primarily terrigenous and increase in thickness of sedimentary sequences	Warming and southward expansion of the Subtropical front and warm subtropical waters	Well-preserved
	Waitakian	25.2 - 21.7		4 (condensed section)		Maximum flooding of Zealandia		Well-preserved
Oligocene (33.9 - 23.03)	Duntroonian	27.3 - 25.2					Full opening of Drake and Tasman Passage and unrestricted Antarctic Circumpolar Current	Well-preserved
	Whaingaroan	34.6 - 27.3				Period of regional erosion		Relatively poor
	Kaiatan	36.7 - 34.6					Cool water flow through the Tasman Passage and formation of proto-Antarctic Circumpolar Current	Relatively poor
	Runangan	39.1 - 36.7	TST		Warmhouse		Formation of the proto-Subtropical front	Well-preserved
	Bortonian	42.6 - 39.1		3				Well-preserved
Eocene (56 - 33.9)	Porangan	45.7 - 42.6					Global increase in heterogeneity between oceanic basins	Well-preserved
	Heretaungan	48.9 - 45.7			Hothouse		Paleocene-Eocene thermal maximum and Early Eocene climatic optimum	Poor
	Mangaorapan	52 - 48.9						Poor
	Waipawan	56 - 52				End of Tasman sea spreading (~52 Ma)		Poor
Paleocene (66 - 56)	Teurian	66 - 56		2	Warmhouse		No evidence of oceanic fronts and Antarctica ice free	Relatively poor

TABLE 2.1. Summary of key climatic, oceanographic, and geological events through the Cenozoic in New Zealand.

Chapter 3

Data and Analysis

3.1. Introduction

The purpose of this Chapter is to discuss the workflows used to generate the main datasets employed within these Chapters and provide details on the computational methods of data analysis. The Chapter is split into two sections: (1) the fossil record datasets used, and (2) methods of data analysis (including code availability).

3.2. Fossil Record Datasets

This thesis is based primarily on three datasets: (1) the synoptic dataset, (2) the collection-based database of fossil occurrences within New Zealand, the Fossil Record Electronic Database (FRED), and (3) the synonymy database. This section provides a description of these datasets the procedures used to generate the main fossil occurrence dataset used in this thesis. Any further, Chapter-specific treatments applied to these datasets are described in detail within each of the main results Chapters (Chapters 5 – 8). The datasets are outline below in Table 3.1.

Dataset name	Sources	Type	Chapters featured
1. FRED Shelfal dataset	(1) FRED, (2) synoptic dataset, (3) synonymy database	Fossil occurrences	5 - 8
2.Synoptic dataset	Synoptic dataset	Fossil first and last appearances	5
3. Functional trait database	(1) Synoptic dataset, (2) WoRMS, (3) Paleobiology Database, (4) K. S. Collins	Fossil Ecospace	7 - 8

TABLE 3.2. Summary of the key datasets discussed within this Chapter.

3.2.1. Synoptic Dataset

The synoptic dataset was compiled by just two paleontologists (A.G. Beu and P.A. Maxwell) over many years. The dataset is very highly vetted and comprises the best estimate of the true stratigraphic ranges (first and last appearances) and paleoecological data for 5,241 New Zealand molluscan species with a Cenozoic fossil record. The dataset used here is an updated version of the synoptic dataset used by Crampton et al. (2006b). Paleoecological traits include information on the following variables.

- Water depth - binary variable based on whether the species is inferred to live at shelfal depths (0-200m water depth - or not).

- Feeding - referring to the primary mode of feeding, e.g., suspension feeders or carnivores etc.

- Tiering - describes the species main vertical life position within the substrate, i.e., whether they are infaunal (live within the substrate) or epifaunal (live on top of the substrate).

- Size – based on the maximum linear dimension of individual species, separated into 5 different size classes.

- Mineralogical composition of shell – binary variable indicating the composition of the shell (aragonitic or calcitic). Calcitic taxa are classed as such if they have a shell component that could, in the absence of aragonite, be identified to species level (see Crampton et al. 2006b and Foote et al. 2015 for further details). Importantly, mineralogical shell composition affects the preservation potential of shells and, therefore, possibly their representation in the fossil record (e.g., Foote et al. 2015).

The stratigraphic ranges of individual species as recorded in the synoptic dataset were used to constrain the ages of the downloaded FRED dataset (see below) and in Chapter 5 for quantifying the impact of the Pull of the Recent on New Zealand's Cenozoic shallow-marine fossil record. The paleoecological traits within the synoptic dataset form the basis of the functional trait database that was used to examine ecospace occupation (Chapters 7 and 8). The ecospace was created from the feeding and tiering traits provided in the synoptic dataset (including several additional updated feeding and tiering traits) and separately sourced traits for locomotion. The additional information for feeding, tiering and locomotion traits was sourced predominately from the input of malacologists (primarily K.S. Collins) and from two other online databases: World Register of Marine Species (WoRMS Editorial Board 2021, https://www.marinespecies.org) and the Paleobiology Database (Peters and McClennen 2016).

3.2.2. Fossil Record Electronic Database (FRED)

FRED is a database of fossil occurrences primarily from New Zealand. The database is well described elsewhere, including the paper by Clowes et al. (2020) (see Chapter 1). As a result, a detailed account of this database is not included in this thesis, but I do provide a summary here for context. FRED forms one of the oldest, if not the oldest, fossil record database globally. The database has existed in paper form since 1946 and was digitized from the 1970s onwards (Crampton et al. 2003). The database is collection-based, where a collection represents a record of "all taxa collected and identified by one or more paleontologists from a single locality on a single occasion" (Crampton et al. 2006a). The structure of FRED differs from more recent databases such as the Paleobiology Database, which is primarily based on fossil lists from published literature. FRED does contain fossil lists from published literature, but the majority of fossil lists are sourced from field- and laboratory-based observations first, that may have been subsequently published. This distinction is important as fossil lists obtained from published literature are subject to several additional analytical biases (see below). However, published literature has several advantages, particularly the additional phases of vetting. Whereas the taxonomy of molluscan fossil species is partly subjective and dependent on the identifying paleontologists; 76% of the molluscan records within FRED were identified by a few, closely collaborating paleontologists (A.G. Beu, J. Marwick, C.A. Fleming) (Crampton et al. 2006b), producing a generally high level of consistency between records and collections. Notwithstanding this, data contained within FRED does vary in terms of quality and vintage (particularly the use of outdated taxonomic names and uncertain stratigraphic ages for individual collections) and requires vetting and quality control.

Much of the analysis in this thesis is based on a download of the FRED database from GNS Science, formerly the New Zealand Geological Survey (https://fred.org.nz/). The

Cenozoic marine molluscan dataset was downloaded in September 2017 and was subjected to several iterations of automated and manual checking and updating to generate the final dataset used here. These procedures are:

1. Prior to downloading the dataset, the existing synonymy list was manually updated to include revised taxonomy based on recent publications, particularly several papers covering molluscs of the last 2 Myrs (Beu et al. 2004, Beu 2006, Beu and Raine 2009, Beu 2011, 2012).

2. The downloaded file was processed through a custom Java program called BioCount that was created by Iain Matcham, formerly of GNS Science. This program allowed the application of customized filters and updated taxonomic synonyms. Several filters were applied to omit uncertain identifications (including prefixes cf. and aff. and additional characters with specific meanings in the FRED database). This resulted in the removal of 2,792 occurrences. For comparison, a second dataset was also generated that retained uncertain identifications. Both downloads incorporated the newly compiled synonym list and updated any outdated taxonomic names with their currently accepted names.

3. The output from step 2 was checked manually for errors and, where necessary, corrected. These errors were related primarily to spurious prefixes not removed by BioCount and the inclusion of several non-marine gastropod genera, which were also removed.

4. An additional cleaning protocol was employed to remove fossil occurrences recorded outside their known biostratigraphic ranges, as documented in the synoptic dataset.

5. Lastly, the dataset was further restricted to level-bottom shelf-dwelling taxa (except for Chapter 5 where both shelf and non-shelfal datasets are considered), based on the binary identification of shelf level provided in the synoptic dataset. This final dataset includes 26,873 species-level occurrences of 1713 species from 4173 collections, herein referred to as the FRED shelfal dataset.

3.3. Data Analysis

Most of the data analysis within this thesis was undertaken in the statistical computing language R (R Core Development Team 2021) using a windows/linux operating system. All data were manipulated and transformed within R (with the exception of the analyses presented in Chapter 5) to create the data subsets and partitions used in each of the main analysis Chapters (see subsequent section on code availability). Specific R packages and functions are mentioned at relevant points in the text.

The analyses for Chapters 6 to 8 were undertaken using the high-performance cluster at Victoria University of Wellington, with 80 available cores (Sgees001, see https://wiki.geo.vuw.ac.nz/index.php?title=Sgees001 for further details). The code in these Chapters was optimized to make use of the high-performance cluster. Geographic paleocoordinates used in Chapters 6 to 8 were derived using GPlates (Müller et al. 2018) and QGIS (QGIS Development Team 2021). Macroevolutionary rates were derived using program MARK (for capture-mark recapture models) through the R package *RMark* (Laake 2013) and PyRate (a python-based program) (Silvestro et al. 2014). PyRate was operated in the terminal of the high-performance cluster.

3.3.1. Data and Code Availability

All datasets and code used to generate the main results in Chapters 5 to 8 are included here as digital supplementary material. The data and code (R scripts) are provided as zip files, which include a text file explaining the function and use of the data files and scripts. The R scripts are subdivided into individual scripts for the "core function" – which generates the main results – and a separate script for the analysis and plotting of the output results from the core function. Although it is possible to run the core function R scripts for Chapters 6 to 8 on an standard computer (with the removal, or reduction in the number of clusters used for

parallel computing) – it is not recommended as the computational time is more than 24 hours.

The outputs of the core functions are included within the provided zip files. The code used to

generate the PyRate results is not included within this thesis, but the process is described in

Chapter 4 (section 4.7.5).

Chapter 4

Methods Review

4.1. Introduction

As noted in Chapter 2, this thesis is composed of several published research papers, or manuscripts that will be submitted for publication (Chapters 5 to 8). The purpose of this Chapter is to contextualise, explain and justify methodological decisions. Methods and processes that are described in detail within the following Chapters are not discussed here, this Chapter focuses only on methods that require additional background information or explanation.

The Chapter is split into eight sections: (1) sampling and (2) spatial standardization methods used, (3) estimation of oceanic temperature, (4) estimation of climatic variability, (5) the application of coverage-based subsampling to other measures of biodiversity, (6) measuring macroevolutionary rates, and (7) time-series correlation.

4.2. Controlling for Sample Size

Measures of biodiversity are dependent on sample size, particularly measures of richness. In general, as the size of the sample increases, so does the probability of sampling rarer or hard-to-detect species or functional groups. Three broad methods exist for controlling for spatial or temporal variability in sampling: (1) sample-based rarefaction, (2) coverage-based subsampling (also known as shareholder quorum subsampling and coverage-based rarefaction) and (3) extrapolation. A full review of all available methods is out of the scope of this thesis; however, a broad comparison of the methodologies and rationale for those utilized in this thesis is included here.

Prior to discussing the methods, it is important to consider the nature of the data. Although all the methods discussed below can be applied to both ecological and paleontological datasets, there are conceptual differences in the way they are applied, and the terminology employed. This thesis examines paleontological data, which are sourced primarily from databases of fossil collections (or references in some fossil databases) as presence-absence data. Therefore, the abundance distributions of fossil datasets are usually compiled using collection-level occurrences rather than a measured count of individuals from each sample. Despite this, the abundance of fossil-level occurrence data still reflects relative abundance of different species, i.e., those that are relatively common (i.e., are present in many collections), versus those that are rare (i.e., present in only a small number of collections) (Alroy 2010b, c, 2014, Close et al. 2018, Dunne et al. 2018, Alroy 2020). This characteristic of fossil data allows the use of the same richness estimators utilised in ecology, providing several additional biases are considered and corrected for, if necessary. These additional biases are primarily related to fossil occurrence datasets derived from published

69

literature compilations, centring on three prominent issues, as highlighted by Alroy (2010b, c, 2014) and Close et al. (2018):

1. Publication biases, e.g., publications on new fossils occur at a faster rate than those for already known species, and researchers tend to focus on understudied aspects of the fossil record.

2. The "expanding sampling-universe problem", i.e., as literature is added to the dataset, both the sampling universe and the underlying pool of taxa change.

3. The "monograph problem", i.e., repeated reports of occurrences of common taxa in large monographs.

The New Zealand molluscan fossil record, as captured in the Fossil Record Electronic Database (FRED), circumvents most of these biases as it is primarily derived from field-based collections and observations. As a result, and as discussed later in this Chapter, several commonly applied corrections are not necessary for the Cenozoic molluscan fossil record of New Zealand.

4.2.1. Classical Rarefaction

Whereas "rarefaction" is often used as a general term for all methods that exist for controlling for sampling size. Here, I reserve the term "classical rarefaction" purely for sample-size based methodologies. Classical rarefaction is conceptually based on reducing larger samples to the size of the smallest sample. This creates samples of the same size, but does not accommodate for differences in the abundance distributions of the samples being compared, i.e., the completeness of the sample (Chao and Jost 2012). As a result, values of richness calculated from classical rarefaction can be misleading and may retain a signal of

sample size, rather than being a meaningful measure of biological diversity (Chao and Jost 2012).

4.2.2. Coverage-Based Subsampling

Coverage-based subsampling resamples to equal completeness rather than equal sample size. In this context, sample completeness is estimated by coverage, which represents the "portion of the total number of individuals in a community that belong to the species represented in the sample" (Chao and Jost 2012), i.e., the complement to the estimate of how much of the sample is missing. Coverage is usually estimated using Good's u (Eq. 4.1) as a basis, which is simply 1 minus, "the number of singletons" (occurrences of taxa only recorded once) (s_1) divided by the number of total occurrences (N) (or total abundance)". This estimation of coverage can be modified to take account of other factors, for example, excluding dominant taxa (Alroy 2010c), or the number of doubletons (s_2) (Eq. 4.2).

There are two commonly used algorithmic approaches to estimating richness in this way, as proposed by Alroy (2010a, 2014) and Chao and Jost (2012). Alroy's original method is implemented through randomly drawing occurrences until a set coverage level is reached (see Appendix 1 for a simplified outline of the method). This has been subsequently updated to accommodate for several publication biases, and subsamples collections (or references) rather than individual occurrences (see Appendix 2 for a simplified outline of this revised method). This updated approach calculates Good's u after each occurrence is added and calculates richness every time Good's u meets the target quorum (coverage). The advantage of Alroy's original approach is that it generates a subsampled abundance frequency, which could also be utilized to measure other partitions of biodiversity. This is important and is discussed later in this Chapter.

The method of Chao and Jost (2012), on the other hand, estimates richness at a given coverage, based on an analytical rather than resampling approach (i.e., based on analytical

equations), and allows extrapolation to the set coverage when it is not met (using the Chao1 estimator, which is commonly employed in modern ecology). This method produces identical results to the uncorrected algorithmic approach of Alroy (2014) when limited to interpolation (Close et al. 2018). The advantage of the method of Chao and Jost (2012) is that it allows smaller samples to be included provided extrapolated estimates do not exceed recommended limits (twice the size of the reference sample) and is easily implemented through the R package *iNEXT* (Hsieh et al. 2016). The primary drawback of this method is that it does not allow for the corrections implemented in Alroy's updated method. However, as previously discussed, the New Zealand molluscan fossil record is largely unaffected by these specific biases due to its structure. Regardless of the approach, coverage-based subsampling has been shown to be sensitive to evenness (Hannisdal et al. 2012, Mitchell 2015, Alroy 2020), where evenness describes the uniformity of the relative abundance distribution of a community or sample. This is particularly problematic when the relative abundance distribution is extremely skewed (e.g., strongly dominated by the most common species), and estimated richness may track evenness rather than richness is this scenario (Hannisdal et al. 2012).

Eq. 4.1. Good's *u* (Good 1953)

$$u = 1 - \frac{s_1}{N}$$

(where s_1 is the number of singletons and N is the total number of occurrences)

Eq. 4.2. Updated coverage estimator (Chao and Shen 2010)

$$\hat{C}_n = 1 - \frac{s_1}{N} \left[\frac{(N-1)/s_1}{(N-1)s + 2s_2} \right]$$

(as above for Eq. 4.1., and $\hat{C}_n$ is the coverage of sample *n*, and s_2 is the number of doubletons)

4.2.3. Extrapolators

Both classical rarefaction and coverage-based subsampling seek to interpolate richness values at equal sampling size or equal coverage, respectively. Extrapolation methods seek to estimate the true richness of taxa within each "sampling universe" (Alroy 2020), i.e., an estimate of true richness based on the "relative frequencies of rare species to analytically estimate undetected species from limited samples" (Close et al. 2018). Here, several common methods of extrapolation utilized in ecology and paleobiology are discussed: (1) squares, (2) corrected first-order jackknife (cJ1) and (3) a Poisson sampling model (True Richness estimated using a Poisson Sampling model – TRiPS). The Chao1 estimator, noted above, is not discussed here, as its use in coverage-based subsampling is to predict richness at a set coverage, and its performance is surpassed by cJ1 and squares when used solely as an extrapolator (Alroy 2020). Notwithstanding this, Chao1 shares methodological similarities with squares. Squares (Eq. 4.3) and cJ1 (Eq. 4.4) are both nonparametric and increase true richness estimates when there are numerous singletons; however, cJ1 takes account of the overall sample size, assuming that sampling is Poisson (Alroy 2020). On the other hand, TRiPS is parametric and explicitly models the sampling process (maximum-likelihood Poisson sampling rate) while estimating richness (Starrfelt and Liow 2016, Alroy 2020).

Eq. 4.3. Squares (Alroy 2020)

$$R' = \frac{S + s_1^2 \sum n^2}{N^2 + s_1 n}$$

(where S is the number of species in a subsample s_1 is the number of singletons, n the vector of species counts, N the number of occurrences [sum of n])

Eq. 4.4. cJ1 (Alroy 2020)

(a) cJ1 equation

$$R' = \frac{S + s_1}{(1 - e^{-\lambda}) + \lambda e^{-\lambda}}$$

(b) Poisson sampling equation (utilized in Eq. 4.4a)

$$\left(\lambda = \frac{\log N}{s_1} \right)$$

(where S is the number of species in a subsample, s_1 is the number of singletons and λ the

Poisson sample rate).

4.2.4. Comparison of Methods

In a recent review, Alroy (2020) evaluated the use of cJ1, squares, TRiPS and

coverage-based subsampling based on several basic validation tests ([1] ability to extrapolate;

[2] consistency of estimates; [3] combination – split-analyse-and-sum test; and [4] evenness),

using ecological data to estimate species richness. The extrapolators cJ1 and squares

performed well in all the tests, but neither excelled when sampling was poor. On the other

hand, TRiPS failed on several occasions, particularly in its ability to extrapolate (Alroy

2020). Coverage-based subsampling (following Alroy's 2014 method) generally performed

well but was confounded by variations in evenness (Alroy 2020), as discussed previously. In

a similar review, Close et al. (2018) evaluated the use of several richness estimators using

fossil datasets. Richness estimators tested included several extrapolators, namely Chao1, λ^5

(similar to cJ1 – see Alroy [2017]) and TRiPS, and interpolated coverage-based subsampling

(following the approach of Alroy 2014). Close et al. (2018) also observed that TRiPS did not

extrapolate well, particularly when the underlying frequency distribution was not flat. λ^5

performed well, but similar to squares and cJ1, did not perform well when sampling was poor. Coverage-based interpolators generally produced more stable subsampling diversity estimates, and Close et al. (2018) suggested that richness estimators standardized by coverage are one of the best current methods for reconstructing paleodiversity patterns.

On this basis of these findings, coverage-based subsampling and extrapolators such as cJ1 appear to be the most robust methods for measuring species richness. To test the applicability of these methods for quantifying biodiversity in the Cenozoic New Zealand molluscan shallow-marine fossil record, I compared the results of coverage-based subsampling, cJ1 (based on code in Alroy 2020) and TRiPS (based on the "TRiPS_simple.R" code in Starrfelt and Liow 2016) using the FRED shelfal dataset (see Chapter 3, section 3.2.2) at the resolution of individual NZ stages (Fig. 2.1).

The above comparison and methodology outline the use of these methods for measuring species richness. However, functional richness may also be affected by sampling intensity, as rarer functional groups (defined based on shared functional traits between species – see Chapter 7, section 7.2.3) also have an increased probability of being sampled as sampling intensity increases. As a result, I also test the use of these three methods on the estimation of functional richness (the number of functional groups within a sample), based on the abundance of occurrences from functional groups (see Chapter 7 for more details).

The aim of this test was to differentiate how the methods vary with sampling intensity. Importantly, the spatial and temporal resolution of the FRED shelfal dataset used in the main results Chapters is not the same as tested here. The sampling intensity of each NZ stage, measured as the number of collections, varies throughout the Cenozoic and is particularly poor in the oldest time intervals (as a result these intervals are generally omitted in the research Chapters). For species richness, comparisons between coverage-based subsampling (set to coverage of 0.7) with TRiPS and cJ1, suggest both TRiPS and cJ1

perform similarly (Fig. 4.1, panel E), and that TRiPS and cJ1 likely overestimate species richness at high sampling intensities (Fig. 4.1, panels A and C). Whereas this observation could also be taken to mean that coverage-based subsampling underestimates species richness, the strong correlation (based on first-differences, see section 4.8) between number of collections and species richness estimate using TRiPS/cJ1 (Fig. 4.2, panels C and E) suggest this is not the case ($r_s = 0.67$, $p = 0.007$ for both). Conversely, functional richness showed the opposite pattern, which suggests that either coverage-based subsampling overestimates functional richness or TRiPS and cJ1 underestimate functional richness (Fig. 4.1, panels B and D), particularly at lower sampling intensities. In addition, no significant correlations between the number of collections and functional richness estimates were observed (Fig. 4.2, panels B, D and F). This could be taken to mean that functional richness is less sensitive to variation in sampling intensity. However, there is also evidence that the disparity between the methods results from differences in the underlying abundance distributions between species and functional groups, where most occurrences are within a small number of functional groups (Fig. 4.3). Coverage-based subsampling is known to be confounded by skewed abundance distributions, tracking evenness rather than richness. On this basis, it is likely that coverage-based subsampling overestimates functional richness.

Undoubtedly all the methods discussed here have some flaws, however, sampling standardization using coverage-based subsampling and cJ1 perform well using both ecological and fossil data, based on published comparisons. As explained above, for the Cenozoic shallow-marine molluscan fossil record of New Zealand, I find that coverage-based subsampling is likely the best method for measuring species richness, whereas extrapolators such as cJ1 appear to be better for functional richness. As a result, this thesis focuses primarily on sampling standardization using coverage-based subsampling and cJ1.
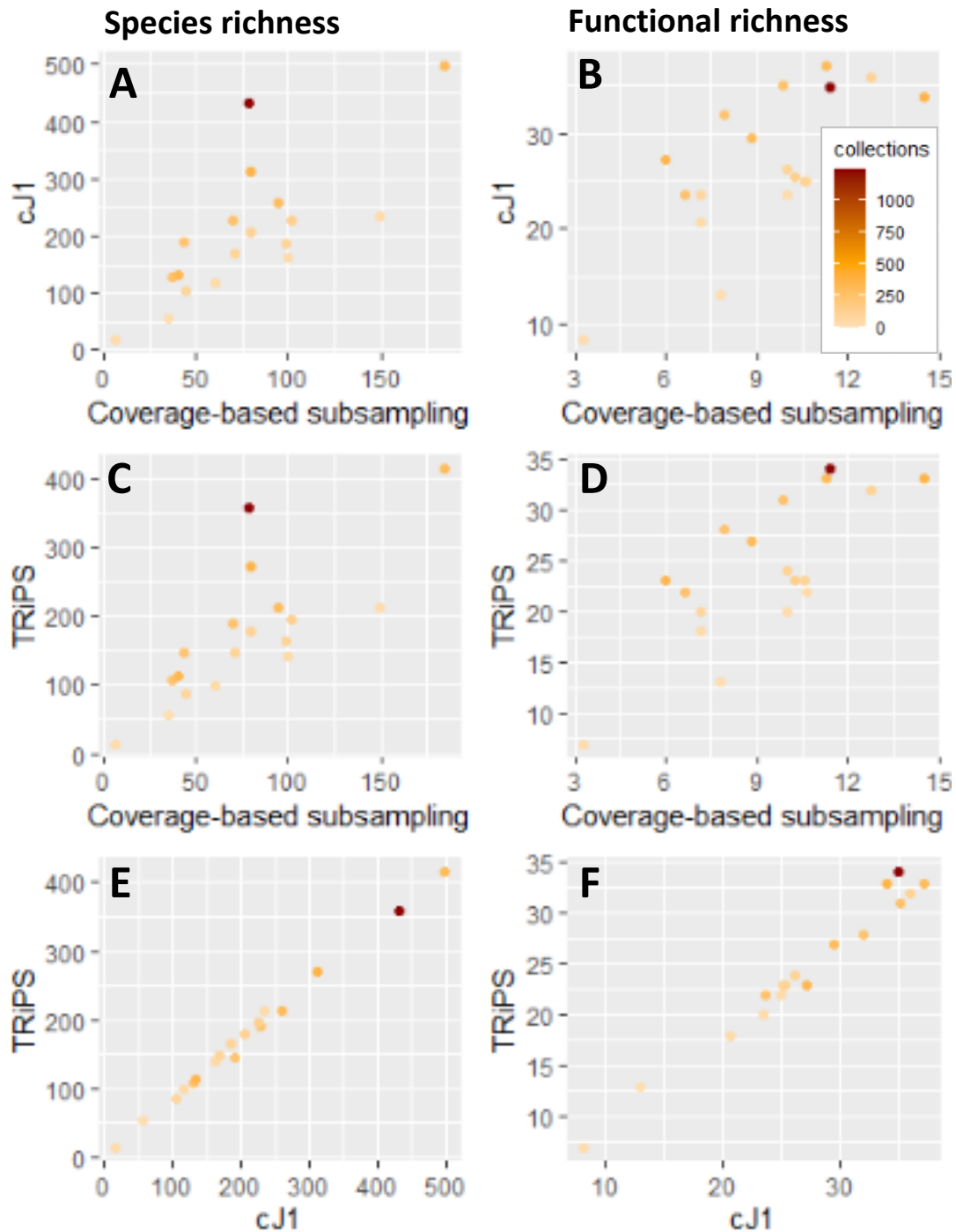
FIGURE 4.1. Comparisons between coverage-based subsampling, cJ1 and TRiPS using the Cenozoic shallow-marine molluscan fossil data at the resolution of the NZ stages. Individual points are colour coded based on the number of collections (i.e., sampling intensity).
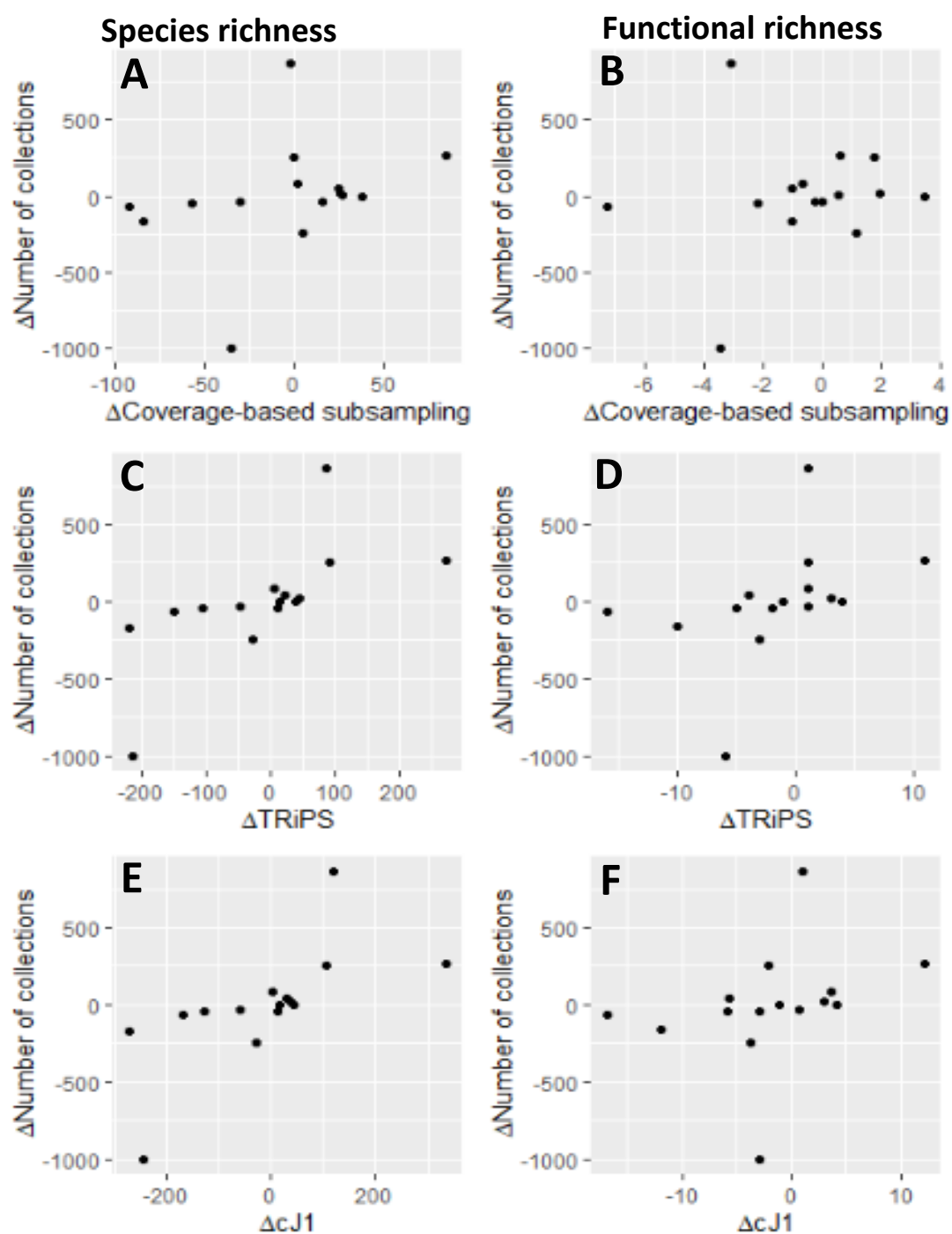
FIGURE 4.2. Comparisons between sampling intensity (number of collections) versus

measures of richness estimated using coverage-based subsampling, cJ1 and TRiPS. Plots are

based on first-differences (indicated by Δ, see section 4.8). Analyses were undertaken using the Cenozoic shallow-marine molluscan fossil data at the resolution of the NZ stages.
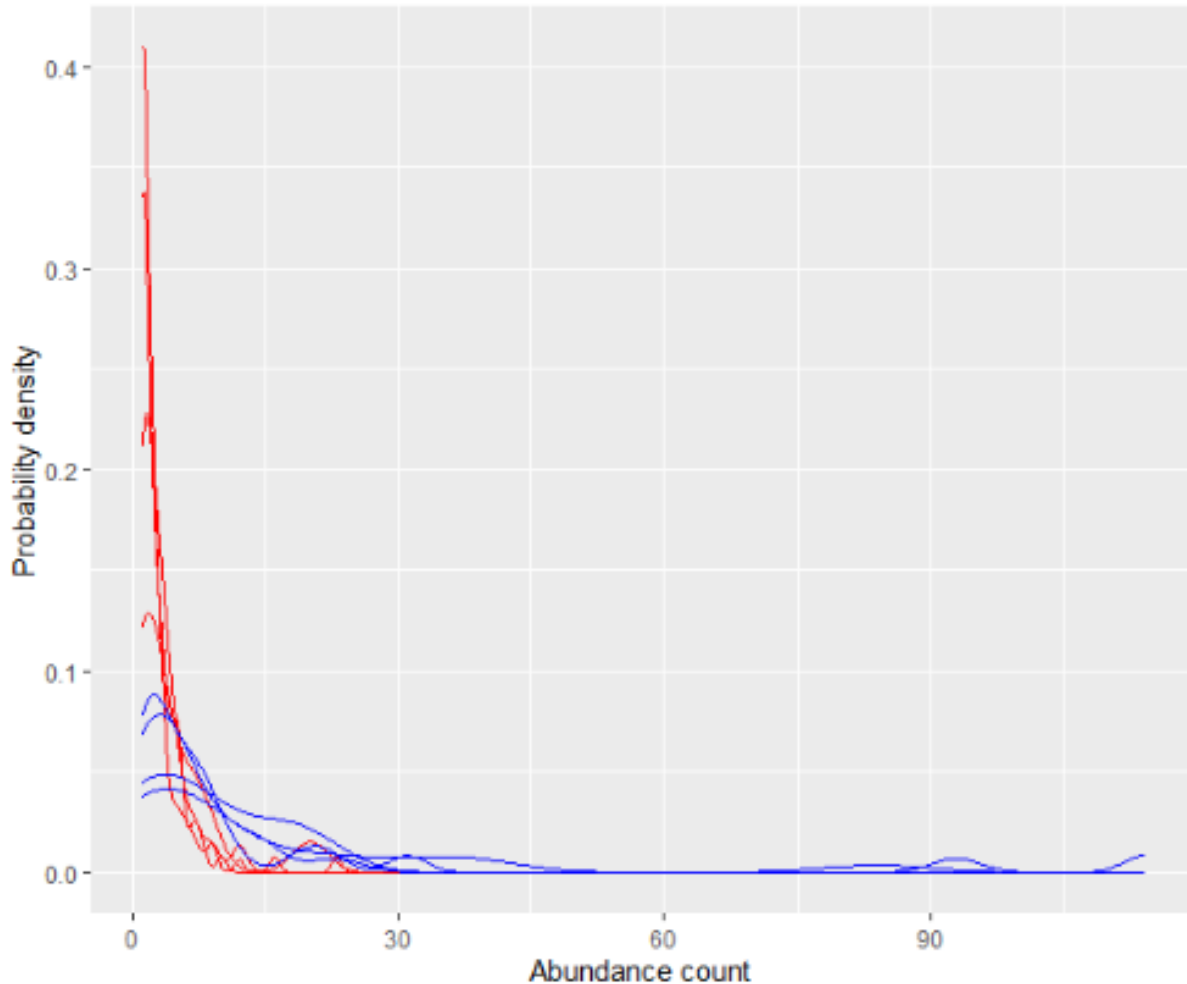
FIGURE 4.3. Probability density of collection-level abundance counts across five individual stages with low numbers of collections (Clifdenian, Lillburnian, Whaingaroan and Kaiatan). Blue lines represent the probability density of the abundance count of species within functional groups, whereas red lines represent species richness. The distribution of functional groups is skewed towards higher abundance counts within each functional group (i.e., less singletons and doubletons and in some cases numerous occurrences of the same functional group). As a result, functional richness estimated using coverage-based subsampling may track evenness rather than richness.

4.3. Controlling for Spatial Area

Measures of biodiversity are dependent on the area sampled; however, the relationships between different partitions of biodiversity and spatial area are not uniform (see Chapter 6). As a result, spatial area needs to be considered when comparing measures of biodiversity from different time intervals. Whereas measuring area in modern ecosystems is conceptually easy (although in reality it is still a fundamental issue in theoretical ecology), i.e., it can be physically measured, this is not possible for fossil ecosystems. Furthermore, the spatial distribution of the fossil record through time is uneven and constrained by tectonic processes and the current configuration of the Earth's continents and sedimentary basins. As a result, estimating spatial area in the fossil record relies on both paleogeographic reconstructions and methods of quantifying spatial extent, rather than area explicitly, that can accommodate for the uneven and patchy distribution of fossil localities.

The primary spatial unit for the fossil record is the geographic coordinates of fossil localities, often restored to inferred paleocoordinates. As a result, the methods discussed here are restricted to analytical techniques based on point data only. Broadly, there are four main ways to measure spatial extent using point data: area, distance, grid-based methods (gridded) and other methods that summarize spatial distribution. Within this section I discuss and compare these methods and assess their suitability for the New Zealand shallow-marine molluscan fossil record.

4.3.1. Area

Even in modern ecology, measuring area from coordinate data of site localities is problematic. This is particularly true for the estimation of the geographic range (home range) of individual species, which shares similarities with measuring geographic area with fossil data, i.e., the patchy and uneven nature of ecological survey data and focus on quantifying area from point data, rather than the area of geographic features. Consequently, the methods applied to geographic range estimation in ecology can be easily adapted to quantifying the geographic area of fossil data. These methods range from relatively simple, e.g., the area of a convex-hull or alpha-hull, to more complex such as those based on graph theory and kernel density estimation.

Perhaps the most common method of quantifying area in the fossil record and in ecology is the convex-hull. The convex-hull is defined as the smallest convex polygon that encloses all points within a set of spatial data. An alpha-hull is similar, but is constructed from the summed area of polygons (termed "alpha shapes"), constructed of piecewise linear simple curves (Darroch and Saupe 2018). The advantage of the alpha-hull is that it allows, in part, for gaps between points, which would be included within a convex-hull. However, a recent review suggests that the alpha-hull is outperformed by the convex-hull when reconstructing geographic ranges from a paleontological perspective (Darroch and Saupe 2018). The convex-hull has been used in a number of terrestrial paleontological studies as a proxy for area (Carrasco et al. 2009, Barnosky et al. 2011, Lagomarcino and Miller 2012, Dunhill and Wills 2015), and has been shown to perform well in the calculation of the species-area relationship in the terrestrial fossil record (Barnosky et al. 2005). Notwithstanding this, the convex-hull has fundamental issues with both underestimation and overestimation of area depending on the spatial distribution of the data (Alroy, written pers. comm. 2020), and cannot account for any variation in the density of the data.

81

Another recent method that could be applied to the fossil record makes use of graph theory (Alroy, written pers. comm. 2020). This method creates a linear graph, termed a "mosaic", that connects all points in the dataset. The area (*A*) is then derived from the length of the mosaic (*M*), the number of edges (*e*) and a constant (*L*) (Eq. 4.5). This allows for, in part, variation in the density of data (e.g., clustered data will increase the number of edges). This method performs well for both empirical and simulated data for quantifying the geographic range of terrestrial ecological data, particularly against more complex methods based on kernel density estimation (Alroy, written pers. comm. 2020).

Eq. 4.5. Mosaic area (Alroy, written pers. comm. 2020)

$$A = L \frac{M^2}{e}$$

Kernel density estimators were first used in modern ecology in the late 20th century (Worton 1989), but have recently been further developed, particularly aided by computational advances (Fleming and Calabrese 2017). The method is non-parametric and estimates the probability density function of the point data, which can yield contours relative to statistical significance that can be used to estimate area. This method allows for variation in the density of the data and can accommodate for sampling size (Fleming and Calabrese 2017). However, this method has not been previously applied to fossil data and its suitability in this regard is unknown.

To assess the validity of the methods discussed for quantifying area in New Zealand's Cenozoic shallow marine molluscan fossil record, I qualitatively compared the graphical outputs of the convex-hull, mosaic and contours from a basic kernel distribution estimator using paleocoordinate data from the youngest shallow-marine molluscan fossil record of New Zealand (1.63 Ma – recent). This interval of the fossil record is relatively well-sampled and provides a good test to assess the suitability of these methods. We were unable to compute

the alpha-hull with all data points due to the uneven distribution of the data (causing the function to fail); however, visual inspection of random subsets of the data indicated they were significantly biased by landmass distribution. The convex-hull was also significantly biased by landmass and clearly overestimates area (Fig. 4.4, panel B). The mosaic was still biased by landmass and overestimates area, but to a much lesser extent than the convex-hull (Fig. 4.4, panel A). The kernel density distribution was tested at several confidence levels (95%, 97.5% and 99%), and did not deal well with the patchy nature of the data and excluded numerous data points (Fig. 4.4, panel C, plotted at the 97.5% confidence interval). As a result, estimates of area derived from the contours at the significance values used here would likely substantially underestimate area. On this basis, none of the methods discussed here appear to be suitable for the New Zealand shallow-marine molluscan fossil record. The mosaic, kernel density distribution contours, convex-hull and alpha-hull were implemented in R using the packages *mosaic* (tgraph function) (Alroy, written pers. comm. 2021), *MASS* (Venables and Ripley 2002) (kde2d function), function chull (R Core Team 2021) and package *alphahull* (Pateiro-Lopez and Rodriguez-Casal 2019) respectively.
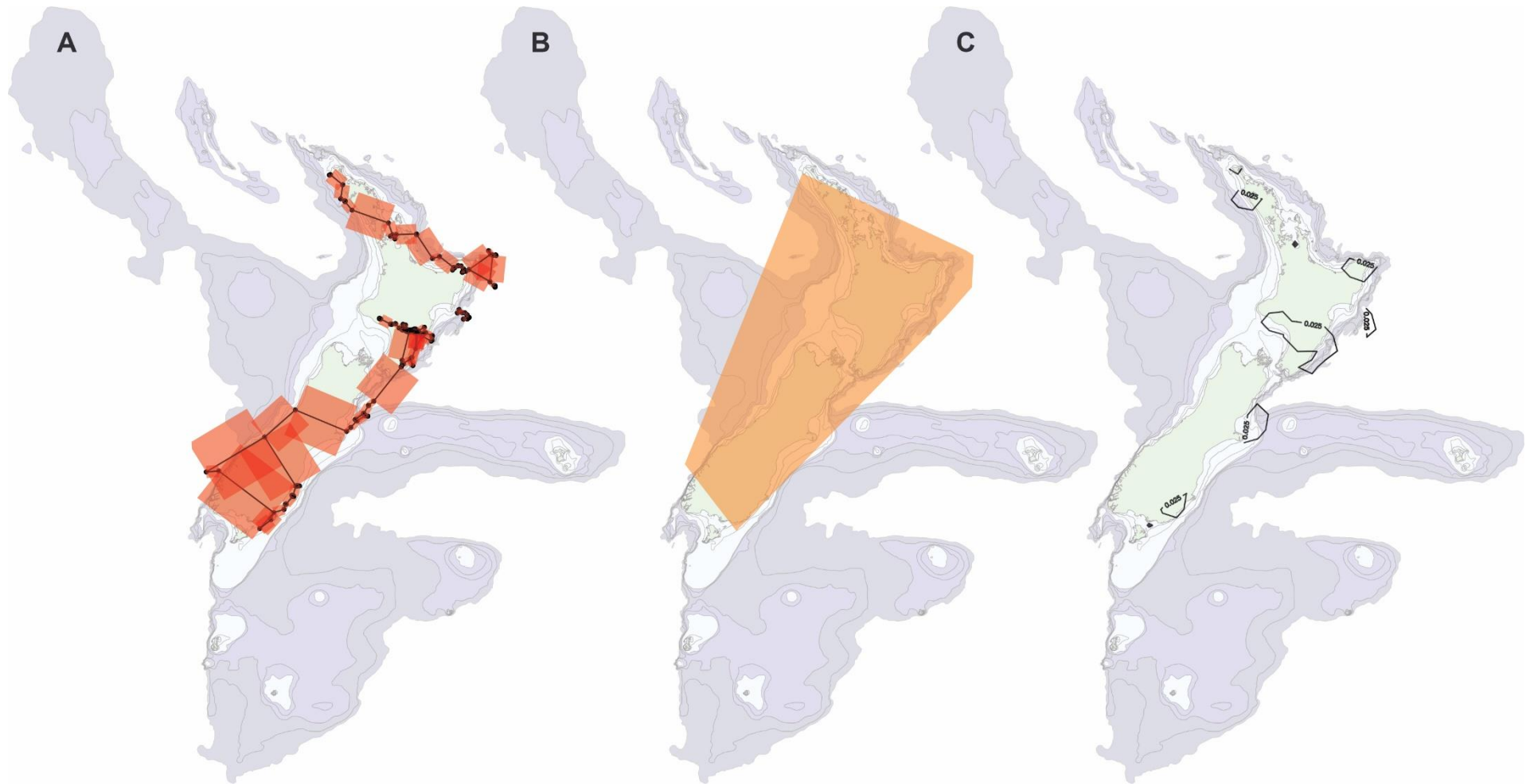
FIGURE 4.4. Comparisons of three graphical outputs for quantifying spatial area using mosaic (A), convex-hull (B) and contours from a basic kernel distribution estimator (C), based on paleocoordinate data from the youngest shallow-marine molluscan fossil record of New Zealand (1.63 Ma – recent) as shown by the black points in map (A). Kernel density estimator contours are based on a confidence interval of 97.5 %. Base map is the modern geography of New Zealand, modified from maps provided by D. Strogen.

4.3.2. Distance

Distance is a much simpler measure of spatial extent and is generally measured as the maximum great-circle distance or the average pairwise great-circle distance between point data. Despite its simplicity the maximum great-circle distance is the standard metric used to quantify geographic range in the marine and terrestrial fossil record using occurrence data (Close et al. 2017). However, distance can only capture one dimension and cannot account for any variation in the density of the data. Considering the observable variation in the density and distribution in paleocoordinates within the New Zealand dataset (Figs. 4.4, panel A, and see Figs. 6.2 and 6.3), distance is not considered to be an appropriate measure of spatial distribution for the data used in this thesis.

4.3.3. Gridded and Spatial Distribution

Gridded and spatial distribution methods are discussed together as they are used in combination to measure spatial area/distribution in the research Chapters of this thesis. This is discussed in detail in Chapter 6, and briefly in Chapters 7 and 8. Nevertheless, I outline the basic principles of both and discuss the merits of combining them here.

Gridding is a common method in both ecological (Lennon et al. 2001, Antão et al. 2020) and paleontological studies (Foote et al. 2008, Crampton et al. 2011, Penny and Kröger 2019), and provides a method to estimate spatial area and mitigate for the effects of finer scale spatial variability in sampling by combining collections within each grid cell. Perhaps the most important factor of gridding is creating an equal-area grid. As the Earth is spherical (although imperfect), it is impossible to apply a square or rectangular grid that does not change in geometry across latitudinal or longitudinal scales. This is not a major problem for studies at local spatial scales but becomes more significant at the spatial scales typically used in paleodiversity studies, i.e., regional, or global. A simple solution to this is to use an equal-

area penta-hexagonal grid. A global penta-hexagonal grid can be implemented using the R package *icosa* (Kocsis 2020), which allows the user to specify the size of each individual grid cell. Whereas the name implies individual grid cells are pentagons or hexagons, pentagonal grid cells are only present at the lowest resolution possible (i.e., where the globe is modelled as a dodecahedron – constructed of 12 equal-area pentagons, see Kocsis [2020] for further information). Consequently, at the scales used in this thesis, grid cells are hexagonal. Spatial area can then be measured as grid cell occupancy, i.e., the number of hexagonal grid cells occupied by fossil collections. However, grid cell occupancy does not capture any information regarding the dispersion of the data - 10 grid cells could be contiguous or widely dispersed across a continent. This problem is highlighted in Chapter 6 (see Fig. 6.2) and can be resolved using the method outlined below.

Distribution methods summarize the variation in spatial distribution of point data. Whereas several methods exist for quantifying spatial distribution (e.g., nearest neighbor-based algorithms), very few have an output that incorporates both the spatial distribution and extent of the data in a tangible unit of measurement. Several recent studies in the fossil record have employed summed minimum spanning tree (MST) length to spatially standardize global fossil data (Alroy 2010b, Close et al. 2017, Close et al. 2018, Antell et al. 2020, Close et al. 2020a, Close et al. 2020b). A MST is an algorithmic method to create the smallest possible incomplete (i.e., not closed) series (tree) of connected vertices between a cluster of points. The sum of the length of the vertices provides the summed MST length, i.e., the minimum distance that connect all points. Therefore, summed MST length incorporates elements of both the extent and distribution of point data, and can be quantified in spatial units (e.g., meters or kilometers). However, one issue with summed MST length is that dense clusters of geographic points can contribute disproportionately to the length of the area relative to the overall extent of the points. This attribute may be an advantage in certain circumstances,

particularly where spatial distribution is the focus, but in the context of attempting to quantify a metric that is close to spatial area, and when the dataset contains clusters of data, it is not. Aggregating fossil collections by grid cell and using the central coordinates removes this problem. The combined use of summed MST length and a hexagrid system is discussed in more detail in Chapter 6 and forms the basis for spatial standardization and quantification of a proxy for area in this thesis. Importantly, the method derived in Chapter 6 is optimized to measure geographic scale in New Zealand's shallow-marine ecosystems, i.e., predominately coastal ecosystems, surrounding a landmass. Modification of this method or other methods may be more applicable for studies focusing on different geographic configurations or taxonomic groups (particularly terrestrial ecosystems).

4.4. Measuring Oceanic Temperature

Oceanic temperature is considered to be one important correlate or control of marine species diversity, both now (Schumm et al. 2019, Gagné et al. 2020) and in the past (Erwin 2009, Mayhew et al. 2012, Valentine and Jablonski 2015). As a result, oceanic temperature is an important explanatory variable to consider in the structuring of New Zealand's shallow-marine biodiversity. However, quantifying relative changes in New Zealand's oceanic temperature is difficult due to: (1) the nature of available temperature proxies, (2) the temporally limited and patchy record of New Zealand specific data, and (3) the position of New Zealand's shallow-marine shelf in relation to the evolving oceanic fronts during the Cenozoic. This section discusses these three issues in relation to the availability and suitability of data for synthesizing a time series of oceanic temperature relevant to New Zealand's shallow-marine fossil record.

4.4.1.   Cenozoic Palaeoceanography of New Zealand

Prior to discussing oceanic temperature proxies, it is important to recap (see Chapter 2, for detailed discussion) the general oceanographic setting of New Zealand during the Cenozoic. This is significant because New Zealand is situated at the interface between the Pacific and Southern Oceans, separating cooler Antarctic-subantarctic waters in the south from cool-subtropical waters in the north. Importantly, the climates of these individual oceanic provinces are not coupled throughout the Cenozoic (Cramer et al. 2009), which is an important consideration for the suitability and selection of proxy data. A detailed summary of the Cenozoic to modern climatic evolution of New Zealand is provided in Chapter 2, here I discuss and reiterate the information relevant to deriving a suitable oceanic temperature time series.

During the Cenozoic New Zealand has remained at relatively static southern mid-latitudes, with the majority of its exposed landmass and shallow-marine shelf area located within cool-subtropical waters and, following its formation in the Late Eocene, north of the subtropical front (STF) (Nelson and Cooke 2001). The only major exception to this is during the last 5.33 Myrs, when the STF has been constrained along Chatham Rise along New Zealand's eastern margin, transecting the Solander Trough to the south of New Zealand (its modern location) (see Fig. 2.2) (Nelson and Cooke 2001). There is also evidence that the STF migrated northwards along the western margin of New Zealand during glacial cycles in the Pleistocene (Bostock et al. 2015). Therefore, in very broad terms, the majority of New Zealand's shallow-marine shelf area during the Cenozoic has been bathed in waters that originate predominantly from the north via the South Pacific Gyre and East Australian Current (and their former equivalents during the Cenozoic). On this basis, it is likely that oceanic temperature proxy data representative of the cool-subtropical water masses is more suitable for deriving a relative time series of oceanic temperature for the New Zealand region during the Cenozoic, than the subantarctic water masses to the south. Nonetheless, I still discuss and compare data from the south of New Zealand and global trends.

4.4.2. Oceanic temperature proxies

There are number of proxies used to derive oceanic temperature. These proxies are generally derived from geochemical indices from fossils (e.g., precipitated shells) or biological signatures (e.g., alkenones from haptophyte algae). The most common and widely accepted data source in the marine geological record is foraminiferal geochemistry, particularly $\delta^{18}O$ or Mg/Ca ratio in benthic and planktonic foraminifera that precipitate calcitic tests. Although the focus of this thesis is on shallow-marine shelfal ecosystems, shallow-marine benthic and planktonic forams are exposed to variability in seasonal, latitudinal and geographical temperature and salinity (particularly $\delta^{18}O$) (Lear et al. 2000).

Whilst this variability may be important over transient timescales and at local spatial scales, this thesis focuses on long-term biodiversity patterns across New Zealand, representative of ecological equilibria. Therefore, a record of oceanic temperature that is representative of average long-term oceanic conditions is more appropriate. The deep sea benthic foraminiferan record provides such a time series as it is largely insulated from this short-term variability in temperature and salinity (Lear et al. 2000) and reflects variation in high latitude oceanic temperatures, i.e., the source waters of deep and bottom waters (Miller et al. 2020).

Both δ^{18}O values and the Mg/Ca ratio are sensitive to temperature, however, δ^{18}O is also affected by seawater geochemistry, which varies as a result of salinity gradients and global ice volume (Barker et al. 2005). Although deep sea foraminifera are partially insulated from the latter as previously discussed. Despite this, global estimates of oceanic temperature using deep sea foram data, particularly during the Phanerozoic and Cenozoic in isolation, are generally based on compilations using δ^{18}O (Cramer et al. 2009, Veizer and Prokoph 2015, Miller et al. 2020, Westerhold et al. 2020). This is largely because δ^{18}O analysis is a more established technique than Mg/Ca and is unrivalled in terms of the temporal and geographic extent of the available data (Cramer et al. 2009, Westerhold et al. 2020). Nevertheless, recent reconstructions of global Cenozoic oceanic temperatures from both δ^{18}O and Mg/Ca are relatively similar (Lear et al. 2000, Cramer et al. 2011). On this basis, and due to the available data (as discussed below), I focus on reconstructing oceanic temperature using deep sea benthic foram δ^{18}O records.

Nevertheless, the deep sea benthic foram δ^{18}O record has some limitations and its insulation from short-term variability can also lead to insensitivity to globally important climatic excursions. One important climatic excursion that is poorly expressed in the global and basinal Cenozoic deep sea benthic foram δ^{18}O record is the transient warming during the Late Miocene/Early Pliocene (labelled PTW in Fig. 4.5). The timing of this warming is

directly after the peak of the continued "Late Miocene cooling" (labelled LMC in Fig. 4.5) (Herbert et al. 2016). This warming is evident in the taxonomic composition of New Zealand's shallow-marine molluscan fossil record, particularly in the Early Pliocene (Opoitian Stage) (see Chapter 2) and is captured by South Pacific surface sea temperature estimated from alkenones ($U^{k'}_{37}$ index) (Fig. 4.5). Whereas this thesis focuses on deep sea benthic foram $\delta^{18}O$ compilations (as discussed below), the South Pacific alkenone sea surface temperature record over the last 12 Myrs helps explain several biological patterns during this interval (discussed further in Chapter 7).
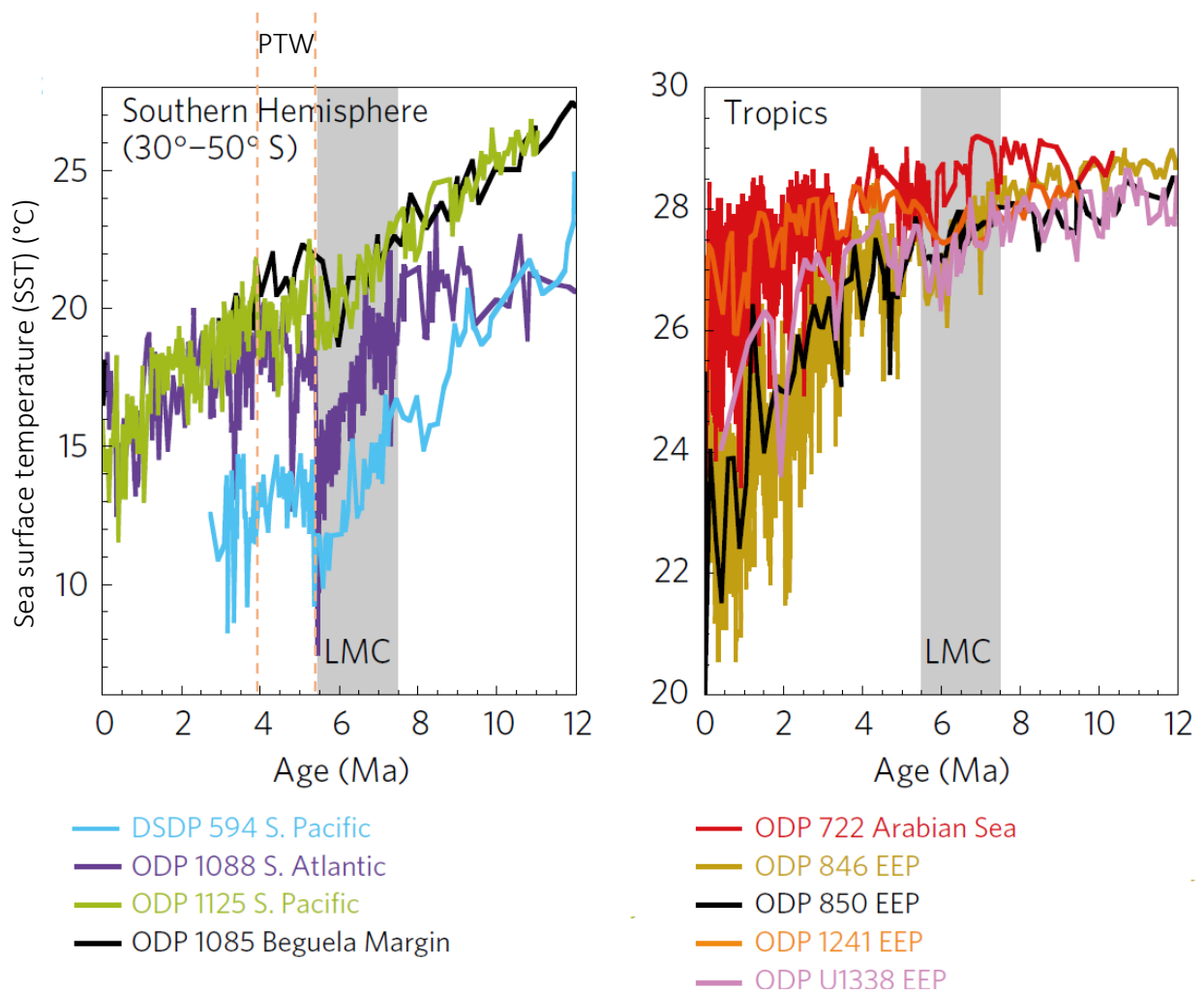
FIGURE 4.5. Sea surface temperatures over the last 12 Myrs estimated using alkenones related to haptophyte algae, based on the $U^{k'}_{37}$ index. Sea surface temperatures are given for the Southern Hemisphere (relevant to New Zealand) and the tropics for comparison. Deep

Sea Drilling Project (DSDP) 594 (light blue) was drilled of the east coast of Otago. LMC refers to the "Late Miocene Cooling". A period of transient warming at the beginning of the Pliocene is clear across South Pacific records (Labelled PTW). Figure modified from Herbert et al. (2016).

4.4.3. Available data sources

The main data sources for global deep sea benthic foram $\delta^{18}O$ records during the Cenozoic are from the publications of the Deep Sea Drilling Project (DSDP), Oceanic Drilling Program (ODP) and Integrated Ocean Drilling Program/International Ocean Discovery Program (IODP). The aim of these drilling programs was to collect deep-sea cores from areas of geological or oceanographic interest around the Earth's oceans, although in practice data are generally more concentrated in the northern hemisphere relative to the southern hemisphere (Fig. 4.6). A number of these boreholes were drilled in close proximity to New Zealand, but only a small proportion of them contain publicly available $\delta^{18}O$ data for the Cenozoic (Table 4.1). Furthermore, several do not have age-depth models and therefore cannot be utilised, and/or the available $\delta^{18}O$ values with age-depth models are derived from several different foram species, which compromises their utility as paleotemperature records because different foram species can have variable isotopic offsets (as discussed below).

In a study by Cramer et al. (2009), deep sea benthic $\delta^{18}O$ and $\delta^{13}C$ data were compiled from global publications over the last 25 years from the DSDP and ODP, grouped by oceanic basin (Pacific Ocean, Southern Ocean, South Atlantic and North Atlantic). This dataset has several advantages to data from individual holes, as ages for the isotopic data are calibrated to the international geological timescale of Gradstein et al. (2004) and values either represent monogeneric *Cibicidoides* spp. $\delta^{18}O$ values, or they have been adjusted to an equivalent value where appropriate. In the following sections I discuss and compare the pros and cons of the

92

New Zealand specific data and use of the global, Pacific and Southern Ocean compilations of
Cramer et al. (2009).
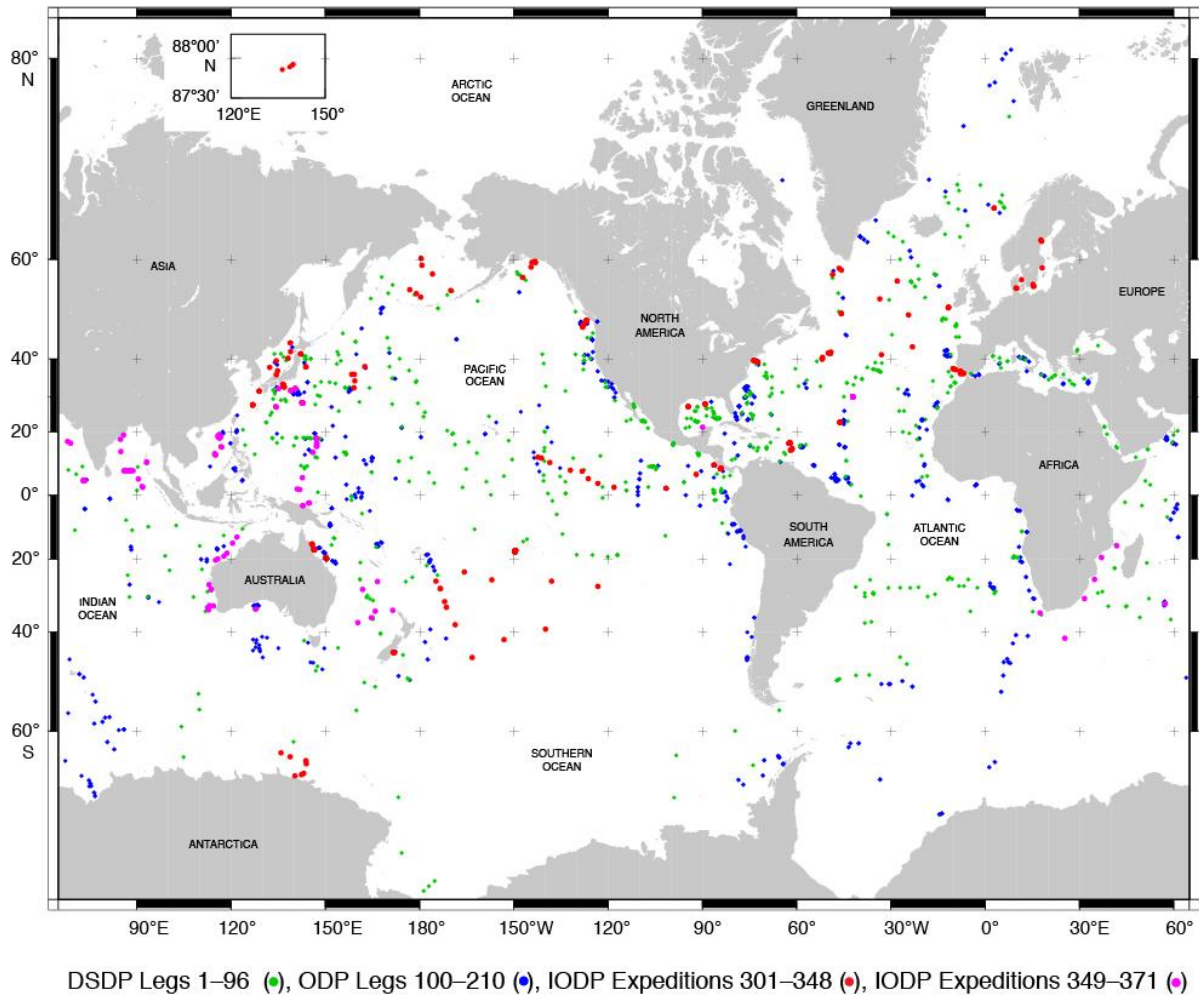
FIGURE 4.6. Map of global borehole locations from DSDP, ODP and IODP expeditions.
(Expedition maps are freely available for reproduction from
http://iodp.tamu.edu/scienceops/maps.html).

Drill Program	Hole no(s).	Age range	Reference	Comments
DSDP	277, 279, 281	Late Paleocene - Pleistocene	(Shackleton and Kennett 1975)	No depth-age model
DSDP	588, 590, 591	Early Miocene - Pleistocene	(Kennett 1986)	Ages provided
DSDP	277, 592, 593	Oligocene	(Murphy and Kennett 1985)	No depth-age model
DPSP	594	Late Pleistocene	(Nelson et al. 1986)	Depth-age model provided
ODP	1170A, 1172A	Miocene	(Ennyu and Arthur 2004)	Ages provided
ODP	1123	Late Pliocene- Pleistocene	Harris (2002)	No depth-age model

TABLE 4.1. Table of DSDP and ODP holes with publicly available deep-sea δ^{18}O records.

4.4.4. Estimating absolute oceanic temperature

Raw δ^{18}O values can be used as a proxy for oceanic temperature; however, I chose to convert δ^{18}O values to temperature in degrees Celsius (°C). This has two distinct advantages: (1) it is more intuitive when considering relative temperature changes and (2) it allows computation of a proxy for climatic variability based on temperature (discussed in the subsequent section). Several transfer functions from δ^{18}O to temperature in °C have been suggested, with the majority following a linear equation (Shackleton 1974, Visser et al. 2003), although quadratic equations have also been suggested (Marchitto et al. 2014). Both the standard and quadratic transfer equations are calibrated to the δ^{18}O of seawater on the Standard Mean Ocean Water (SMOW) scale. A value of SMOW can be estimated in the geological record by considering the average δ^{18}O trend through time (Eq. 4.6) (although this correction is minimal in the Cenozoic). The standard linear (Shackleton 1974, Visser et al. 2003) and quadratic transfer functions are relatively simple (Eq. 4.7 and Eq. 4.8, respectively). The quadratic equation has not been extended to geological estimates, however,

$\delta^{18}O_{ws}$ can be substituted for $\delta^{18}O_{pw}$, similar to the transfer function of Visser et al. (2003) (Eq. 4.7). Whereas there is evidence that linear transfer functions may both overestimate or underestimate oceanic temperature in recent to subrecent deep-ocean sediments (Marchitto et al. 2014), the quadratic transfer function has not been applied to deep time data (i.e., Pleistocene and older). To assess the difference between the linear and quadratic transfer functions over long-time intervals I calculated oceanic temperature using both transfer functions utilizing the Cenozoic Pacific Ocean *Cibicidoides spp.* (or equivalent corrected values) $\delta^{18}O$ values from the Cramer et al. (2009) data set. For both transfer functions I used the proposed Phanerozoic trajectory of SMOW from Veizer and Prokoph (2015). A plot of the derived temperature (°C) values from both transfer functions show the relationship between the two is linear over the majority of the temperature ranges observed, and decoupling at both lower and higher temperatures is relatively minor (Fig. 4.7). There is also an offset between the temperatures, likely a result of the SMOW values derived from the Phanerozoic trajectory, which are smaller than the site specific values in Marchitto et al. (2014). On this basis, I decided to convert $\delta^{18}O$ values to °C using the linear transfer function following the methodology of Veizer and Prokoph (2015), and using only data from *Cibicidoides spp.*

Eq. 4.6. Proposed Phanerozoic trajectory of SMOW (Veizer and Prokoph 2015)

$$\delta^{18}O_{pw}(\text{‰}) = -0.00003\text{‰}t^2 + 0.0046\text{‰}t$$

(where *t* is time in Ma and $\delta^{18}O_{pw}$ is Phanerozoic $\delta^{18}O$ in SMOW)

Eq. 4.7. Linear transfer function for $\delta^{18}O$ to temperature (°C) (Shackleton 1974, Visser et al. 2003)

$$T(°C) = 16.9 - 4(\delta^{18}O - \delta^{18}O_{pw} - 0.27‰)$$

(where T is temperature in degrees Celsius and $\delta^{18}O_{ws}$ in Eq. 4.8. is location specific seawater on the SMOW scale)

Eq. 4.8. Quadratic transfer function for $\delta^{18}O$ to temperature (°C) (for *Cibicidoides spp.)* (Marchitto et al. 2014)

$$T(°C) = \frac{0.245 - \sqrt{0.045461 + 0.0044(\delta^{18}O - \delta^{18}O_{ws})}}{0.0022}$$

(where T is temperature in degrees Celsius and $\delta^{18}O_{ws}$ in Eq. 4.8. is location specific seawater on the SMOW scale)
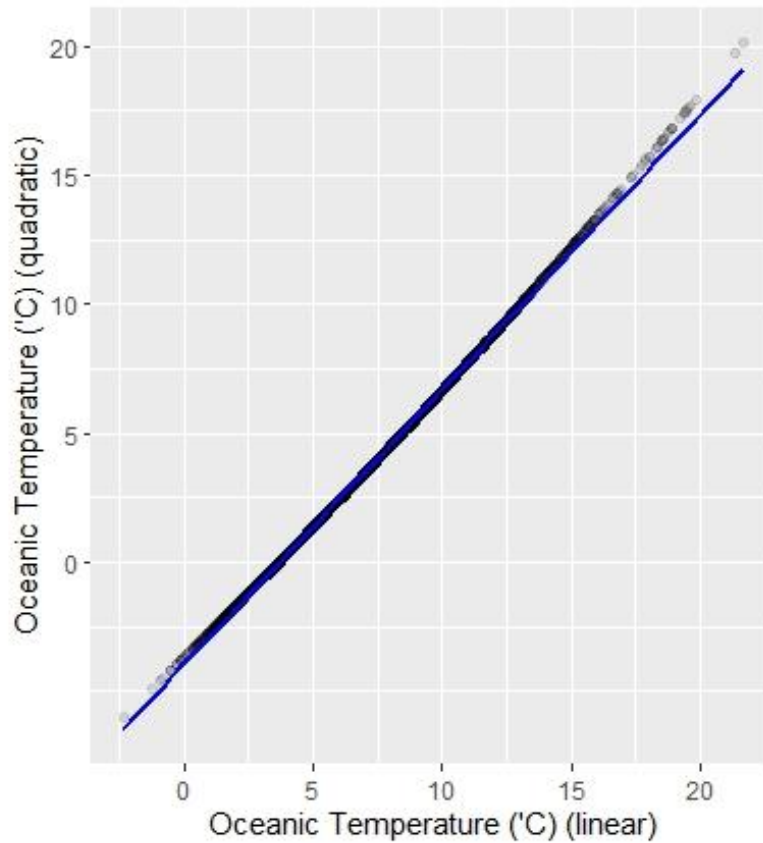
FIGURE 4.7. Plot of estimated oceanic temperatures in (°C) using the Cenozoic Pacific Ocean $\delta^{18}O$ values from the Cramer et al. (2009) dataset. The biplot compares the linear and quadratic transfer functions (Eq. 4.7 and Eq. 4.8). Individual black points represent single data points from deep-sea benthic forams. Solid blue line represents the best-fit reduced major axis linear regression line.

4.4.5. Oceanic temperature time-series

As noted above, the primary datasets for deep-sea benthic foraminifera $\delta^{18}O$ used records in this thesis include individual borehole data from the ODP and DSDP (see Table 4.1) and the global compilation from Cramer et al. (2009). Importantly and as argued previously, based on New Zealand's paleoceanographic history, data representative of cool-subtropical waters through the Cenozoic are likely more suitable in deriving regional oceanic temperature estimates for New Zealand. In this section, I compare global data with ocean

97

specific data (Pacific and Southern Oceans) and regional data from within, or in proximity, to New Zealand's oceanic territory.

The temporal resolution of the FRED dataset corresponds to stages of the New Zealand regional geological time scale (Fig. 2.1) (Raine et al. 2015). Therefore, to compare measures of biodiversity with oceanic temperature the two measures must be at the same temporal scale. The most common methods applied in paleobiological studies comparing oceanic temperature with biodiversity at the temporal scales discussed here include: (1) interpolation using a loess regression at time interval midpoints (Crampton et al. 2006a), or (2) weighted means within individual time intervals (Mayhew et al. 2012, Mannion et al. 2015). These methods differ from studies that focus explicitly on trends in oceanic temperature, as these generally interpolate to a finer resolution (~0.1 Myr), weighted to equally spaced bins. A comparison of the two methods using the Cramer et al. (2009) Pacific Ocean dataset shows the relative difference between an interpolated loess regression and weighted means for individual Cenozoic New Zealand stages is relatively modest through much of the Cenozoic (Fig. 4.8). The largest differences between the two methods are in the oldest time bins, which as discussed in subsequent Chapters (see Chapter 5 and 7) are not included due to the limited fossil data. On this basis, I decided to use interpolated loess regression as this method allows for the quantification of climatic variability, as discussed in the next section.
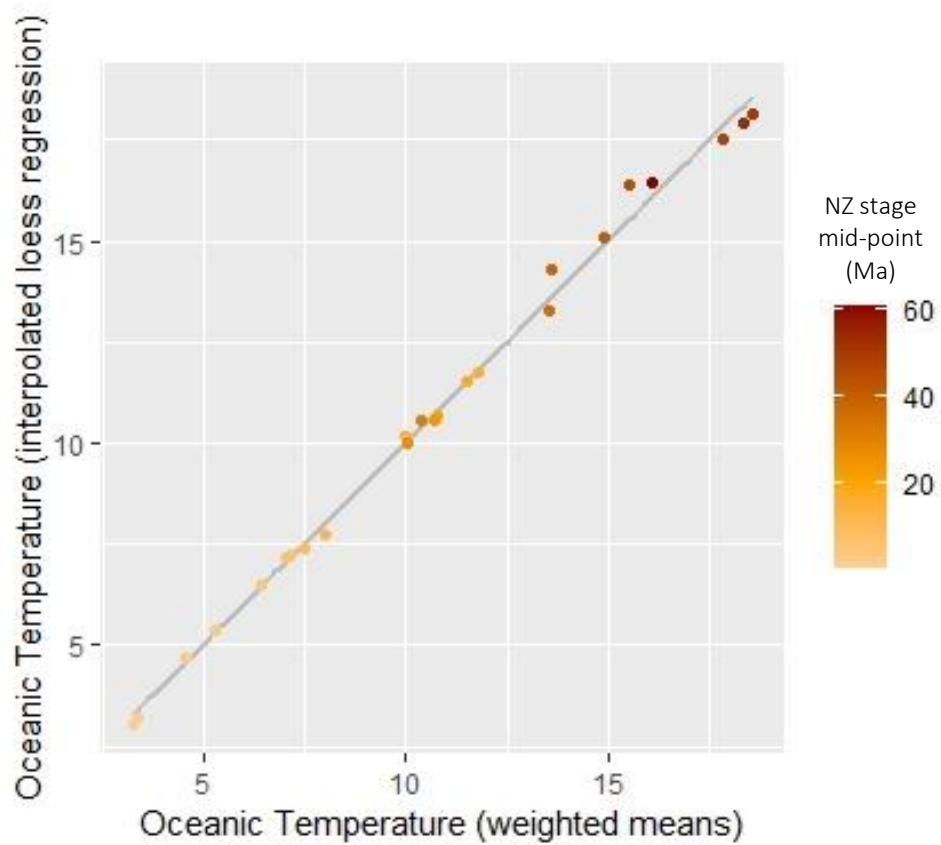
FIGURE 4.8. Comparison between representative oceanic temperatures for individual

Cenozoic New Zealand stages derived from interpolated loess regression (span = 0.1) and

weighted means of the Cramer et al. (2009) Pacific Ocean dataset. Solid grey line represents

the best-fit reduced major axis regression line. Individual points are colour coded to their age

based on the mid-points of the NZ stages they represent. Importantly, the rank of different

stages will not change between the two methods, so it is unlikely to have any significant

effect on derivative interpretations.

Qualitative comparisons between the derived oceanic temperatures from individual oceanic basins and the regional New Zealand data show broadly consistent patterns (Fig. 4.9). Time-series biplots of the interpolated loess regression for each dataset (based on first-differences, see section 4.8) illustrates that interpolated oceanic temperatures in the Pacific Ocean, Southern Ocean (north – as defined by Cramer et al. 2009) and full dataset (i.e. global) are also relatively similar, particularly the Pacific Ocean and global signal (Fig. 4.10, panel B). However, the timing and magnitude of warming during the Miocene climatic optimum is variable. The global and Southern Ocean (north) records suggest that the warming was in the middle Miocene (~17 – 14 Ma), whereas the New Zealand regional record suggests that warming was earlier and occurred between the Otaian and Altonian Stages (21.7 – 15.9 Ma). The Pacific Ocean record also supports an earlier warming between the Otaian and Altonian Stages (21.7 – 15.9 Ma) but suggests peak warmth (by only a small increase +0.22 °C) was in the Clifdenian Stage (15.9 – 15.1 Ma), whilst the New Zealand record suggests cooling immediately after the Altonian Stage (18.7 – 15.9 Ma). Although peak warmth in the Altonian (i.e., middle Miocene) seems to be consistent across the New Zealand data, the magnitude of the subsequent cooling appears to be confounded by the locations of the boreholes. Boreholes from the ODP (1170A and 1172A) lie close to the STF boundary (1772A) or south of the STF (1170A) at the time of the Miocene climatic optimum (Fig. 4.11). These two boreholes suggest a short-lived warming and more rapid cooling of a higher magnitude relative to the DSDP holes located north of the STF (588/588A-C, 590A-B and 591/591B) (Fig. 4.11). Considering that New Zealand was likely almost entirely north of the STF during the Miocene climatic optimum (Nelson and Cooke 2001), this magnitude of cooling is unlikely to be representative of average conditions across New Zealand. In addition, there is also a difference between the magnitude of decrease in the regional New Zealand record during the Plio-Pleistocene relative to the Pacific Ocean dataset. However,

this appears to be a result of the low quantity of data during several stages within this interval in the New Zealand dataset. A similar disparity related to the low quantity of data in the Southern Ocean (north) record is also observed in the Oligocene (Fig. 4.10, panel A). On this basis, oceanic temperature derived from the Pacific Ocean dataset of Cramer et al. (2009) appears to be the most suitable throughout the Cenozoic and is used in Chapters 7 and 8 for correlation with measures of biodiversity.
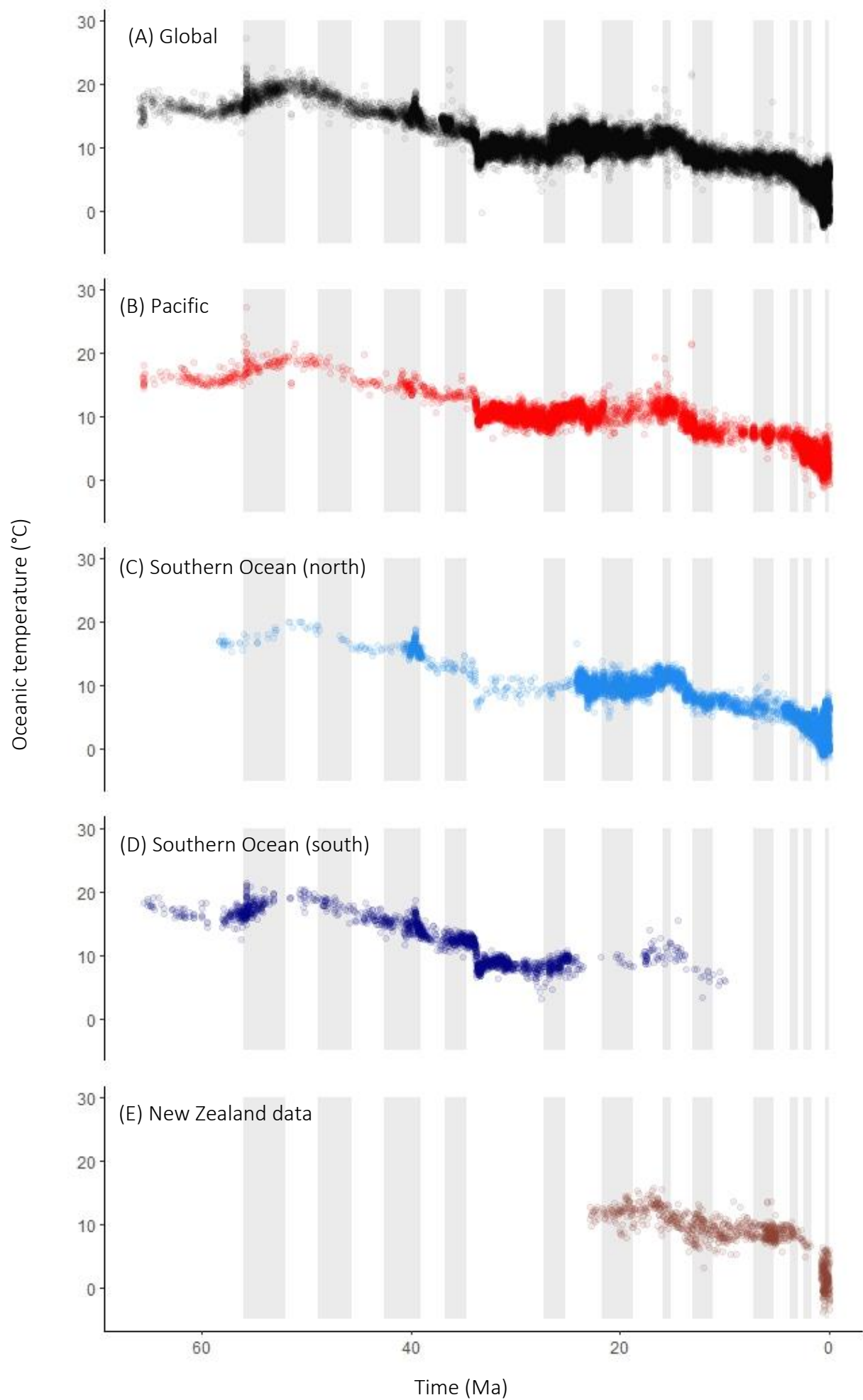
FIGURE 4.9. Comparison of deep-sea oceanic temperature trends from the global record and for individual oceanic basins. Individual data points represent individual deep-sea foram temperatures estimated from $\delta^{18}O$ converted to °C using the linear transfer function following the methodology of Veizer and Prokoph (2015). Shaded and non-shaded bars represent individual New Zealand Stages (after Raine et al. 2015).
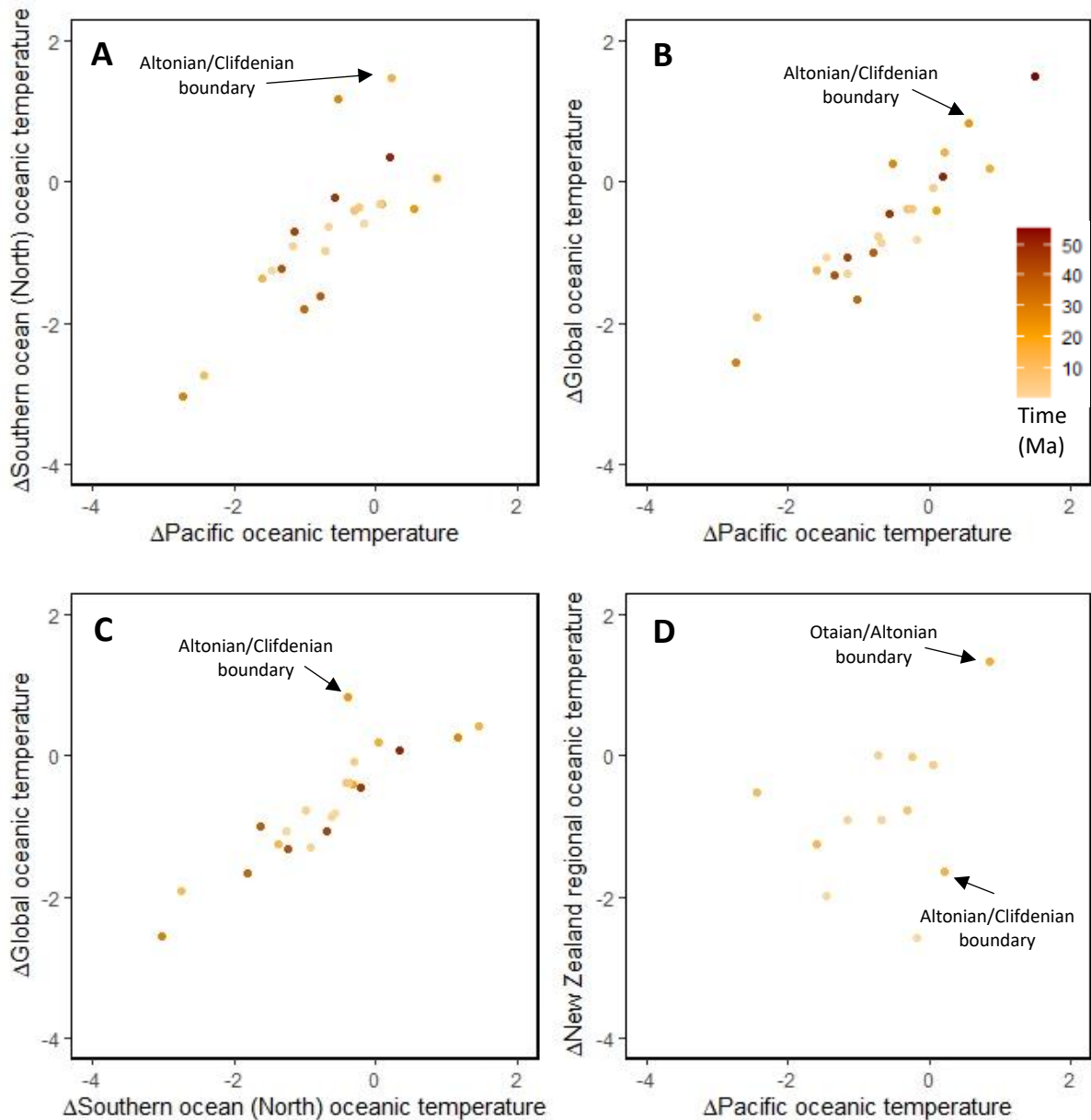
FIGURE 4.10. Comparisons between the global deep-sea oceanic temperature records with the interpolated Pacific and Southern Ocean (north) and regional New Zealand data. Individual data points represent interpolated oceanic temperatures at the mid-points of individual New Zealand stages. Plots are based on first differences (see section 4.8). Individual points are colour coded to the New Zealand stage boundary they represent (not mid-points as the oceanic temperature data are differenced). Boundaries between stages during the Miocene climatic optimum are highlighted based on disparities discussed within the main text.

104
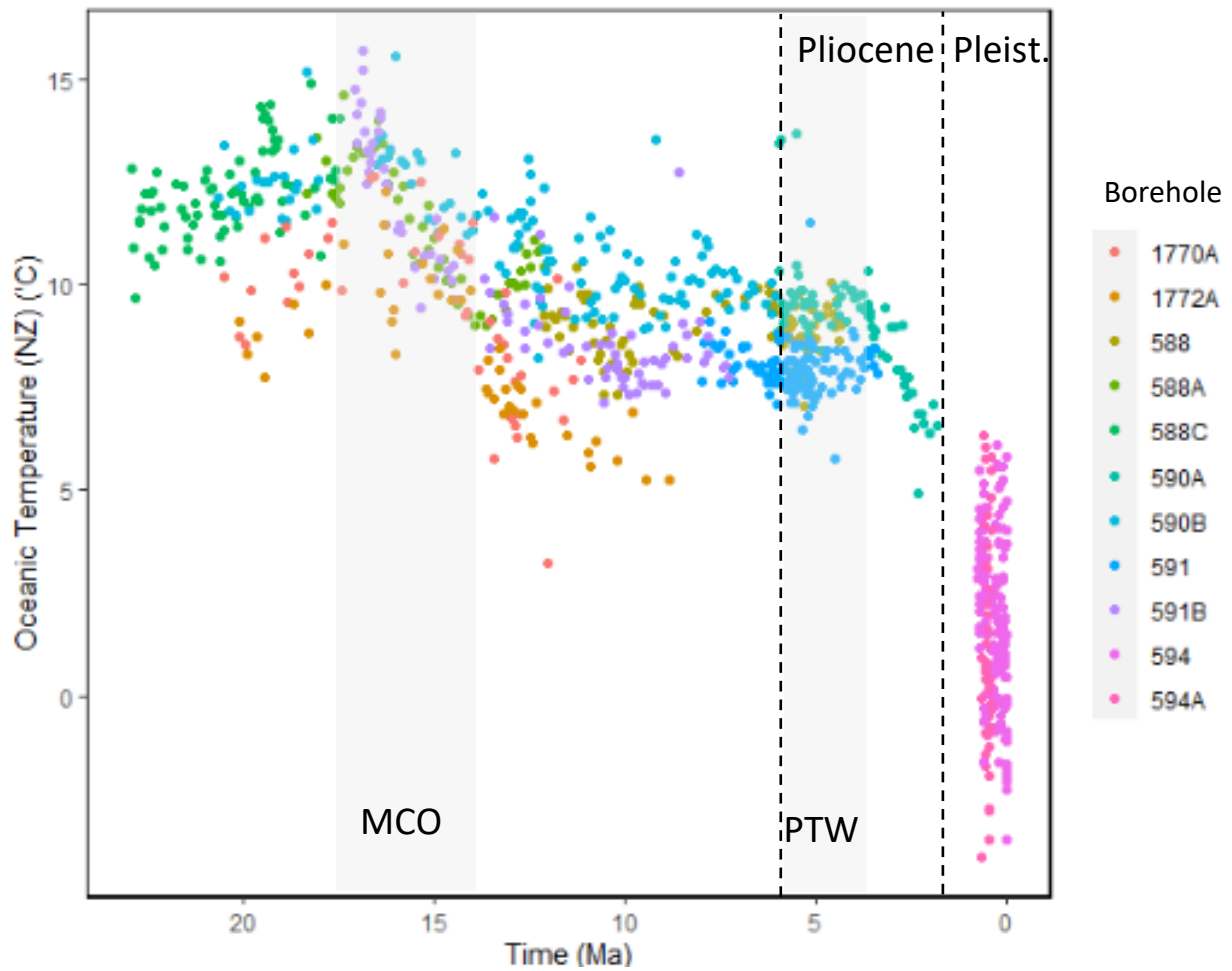
FIGURE 4.11. New Zealand deep-sea oceanic temperature colour coded to individual ODP/DSDP boreholes. Individual data points represent deep-sea foram temperatures estimated from $\delta^{18}O$ converted to °C using the linear transfer function following the methodology of Veizer and Prokoph (2015). Pliocene and Pleistocene (Pleist.) boundaries are delineated by vertical dashed lines. Miocene climatic optimum (based on the global timing, i.e., ~17 – 14 Ma) (MCO) and a phase of transient warming in the Pliocene (PTW) are indicated by the labelled shaded grey bars.

4.5. Measuring Climatic Variability

For the purposes of this study, I define climatic variability as the short-term temporal fluctuation in large scale regional to global climate and oceanic temperature. Perhaps the best example of increased climatic variability is the Milankovich-paced glacial-interglacial cycles that characterize the global Icehouse climatic state of the past ~3.3 Myrs (see Chapter 2) (Westerhold et al. 2020). Within this period global oceanic temperature fluctuated in the order of several degrees (depending on the location and temperature proxy used). Conversely, during the late Eocene until the Eocene/Oligocene boundary (~ 38 - 34 Ma), the global climate system is thought to have remained relatively stable (Mudelsee et al. 2014). As discussed in Chapter 2, climatic stability is thought to be an important control on biodiversity and macroevolution, particularly in the formation of the modern latitudinal diversity gradient (LDG). This is largely based on the theory that tropical latitudes are more climatically stable and as a result may have increased origination rates or decreased extinction rates (Jablonski et al. 2017). However, the temporal effects of increased climatic variability or climatic stability on mid latitude marine biodiversity and macroevolutionary rates are poorly understood. Part of this problem is the difficulty in quantifying climatic variability.

As noted above, the deep-sea Pacific Ocean temperature is thought to be the most applicable record for quantifying long-term fluctuations in oceanic temperature relevant to New Zealand. Therefore, any measure of climatic variability should also be based within the same oceanic basin, particularly as there is a degree of heterogeneity between oceanic basins from the Eocene onwards (see Chapter 2 and Cramer et al. 2009). The difficulty in measuring climatic variability is primarily a result of separating long-term climatic changes from short-term oscillations in climate. This problem is compounded by the change in global climatic states, such that the temporal pace of climatic oscillations varies between global climatic

states and even within individual states (i.e. 40 Ka to 100 Ka shift in glacial cyclicity during the Pleistocene).

Considering the above issues and limitations of the available datasets, the most intuitive measure of climatic variability appears to be the temporal variance of the individual data points within each time bin. Qualitatively, the increase in variance during Plio-Pleistocene, and particularly the Pleistocene, is visually distinctive (Fig. 4.9). Furthermore, the resolution of the Pacific Ocean dataset is relatively even from the middle Eocene (~45 Ma) onwards, which forms the focal time period of this thesis. Conversely, the density of the Southern Ocean records (both north and south – as defined by Cramer et al. 2009) varies significantly through the Cenozoic. As a result, the derived variance with each individual time bin may reflect the density of the data rather than any meaningful change in climatic variability. For this reason, only the variance of the Pacific Ocean dataset is used.

The variance of the raw Pacific Ocean data still includes the signal of the long-term trend in oceanic temperature. To adjust for this, I measure the variance of the residuals from the interpolated loess regression used to quantify average Cenozoic oceanic temperatures (described in the previous section). Residual plots suggest several outliers may be present within each time bin. These may represent true data points but may also represent more local climatic conditions at the specific borehole location or erroneous measurements. For the final dataset used in Chapters 6 and 7 I opt to remove outliers that fall outside the 95th percentile.

4.6. Extension of coverage-based subsampling

Coverage-based subsampling allows sampling standardization to equal completeness (coverage) rather than sampling size, as discussed earlier in this Chapter. Whereas the methods used to derive individual spatial partitions of biodiversity (see Chapter 2) are discussed in detail in the relevant main results Chapters of this thesis (Chapters 6 to 8), there are several new methods developed here that are described only briefly within following Chapters. These methods are based on the application of coverage-based subsampling to other spatial partitions of diversity, but also to other methods of quantifying evenness and temporal turnover. The aim of this section is to provide additional detail about these newly developed methods.

4.6.1. Beta diversity

Within this thesis, beta diversity is measured as the dissimilarity between multiple sites (where each site represents an individual grid cell, see Section 4.3.3) using the R package *betapart* (Baselga and Orme 2012) (see Chapter 5). However, this approach does not allow for any temporal variation in sampling between time intervals. Several studies of the fossil record have applied the principles of coverage-based subsampling to beta diversity, but these have only been adapted to pairwise methods of beta diversity (Brocklehurst et al. 2018, Penny and Kröger 2019). For this thesis I adapted the multi-site method of Baselga and Orme (2012) based on the original implementation of the coverage-based subsampling by Alroy (2010a) (see Appendix 1) using the code of Holland (2015) as a basis. This allows for the generation of a subsampled abundance frequency at a fixed level of coverage which can be used to compute beta diversity between sites. The workflow is summarised below:

1. For each individual time bin, fossil collection data are aggregated by grid cell (individual hexagrid cell), where each grid cell is treated as a sample.

2. The estimated coverage of each sample is calculated based on Good's *u*. Any sample that does not meet the set coverage level is omitted from the analysis.

3. For each sample that meets the set coverage level, the subsampled abundance frequency is calculated based on the code of Holland (2015) and converted to binary incidence data.

4. Multi-site beta diversity metrics are measured based on the subsampled abundance frequency vectors (reduced to incidence data) of all sites.

5. This subsampling process is repeated a number of times (100 times within Chapter 8 and 7) to provide an average measure of multi-site beta diversity at a fixed level of quota. The output includes a multi-site measure of Sorensen and Simpson dissimilarity and nestedness.

4.6.2. Functional evenness

Within this thesis functional evenness is defined as the distribution of species amongst functional groups. To measure evenness, I follow the method of Edie et al. (2018b), i.e., the inverse Simpson's index, normalized to the number of functional groups (Eq. 4.9). As discussed earlier, functional groups are also confounded by sample size, such that as the size of the sample increases, so does the probability of sampling rarer or hard-to-detect functional groups. This likely also confounds measures of functional evenness. Although cJ1 appears to be the best available measure of functional richness (see section 4.2.4), it cannot generate an abundance distribution that can be used to measure functional evenness. I show in Chapter 6 that once the data are sampling and spatially standardized, measures of functional richness derived from coverage-based subsampling are similar to functional richness derived from cJ1 at a fixed spatial scale. On this basis, I standardize functional evenness using coverage-based subsampling following the workflow summarized below.

1. For each individual time bin, species are assigned to functional groups based on their unique combination of functional traits within the ecospace dataset.

2. An abundance vector of species within functional groups is created based on individual occurrences at the collection-level. In Chapters 6 and 7 the abundance vector is based on a fixed spatial scale (using summed MST length).

3. The subsampled frequency of the abundance distribution of the species within functional groups is calculated based on code of Holland (2015). Functional evenness is calculated as the inverse Simpson's index divided by the number of functional groups (within the subsampling trial).

4. Th subsampling process is repeated a number of times (100 times within Chapter 6) to provide an average measure of functional evenness at a fixed coverage level.

Eq. 4.9. Functional evenness (Edie et al. 2018).

$$\widehat{FE} = \frac{1/\widehat{D}}{\widehat{FG}}$$

(where $\widehat{FE}$ is functional evenness, $\widehat{D}$ is Simpson's index and $\widehat{FG}$ is the number of functional groups within a single subsample)

4.6.3. Functional group temporal turnover

Temporal turnover measures the similarity or dissimilarity of assemblages through time. This can be applied taxonomically to determine how species composition changes through time but can also be used in the same way to measure how the composition of functional groups changes. Temporal turnover can be measured using the Bray-Curtis dissimilarity index, which quantifies compositional similarity between assemblages based on their abundance distributions. Whereas temporal turnover is unidirectional (i.e. forward in time) and Bray-Curtis dissimilarity is bidirectional, this metric has been shown to be a good estimate of temporal turnover using abundance data (Magurran and Henderson 2010). Similar

to beta diversity (see Chapter 6 for details), Bray-Curtis dissimilarity can also be partitioned into additive components: balanced variation (i.e. substitution, analogous to spatial turnover in beta diversity) and abundance gradients (i.e. subsets, analogous to nestedness in beta diversity) (Baselga 2017). As noted earlier, functional groups are also confounded by sample size, and may skew measures of dissimilarity based on raw abundance distribution. On this basis, I standardize functional group temporal turnover following the process summarized below.

1. For two consecutive time bins, species are assigned to functional groups based on their unique combination of functional traits within the ecospace dataset.

2. An abundance vector of species within functional groups is created based on individual occurrences at the collection-level for each individual time bin.

3. The subsampled frequency of the abundance distribution of species within functional groups is calculated based on code of Holland (2015) for each individual time bin.

4. Temporal turnover is calculated as the Bray-Curtis dissimilarity (using the R package *betapart*, see Baselga and Orne 2012) between the subsampled abundance distribution frequency for each time bin.

5. This subsampling process is repeated a number of times (1000 times within Chapter 8) to provide an average measure of Bray-Curtis dissimilarity at a fixed level of quota. The output includes a measure of total dissimilarity and the additive components that measure balanced-variation and abundance-gradients.

4.7. Macroevolutionary rates

Macroevolutionary rates quantify the origination and extinction of taxa, and as discussed within Chapter 2 and Chapter 7, form an essential component of paleontological studies and our understanding the driving mechanisms of macroevolution and diversity dynamics (Foote 2000, Foote 2003, Alroy 2008, Liow and Nichols 2010, Silvestro et al. 2014). However, quantifying origination and extinction rates using data from fossil occurrences or stratigraphic ranges is confounded by the incompleteness of the fossil record and temporal and spatial variation in preservation and sampling effort. As a result, any method quantifying origination and extinction must allow for these confounding effects to provide accurate rate estimates. Within this section, I discuss several methods of quantifying macroevolutionary rates in the fossil record, and their suitability for the shallow-marine molluscan fossil record of New Zealand. A full review is out of the scope of this thesis and is absent from the published literature.

4.7.1. Foote's Per-Capita Rates

The macroevolutionary rates of Foote (2000) built on earlier works of paleontologists (e.g., Raup and Sepkoski 1982, Van Valen 1984), to recalibrate taxonomic rates to simultaneously estimate origination, extinction, and preservation rates based on the first and last appearances of fossil taxa within a focal time interval. The method of Foote (2000) is based on forward and backward survivorship probabilities, i.e., the probabilities that a taxon will have a first or last appearance within an interval and omits single-interval taxa (i.e., singletons) that have several undesirable effects for modelling origination and extinction (see Foote 2000). The forward and backward survivorship probabilities are based on the number of taxa that cross the bottom boundary of the focal interval (bottom-crossers – N_b), the

number of taxa that cross the top boundary of the focal interval (top-crossers - N_t), and the number of taxa that cross both (N_{bt}) (see Fig. 4.12). The derived per-capita rates of origination ($\hat{p}$) and extinction ($\hat{q}$) are calculated as the inverse logarithmic of the ratio between taxa that cross both boundaries and taxa that only cross the top (origination) or bottom (extinction) boundaries of the focal interval, divided by the duration of the time interval. (Eq. 4.10 and 4.11).
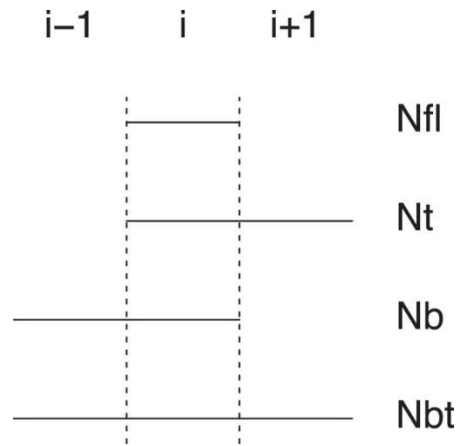
FIGURE 4.12. The four fundamental classes of taxa used by Foote (2000) to derive per-capita rates of origination and extinction. Relative time intervals are shown along the top, where i is the focal interval and i-1 and i+1 are the preceding and succeeding intervals, respectively. N_t are termed top-crossers, N_b bottom-crossers and N_{bt} top and bottom crossers. N_{fl} refers to taxa that are recorded within a single interval only.

Eq. 4.10. Per-capita rate of origination (Foote 2000)

$$\hat{p} = \frac{-\ln\left(N_{bt}/N_t\right)}{\Delta t}$$

(where N_{bt} = number of bottom and top crossers, N_t = number of top crossers and Δt = time bin duration).

Eq. 4.11. Per-capita rate of extinction (Foote 2000)

$$\hat{q} = \frac{-\ln(N_b/N_t)}{\Delta t}$$

(see description above for Eq. 4.10)

4.7.2. Alroy's Three-Timer and Gap-Filler Rates

Alroy (2008, 2014) adapted the per-capita origination and extinction rates of Foote (2000) to be used on occurrence data, arguing that rates that consider age ranges are distorted by the Pull of the Recent (see Chapter 5), the Signor-Lipps effect (which is based on the observation that the true and first and last appearance of a taxon will not be recorded as a fossil) and simple edge effects. Rather than focus on the bottom-crossers and top-crossers, Alroy's three-timer (Alroy 2008) and gap-filler rates (Alroy 2014) are based on the occurrence of taxa within the previous interval and subsequent two intervals relative to the focal interval (i.e. *i-1; i; i+1;* and *i+2*, where *i* represents the focal interval). Within this framework there are four categories of taxa: (1) two-timers (t2), which represent taxa that are present in at least the preceding and focal interval (*i*-1 and *i*); (2) three-timers (t3), which represent taxa that are present in the preceding, focal and succeeding intervals (*i*-1, *i,* and *i*+1). These two categories where initially used by Alroy (2008) to create the three-timer rates of origination and extinction. Subsequently, Alroy (2014), (3) denoted gap-fillers (*g*) as those taxa that are present in the preceding interval (*i*-1) and last interval (*i+2*) but missing from the immediately succeeding interval (*i*+1). Lastly, Alroy identified (4) part-timers (*p*), which include taxa that are missing from the focal interval (*i)* but present in the preceding (*i*-1) and proceeding interval (*i*+1). The three-timer method was updated to include gap-fillers, which Alroy (2014) suggests are important because they allow for the distinction between real extinctions and extinctions from incomplete sampling. The three-timer and gap-filler equations are shown below in Eq. 4.12 to Eq. 4.13.

Eq. 4.12. Gap-filler rate of origination (Alroy 2014)

$$\mu = \log\left(\frac{t_2 + p}{t_3 + p + g}\right)$$

(where t_2 = two-timers, t_3 = three-timers, p = part-timers and g = gap-fillers – see Fig. 4.13)

Eq. 4.13. Gap-filler rate of extinction (Alroy 2014)

$$\lambda = \log\left(\frac{t_{2,n+1} + p}{t_{3,n} + p + \hat{g}}\right)$$

(where $t_{2,n+1}$ = two-timers in the following interval (i.e., $i+1$) and $\hat{g}$ = mirror image of gap-fillers – see g Fig. 4.13), and see description for Eq. 4.12).
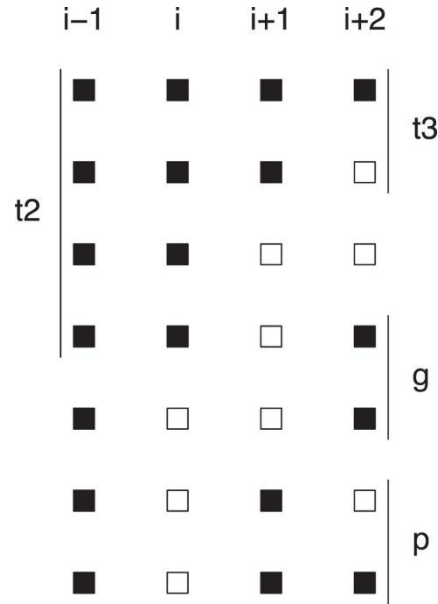
FIGURE 4.13. The five classes of sampling history used by Alroy (2008, 2014) to derive per-capita rates of origination and extinction. Relative time intervals are shown along the top, where i is the focal interval and i-1, i+1, and i+2 refer to the intervals preceding (-1) and succeeding (+1 and +2) the focal interval i. The three-timer rates (Alroy 2008) are based on the two-timer (t2) and three-timer (t3) sampling histories. The gap-filler rates (Alroy 2014) also include gap-filler (g) and part-timers (p).

4.7.3. Capture-Mark Recapture

Capture-mark recapture models were originally created to study ecological populations, particularly growth rates. This method was originally adapted to paleontological data by Nichols and Pollock (1983), although its potential had been recognized prior to this. More recently it has been revisited by Liow and Nichols (2010), particularly aided by the R package *RMark* (Laake 2013), which provides an interface between R and the MARK program, which is used to run capture-mark recapture models. The most common capture-mark recapture model applied to fossil data is the Pradel seniority model (Pradel 1996), which uses fossil occurrence data as 'sighting records' to create an encounter history (i.e., sampling history in the fossil record) (Connolly and Miller 2001, Liow and Nichols 2010, Liow and Finarelli 2014). Within the Pradel seniority model sampling histories of individual taxon can be used to calculate sampling estimates and survival and seniority probabilities (Liow and Finarelli 2014). Survival probability is based on the probability of remaining alive between time interval i and $i+1$. Conversely, seniority probability works backwards through the encounter history and estimates the probability of entering the population between the interval i and $i-1$. Within this framework, extinction probability is estimated as the compliment to survival probability (1 – survival probability, i.e. the probability of leaving the population) and origination probability, the compliment of seniority probability (1 – seniority probability, i.e., the probability of entering the population) (Liow and Finarelli 2014). Consequently, rather than estimating macroevolutionary rates within a time interval, the Pradel seniority model measures the origination and extinction probabilities between intervals.

4.7.4. PyRate

PyRate is a python based program that allows the calculation of origination, extinction, diversification and sampling rates from fossil occurrence data within a Bayesian framework, adapted from phylogenetic methods (Silvestro et al. 2014). PyRate analyses can be constrained by several models of preservation (e.g., heterogenous or homogenous) which can be optimized for the dataset used. PyRate differs from the other measures discussed here as it allows the timing of rate shifts to be estimated, rather than estimating rates within specific time bins, although this option is also possible (Silvestro et al. 2014).

4.7.5. Application of methods

All the methods for estimating macroevolutionary rates discussed here are theoretically workable and none can be discounted. Furthermore, limited research has been undertaken to compare these methods. As a result, it is difficult to assess which method is the most suitable or appropriate for the Cenozoic shallow-marine molluscan fossil record of New Zealand. One of the few reviews available compares the per-capita (Foote 2000), three-timer (Alroy 2008), and capture-mark recapture rates on detecting diversification rates in relation to preservation and tectonic history from simulated fossil records (Smiley 2018). Smiley (2018) found that the three metrics behaved differently under variable preservation scenarios, concluding that three-timer and capture-mark recapture methods produced more accurate rate estimates, but per-capita rates were better at reproducing true shifts in origination rates.

To compare the methods I used the FRED shelfal dataset (see Chapter 3, section 3.2.2) at the temporal resolution of the time bins used within Chapter 7 and 8 (which combines some New Zealand stages to accommodate for low collection numbers in certain intervals, yielding 16 time bins). I estimated the origination and extinction rates of Foote (2000) and Alroy (2008, 2014) using the R package *DivDyn* (Kocsis et al. 2019). Capture-

mark recapture rates were implemented using the R package *RMark* (Laake 2013), using the Pradel seniority model, following the method of Liow and Finarelli (2014).

For PyRate the rates were based on resampling fossil ages by site, with 10 subsamples taken. This resampling is based on collections, so that occurrences within the same interval are resampled as the same age. Without this protocol individual species from the same collection can have different resampled age ranges. Based on a subset of the data, the best-fit model was the time-variable Poisson process, which assumes that preservation rates are constant within a predefined time frame (i.e., the resolution of the time bins used here). In addition, I allowed for the preservation rate parameter prior to be estimated from the data. For each of the 10 samples I initially ran the reverse jump MCMC algorithm (see Silvestro et al. 2019) for 10 million iterations (sampling of the MCMC chain every 1000 iterations). This was increased to 40 million iterations (sampling of the MCMC chain every 5000 iterations) based on poor convergence of the MCMC trail. However, even at 40 million iterations (which took several days on a high performance cluster) the convergence was poor in each individual subsample and after the individual samples where combined. On this basis, I do not consider the PyRate analysis in our comparisons here.

To compare the capture-mark recapture rates to the Foote (2000) and Alroy (2008, 2014) rates I converted the survival and seniority probability into instantaneous rates of origination and extinction (i.e., -log), following Smiley (2018). To allow direct comparisons between the rates I temporally shift the capture-mark recapture instantaneous origination and extinction rates to the mid-points of the succeeding and preceding time intervals. As the origination rates are based on the probability of seniority between time intervals i and i-1, I shift the origination rate to the preceding interval mid-point. Conversely, as the extinction rates are based on the survival probability between time intervals i and i+1, I shift the extinction rate to the proceeding interval mid-point. We find that the Foote (2000) and Alroy

(2008, 2014) origination and extinction rates are very similar over the last 45 Myrs (Figs. 4.14 and 4.15). The capture-mark recapture rates are broadly similar to both the rates of Foote (2000) and Alroy (2008, 2014), but there are two prominent differences. Firstly, the capture-mark recapture extinction rates suggest an increase in extinction during the Oligocene (Whaingaroan Stage: 34.6 – 27.3 Ma), whereas the other rates suggest a decrease. Second, the capture-mark recapture origination rates suggest that the post-Oligocene peak in origination is during the Miocene climatic optimum (Altonian Stage: 18.7 – 15.9 Ma), whereas the other rates suggest origination rates during the Miocene climatic optimum and early Late Miocene (Tongaporutuan Stage: 11.04 – 7.2 Ma) are similar. There is no clear evidence as to which method provides the most accurate macroevolutionary rates. However, based on the broad similarities between the methods I opt to use the capture-mark recapture rates using the Pradel seniority model to quantify origination and extinction in the Cenozoic shallow-marine fossil record of New Zealand. This is based on CMR's full use of the encounter history (i.e., full sampling history) and the loss of data from only one time interval (i.e., macroevolutionary rates not calculated for first/last time interval). Alroy's methods do capture a portion of the sampling history (i.e., $i-1$, i, $i+1$ and $i+2$), but cannot calculate macroevolutionary rates in three time intervals due to the analytical equations used.
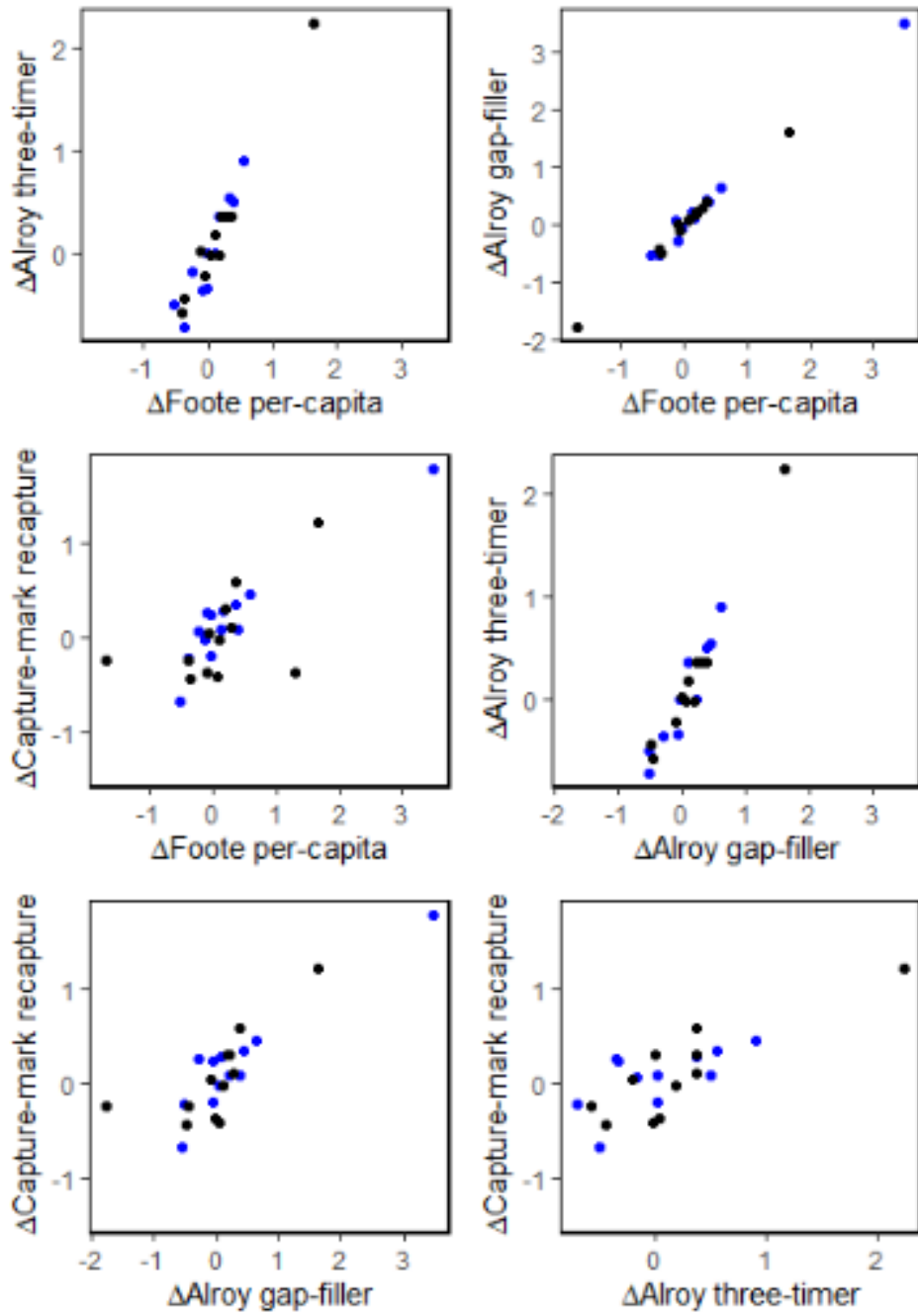
FIGURE 4.14. Comparison between macroevolutionary rates (first differences – see section 4.8) following the methods Alroy (2008) (three-timer), Alroy (2014) (gap-filler), Foote (2000) (per-capita) and Liow and Finarelli (2014)/Smiley (2018) (capture-mark recapture). Blue points represent origination rates and black points extinction rates.
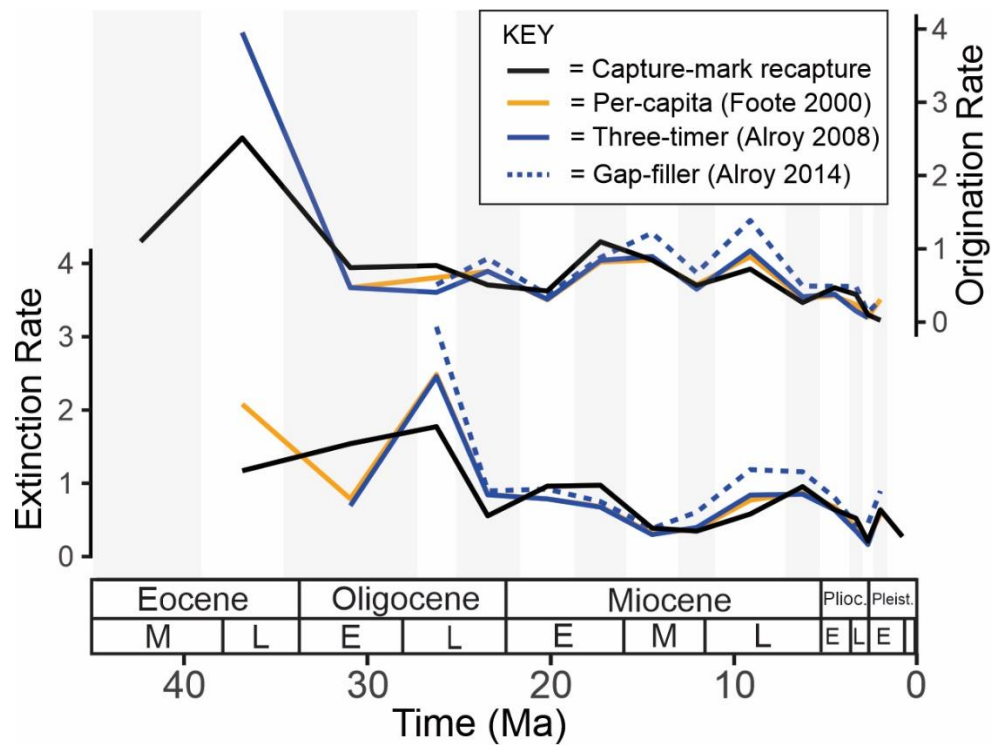
FIGURE 4.15 Time-series comparison between macroevolutionary rates following the methods Alroy (2008) (three-timer) (blue), Alroy (2014) (gap-filler) (blue-dotted), Foote (2000) (per-capita) (orange) and Liow and Finarelli (2014) (capture-mark recapture) (black). Shaded and non-shaded bar show the time-bins used within the analyses. For clarity error bars are not shown here but are shown in the capture-mark recapture analyses in Chapter 8.

4.8. Time-series correlation

Temporal data can be impacted by correlation within and between time series. Autocorrelation (or serial correlation) refers to the correlation of a time series with itself at different time lags. Cross-correlation refers to the correlation between time series at different time lags. Both can lead to spurious results when considering trends through time or correlations between variables (i.e., correlation is not causation) and can inflate type I errors. Ultimately, paleobiologists deal with considerable uncertainty and low-resolution time-series. This further complicates correlation within and between time-series, and a review of methods suitable for paleontological data is largely absent from the literature. As a result, this section focuses on three commonly used methods within the literature that reduce the likelihood of spurious correlation between time-series and provides context to their use within the following Chapters. These include: (1) detrending based on first-differences, (2) generalized-least squares autoregressive models, and (3) pairwise correlations using phase randomized surrogate data. The purpose of this section is to provide additional details on these specific methods and their suitability for the data used within this thesis.

Detrending broadly refers to the removal of a trend from a time series. Perhaps the simplest method of detrending, and most commonly used method in paleontology, is differencing. Differencing is simply the difference between values within a time series and can be lagged to varying degrees. In paleobiology the most widely applied lag is the first-difference. Other lags may be more suitable, particularly if cyclicity is suspected, and this can be assessed visually using the correlogram of an autocorrelation function (ACF). The co-variation between two variables based on first-differencing focuses on short-term, bin-to-bin variation, and the correlation can be assessed through pairwise correlation coefficients, i.e. the Spearman-rank order correlation coefficient. This is particularly useful for establishing

122

the impact of confounding effects in the fossil record on other biological or ecological measures. For example, we may expect apparent trends of regional species richness (i.e., not spatially and/or sampling standardized) to be confounded by short-term bin-to-bin variation in sampling intensity (e.g., number of collections). First-differences provides an effective way to see if the apparent diversity trend in species richness is conflated by sampling intensity between successive time bins. However, focusing on bin-to-bin variation removes any long-term trends within the data and may also remove real signals that are of biological significance and interest. For example, we may suspect that absolute oceanic temperature will have a positive effect on regional species richness in temperate ecosystems through the Cenozoic. First-differencing of oceanic temperature assumes that an equal decrease between different intervals should have an equal effect on biodiversity. However, an equal decrease in oceanic temperature within a climatic state during the Cenozoic, e.g., the Warmhouse climatic state, is likely to impact diversity differently than an equal decrease between climatic states, for example the transition from a Warmhouse to a Coolhouse climatic state. This pattern may be obscured in data that have been first-differenced.

Generalized-least squares offers a different approach to account for the dependence due to cross-correlation between time series (Hunt and Roy 2006), and can be extended to multiple regression. This method allows relaxation of some of the assumptions of ordinary-least squares, in particular, unequal error variances and the correlations between different errors (Fox and Weisberg 2018). For time series correlation, the regression residuals can be modelled as an ARMA (autoregressive moving average) process. The best fit ARMA process for an individual time-series can be evaluated using the *auto.arima* function in the *forecast* package (Hyndman and Khandakar 2008, Hynman et al. 2020) in R (based on AIC or AICc model comparison), and also through inspection of the autocorrelation (ACF), partial autocorrelation (PACF) and cross-correlation function (CCF) correlograms. This process can

be included within the generalized-least squares regression model, thus accounting for dependence among observations (Hunt and Roy 2006). This method has been implemented with several paleontological datasets for both pairwise and multiple regression. For example, Hunt and Roy (2006) used generalized least squares with an autoregressive process to demonstrate a statistically significant negative relationship between mean ostracod valve length and interpolated oceanic temperature over the last 40 Myrs, i.e., Cope's rule. In another study, Cleary et al. (2015) used an additive generalized least squares with an autoregressive process to assess the role of sampling proxy effects on skeletal completeness in the ichthyosaur fossil record during the Triassic, Jurassic and Cretaceous Periods. Similar to both Hunt and Roy (2006) and Clearly et al. (2015), I find the autoregressive AR(1) (i.e., autoregressive model of the order p = 1) model is generally the best fit for the time-series examined herein.

The final method discussed here uses phase randomized surrogate data to estimate the significance in pairwise parametric/non-parametric correlation (e.g., Pearson, Spearman and Kendall correlation coefficients) for serially correlated data. This method compensates for potential serial correlation of data, particularly for AR1 and AR2 autocorrelation structures, by adjusting the significance value (*p*). This method can also incorporate first-differencing to focus on bin-to-bin short-term trends and can be implemented in R using the *surrogateCor* function in the *astrochron* (Myers 2014) package.

All three methods discussed here are used throughout this thesis. In general, for pairwise correlations between time-series that include measures of apparent or standardized biodiversity and indices for the quality of the fossil record (i.e., number of collections or number of sedimentary rock packages), I focus on short-term bin-to-bin variation employing first-differences. We measure the strength and significance of pairwise correlations using the non-parametric Spearman rank-order correlation coefficient, with significance estimated as

standard, or using phase randomized surrogate data (particularly in Chapters 7 and 8). For pairwise time-series correlations where I anticipate that first-differencing is not appropriate, I use and compare the Spearman rank-order correlation coefficient estimating significance using phase randomized surrogate data and generalized-least squares regression with an autoregressive AR(1) structure. For multiple regression between time-series I employ generalized-least squares regression with an autoregressive AR(1) structure, and derive the best fit model based on the AICc weighting (see Chapters 7 and 8 for details).

Chapter 5

The Pull of the Recent revisited – negligible species-level effect in a regional marine fossil record

"It is time for us to move on from bias and focus and structure."

(*Steven M. Holland,* 2017)

Preface

The fossil record is our only direct evidence of the history of life on Earth, but it is also imperfect. This imperfection is primarily the result of (1) processes that lead to the generation of the available fossil record and (2) processes involved in the collection and taxonomic identification of the available fossil record. These attributes of the fossil record were predominately a theoretical problem during the early evolution of paleobiology as a subject, but became an analytical problem during the rise of quantitative paleobiology and the integration of biology and ecology to quantifying and interpreting diversity patterns in the fossil record (Valentine 1969, Raup 1972, Valentine 1973, Bambach 1977, Raup 1979, Sepkoski et al. 1981, Sepkoski 2012). The problem centers on whether the observed global diversity patterns through the Phanerozoic, particularly a two- to four-fold increase during the Cenozoic to recent, should be taken at face value or, alternatively, is the result of confounding effects in the fossil record. A prominent explanation for the observed Cenozoic increase was named the 'Pull of the Recent'. The term was originally conceived to describe several aspects of the fossil record that could increase apparent diversity toward the recent, particularly more complete sampling in the recent. Whereas a partial consensus to this debate was reached for the overall Phanerozoic diversity pattern based on the statistical similarities between independently derived diversity estimates (Sepkoski et al. 1981), the nature and magnitude of the increase in global biodiversity toward the recent, and consequently the impact of Pull of the Recent, was not resolved.

The Pull of the Recent is in part an analytical bias, as it is primarily caused by the inclusion of Lazarus taxa when diversity is quantified based on the biostratigraphic ranges of taxa. This can be negated through sampled-in-bin estimates, which are now broadly regarded as the best method of quantifying diversity where incidence data is available. However, this

127

analytical approach also entirely eliminates Lazarus taxa from analysis and therefore removes data and potentially erodes real biological signal (Sahney and Benton 2017). Given the widespread use of range-through diversity measures in previous studies, and the richness of data compilations that record taxonomic first and last occurrences, it is important to quantify the magnitude of the Pull of the Recent in the fossil record, particularly due to the contrasting results on the magnitude of diversity increase toward the recent both globally, and regionally in New Zealand (Beu 1990; Crampton et al. 2006b, 2011; Hendy 2009). Furthermore, some modern methods of quantifying origination and extinction rates are based on biostratigraphic ranges (Foote 2000, Silvestro et al. 2014), which may be confounded by the Pull of the Recent.

This chapter was inspired by the study of Jablonski et al. (2003), quantifying the impact of the Pull of the Recent in the Phanerozoic marine fossil record, using bivalves at the generic level. Their results suggest that the Pull of the Recent only accounts for 5 % of the Cenozoic increase in bivalve diversity, implying that the apparent diversity increase may be a genuine biological pattern. This chapter focuses on testing the impact of the Pull of the Recent on New Zealand's shallow-marine Cenozoic fossil record following the methodology of Jablonski et al. (2003), and is a modified version of the original paper submitted on this subject (Womack et al. 2020). The paper takes several novel approaches to assess the impact of the Pull of the Recent, including (1) quantifying the Pull of the Recent at species-level, (2) focusing on the Cenozoic Era, (3) considering smaller time intervals, and (4) by examining if particular attributes of taxa control their likely occurrence in the youngest fossil record. The findings of this paper are in agreement with global estimates of the Pull of the Recent (Jablonski et al. 2003). Co-authors to this paper include the primary and secondary supervisor of this thesis, whose input was provided through regular discussion and advice at meetings and reviewing the manuscript.

Whereas the primary aim of this chapter was to quantify the impact of the Pull of the Recent, it also focuses on the raw data, which serves as an important primer, regardless of the analytical methods used to quantify biodiversity, to understand the structure of the fossil record. Rather than focusing and emphasizing bias, it is this structure that should be utilized to guide and frame our analyses (Holland 2017). The results discussed within this chapter, further corroborate the quality of the New Zealand shallow-marine fossil record (Cooper et al. 2006, Crampton et al. 2006a, Crampton et al. 2006b, Foote et al. 2007, Clowes et al. 2020) and sets the scene for the subsequent chapters which focus on other structural attributes of the New Zealand fossil record (Chapter 6) and their interpretation (Chapters 7 and 8).

5.1. Introduction

The magnitude of a putative increase in global biodiversity toward the recent is still debated (Sepkoski et al. 1981, Smith 2007, Alroy et al. 2008, Holland and Sclafani 2015). Some argue there is a significant increase in Cenozoic to recent biodiversity in the marine (Valentine 1969, Bambach 1999, Jablonski et al. 2003, Markov and Korotayev 2007, Stanley 2007) and terrestrial (Benton 1995, Sahney et al. 2010, Sahney and Benton 2017) realms, whereas others argue there is not (Raup 1972, Alroy et al. 2001, Alroy et al. 2008). This question has major implications for all of biology and paleontology and centers on the magnitude and nature of several well-known attributes of the fossil record, including the Pull of the Recent. The Pull of the Recent was originally attributed to three specific factors that increase apparent diversity of the fossil record toward the recent (Raup 1979), namely more complete sampling in the recent, misassignment of fossils to recent taxa, and more fully developed taxonomic understanding of recent taxa. The latter two factors are the least important and can be mitigated through the use of revised and updated taxonomy. The most prominent factor, and focus of this paper, is the comparatively complete knowledge and sampling of the living biota in the recent compared with fossil biotas of the past, which extends the biostratigraphic ranges of fossil taxa and therefore increases range-through diversity estimates toward the recent through the inclusion of living "Lazarus" taxa (i.e., those with gaps in their stratigraphic records). Whereas this effect can be negated using sampled-in-bin estimates of diversity (Alroy et al. 2008, Alroy 2010), this approach also eliminates Lazarus taxa from analysis and therefore removes data and potentially erodes real biological signal (Sahney and Benton 2017).

Previous studies have argued that the effect of the Pull of the Recent is minimal (<5%) in the case of global Cenozoic marine bivalve genera (Jablonski et al. 2003) and

tetrapod genera (Sahney and Benton 2017), and therefore apparently large increases in diversity toward the recent are not simple consequences of this effect and may be real. There is a growing body of literature debating the appropriateness of using genera as proxies for species diversity (Hendricks et al. 2014, Xing et al. 2014, Wiese et al. 2016), but due to limitations of the available data, the effect of the Pull of the Recent on species-level data has not been quantified previously.

Here we test the effect of the Pull of the Recent at the regional scale and at both genus and species levels using the exemplary, well-studied, and relatively complete fossil record of Cenozoic mollusks from New Zealand (see Chapter 3, section 3.2) (Crampton et al. 2003, Crampton et al. 2006a, Crampton et al. 2011, Clowes et al. 2020). The record is highly resolved taxonomically at both species and genus levels, with approximately 40% of Holocene species known as fossils (Foote and Raup 1996, Foote 1997, Paul and Donovan 1998, Cooper et al. 2006, Crampton et al. 2006a). In addition, New Zealand has remained geographically isolated for the duration of the Cenozoic at relatively static southern midlatitudes (see Chapter 2, section 2.2) (King 2000, Crampton et al. 2006b). As a result, the majority of its Cenozoic to recent molluscan species are endemic (Beu and Maxwell 1990, Spencer et al. 2017), with <0.1% of the pre-Haweran (pre–0.34 Ma, see Fig. 5.1) shelfal molluscan species known to range beyond New Zealand (Crampton et al. 2010).

5.2. Materials and Methods

5.2.1. Materials

This study focuses on Cenozoic marine mollusks (bivalves, gastropods, and scaphopods) from New Zealand. Analyses within this study are reported for both genera and species, where subgenera are operationally elevated to genera and subspecies are excluded. Geologic ages are given in terms of New Zealand Cenozoic stages (Fig. 5.1) and reported analyses were undertaken at the resolution of these stages, yielding 24 time bins with an average duration of 2.75 Myr.

FIGURE 5.1. New Zealand Cenozoic timescale (Raine et al. 2015) with units in Ma. Analyses are undertaken at the resolution of the New Zealand stages shown here.

Three datasets were used within this study to generate biostratigraphic ranges for molluscan species and genera and determine their last known occurrences within the fossil record: (1) synoptic dataset (see Chapter 3, section 3.2.1); (2) the checklist of the living Mollusca recorded from the New Zealand Exclusive Economic Zone; and (3) the Fossil Record Electronic Database (FRED) (see Chapter 3, section 3.2.2). Details of these three datasets follow.

Synoptic Dataset.—The synoptic dataset provides biostratigraphic age ranges from the records of Cenozoic molluscan species from the New Zealand region and their inferred biological and ecological traits (Beu and Maxwell 1990, Crampton et al. 2003, Cooper et al. 2006, Crampton et al. 2006a, Beu and Raine 2009).

The Checklist of the Living Mollusca Recorded from the New Zealand Exclusive Economic Zone.—This dataset provides a list of all described species and subspecies of recent and subrecent mollusks (5187 species) recorded within the New Zealand Exclusive Economic Zone (New Zealand's oceanic territory), as compiled and described by Spencer et al. (2019). Biostratigraphic ranges of mollusks in the synoptic dataset were updated using this dataset. Only species native to New Zealand and genera that comprise at least one native species were included so as to exclude introduced genera and species.

FRED.—This is a collection-based database of fossil occurrences within New Zealand, and it has been well described elsewhere (Crampton et al. 2011, Clowes et al. 2020). Data relating to Cenozoic Mollusca were downloaded from FRED in September 2018 and subsequently revised with a newly updated version of the synonymy list employed by Crampton et al. (2006b, 2011), which was modified to take account of relevant recent literature. Ambiguous identifications were removed from the dataset (including removal of taxa with the prefixes "cf." and "aff."), and only occurrences from collections restricted to a

single time bin were retained. The FRED dataset was used to determine the last fossil occurrences of the genera and species used within this study. Taxa listed in the synoptic dataset but not identified within the FRED dataset were assumed to have a last fossil occurrence within their stage of origination as recorded in the synoptic dataset.

The final data analyzed here comprise the ranges of 1182 genera and 5223 species recorded in the synoptic dataset and their stage-by-stage occurrences as recorded in the FRED database.

5.2.2. Methods

We follow the methodology of Jablonski et al. (2003) at genus and species level and repeat analyses to consider the effects of habitat and lithification on the Pull of the Recent. To allow for the disproportionately low representation of non-shelfal taxa in the New Zealand Cenozoic record (Beu and Maxwell 1990, Rex et al. 1999, Crampton et al. 2003), we consider the effect of the Pull of the Recent for all habitats and, separately, for shelfal taxa only. Taxa are considered shelfal if they are inferred to have been confined to, or to have ranged into, shelf water depths (< ~200 m water depth) (Crampton et al. 2006b). To allow for the relative decrease of lithified versus unconsolidated collections in the Plio-Pleistocene (Crampton et al. 2006b, Hendy 2009), we consider the effect of the Pull of the Recent on a subset of FRED collections for which matrix lithification information is available (approximately one-third of the collections used within this study). Matrix lithification is classified into four subjective categories within FRED: "hard," "moderately hard," "moderately soft," and "unconsolidated." Here we dichotomize the data, such that "unconsolidated" is considered unlithified and the remaining categories are aggregated into the lithified category. We limit analyses based on lithification to only genera and species recorded within FRED collections with matrix lithification data, using biostratigraphic ranges

derived from the synoptic dataset. We also consider whether species duration has a

significant effect on presence in the youngest fossil record. Species with a long duration are

considered to have a duration >3.70 Myr; that is, those that originate before the end of the

Opoitian Stage. Species with duration equal to or <3.70 Myr are termed "short duration."

Analyses were repeated for the three taxonomic classes combined and with bivalve

genera only to allow direct comparison with results from Jablonski et al. (2003). For each

time bin, the range-through diversity was calculated, defined here as the total number of

unique taxa that range through or have a first or last occurrence within the focal time bin.

Measures of diversity reported here are not sampling standardized and are referred to as

"apparent" diversity; these measures are not intended to depict "true," sampling-standardized

diversity history. The data were filtered to identify which taxa have been recorded in the last

5.33 Myr, similar to Jablonski et al. (2003), and repeated at three progressively shorter

intervals (resolving to New Zealand stages): 3.7 Ma (Mangapanian), 2.4 Ma (Nukumaruan),

and 1.63 Ma (Castlecliffian) to recent. This is intuitively reasonable, as the average duration

of species is shorter than that of genera, so the effect of the Pull of the Recent is expected to

manifest at shorter timescales.

To determine whether particular attributes of taxa control their likely occurrence in

the youngest fossil record, we examined a number of biological traits recorded within the

synoptic dataset. Traits include shell mineralogical composition (aragonitic or calcitic), shell

size (dichotomized as <10 mm or >10 mm maximum linear dimension), habitat (shelfal vs.

non-shelfal, see previous definition of shelfal), and taxonomic class (bivalve, gastropod, or

scaphopod). Calcitic taxa are classed as such if they have a shell component that could, in the

absence of aragonite, be identified to species level (see Crampton et al. [2006b] and Foote et

al. [2015] for further details). Shell size is based on the maximum skeletal dimension of an

average adult (Cooper et al. 2006). In cases for which species-specific traits have not or could

not be determined, they were established at the genus level so long as all species within that genus displayed the same traits. The traits of the taxa with and without a fossil record within the focal time interval, where known, were tabulated in 2×2 contingency tables (excluding scaphopods from taxonomic class due to low number of known species). This was repeated for shelfal mollusks only and for bivalves and gastropods separately. We report both the maximum-likelihood chi-square (ML chi-sq., or G) test statistic and the log odds ratio to test for independence and association between the taxa missing and the taxa present within the specified intervals. The standard continuity correction was applied during the calculation of the log odds ratio to allow for zero entries (Agresti and Kateri 2011). Family-wise error was corrected using the false discovery rate, following the procedure of Curran-Everett (2000). See Agresti and Kateri (2011) for a detailed description of these analytical methods. All analyses and computations reported here were undertaken using R (R Development Core Team 2020).

5.3. Results

Of the 401 extant molluscan genera with a fossil record in the Cenozoic, 98% (100% for bivalves) are represented in the fossil record between the Opoitian (5.33 Ma) to recent, decreasing to 88% in the Castlecliffian (1.63 Ma) to recent (Fig. 5.2, Table 5.1). Of the 673 extant molluscan species with a fossil record in the Cenozoic, 98% are represented in the fossil record between the Opoitian (5.33 Ma) to recent, decreasing to 80% in the Castlecliffian (1.63 Ma) to recent (Fig. 5.2). This decrease in representation is more pronounced in molluscan species with long durations (>3.70 Myr), 74%, an effect that relates to poor exposure of deep-water facies in the youngest part of the stratigraphic record and consequent absence of preexisting, deep-water species from the youngest fossil collections. Excluding non-shelfal taxa reduces the effect of the Pull of the Recent slightly for genera and species and for species with long durations (see Table 5.1). Lithification has minimal effect on the Pull of the Recent for the subset of FRED collections with matrix lithification data, with <1% difference in the percentage of taxa (both genera and species) with a fossil record between analyses based on lithified collections only and those based on all collections with matrix lithification data (see Table 5.1 and Fig. 5.3).
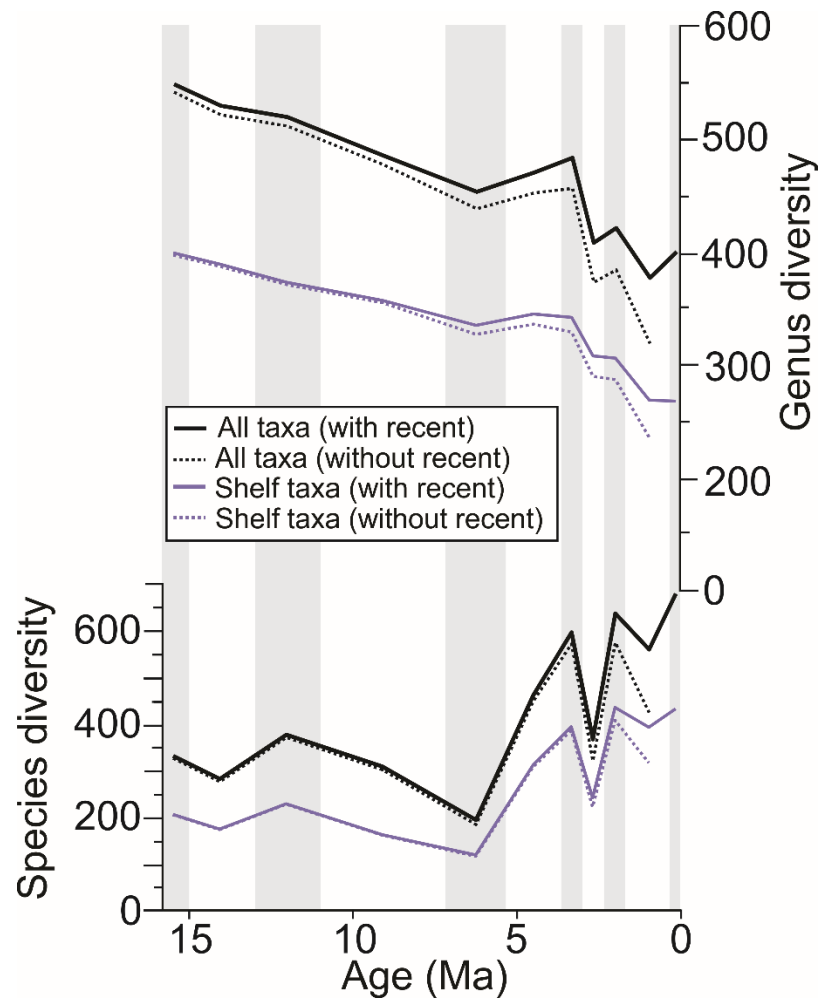
FIGURE 5.2. Genus- and species-level range-through diversity over the last 15.9 Myr (Clifdenian to recent) for all mollusks and shelfal mollusks only, with (solid line) and without (dashed line) the Pull of the Recent. Shaded bars show the duration of the time bins utilized, equivalent to the New Zealand stages.
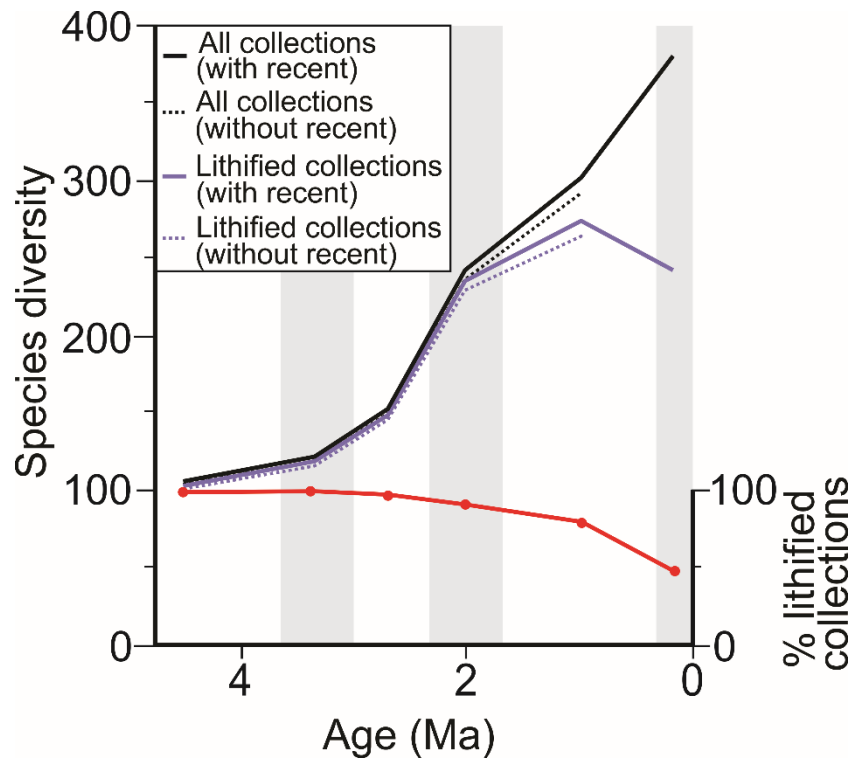
FIGURE 5.3. Species-level range-through apparent diversity over the last 5.33 Myr (Plio-Pleistocene) for molluscan species from a subset of the FRED database with matrix lithification data available (approximately a third of collections used within this study) and for lithified collections only, with (solid line) and without (dashed line) the Pull of the Recent. Relative percentage of lithified collections is based on the relative percentage of lithified to unlithified collections from the subset of FRED collections with matrix lithification data (see "Methods"). Shaded bars show the duration of the time bins utilized, equivalent to the New Zealand stages.

| Taxa | Habitat | Percentage of taxa with a fossil record (%) | | | | Total taxa that extend to recent |
		Opoitian (5.33 Ma) to recent	Waipipian (3.7 Ma) to recent	Nukumaruan (2.4 Ma) to recent	Castlecliffian (1.63 Ma) to recent	
All molluscan genera	All	98.0	96.5	93.3	88.0	401
	shelfal	98.5	98.2	95.6	90.4	271
Bivalve genera	All	100	-	-	-	139
	shelfal	100	-	-	-	133
All species	All	98.1	96.3	90.6	79.5	673
	shelfal	99.2	98.0	92.9	83.4	397
Species with long durations*	All	89.3	79.5	77.9	73.8	122
	shelfal	96.1	89.6	87	81.8	77
Species with short durations*	All	-	-	93.5	80.8	551
	shelfal	-	-	94.4	83.8	320
Genera (all collections)**	All	90.8	90.1	88.4	84.9	324
Genera (lithified collections)**		90.3	89.7	87.6	83.7	318
Species (all collections)**	All	96.6	97.1	96.9	90.2	364
Species (lithified collections)**		96.5	97.0	96.8	89.8	344

TABLE 5.1. The impact of the Pull of the Recent, expressed as the percentage of taxa with a fossil record within the listed time intervals. Results are reported in seven categories: (1) all molluscan genera; (2) bivalve genera, to allow direct comparison with the results of Jablonski et al. (2003); (3) all molluscan species; (4) all molluscan species considered to have a long duration; (5) molluscan species considered to have a short duration; and (6) genera and (7) species from all FRED collections with matrix lithification data and from lithified collections only. Results reported for categories 1-5 also consider the impact of inferred habitat, providing results for taxa from all habitats and shelfal taxa only.

* Species with a long duration are defined as species that have a first occurrence within or prior to the Opoitian and species with short durations are considered to have a have a first occurrence after the Opoitian. Therefore, the Opoitian and Waipipian are omitted from species with short durations as no species will have a first occurrence in the Opoitian and all species with a first occurrence in the Waipipian will have a fossil record within the Waipipian.

** Only includes genera or species within a subset of the FRED data that has matrix lithification data.

We report the results for maximum-likelihood chi-square test and the log odds ratio for the traits of species missing and present from the Nukumaruan (2.4 Ma) and Castlecliffian (1.63 Ma) to recent in Table 5.2. Genera and longer time intervals were not considered due to low numbers of missing taxa with known traits. There is equivocal evidence to suggest that taxonomic class, size (bivalves only), and habitat have statistically significant effects on representation in the youngest fossil record (see Table 5.2 for relevant time intervals and strength and direction of associations). However, in all cases the effect size is not large according to the classification of Chen et al. (2010), wherein a large effect size is classified as a log (log base 10) odds ratio > ~0.8, and p-values are not statistically significant at the 5% level once the effects of family-wise error are taken into account.

Time interval	Taxonomic Class	Habitat	Variable	Log odds ratio	ML chi sq. test	
					Chi-square	p
Castlecliffian to recent (1.63 - 0 Ma)	All	All	Composition	-0.083	0.327	0.567
			Size	-0.164	3.402	0.065
			Shelf level	0.194	3.654	0.056
			Class	0.233	5.910	0.015
		Shelfal	Composition	0.050	0.003	0.956
			Size	-0.185	2.290	0.130
			Class	0.149	1.525	0.217
	Bivalvia	All	Composition	0.122	0.257	0.612
			Size	-0.400	5.277	0.022
			Shelf level	0.150	0.406	0.524
	Gastropoda		Composition	0.001	0.575	0.448
			Size	-0.050	0.239	0.625
			Shelf level	0.196	2.786	0.095
Nukumaruan to recent (2.4 - 0 Ma)	All	All	Composition	0.013	0.012	0.911
			Size	-0.153	1.464	0.226
			Shelf level	0.271	3.834	0.050
			Class	0.226	2.903	0.088
		Shelfal	Composition	0.130	0.025	0.874
			Size	-0.053	0.087	0.768
			Class	0.269	2.317	0.128
	Bivalvia	All	Composition	0.232	0.546	0.460
			Size	-0.232	0.877	0.349
			Shelf level	0.463	2.679	0.102
	Gastropoda		Composition	0.252	0.338	0.561
			Size	-0.084	0.332	0.564
			Shelf level	0.195	1.440	0.230

TABLE 5.2. Values for the maximum likelihood chi-square test and log odds ratio. Calculated from 2x2 contingency tables of traits of extant species that have a fossil record within or missing from Nukumaruan (2.4 Ma) and Castlecliffian (1.63 Ma) to recent. None of the *p*-values are significant at the 5 % level once family-wise is taken into account (Curran-Everett 2000). Log odds ratios significantly above or below 0 indicate association between factors, whilst a value of 0.0 would indicate independence between factors. The direction of the interaction is determined by the ordering of variables. For the analyses reported here, a negative log odds ratio suggests that the following attributes for each variable are more

strongly associated with missing fossil record: composition – calcitic; size – larger (> 10 mm); shelf level – shelfal; class – bivalves. A positive log odds ratio would indicate the corresponding paired attribute for each variable is more strongly associated (e.g., aragonitic for composition).

5.4. Discussion and Summary

5.4.1. Trends in Apparent Diversity

The apparent range-through diversity for New Zealand molluscan genera and species shows a general pattern of long-term increase through the Cenozoic (Fig. 5.4), as observed in previous studies of global range-through diversity (Jablonski et al. 2003). In contrast to these earlier studies, however, the apparent diversity of New Zealand mollusks peaks in the Otaian and Altonian Stages (21.7–15.9 Ma) rather than in the recent. This peak in apparent diversity is associated with the influx of a large number of genera from warmer areas at this time (Beu and Maxwell 1990) and coincides with peak warmth during the early Miocene climatic optimum (see Chapter 2, section 2.4.3 for discussion on previous studies in New Zealand). In addition, the apparent range-through diversity of species is lower than that of genera (Fig. 5.4) for the majority of the Cenozoic. This is due to the use of the range-through diversity metric, which results in greater range extension with increasing taxonomic rank. This phenomenon is not observed when considering raw within-bin richness for genera and species and is simply a consequence of genera not being represented by named species within a given time bin (result not shown). Patterns of apparent species range-through diversity shown in Figure 5.3 differ from those of Figure 5.2, because species included are limited to a subset of FRED collections with matrix lithification data. Notwithstanding this, excluding unlithified collections does notably decrease apparent range-through diversity in the Castlecliffian (1.63–0.34 Ma) and Haweran (0.34–0 Ma) stages, but is unlikely to significantly alter the apparent Cenozoic diversity trajectory described previously. Detailed discussion on diversity patterns are out of the scope of this study and have been reviewed elsewhere for the Cenozoic New Zealand molluscan fossil record (Crampton et al. 2006b, 2011).
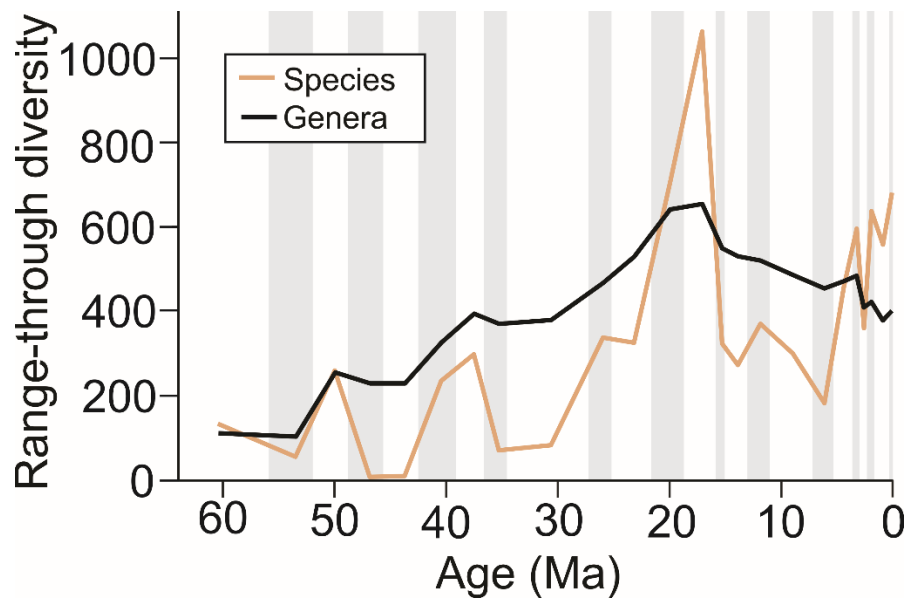
FIGURE 5.4. Cenozoic genus and species level range-through diversity for all mollusks. Shaded bars show the duration of the time bins utilized, equivalent to the New Zealand stages.

5.4.2. Impact of the Pull of the Recent

The Pull of the Recent effects <2% of genera when considering taxa missing from the youngest ~5 Myr of the New Zealand fossil record. This result is in agreement with the global estimate of the Pull of the Recent for marine mollusks reported for a similar time interval (Jablonski et al. 2003). The impact of the Pull of the Recent on species has not been examined previously, and we show here that this effect is also minor (<4%) when considering longer time intervals, similar to that of consolidated time bins used in previous studies of New Zealand shelfal mollusks (Crampton et al. 2011). The impact of the Pull of the Recent increases significantly (up to 20.5%) when considering smaller time intervals (<1.63 Ma). The taxonomic composition of mollusks missing from the youngest part of the fossil record cannot easily be explained by effects related to shell mineralogical composition, body size,

habitat, or taxonomic class. Lithification has minimal effect on the impact of the Pull of the Recent, but does have a notable effect on apparent range-through diversity in the Pleistocene.

5.4.3. Summary

Given the widespread use of range-through diversity measures in previous studies, and the richness of data compilations that record taxonomic first and last occurrences, it is important to quantify the magnitude of the Pull of the Recent in the fossil record, particularly due to the contrasting results on the magnitude of diversity increase toward the recent both globally, and regionally in New Zealand (Beu 1990; Crampton et al. 2006b, 2011; Hendy 2009). Here we demonstrate that, for a tectonically active region and for genera and species, the Pull of the Recent does not exert a strong effect on apparent diversity patterns over the Cenozoic. This implies that the apparent rise in regional marine diversity either represents a true biological signal, or results from other confounding effects not considered here, or reflects a combination of factors.

Chapter 6

The Spatial Scaling of Beta Diversity in the Shallow-Marine Fossil Record

"The crust of the Earth is a vast museum"

(*Charles Darwin,* On the Origin of Species, 1859)

Preface

The spatial distribution of the fossil record through time is uneven and constrained by tectonic processes and the current configuration of the Earth's continents. As a result, true biological patterns in the fossil record may be overprinted by variability in the spatial distribution of fossil ecosystems through time, and therefore, requires consideration in paleodiversity studies. Despite this, the impact of spatial bias on temporal diversity patterns in the fossil record has often been overlooked, and studies that have assessed the impact of the spatial distribution of paleontological samples have generally focused on global patterns of biodiversity (Alroy 2010b, Vilhena and Smith 2013, Close et al. 2017, Close et al. 2020a, Close et al. 2020b). Consequently, the impact of spatial biases at smaller spatial scales and for other partitions of biodiversity are largely unknown. This is particularly true for beta diversity, which quantifies the spatial structuring of ecological communities and is a fundamental partition of biodiversity. Furthermore, despite its common application in ecology, studies of beta diversity in the fossil record are relatively limited at regional spatial scales that provide an important intermediate spatial scale for understanding macroevolutionary processes.

Measuring beta diversity is a controversial topic, and there are a plethora of metrics (Koleff et al. 2003) with no overall consensus on which metrics should be used for addressing particular ecological questions (Anderson et al. 2011) or what actually constitutes true beta diversity (Tuomisto 2010a, b, Anderson et al. 2011). In addition, unlike alpha and gamma diversity, beta diversity cannont be physically measured and is a derived quantity. As a result, the relationship between spatial scale and beta diversity is complex, and is still debated in current ecological research (Barton et al. 2013, Antão et al. 2019). Similarly, quantifying spatial variation in the fossil record poses a problem, as different methods emphasize

149

different aspects of the distribution of fossil localities and there is often uncertainty in paleogeographical reconstructions. This is particularly true for New Zealand, where global tectonic models often place their uncertainty in New Zealand due to its unique tectonic setting, straddling the obliquely convergent boundary between the Pacific and Australian plates.

This chapter focuses on this knowledge gap and tests the spatial scaling of beta diversity in the shallow-marine fossil record, utilizing the fossil dataset generated in Chapter 3 (FRED shelfal dataset), and is modified version of the original paper submitted on this subject (Womack et al. 2020). The aims of the paper were to (1) derive suitable paleocoordinates for fossil localities, (2) assess the most suitable measures of beta diversity and spatial distribution, (3) quantify the spatial scaling of beta diversity, and (4) create a method that can accommodate for both uneven sampling and spatial distribution of fossil localities through time. Co-authors to this paper include the primary and secondary supervisor of this thesis, whose input was provided through regular discussion and advice at meetings and comments on the manuscript.

The results of this chapter suggest that beta diversity is spatially dependent at local to regional spatial scales in the Cenozoic shallow-marine molluscan fossil record of New Zealand and that uneven spatial sampling can influence recovered temporal trends in beta diversity. The methodological approach generated here is used and refined in subsequent chapters that focus on the interpretation of the resultant diversity patterns (Chapters 7 and 8).

6.1. Introduction

The correct interpretation of past patterns of biodiversity is fundamental to our understanding of how life responds to environmental change and provides a valuable context for the current biodiversity crisis (Urban 2015, Blowes et al. 2019). Biodiversity can be subdivided broadly into three main partitions: alpha, beta, and gamma diversity (Whittaker 1960). Alpha is diversity at the finest scale—the point sample—and gamma is diversity at the largest scale of observation (Patzkowsky and Holland 2012). Beta diversity represents spatial variation of species among sites at the scale of observation, and in a sense, it links alpha and gamma diversities (Anderson et al. 2011). The spatial scale of observation is contextual, and gamma diversity may reflect diversity at relatively small to global scales. Alpha and gamma diversity are, in theory, directly measurable in ecology, that is, through exhaustive investigation, whereas beta diversity reflects a derived quantity. Beta diversity is a fundamental partition of biodiversity and is central to many macroecological phenomena (Lennon et al. 2001). It has been shown to be a significant contributor to gamma diversity in ecological (Lennon et al. 2001, McKnight et al. 2007, Veech and Crist 2007) and paleoecological studies (Holland 2010, Na and Kiessling 2015) and an important factor in understanding the drivers of biodiversity (Aberhan and Kiessling 2012, Patzkowsky and Holland 2012, Hofmann et al. 2019). Previous studies within the fossil record have highlighted that there is a need for analyses at finer spatial and geographic scales (Aberhan and Kiessling 2012, Patzkowsky and Holland 2012), particularly at the regional scale, which has been identified as a biogeographically important scale for understanding macroevolutionary patterns (Vermeij and Leighton 2003). Studies of beta diversity at the regional scale are relatively limited, and even scarcer are studies on the spatial dependence of beta diversity (Barton et al. 2013). This is particularly important in the fossil record, where

temporal variation in spatial distribution of fossil localities is a known and pervasive bias (Vilhena and Smith 2013, Close et al. 2017, Close et al. 2020a, Close et al. 2020b).

Previous studies of beta diversity in the fossil record at large spatiotemporal scales have allowed for spatial variability in sampling by combining collections based on either a grid system (Crampton et al. 2011, Brocklehurst et al. 2018, Penny and Kröger 2019) or a specified distance (He et al. 2018). However, these approaches do not necessarily account for temporal variation in spatial distribution or area, and time series of beta diversity may be affected by this (see Chapter 4, section 4.3). Broadly speaking, beta diversity is commonly measured by the direct relationship between alpha and gamma components, true beta diversity *sensu stricto* (see Tuomisto 2010a, b), or by measuring compositional dissimilarity between point samples (Anderson et al. 2011). Conceptually and in ecological terms, assuming a fixed grain size (the spatial scale of individual units of observation) and variable extent (the spatial scale encompassing all units of observation), we would expect beta diversity, measured from the direct relationship of alpha and gamma, to increase with increasing area (Barton et al. 2013). The rate of increase is expected to decline from local to regional scales as fewer new species are encountered and one approaches total regional (gamma) diversity, in the manner shown by the triphasic species–area relationship (Rosenzweig 1995). Intuitively, this would be applicable to the fossil record, as we anticipate the same relationship (Sclafani and Holland 2013). The relationship between spatial scale and compositional dissimilarity is, however, more complex and dependent on taxa, grain and extent size (Barton et al. 2013) and is still debated in current ecological research (Barton et al. 2013, Antão et al. 2019).

Quantifying spatial scales in the fossil record is further complicated by uneven sampling and uncertainties in paleogeographic configuration (see Chapter 4, section 4.3). Thus, spatial distribution of the sampled fossil record may be a result of uneven sampling—

patchy preservation of the rock record and sampling of that record, for instance—or true changes in the spatial structuring of environments, such as increase, decrease, or changing location of continental shelf area or landmass. To complicate matters, the confounding effects of sampling and the quality of the fossil record may have been simultaneously affected by latent common-cause factors (Smith 2001, Crampton et al. 2011, Peters and Heim 2011). Understanding the spatial variation in beta diversity provides a unique perspective on this problem.

To our knowledge, no studies have previously focused on or corrected for spatial scaling of beta diversity in the fossil record. However, recent studies on the species–area effect in the fossil record have shown that there is a conspicuous difference in measures of global richness once standardization to spatial regions of equal size is used (Close et al. 2017, 2020a, b). These studies employ spatial standardization based on the summed minimum spanning tree (MST) length, defined as the shortest possible total distance of segments that connect a set of fossil localities (Close et al. 2017). Here we focus primarily on testing the spatial dependence of different measures of beta diversity at the regional scale and species level using the exemplary, well-studied, and relatively complete fossil record of Cenozoic mollusks from New Zealand (Crampton et al. 2006a, b, 2011). In addition, we provide a spatially standardized time series of beta diversity using the same data.

6.2.   Material and Methods

6.2.1. Materials

This study focuses on New Zealand's rich record of Cenozoic marine mollusks (bivalves, gastropods, and scaphopods), using occurrence data derived from the Fossil Record Electronic Database (FRED). FRED is a collection-based compilation of fossil occurrences within New Zealand, and it is well described elsewhere (Clowes et al. 2020). Data relating to Cenozoic Mollusca were downloaded from FRED in September 2018 and subsequently revised with a newly updated version of the synonymy list (employed by Crampton et al. 2006b, 2011; Womack et al. 2020), modified to take account of relevant recent literature (Beu et al. 2004; Beu 2006, 2011, 2012; Beu and Raine 2009) (i.e., the FRED Shelfal dataset - see Chapter 3, section 3.2.2). Analyses within this study are reported at the species level, excluding subspecies. Ambiguous identifications were removed from the dataset (including removal of taxa with the prefixes "cf." and "aff."). An additional cleaning protocol was employed to remove fossil occurrences recorded outside their known biostratigraphic ranges as documented in the synoptic dataset (Beu and Maxwell 1990; Crampton et al. 2003, 2006a; Cooper et al. 2006; Beu and Raine 2009). The synoptic dataset is separate from FRED and comprises expert interpretations of the biostratigraphic ranges of mollusk species; it is taxonomically highly vetted, having been synthesized from the known fossil record by just two paleontologists (A. G. Beu and P. A. Maxwell) over many years. In contrast, the FRED dataset, although also vetted here, may still contain spurious or erroneous identifications that could bias measures of beta diversity based on dissimilarity (see Chapter 3, sections 3.2.1 and 3.2.2).

We restrict our analysis to level-bottom shelf-dwelling taxa to allow for the disproportionately low representation of non-shelfal taxa and environments in the New

Zealand Cenozoic stratigraphic record (Beu and Maxwell 1990). Shelfal taxa are defined here as those inferred to have been confined to, or to have ranged into, shelf water depths (< ~200 m water depth) (Crampton et al. 2006b). We exclude data from the Chatham Islands due to their large distance from New Zealand's mainland for most of the Cenozoic (> ~700 km). Geological ages are given in terms of New Zealand Cenozoic Stages (Fig. 6.1), and time bins used within this study are based on individual and combined stages, yielding 13 time bins (Fig. 6.1, Table 6.1) with average duration of 3.5 Myr. Our analyses are restricted to occurrences from collections restricted in age to a single time bin. The final dataset analyzed here comprises 26,873 occurrences of 1713 species from 4173 collections.

FIGURE 6.1. New Zealand Cenozoic timescale (after Raine et al. 2015), with units in Ma.

Most analyses are undertaken at the resolution of the highlighted time bins shown here

(shaded and unshaded blocks in the right-hand column); a few analyses for the past 5 Myr are

undertaken at the resolution of individual New Zealand Plio-Pleistocene stages.

Time bin	Base and top age (Ma)	Midpoint (Ma)	Duration (Ma)	NZ Stages	Total collections	Total occupied grid cells	Maximum summed MST length (km)
1	0 – 2.40	1.20	2.40	Recent, Haweran, Castlecliffian	1705	76	4640
2	2.40 – 5.33	3.87	2.93	Nukumaruan, Mangapanian, Waipipian, Opoitian	781	68	3863
3	5.33 – 7.20	6.26	1.87	Kapitean	248	31	2158
4	7.20 – 11.04	9.12	3.84	Tongaporutuan	333	41	2571
5	11.04 – 13.05	12.04	2.01	Waiauan	89	24	2219
6	13.05 – 15.90	14.47	2.85	Lillburnian, Clifdenian	95	22	2290
7	15.90 – 18.70	17.30	2.80	Altonian	308	51	3430
8	18.70 – 21.70	20.20	3.00	Otaian	142	31	2824
9	21.70 – 25.20	23.45	3.50	Waitakian	104	28	2322
10	25.20 – 27.30	26.25	2.10	Duntroonian	108	32	3069
11	27.30 – 34.60	30.95	7.30	Whaingaroan	40	16	1792
12	34.60 – 39.10	36.85	4.50	Runangan, Kaiatan	58	16	1349
13	39.10 – 45.70	42.40	6.60	Bortonian, Porangan	116	20	2818

TABLE 6.1. Time bins used for measuring the spatial scaling of beta diversity and the corresponding New Zealand Stage names (after Raine et al 2015). Total collections are based on the number of unique collections and total occupied grid cells the number of occupied grid cells within each individual time bin. Maximum summed MST lengths are calculated using the central coordinates of all occupied equal area grid cells in each time bin. Unstandardized collection-based analyses are calculated using all collections and grid-based analyses the total occupied grid cells in each individual time bin.

6.2.2. Paleocoordinates

To derive measures of beta diversity that reflect true spatial and temporal variation in biodiversity, we calculated paleocoordinates for individual collections based on their modern latitudes and longitudes. This is particularly important for New Zealand, as it straddles the obliquely convergent boundary between the Pacific and Australian plates, with reconstructions suggesting up to ~800 km of relative plate motion since ~43 Ma and 80°–90° clockwise rotation on the still-active subduction zone on the east coast of New Zealand's North Island since ~20 Ma (Lamb 2011) (see Chapter 2, Fig. 2.5). Paleocoordinates were reconstructed using the software GPlates (Müller et al. 2018). Reconstructions were based on a rigid-block model of New Zealand using relative finite rotations, continental polygons, and coastlines (Matthews et al. 2016) placed in the global paleomagnetic reference frame (Torsvik et al. 2012). The use of a paleomagnetic reference frame for New Zealand has been shown to place strong constraints on both the tectonic evolution of the plate-boundary zone and the dynamical controls of crustal evolution (Lamb 2011). See Figure 6.2 for examples of modern-day and reconstructed paleocoordinates for the Altonian Stage (18.7–15.9 Ma), and paleocoordinates for the Tongaporutuan Stage (11.04–7.2 Ma); Figure 6.3 shows reconstructed paleocoordinates for all time bins (and see Fig. 2.5 for reconstructed paleobathymetry through the Cenozoic).

A Recent geography (Altonian collections)

- • Collection locality
- Australian plate
- Pacific plate

B Altonian (15.9 - 17.3 Ma)

$n_{collections} = 308$
$n_{grid} = 51$

- • Reconstructed collection locality
- Occupied grid cell
- — Plotted MST tree

C Tongaporutuan (7.2 - 11.04 Ma)

$n_{collections} = 333$
$n_{grid} = 41$
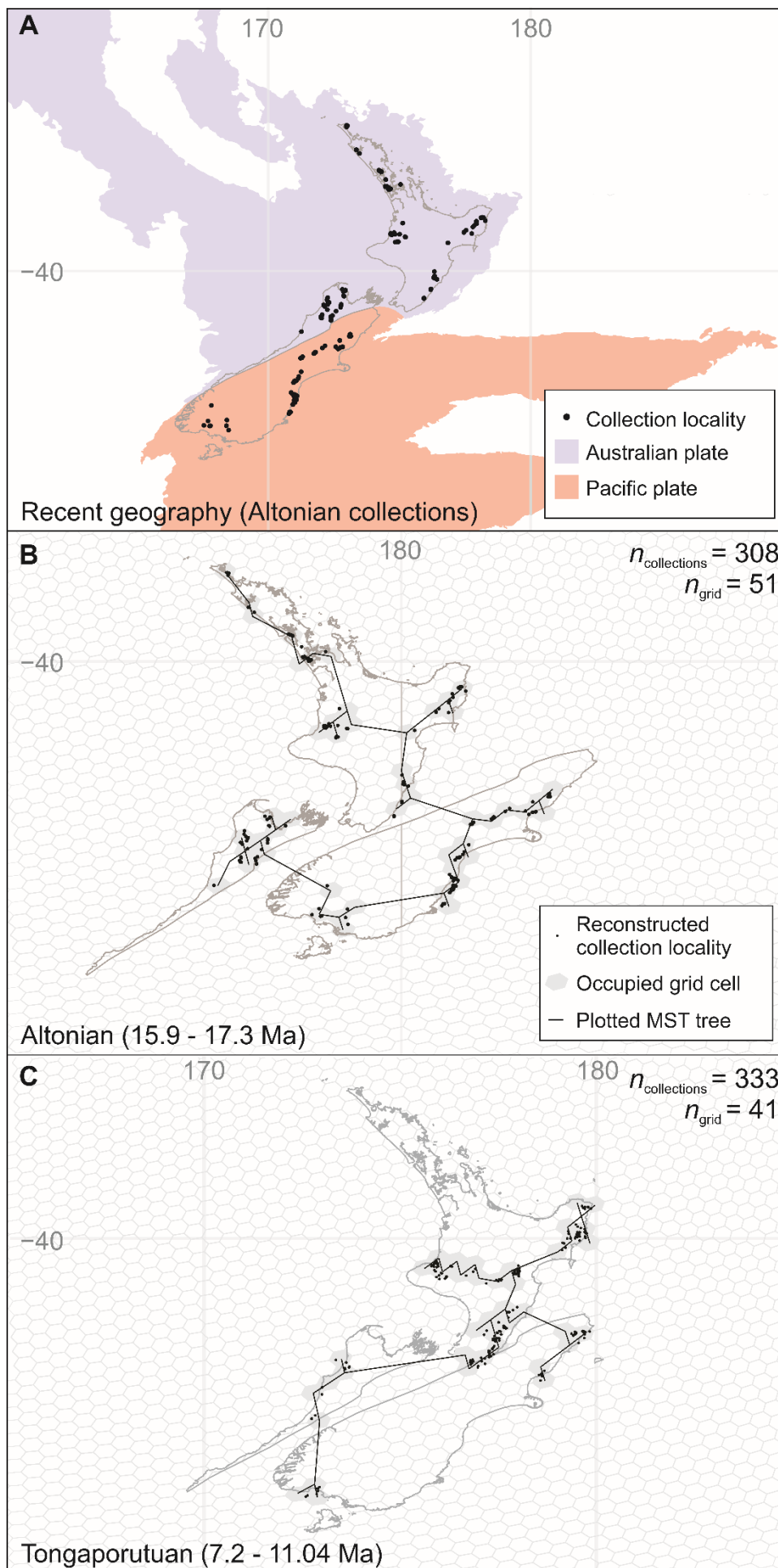
FIGURE 6.2. Recent and example paleogeographic maps of New Zealand showing fossil collection localities. (A) Recent geography of New Zealand showing the position of continental fragments lying on the Australian and Pacific plates and Altonian (Pl) (18.7–15.9 Ma) fossil collection localities. (B,C) Reconstructed paleocoordinates with the rotated fragments of the modern New Zealand coastline outline for reference and Altonian (Pl) (18.7–15.9 Ma) (B) and Tongaporutuan (Tt) (11.04–7.2 Ma) (C) fossil collection localities. Maps (B, C) show the equal-area penta-hexagonal grid with occupied grid cells shaded and minimum spanning tree (MST). The example MSTs are based on the central coordinates of all occupied grid cells.
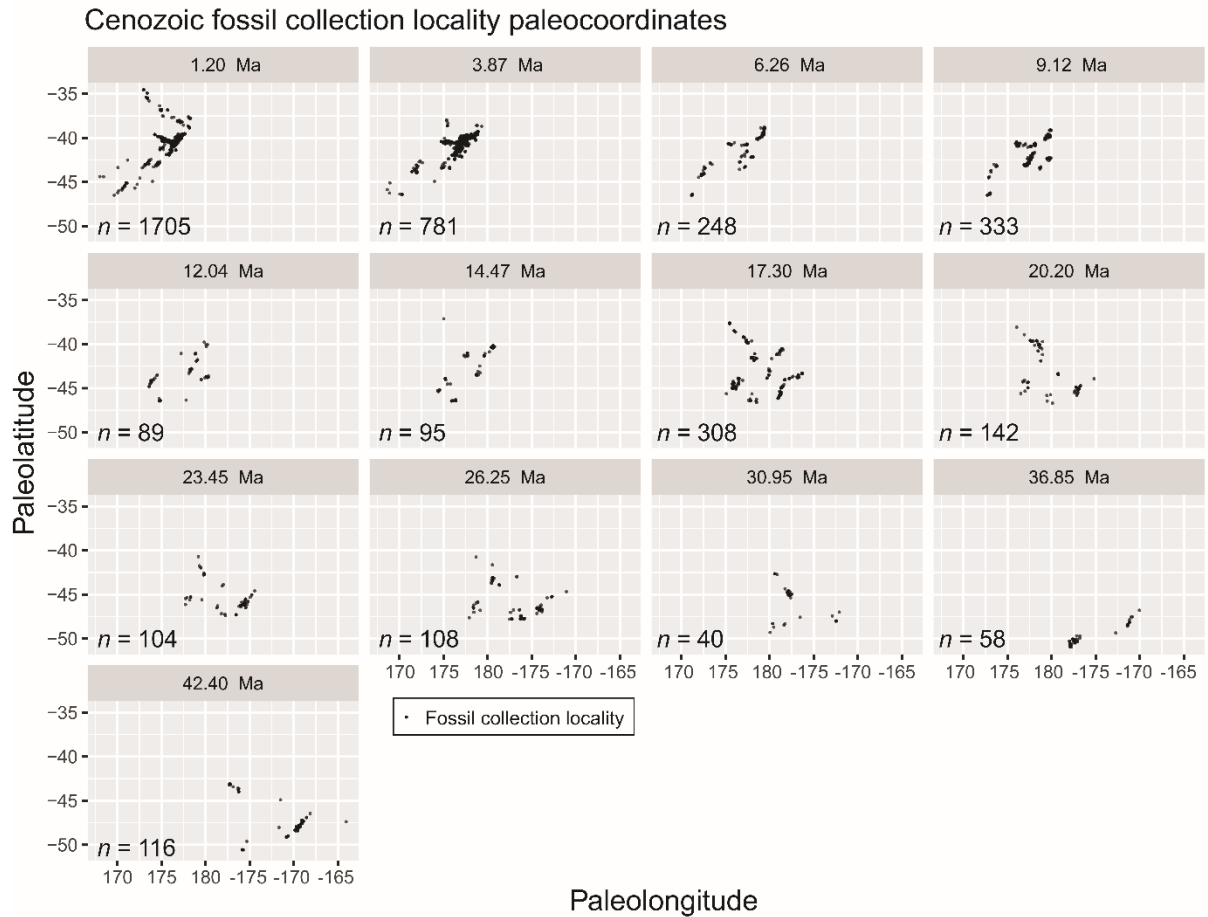
FIGURE 6.3. Cenozoic fossil collection locality paleocoordinates for each time bin, with number of collections denoted by *n*. Time bins are labelled by the midpoint of the bin (see Fig. 6.1 and Table 6.1).

6.2.3. Measuring Spatial Variation

Quantifying spatial variation in beta diversity requires some measure of distance, area, or geographic dispersion. This poses a problem, as different methods emphasize different aspects of the distribution of fossil localities (Close et al. 2017). Principally, we are concerned with variation in area. Previous studies have focused on measuring spatial variation using the area of the convex-hull, great-circle distances, summed MST length (Alroy 2010b; Close et al. 2017, 2020a,b), or a grid-based system (Crampton et al. 2011;

161

Penny and Kröger 2019). Our data represent shelfal fossil communities; ideally, estimates of spatial area should represent shelfal area. Estimating shelfal area for New Zealand is difficult, as there are no spatially or temporally well-resolved, palinspastic paleogeographic maps for New Zealand. Based on the assumption that paleocoordinates for New Zealand shelfal fossil collections approximately fringe the paleocoastline, particularly at larger spatial scales, the area of the convex-hull would be biased substantially by the area formerly occupied by land (see Chapter 4, section 4.3). In addition, the convex-hull has been shown to be highly sensitive to sampling size and to consistently overestimate species geographic ranges (Burgman and Fox 2003). We therefore adopt a grid system as a proxy for spatial area, based on an equal-area penta-hexagonal grid generated using the R (R Development Core Team 2020) package icosa (Kocsis 2020). Studies of beta diversity in the fossil record at the global scale have opted for a grid of side 111 km (Penny and Kröger 2019). We adopt a smaller grid of side 23 km, because our data are at the regional scale. The grid system provides a measure of area occupied by collections, but does not capture any information regarding the geographic dispersion of the data—10 grid cells could be contiguous or widely dispersed across a continent. Therefore, we also consider the summed MST length as a measure of paleogeographic dispersion, calculated using the R functions of Close et al. (2017). Our calculations of summed MST length are based on the central coordinates of the individual grid cells, reducing biases related to clustering of collections in well-sampled areas. Whereas summed MST length may still be partially biased by landmass, it has been shown to correlate well with grid-cell occupation in the fossil record, can approximate complex distribution shapes (e.g., surrounding a landmass), and provides a good compromise in summarizing the spatial distribution of fossil localities based on their coordinates (Alroy 2010; Close et al. 2017).

6.2.4. Measuring Beta Diversity

Measuring beta diversity is a controversial topic, and there exist a plethora of metrics (Koleff et al. 2003) with no overall consensus on which metrics should be used for addressing particular ecological questions (Anderson et al. 2011) or what actually constitutes true beta diversity (Tuomisto 2010a,b; Anderson et al. 2011). Conceptually, the spatial relationship between alpha and gamma diversity is well documented in ecology (Rosenzweig 1995, MacArthur and Wilson 2001, Scheiner 2003, Drakare et al. 2006) and paleontology (Sepkoski 1976, Barnosky et al. 2005, Sclafani and Holland 2013, Close et al. 2017). Conversely, the spatial dependence of beta diversity is relatively unknown (Barton et al. 2013). The problem is compounded when applying beta diversity to the fossil record due to confounding effects of uneven preservation and sampling of the rock and fossil records through time. A full review of beta diversity is out of the scope of this paper, and we follow Anderson et al. (2011) in their broader definition of beta diversity, discussed in the following paragraph.

Beta diversity can be separated into two types, turnover and variation (Vellend 2001; Anderson et al. 2011). "Turnover" measures the change in community composition and structure from one sampling unit to another, spatially, temporally, or over an environmental gradient; whereas "variation" is measured among all possible pairs of units, without reference to any particular gradient or direction (Anderson et al. 2011). Turnover is often measured as similarity, but is also commonly expressed as dissimilarity (simply 1-similarity or by rearranging the formula; see Koleff et al. 2003). In addition, turnover can be subdivided into species replacement (commonly termed "spatial turnover") and nestedness of assemblages (species loss) (Baselga 2010). For example, take two sites of varying species richness. The first site has a higher species richness and the second site is compositionally a subset of the first site's species. The variation in species composition between the two sites is due to

163

differences in species richness, not species replacement, and forms the core concept of nestedness (Ulrich and Almeida-Neto 2012). Subdividing turnover into species replacement and nestedness components can help reveal underlying drivers of assemblage difference that may be obscured when they are combined (Wright et al. 1997; Penny and Kröger 2019). We are interested in both the overall variation in beta diversity and the turnover across all shelfal gradients, and therefore both concepts are useful and applicable.

Fossil collection data contained in FRED are generally binary (presence–absence) and are treated as incidence data, and therefore we only discuss methods relevant to this data type. On this basis, we consider three main classes of measures of beta diversity for incidence data and their spatial scaling (Table 6.2).

1. Classical metrics, which are derived directly from the relationship between alpha and gamma diversity and are more suited to measuring overall variation. We adopt the original multiplicative (β_{Whit}) (Whittaker 1960) and additive (β_{Add}) (Lande 1996) definitions of beta diversity as measures. Both and additive and multiplicative measures of beta diversity have been shown to be effective in the fossil record (Holland 2010; Aberhan and Kiessling 2012).

2. Pairwise measures, which are based on similarity indices between a pair of sites or an average of all pairs, and quantify turnover (Anderson et al. 2011). We adopt pairwise measures of the Sørensen (β_{Sor}) and Simpson (β_{Sim}) dissimilarity as expressed by Koleff et al. (2003).

3. Multiple-site measures, which are derived from multiple-site equivalents of the matching components used in certain pairwise measures and also quantify turnover (Anderson et al. 2011). We adopt multisite implementations of the Sørensen (β_{MSor}) and Simpson (β_{MSim}) pairwise dissimilarity measures (Baselga 2010; Baselga and Leprieur 2015).

Sørensen dissimilarity measures compositional differences arising from species replacement and species loss, whereas Simpson dissimilarity measures only species replacement. The nested component of turnover can be calculated simply by subtracting the Simpson dissimilarity measure from the Sørensen dissimilarity measure (β_{Nest} for pairwise and β_{MNest} for multisite) (Baselga 2010). Pairwise measures are calculated in R using the package vegan (Oksanen et al. 2007) and multisite measures using betapart (Baselga and Orme 2012). To accommodate for local variation in preservation and sampling of collections and facilitate calculation of area and distribution, we assign collections in each time bin to their corresponding grid cell based on their paleocoordinates. Collections in each grid cell are aggregated and reduced to binary incidence data and treated as a single sample; this protocol reduces the impact of local clustering of data around richly sampled localities.

Beta diversity measure	Notation (this paper)	Equation	References
(1) Classic measures			
Multiplicative	β_{Whit}	$\dfrac{S}{\bar{\alpha}}$	(Whittaker 1960)
Additive	β_{Add}	$S - \bar{\alpha}$	(Lande 1996)
(2) Pairwise measures			
Sørensen dissimilarity index	β_{Sor}	$\dfrac{b + c}{2a + b + c}$	(Koleff et al. 2003, Lennon et al. 2001, Sorensen 1948)
Simpson dissimilarity index	β_{Sim}	$\dfrac{\min(b, c)}{a + \min(b, c)}$	(Koleff et al. 2003, Lennon et al. 2001, Simpson 1943)
Nestedness-resultant dissimilarity	β_{Nest}	$\dfrac{\max(b, c) - \min(b, c)}{2a + \min(b, c) + \max(b, c)} \times \dfrac{a}{a + \min(b, c)}$	(Baselga 2010)
(3) Multi-site measures			
Sørensen dissimilarity index	β_{MSor}	$\dfrac{\left[\sum_{i<j} \min(b_{ij}, b_{ji})\right] + \left[\sum_{i<j} \max(b_{ij}, b_{ji})\right]}{2[\sum_i S_i - S_T] + \left[\sum_{i<j} \min(b_{ij}, b_{ji})\right] + \left[\sum_{i<j} \max(b_{ij}, b_{ji})\right]}$	(Baselga 2010)
Simpson dissimilarity index	β_{MSim}	$\dfrac{\left[\sum_{i<j} \min(b_{ij}, b_{ji})\right]}{[\sum_i S_i - S_T] + \left[\sum_{i<j} \min(b_{ij}, b_{ji})\right]}$	(Baselga 2010)
Nestedness-resultant dissimilarity	β_{MNest}	$\dfrac{\left[\sum_{i<j} \min(b_{ij}, b_{ji})\right] - \left[\sum_{i<j} \max(b_{ij}, b_{ji})\right]}{2[\sum_i S_i - S_T] + \left[\sum_{i<j} \min(b_{ij}, b_{ji})\right] + \left[\sum_{i<j} \max(b_{ij}, b_{ji})\right]} \times \dfrac{[\sum_i S_i - S_T]}{[\sum_i S_i - S_T] + \left[\sum_{i<j} \min(b_{ij}, b_{ji})\right]}$	(Baselga 2010)

TABLE 6.2. Equations and notations for measures of beta diversity adopted within this chapter.

(1) S = total species richness (gamma diversity) and $\bar{\alpha}$ = alpha diversity (i.e., collection or within cell diversity).

(2) Components *a, b* and *c* refer to the spatial distribution of a species across a pair of samples as defined by Koleff et al. (2003): a = total species shared by both samples; b = total number of species within the neighbouring sample not shared with the focal sample; c = total number of species within the focal sample not shared with the neighbouring sample.

(3) Multiple site analogues of *a*, *b* and *c* as defined by Baselga (2010): S_i = total number of species in site *i*; S_T = total number of species in all sites; bij = the number of species exclusive to sites *i* and *j*.

We run our analyses with all available collections and a resampling procedure to reduce the impact of uneven numbers of collections per time bin. This resampling procedure is based on 40 collections resampled 1000 times, without replacement. The number of collections per trial is based on the minimum number of collections in the most poorly sampled time bin (40 collections in the Whaingaroan Stage [Lwh]: 34.6–27.3 Ma) (see Fig. 6.4 for a diagram that explains the workflow). No spatial constraints are applied to the resampling procedure, allowing for spatial standardization to be calculated afterward at any chosen level. For each trial, we assign the collections to their grid cells, as described earlier, forming the sampling units. We select one of these grid cells as the focal sample and calculate the geodesic distance to all other sites, and then repeat the process, treating each occupied grid cell as the focal sample in turn. During each iteration, we calculate measures of beta diversity and measures of spatial variation incrementally, moving outward from the focal sample, adding one sample at a time until the most distant of the 40 samples is included. This allows beta diversity to be calculated in reference to a spatial gradient, rather than randomly, thereby avoiding measuring beta diversity for samples where geographic distance and calculated spatial area are essentially decoupled. For each increment, alpha diversity is calculated as the average raw species richness per sample, and gamma diversity as the raw species richness for all samples. Coverage-based subsampling is not applied to alpha and gamma diversity, because we have standardized the number of collections to 40, but could be implemented if spatially standardizing to a fixed spatial scale. Data are reported as the mean and standard error of successful trials out of 1000, based on either number of grid cells or binned summed MST lengths. Successful trials are defined as those that reach the predetermined threshold of grid cells or binned summed MST lengths. Unsuccessful trials occur because larger spatial scales will not be encountered in every resampling trial. Following the resampling procedure, spatial standardization can be achieved by taking the

167

average measures of beta diversity for successful trials within a singular time bin at a fixed number of cells or range of summed MST lengths.

From these calculations we generate spatially and sampling standardized beta-diversity time series. We estimate measures of beta diversity at a fixed number of cells (13) and summed MST range (1000–1100 km). The level of spatial standardization is based on the minimum number of grid cells and minimum summed MST observed across all time bins after standardizing to 40 collections. For comparison, we also calculate unstandardized measures of beta diversity based on all collections and combined collections based on the penta-hexagonal grid system.
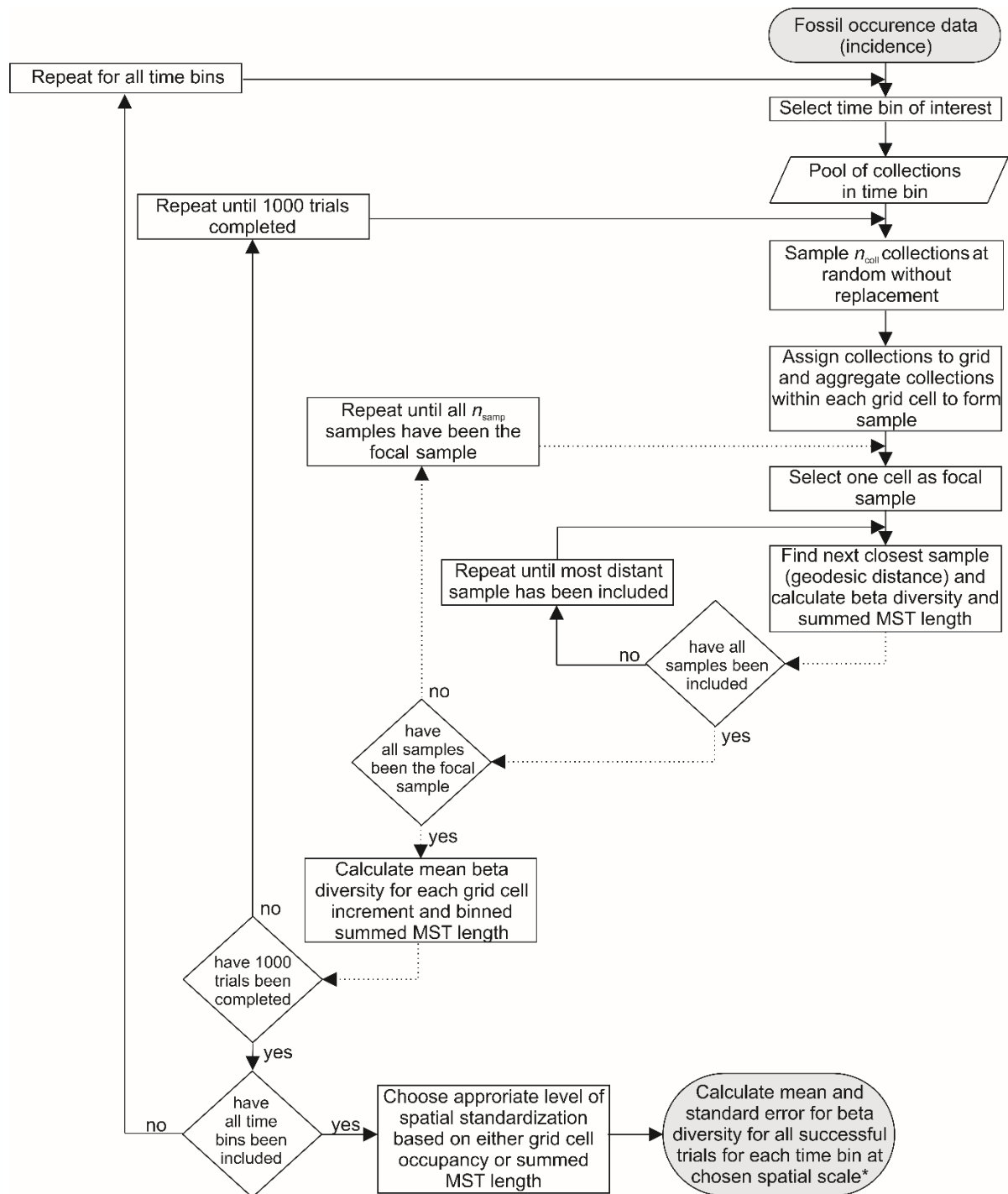
FIGURE 6.4. Flowchart highlighting the methodical resampling protocol to produce measures of beta diversity for each time bin at increasing spatial scales. For this study n_{coll} is set at 40. Dashed arrow lines indicate the loop relating to what is in the text referred to as a single trial. *Mean and standard errors are based on the number of successful trials for the spatial scale of interest, that is, larger spatial scales will not be encountered in every resampling trial.

6.3. Results

Measures of variation (β_{Add} and β_{Whit}) generally conform to the a priori expectation that beta diversity increases with increasing spatial scale and that this rate of increase declines at regional spatial scales (Fig. 6.5, and see Figs. 6.6 and 6.7). This relationship is also seen in both pairwise (see Fig. 6.8) and multisite measures of average Sørensen dissimilarity (Fig. 6.5, panels A–C, and see Figs. 6.9, 6.10). There is a large disparity between the sampling standardized measures of β_{Whit} and β_{Add} beta diversity and their counterparts based on all available data, highlighting the importance of sampling and spatial standardization for these metrics (Fig. 6.5, panels D and E). Overall, we find that the multisite measures of beta diversity (β_{MSor}, β_{MSim} and β_{MNest}) are more stable across spatial scales and are less affected by sample size in comparison to other measures of beta diversity (Figs. 6.6–6.10). Our results and interpretations are therefore based primarily on the multisite measures of beta-diversity turnover and their derived components. Henceforth, we use the following labels for the different data treatments: "collection-based" for unstandardized data taken at face value; "grid-based" for unstandardized data with collections aggregated by grid cells; and "standardized" for the sampling and spatially standardized data, where spatial standardization is achieved using either a fixed quota of cells or a fixed summed MST distance.
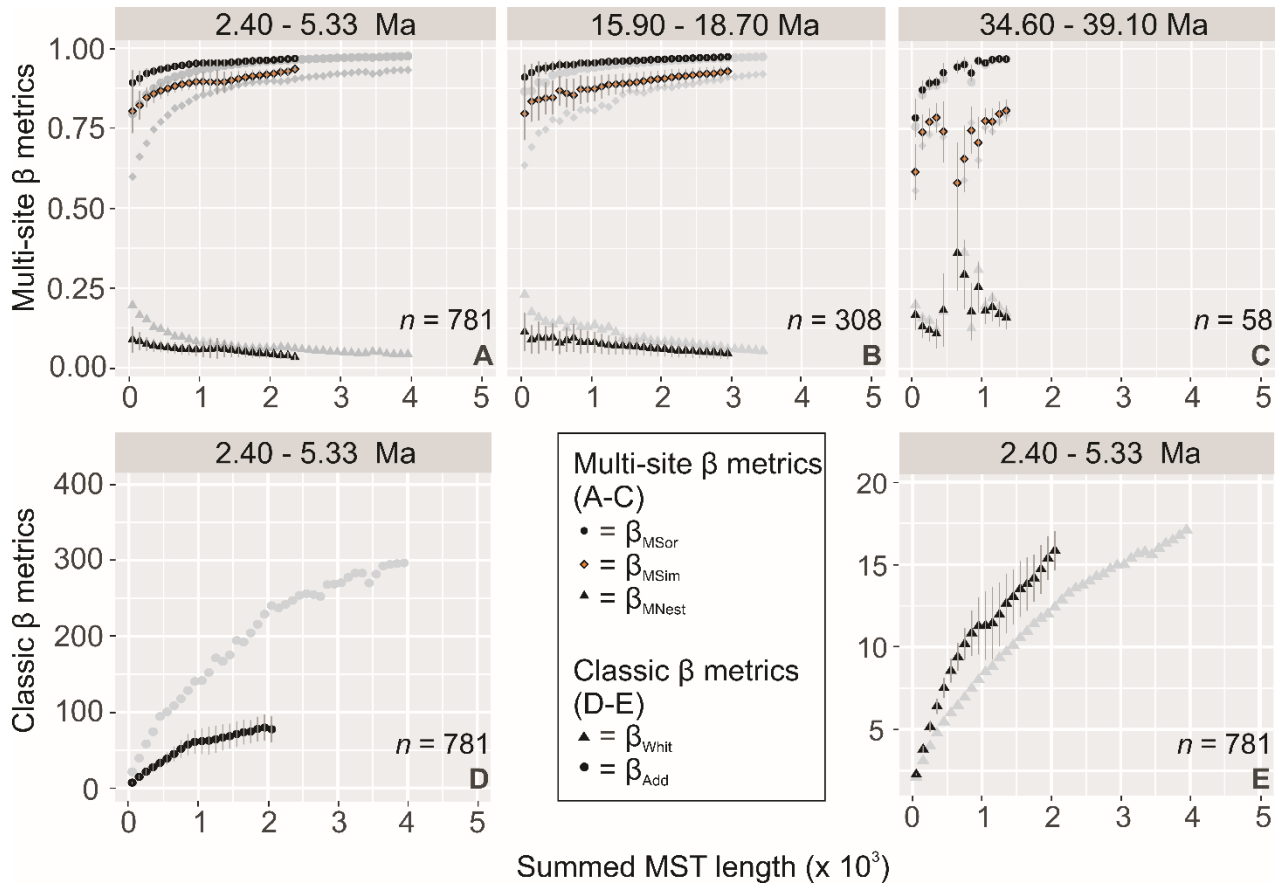
FIGURE 6.5. Measures of beta diversity plotted against summed minimum spanning tree (MST) length, for a representative selection of time bins. Unstandardized results are shown in gray, based on all collections (aggregated by grid cells). Error bars are presented for sampling standardized results only and are based on the standard error of successful trials (maximum 1000) aggregated by summed MST length bins used to calculate average beta diversity (where summed MST was aggregated into 100 km bins). See Fig. 6.4 for detailed method of resampling protocol. (A–C) Multisite measures of beta diversity (β_{MSor}, β_{MSim}, and β_{MNest}) for three time bins: (A), Mangapanian–Opoitian (Wm-Wo) (5.33–2.4 Ma); B, Altonian (Pl) (18.7–15.9 Ma); and (C), Runangan–Kaiatan (Ar-Ak) (39.1–34.6 Ma). Graphs (A, B) highlight time bins with well-ordered data, contrasting with (C), where β_{MSim} and β_{MSor} do not conform to a simple logarithmic trend. This is likely a result, at least in part, of geographic clustering of grid cells within this time bin into two confined groups (see Fig. 6.1). (D, E) Classic measures of beta diversity: additive (β_{Add}) (D) and multiplicative (β_{Whit}) (E) plotted

171

against summed MST length for Mangapanian–Opoitian (Wm-Wo) (5.33–2.4 Ma). The

number of collections per time bin is denoted by *n*.
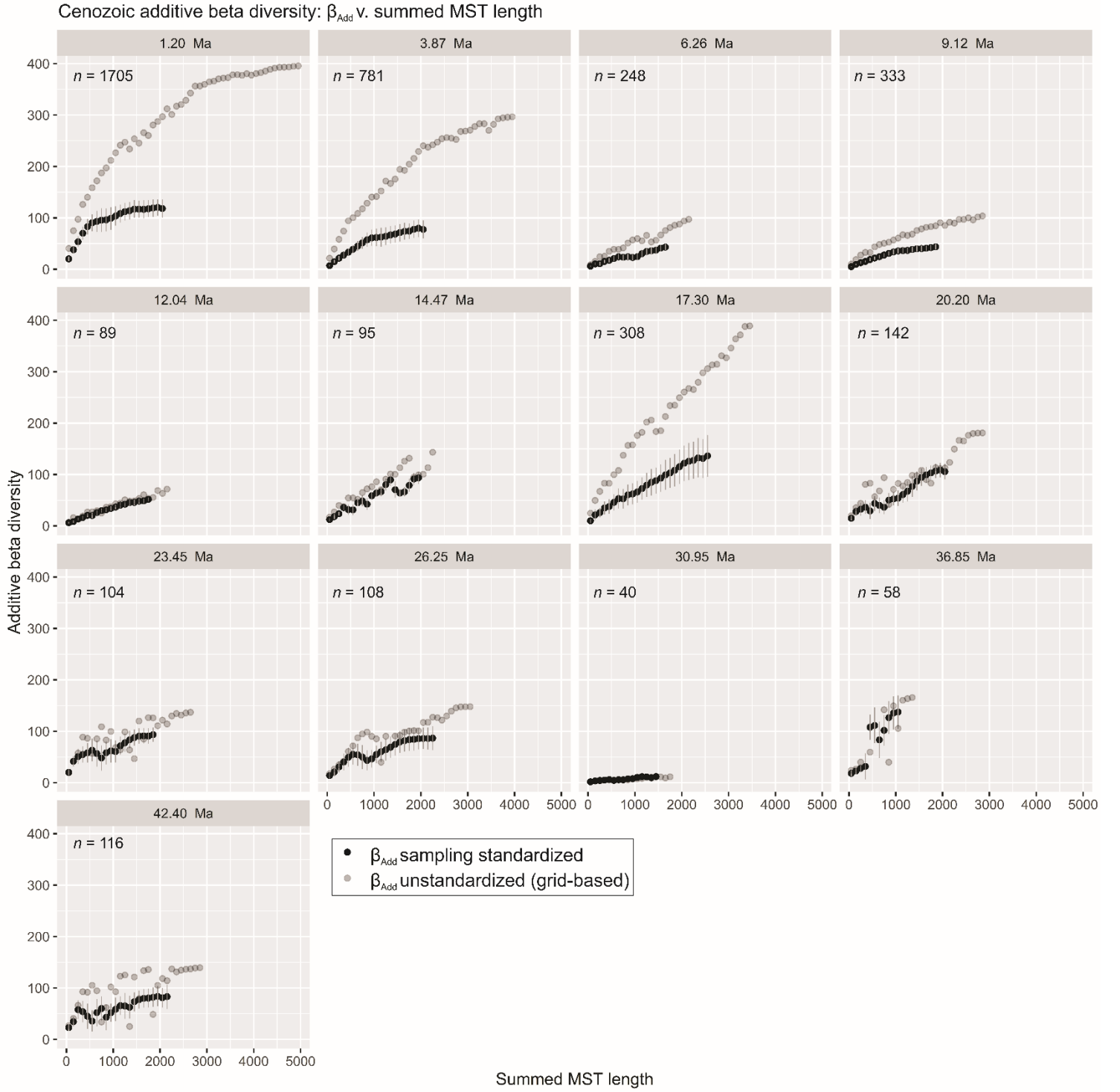
FIGURE 6.6. Sampling standardized (40 collections, resampled 1000 times) and

unstandardized (aggregated by grid cells) Cenozoic additive (β_{Add}) beta diversity plotted

against summed MST length for each time bin. Time bins are indicated by their midpoints

(Ma) (see Fig. 6.1 and Table 6.1) and the number of collections per time bin is denoted by *n*.

Error bars represent standard error (based on number of successful trials) and are given for

standardized measures only. Unstandardized measures are not visible for the Whaingaroan (Lwh) (27.3 – 34.6 Ma: midpoint 30.95 Ma) as the number of collections is the same as the resampled number of collections (40).
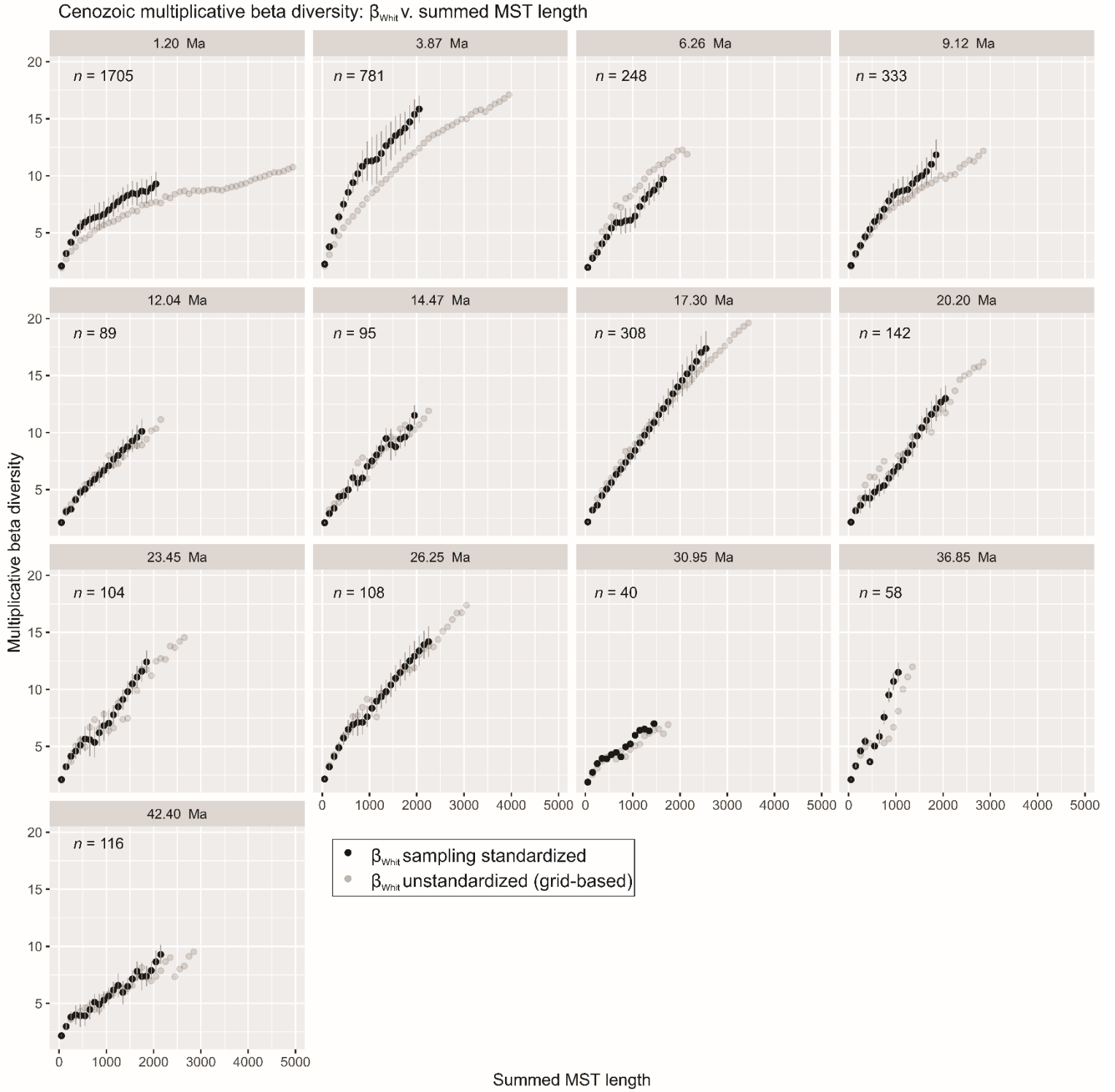
FIGURE 6.7. Sampling standardized (40 collections, resampled 1000 times) and unstandardized (aggregated by grid cells) Cenozoic multiplicative (β_{Whit}) beta diversity plotted against summed MST length for each time bin. Time bins are indicated by their midpoints (Ma) (see Fig. 6.1 and Table 6.1) and the number of collections per time bin is denoted by n. Error bars represent standard error (based on number of successful trials) and

are given for standardized measures only. Unstandardized measures are not visible for the

Whaingaroan (Lwh) (27.3 – 34.6 Ma: midpoint 30.95 Ma) as the number of collections is the

same as the resampled number of collections (40).
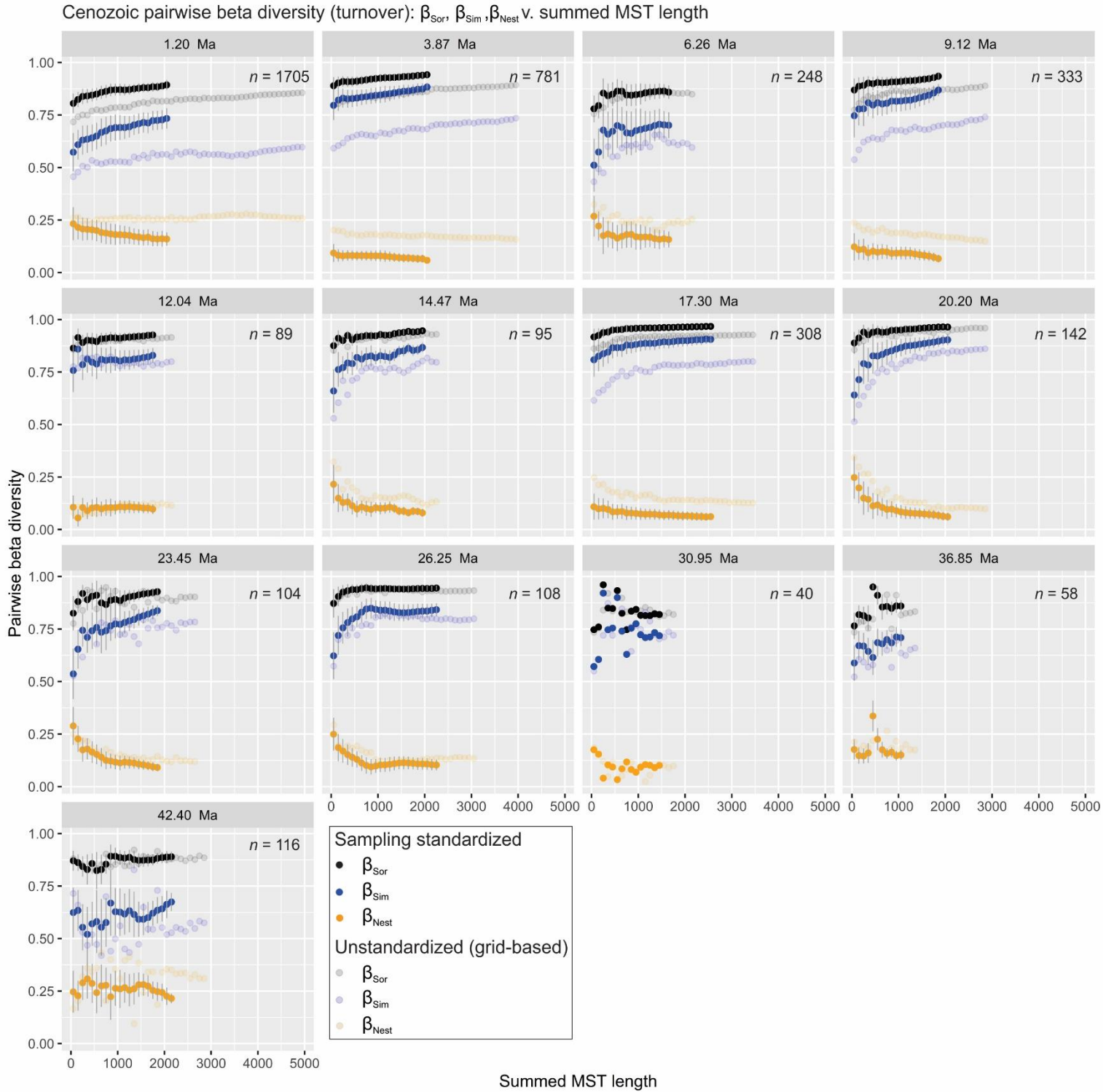
FIGURE 6.8. Sampling standardized (40 collections, resampled 1000 times) and

unstandardized (aggregated by grid cells) Cenozoic pairwise (β_{Sor}, β_{Sim} and β_{Nest}) beta

diversity plotted against summed MST length for each time bin. Time bins are indicated by

their midpoints (Ma) (see Fig. 6.1 and Table 6.1) and the number of collections per time bin

is denoted by *n*. Error bars represent standard error (based on number of successful trials) and

are given for standardized measures only. Unstandardized measures are not visible for the

Whaingaroan (Lwh) (27.3 – 34.6 Ma: midpoint 30.95 Ma) as the number of collections is the

same as the resampled number of collections (40).
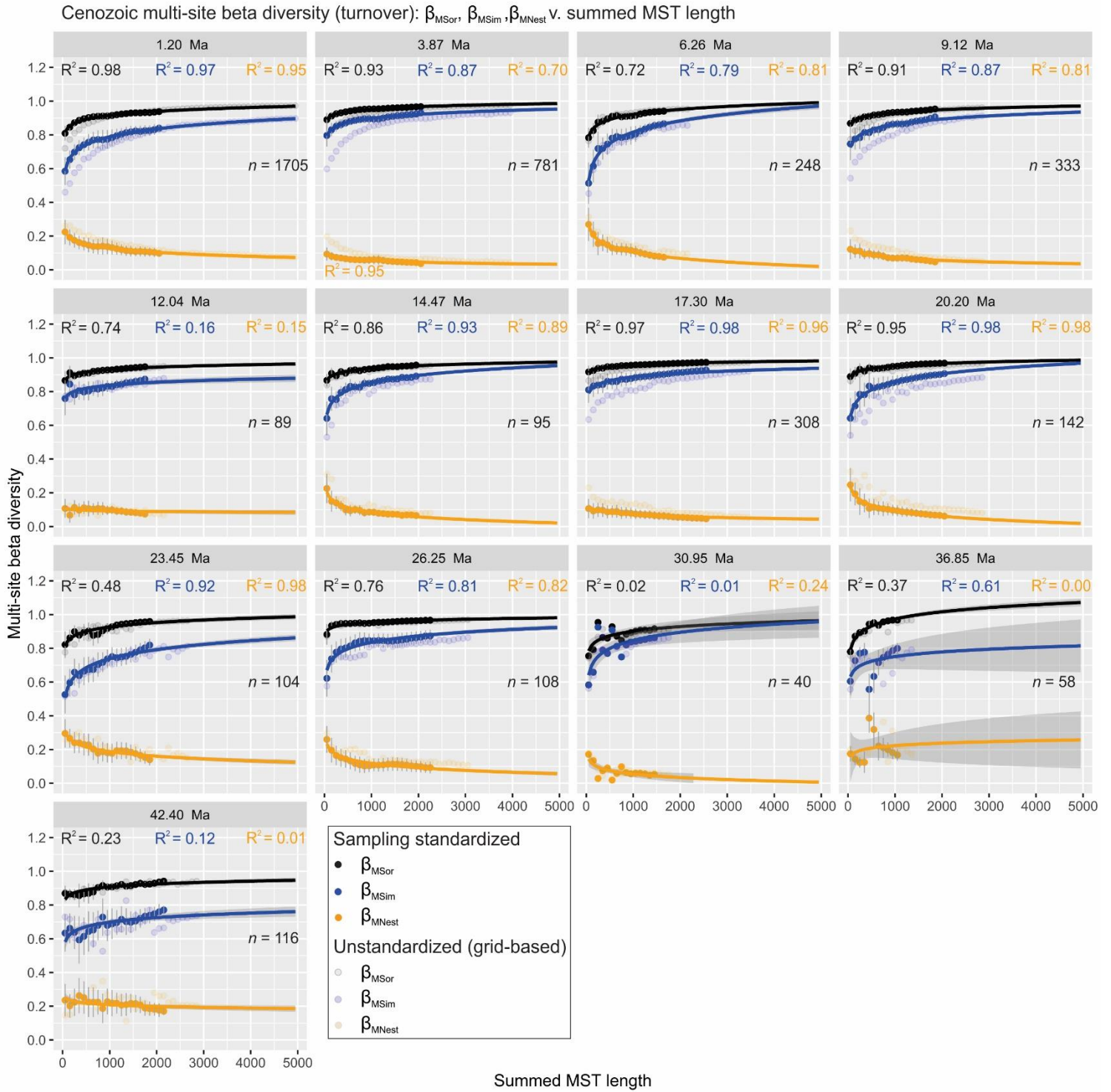
FIGURE 6.9. Sampling standardized (40 collections, resampled 1000 times) and unstandardized (aggregated by grid cells) Cenozoic multi-site (β_{MSor}, β_{MSim} and β_{MNest}) beta diversity plotted against summed MST length with fitted logarithmic trendlines (with 95% confidence interval highlighted) for each time bin. R^2 values for logarithmic trendlines are displayed at the top of each plot. Values of multi-site beta diversity > 1 are not possible and

179

logarithmic trendline extrapolations are for illustrative purposes. Time bins are indicated by their midpoints (Ma) (see Fig. 6.1 and Table 6.1) and the number of collections per time bin is denoted by *n*. Error bars represent standard error (based on number of successful trials) and are given for standardized measures only. Unstandardized measures are not visible for the Whaingaroan (Lwh) (27.3 – 34.6 Ma: midpoint 30.95 Ma) as the number of collections is the same as the resampled number of collections (40).

FIGURE 6.10. Sampling standardized (40 collections, resampled 1000 times) and unstandardized (aggregated by grid cells) Cenozoic multi-site (β_{MSor}, β_{MSim} and β_{MNest}) beta diversity plotted against grid cell occupancy with fitted logarithmic trendlines (with 95% confidence interval highlighted) for each time bin. R^2 values for logarithmic trendlines are displayed at the top of each plot. Values of multi-site beta diversity > 1 are not possible and

logarithmic trendline extrapolations are for illustrative purposes. Time bins are indicated by their midpoints (Ma) (see Fig. 6.1 and Table 6.1) and the number of collections per time bin is denoted by n. Error bars represent standard error (based on number of successful trials) and are given for standardized measures only. Unstandardized measures are not visible for the Whaingaroan (Lwh) (27.3 – 34.6 Ma: midpoint 30.95 Ma) as the number of collections is the same as the resampled number of collections (40).

We find that standardized measures of β_{MSor} and β_{MSim} are slightly elevated at smaller spatial scales, and β_{MNest} is slightly smaller compared with grid-based measurements (Figs. 6.9 and 6.10). This disparity is more pronounced when measuring spatial scales using grid-cell occupancy (cf. Figs. 6.9 and 6.10). Regardless of whether standardization is implemented, both β_{MSor} and β_{MSim} increase with spatial scale, and β_{MNest} decreases (Figs. 6.9 and 6.10), with the exception of the combined Runangan and Kaiatan (Ak-Ar) time bin (39.1–34.6 Ma), which shows a variable pattern of β_{MSim} and β_{MNest}. Collections in this time bin are split into two distinct geographic clusters (see Fig. 6.3), which are likely to affect the derived spatial measures and components of turnover. The general pattern of decrease in β_{MNest} with increased spatial scale suggests that nestedness is decreasingly important from local to regional spatial scales, and at the regional scale, species replacement (β_{MSim}) is the primary component of compositional dissimilarity.

We experiment with several regression trendlines (linear, power, logarithmic, and general additive model) to interpolate a spatially standardized measure of multi-site beta diversity. A logarithmic trendline shows the best fit for summed MST length and grid-cell occupancy versus average values of β_{MSor} and, to a lesser extent β_{MSim} and β_{MNest} (see Figs. 6.9, 6.10), based on both fit to the data and visual assessment of the plausibility of the

extrapolation. However, this relationship diminishes in older time bins, and given the close match between the logarithmic fits and measured average values of beta diversity in younger time bins at our chosen standardization quotas, we avoid this additional computational step and use the measured values of beta diversity to derive our time series of Cenozoic beta diversity (Fig. 6.11). Likewise, we avoid extrapolation using logarithmic regression, because in some cases, the trendlines yield values of β_{MSor} greater than 1.0, particularly for time bins with low spatial measures and when using grid-cell occupancy as the spatial unit (see Fig. 6.10).

Standardized measures of β_{MSor}, β_{MSim}, and β_{MNest} show substantially greater proportional volatility through the Cenozoic than unstandardized collection-based and grid-based measures (Figs. 6.11 and 6.12). Collection-based measures suggest beta diversity (β_{MSor}) increases steadily through the Cenozoic, with a similar steady increase in species replacement (β_{MSim}) and decrease in nestedness (β_{MNest}). Conversely, standardized measures suggest a declining trend in beta diversity (β_{MSor}) in the early Miocene and a large increase in β_{MNest} in the Waitakian (Lw) (25.2–21.7 Ma), coinciding with a large decrease in β_{MSor} and β_{MSim}. When spatial and sampling standardization is employed for ungrouped Plio-Pleistocene stages (see Fig. 6.1), the same downward trend is observed (Figs. 6.13 and 6.14). Grid-based measures produce results that are more similar to standardized measures than collection-based measures, although there are still notable differences in amplitude in peaks and overall trend. Similar disparities are observed between unstandardized and standardized time series of beta-diversity variation (β_{Add} and β_{Whit}) (Fig. 6.15).
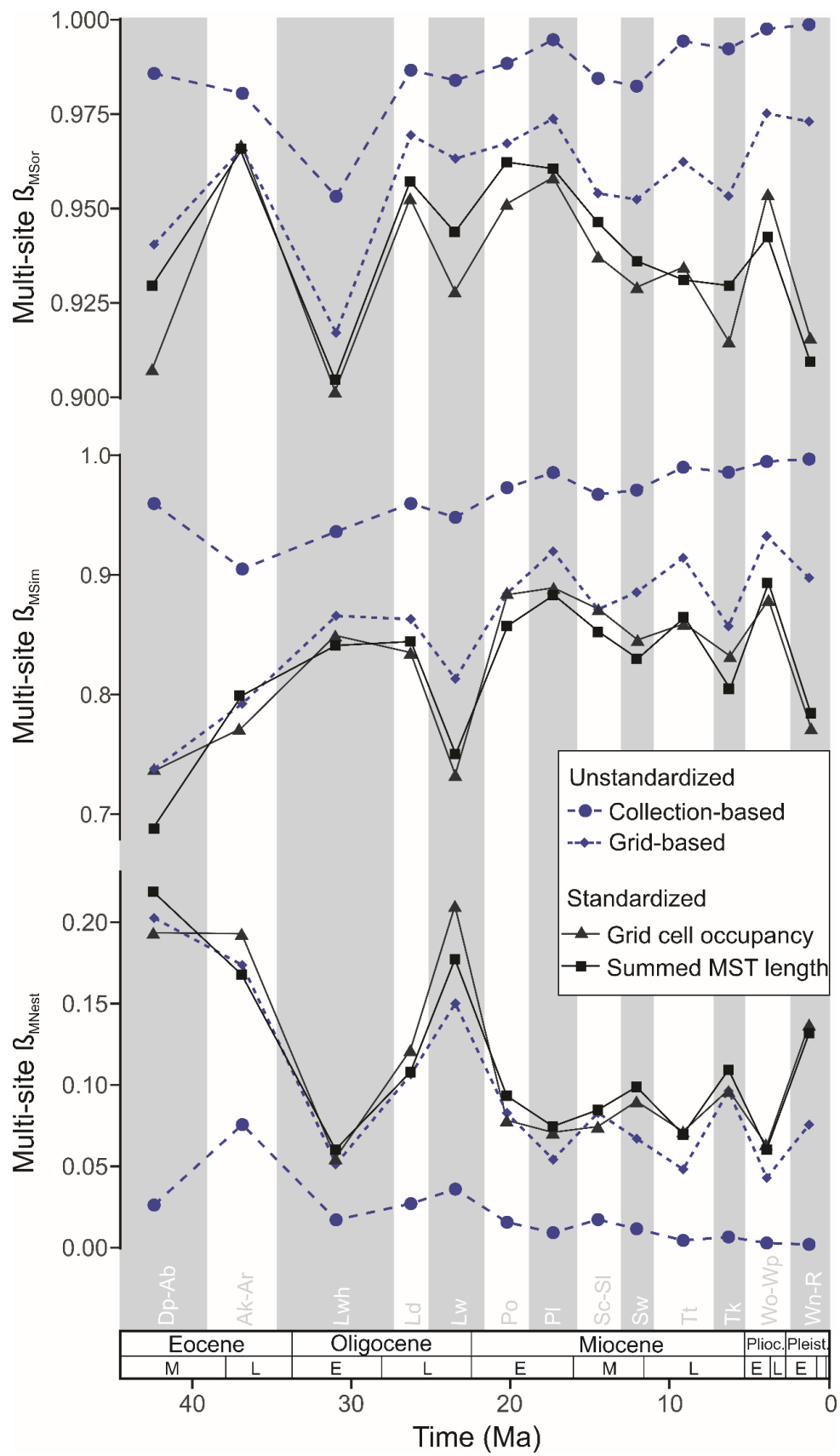
FIGURE 6.11. Cenozoic time series of multisite beta diversity (β_{MSor}, β_{MSim}, and β_{MNest}). We present two unstandardized measures: collection-based and grid-based (collections aggregated by grid cell) (dashed lines); and two spatially and sampling standardized time series using grid-cell occupancy and summed minimum spanning tree (MST) length (solid lines). Analyses are undertaken at the resolution of the highlighted time bins shown here (see abbreviated stage labels and Fig. 1 for reference), and points are plotted at time-bin midpoints (see Fig. 6.1, Table 6.1). For clarity, standard error bars are omitted; see Fig. 6.12 for plots with error bars.
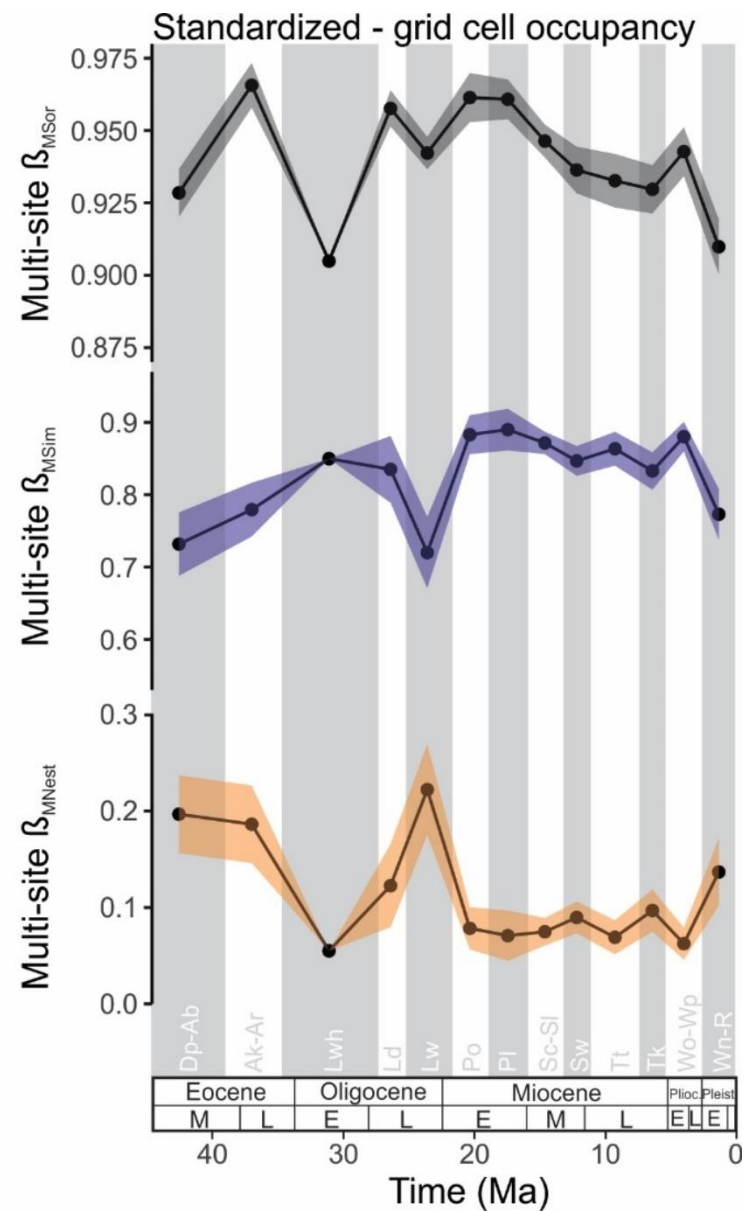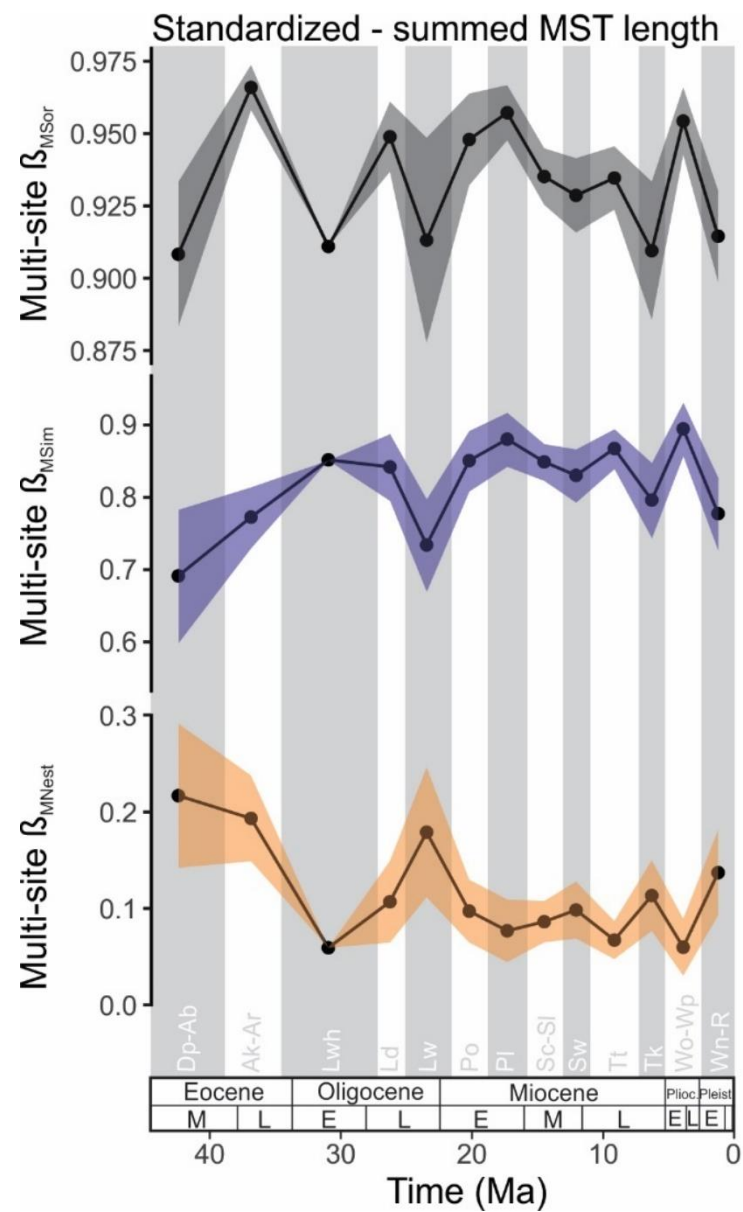
FIGURE 6.12. Cenozoic time series of spatially standardized (summed MST length and grid cell occupancy) multi-site beta diversity (β_{MSor}, β_{MSim} and β_{MNest}). Analyses are undertaken at the resolution of the highlighted time bins indicated by the shaded bars and abbreviated stage labels (and see Fig. 6.1) and points are plotted at time bin mid-points (See Fig. 6.1 and Table 6.1). Error ribbons represent standard error.
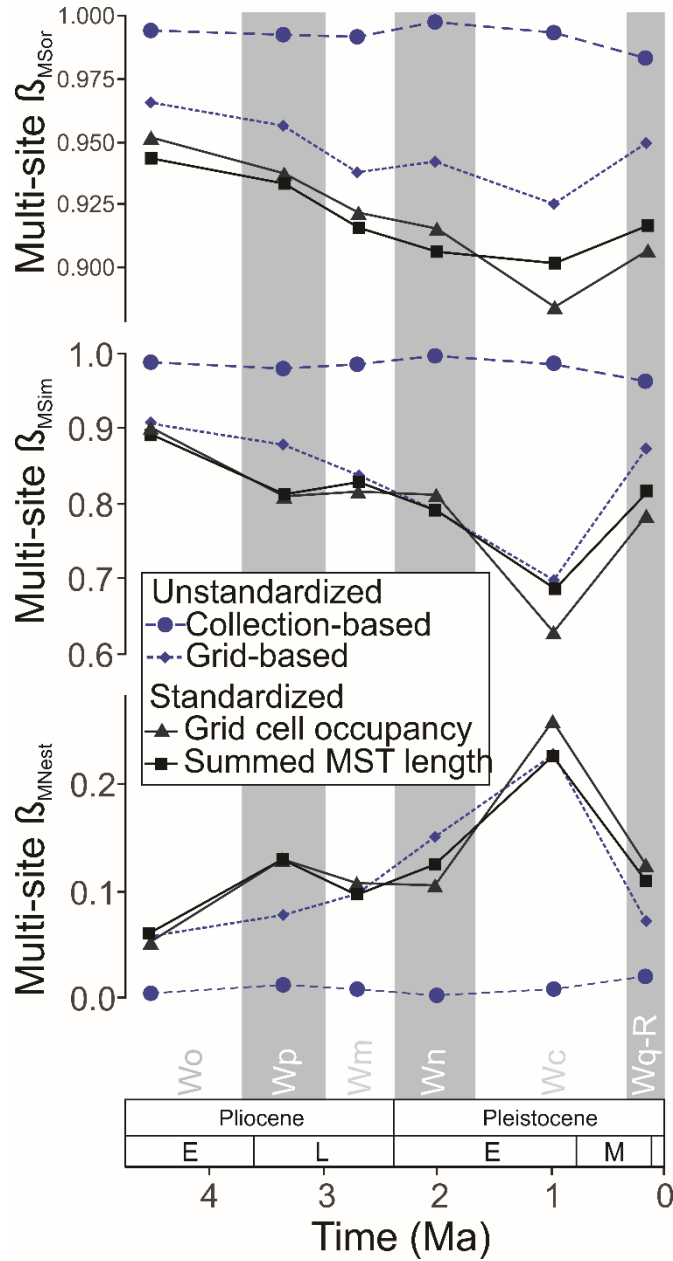
FIGURE 6.13. Plio-Pleistocene time series of multisite beta diversity (β_{MSor}, β_{MSim}, and β_{MNest}).

We present two unstandardized measures: collection-based and grid-based (collections

aggregated by grid cell) (dashed lines); and two spatially and sampling standardized time

using grid-cell occupancy and summed minimum spanning tree (MST) length (solid lines).

Analyses are undertaken at the resolution of the Plio-Pleistocene New Zealand Stages shown

here (see abbreviated stage labels and Fig. 1 for reference), and points are plotted at time-bin

midpoints (see Fig. 6.1, Table 6.1). For clarity, standard error bars are omitted; see Fig. 14 for

plots with error bars.
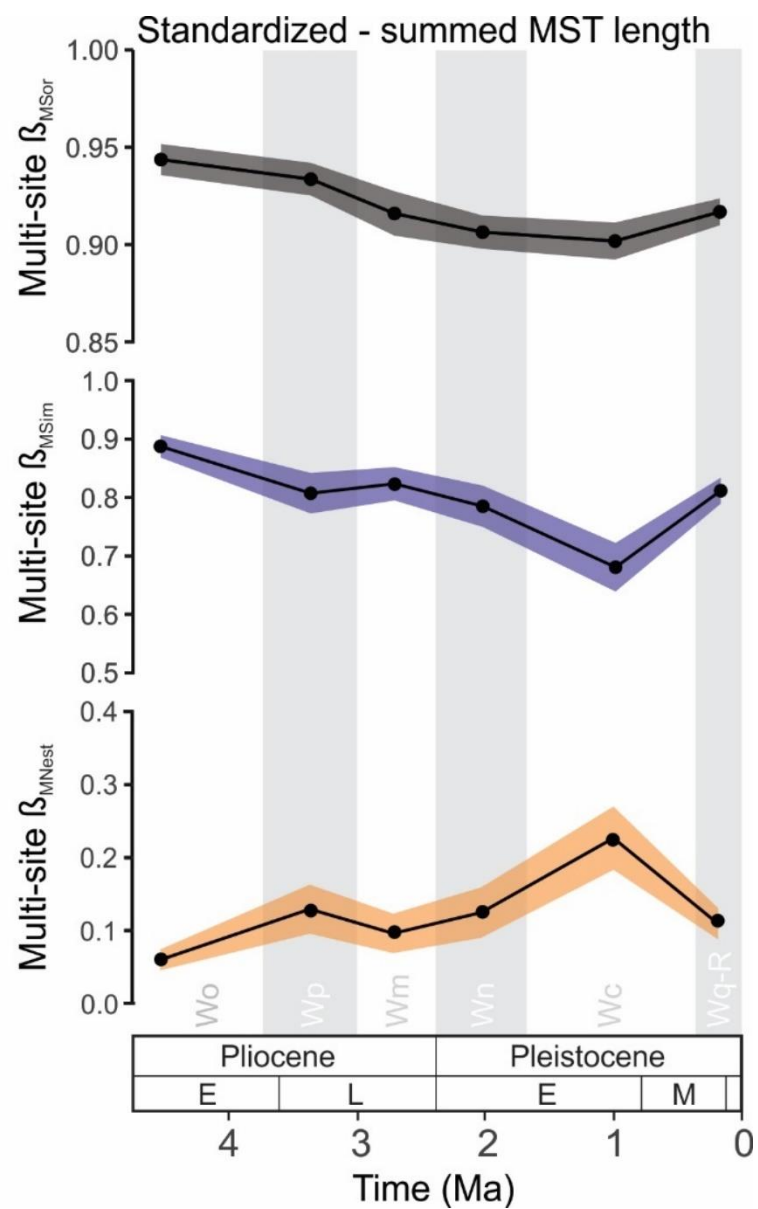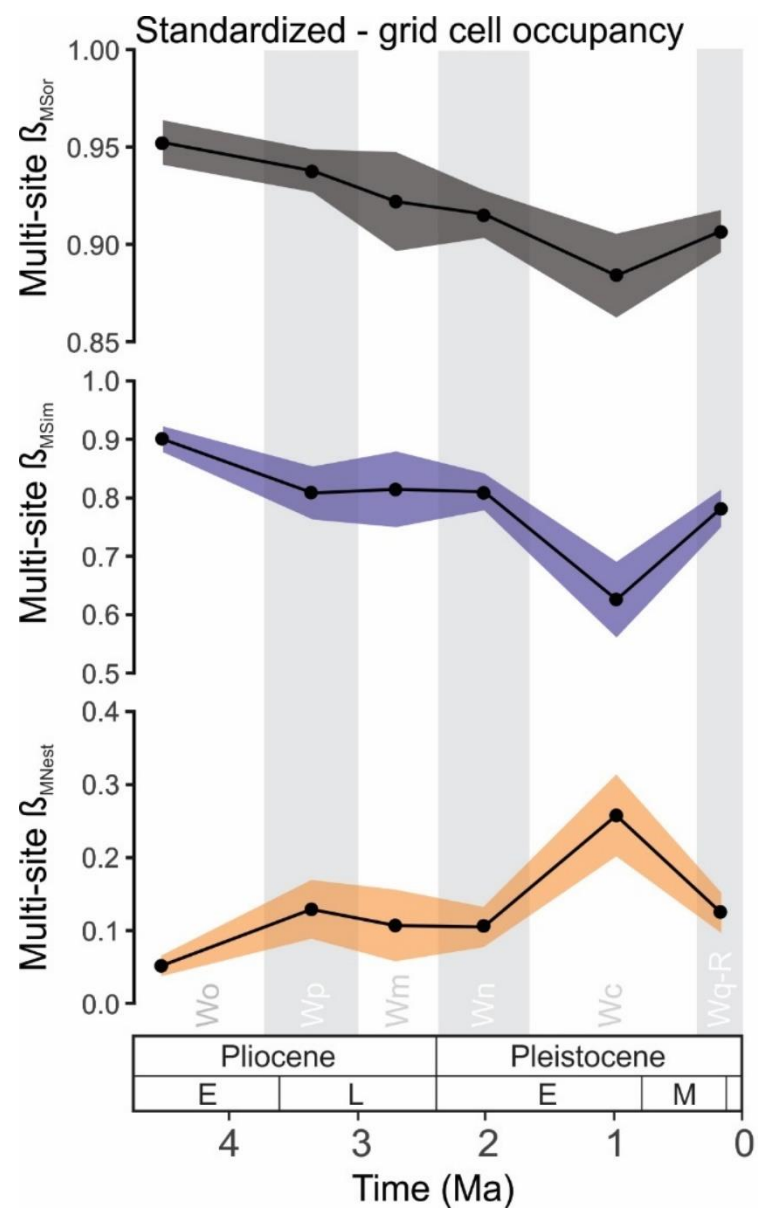
FIGURE 6.14. Plio-Pleistocene time series of spatially standardized (summed MST length and grid cell occupancy) multi-site beta diversity (β_{MSor}, β_{MSim} and β_{MNest}). Analyses are undertaken at the resolution of the Plio-Pleistocene New Zealand Stages indicated by the shaded bars and abbreviated stage labels (and see Fig. 6.1) and points are plotted at time bin mid-points (See Fig. 6.1 and Table 6.1). Error ribbons represent standard error.
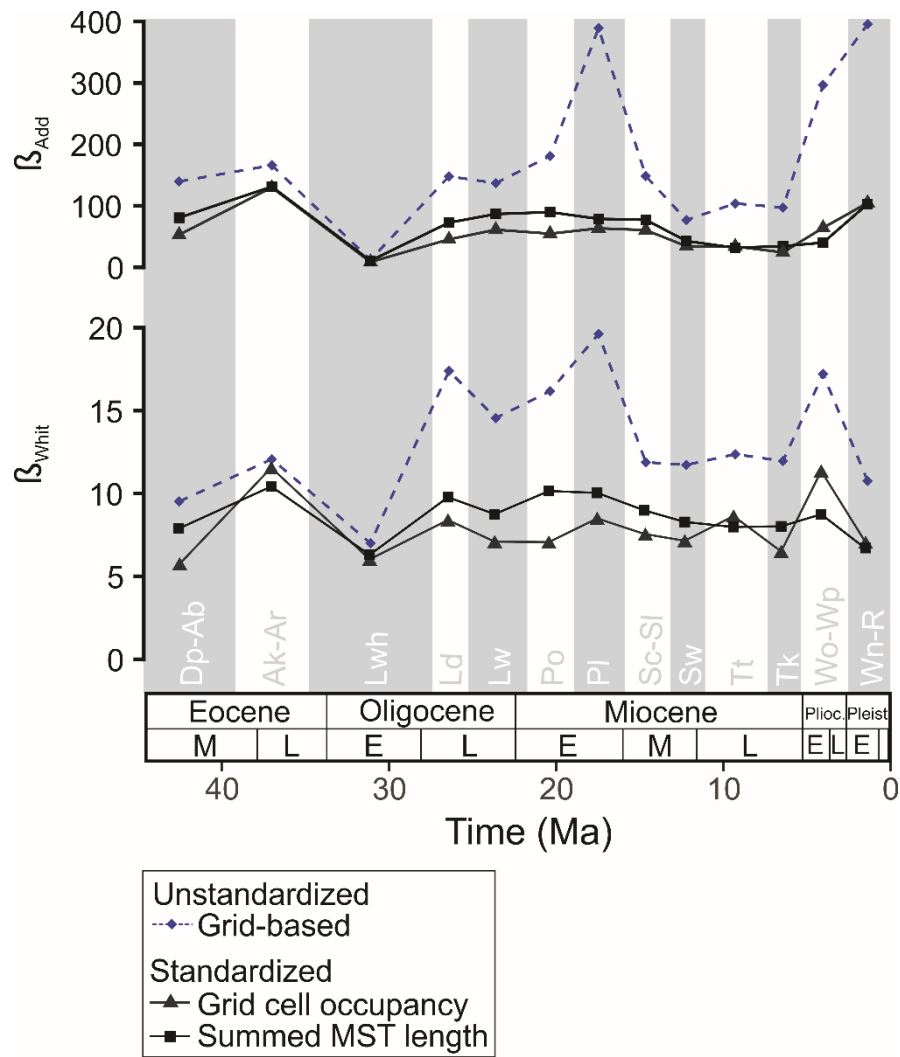
FIGURE 6.15. Cenozoic time series of multiplicative and additive beta diversity (β_{Whit} and β_{Add}). We present two spatially and sampling standardized time series using grid cell occupancy and summed MST length (solid lines) and a unstandardized grid-based time series (dashed lines). Collection-based time series are omitted to allow legibility of standardized measures. Analyses are undertaken at the resolution of the highlighted time bins shown here (see abbreviated stage labels and Fig. 6.1 for reference) and points are plotted at time bin mid-points (See Fig. 6.1 and Table 6.1).

191

To assess the effect of spatial standardization, we calculate the correlation between time series of multisite beta-diversity turnover and total per–time bin measurements of summed MST length and grid-cell occupancy. Correlations are calculated for three levels of standardization: (1) unstandardized based on raw collections; (2) unstandardized based on collections aggregated by grid cell; and (3) sampling and spatially standardized. We employ Spearman's rank-order correlation coefficient (r_s) based on first differences (denoted by the symbol Δ). We find that correlation coefficients for β_{MSor} are strong and significant and β_{MSim} are moderately strong and significant (marginal in the case of grid-cell occupancy) when unstandardized based on collections for both summed MST length and grid-cell occupancy (Table 6.3). Effect size decreases when collections are aggregated by grid cell, but are still moderately strong and significant for β_{MSor} and moderate for β_{MSim} versus for grid-cell occupancy. When sampling and spatial standardization are employed, effect size is negligible for summed MST length, but still slightly elevated for grid-cell occupancy. This suggests that both methods are suitable for spatial standardization and summed MST length may be more effective.

Beta diversity component	Standardization method	Δ summed MST length		Δ grid cell occupancy	
		r_s	p	r_s	p
Δ β_{MSor}	*(1) Unstandardized (based on collections)*	0.692	0.016	0.830	< 0.001
Δ β_{MSim}		0.664	0.022	0.574	0.051
Δ β_{MNest}		-0.350	0.266	-0.385	0.216
Δ β_{MSor}	*(2) Unstandardized (aggregated by grid cell)*	0.371	0.193	0.627	0.029
Δ β_{MSim}		0.406	0.193	0.536	0.073
Δ β_{MNest}		-0.112	0.733	-0.252	0.429
Δ β_{MSor}	*(3) Sampling and spatially standardized*	0.182	0.573	0.396	0.203
Δ β_{MSim}		0.098	0.766	0.301	0.341
Δ β_{MNest}		0.028	0.939	-0.112	0.729

TABLE 6.3. Correlations between multi-site beta diversity versus summed MST length and grid cell occupancy. Three standardization methods are considered: (1) unstandardized, based on fossil collection localities and (2) aggregated by grid cell; and (3) sampling and spatially standardized. Correlations are based on first differences denoted by Δ.

It is also possible that beta diversity is dependent, in part, on time-bin duration. Collections in long time bins may be more separated in time compared with those in short time bins, and this might bias measures of dissimilarity. To test this, we calculate the correlation between time series of standardized measures of multisite beta-diversity turnover and time-bin duration, using Spearman's rank-order correlation coefficient based on first differences. We find that all values are nonsignificant and correlation coefficients for β_{MNest} are low and moderate to moderately low for β_{MSor} and β_{MNest}, respectively (Table 6.4). Whereas correlation coefficients for β_{MSor} and β_{MNest} are slightly elevated, they are nonsignificant and, importantly, show the opposite trend that would be anticipated if time-bin duration biased beta diversity, and are therefore are likely a result of spurious correlation. This suggests volatility in beta diversity is not strongly influenced by time-bin duration.

	r_s	p
$\Delta\ \beta_{MSor}$ v. Δ time bin duration	-0.503	0.099
$\Delta\ \beta_{MSim}$ v. Δ time bin duration	0.070	0.834
$\Delta\ \beta_{MNest}$ v. Δ time bin duration	-0.343	0.276

TABLE 6.4. Correlations between standardized multi-site beta diversity and time bin duration. Correlations are based on first differences denoted by Δ.

6.4. Discussion and Summary

6.4.1. Spatial Scaling of Beta Diversity

All measures of beta diversity (with the exception of β_{MNest}) increase with spatial scale and show declining rates of increase at regional spatial scales. For the New Zealand Cenozoic mollusk data, multisite measures of turnover and their components (β_{MSor}, β_{MSim}, and β_{MNest}) show an approximately logarithmic relationship with summed MST length and grid-cell occupancy, particularly in younger time bins (27.3–0 Ma) (Figs. 6.9 and 6.10). The relationship between spatial scale and beta diversity measured as compositional turnover is complex and dependent on a number of factors, including taxa, grain and extent size, and the metric itself. It is possible that the logarithmic relationship observed may, in part, be related to the species–area relationship, where the decreasing number of new species encountered at increasingly larger spatial scales may limit compositional dissimilarity.

6.4.2. Temporal Patterns of Beta Diversity

Time series derived by standardizing on summed MST length and grid-cell occupancy reveal similar patterns, but summed MST length is considered to be the superior approach, because it captures the geographic dispersal of the data. For example, in the Altonian (Pl) (17.3–15.9 Ma), collection locations are spread relatively evenly in paleogeographic coordinates, whereas Tongaporutuan (Tt) (11.04–7.2 Ma) collections are clumped predominately in the north (Fig. 6.2). Both time bins have similar numbers of collections, and the Altonian (Pl) has 19% more occupied grid cells relative to the Tongaporutuan (Tt) (Fig. 6.2). Based on the average summed MST length and grid-cell occupancy for 40 collections (for all 1000 trials), grid-cell occupancy increases by 17% from the Tongaporutuan (Tt) to the Altonian (Pl), whereas mean summed MST length increases by 28%, a value that captures the observable increase in geographic dispersal. These relative differences in measures of

195

geographic extent explain, in part at least, the differences we see in time series based on the two methods, in particular using β_{Whit} (Fig. 6.13). Studies at larger spatial scales or with less evenly spaced data should consider this potential bias for spatial standardization.

Partitioning turnover (β_{MSor}) into nestedness (β_{MNest}) and species replacement (β_{MSim}) facilitates a greater understanding of patterns in beta diversity. Species replacement (β_{MSim}) is the dominant component of compositional dissimilarity and increases with spatial scale, whereas nestedness (β_{MNest}) decreases. These patterns are logical consequences, in part at least, of environmental and ecological factors that operate at different spatial scales. Hence, as spatial scale increases, climate and environmental contrasts related to latitude will also increase and dispersal limitation of species will become more significant, both factors that are likely to increase species replacement and decrease nestedness.

Patterns of beta diversity should be considered in relation to other standardized components of diversity (alpha and gamma) and measures of environmental heterogeneity (e.g., see Crampton et al. 2011), and detailed interrogation of these New Zealand Cenozoic mollusk data will be the subject of a separate paper. Importantly, for the New Zealand Cenozoic mollusk data, time series of beta diversity that employ spatial and sampling standardization reveal different temporal trends and much greater proportional volatility than those based on unstandardized measurements, displaying contrasting results to previous estimates (Crampton et al. 2006b, 2011). These standardizations are neglected in many studies of beta diversity, but clearly may have a significant impact on paleobiological interpretations, as illustrated in the following two examples from our data.

Standardized β_{MSor} shows a decline from ~20 Ma, whereas unstandardized measures (particularly collection-based) display an increase (Fig. 6.5). Given that this period spans the development of the modern plate boundary through New Zealand and also global cooling,

one could draw very different conclusions about one or both these factors in driving patterns of beta diversity in standardized versus unstandardized measurements.

Standardized β_{MSor} and β_{MSim} both display a prominent downward excursion in the Waitakian Stage (Lw) (25.2–21.7 Ma) that is not observed in unstandardized measures (Fig. 6.5). The Oligocene marks the maximum flooding of the Zealandia continent, peaking in the late Oligocene (Waitakian) (Fleming 1975; Cooper and Cooper 1995; King et al. 1999; King 2000). The relationships between sea level, environmental heterogeneity, and beta diversity are of significant interest, and interpretations based on unstandardized measures are likely to be misleading.

6.4.3. Summary

We show that beta diversity is spatially dependent at local to regional spatial scales in the Cenozoic shallow-marine molluscan fossil record of New Zealand and that uneven spatial sampling can influence recovered temporal trends in beta diversity. Our results suggest that multisite beta-diversity turnover and components (β_{MSor}, β_{MSim}, and β_{MNest}), standardized spatially using summed MST length, are the most suitable for elucidating patterns of beta diversity in the fossil record.

Chapter 7

A positive relationship between functional redundancy and temperature in Cenozoic marine ecosystems

"Chaos is found in greatest abundance wherever order is being sought. It always defeats order, because it is better organized."

(*Terry Pratchett,* Interesting Times, 2002)

Preface

The modern distribution of marine life is commonly described by two spatial patterns: (1) the species-area relationship and (2) the latitudinal diversity gradient. The species-area relationship is one of the most prominent and documented patterns in modern ecology and is also evident in the fossil record (Sepkoski 1976, Barnosky et al. 2005), typically conforming to the power function model (i.e., a linear relationship in log-log space). Recent evidence suggest that a similar relationship exists for functional richness, termed the functional diversity-area relationship (Whittaker et al. 2014, Karadimou et al. 2016). However, the functional diversity-area relationship has not been examined in the fossil record. Observations from the modern latitudinal diversity gradient suggest two important biogeographic patterns in the marine realm. Firstly, there is evidence that taxonomic and functional richness of bivalves show a parallel poleward decrease in richness (Edie et al. 2018, Schumm et al. 2019). Second, the spatial relationship between the distribution of the number of bivalve functional groups versus the number of species is asymptotic, implying functional saturation at a certain spatial scale. The gradient of this relationship appears to be steeper at temperate latitudes relative to the tropics (**Fig.** 7.1). This implies: (1) temperate ecosystems accumulate functional groups across spatial gradients at a greater rate than tropical ecosystems, and (2) functional redundancy – the duplication of functional roles performed by the same species – is higher in the tropics. Although the exact cause of the latitudinal diversity gradient is debated, latitudinal gradients of climate and temperature are commonly invoked as the most important factors in controlling the distribution of marine organisms (Edie et al. 2018, Gagné et al. 2020). If this persists over long timescales, we may expect this spatial pattern to be expressed temporally in relatively static mid-latitude regional ecosystems (i.e., New Zealand).
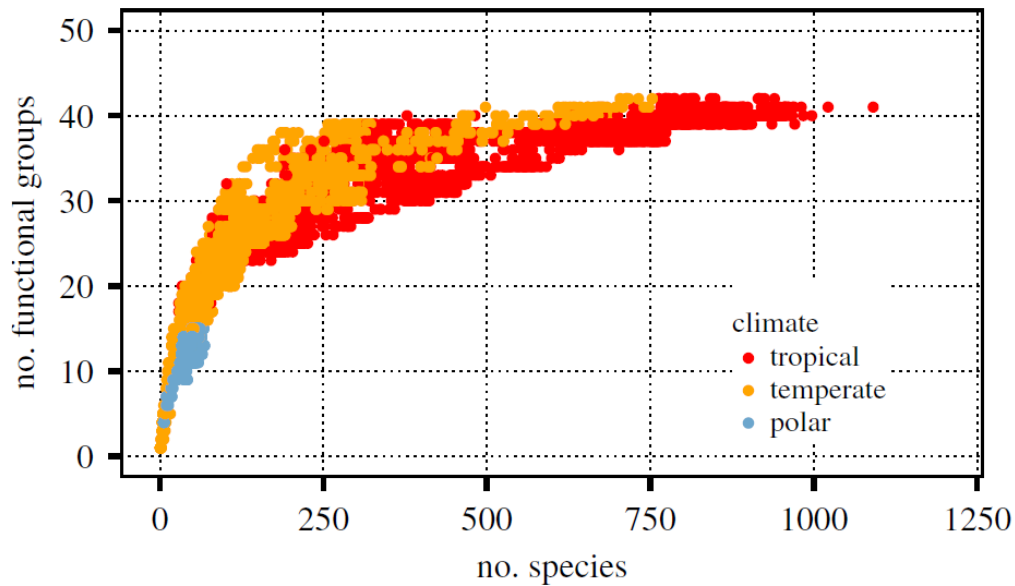
FIGURE 7.1. Plot of the number of bivalve functional groups versus number of species, split into major climatic zones (tropical, temperate, and polar). The plot is constructed from the cumulative number of functional groups and species based on sampling equal-area cells (based on a square-rectangular grid ~ 1° latitude-longitude at the equator. The figure illustrates the asymptotic relationship between the number of functional groups and number of species and shows this relationship is steeper in temperate climates relative to the tropics. Figure from Schumm et al. (2019).

Within this chapter I examine the species-area relationship and functional diversity-area relationship of New Zealand's shallow-marine ecosystems through the Cenozoic. I examine whether they conform to similar patterns observed in modern ecology and whether the interrelationship between the two, here termed the "functional redundancy relationship", is temperature dependent. This chapter is a modified version of the original paper submitted on this subject (Womack et al. in review). Co-authors to this paper include the primary and secondary supervisor of this thesis, and K. S. Collins. Input from all three was provided through regular discussion, advice, and reviewing the manuscript. In addition, K. S. Collins

assisted in the compilation of the functional trait dataset (ecospace), used to quantify

measures of ecosystem function within this chapter. The methods of quantifying spatial

distribution derived within this chapter build on those developed in Chapter 6. Furthermore,

the methods of sampling standardization and the spatial algorithm (to determine the

species/functional diversity-area relationship) form the basis of the analytical methods

employed in Chapter 8.

7.1. Introduction

One of the primary objectives of paleobiology is to determine first-order controls on the spatial and temporal distributions of biodiversity, in order to inform models of and predict future ecological impacts of current climate warming. Understanding the spatial structuring of biodiversity and its components is fundamental to achieving this objective. Biodiversity in the fossil record can be measured in terms of taxonomic and functional diversity, typically captured as species and functional richness, the latter measured as the number of unique functional groups (defined based on common functional traits shared between species) – here used as a proxy for ecological niche. The spatial scaling of species richness, described by the species-area relationship, has been documented in the fossil record (Barnosky et al. 2005), and conforms to the power function model observed in modern ecology ecology ($S = cA^z$, where S is species richness, c and z constants, and A area) (Rosenzweig 1995). The consistent form of the species-area relationship has been used widely to assess the impact of anthropogenic climate change and habitat loss on modern ecosystems (Thomas et al. 2004, Lewis 2006, Halley et al. 2014). On the other hand, spatial scaling of functional richness, herein referred to as the functional diversity-area relationship, has not been quantified in the fossil record and is understudied in modern ecology (Karadimou et al. 2016). Importantly, it is not known whether the relationships between species and functional richness at increasing spatial scales are conserved through time, or whether the two respond differently to factors that might influence the geographic distribution of modern organisms in response to future climate change.

Recent research across the modern marine latitudinal diversity gradient records high levels of functional redundancy – the duplication of functional roles by multiple species - at tropical-subtropical latitudes that decrease in temperate latitudes (Edie et al. 2018, Schumm

et al. 2019). This pattern suggests that functional redundancy may be driven by latitudinal gradients of climate and temperature and, in the case of marine molluscs, is likely controlled by total resource abundance (Schumm et al. 2019). A similar pattern of functional redundancy relative to species richness has also been observed at regional scales in modern mollusc ecosystems (Collins et al. 2019). If functional redundancy is mediated by climate, and in particular temperature, we would anticipate that the spatial relationship between species and functional richness would vary through time and be correlated with climate proxies over geological timescales, assuming these relationships reach an ecological equilibrium. This is important as evidence from modern ecology suggests functional redundancy may increase ecosystem resilience to future environmental change (Biggs et al. 2020).

We quantify and compare the species-area relationship and functional diversity-area relationship using the exemplary Cenozoic (**Fig.** 7.2) fossil record of New Zealand shallow-marine molluscs (Clowes et al. 2020) (see Chapter 3, section 3.2.2) and assess their relationships to regional climate, measured as oceanic temperature (see Chapter 4, section 4.4), and several other potentially important explanatory variables including climatic variability (see Chapter 4, section 4.5), number of sedimentary rock packages, and sampling effort (see methods). Importantly, New Zealand has remained biogeographically isolated for the duration of the Cenozoic with a high degree of endemism, and although there are known periods of immigration of cool and warm water genera to New Zealand, for the large part New Zealand has remained a "closed" system, whereby the majority of changes in species and functional richness observed here are a result of in-situ processes (Crampton et al. 2006, Marshall and Quental 2016) (and see Chapter 2). Furthermore, oceanic temperature is considered to be one important correlate or control of marine species diversity through space and time (Erwin 2009, Valentine and Jablonski 2015, Edie et al. 2018, Antão et al. 2020)

(and see Chapter 2, sections 2.3.1 and 2.3.2) and current climate models project significant changes in the state of both the surface and deep ocean temperatures over the coming century (Bindoff et al. 2019).

We explicitly standardize our data for the uneven distribution of fossil collections temporally (Holland 2015, Alroy 2020) (see Chapter 4, section 4.2) and spatially using summed minimum spanning tree (MST) length (Close et al. 2020a, Womack et al. 2020b) (see Chapter 4, section 4.3), and account for autocorrelation between time series using two methods (see Chapter 4, section 4.8). For simplicity, we refer to the relationship between the species-area relationship and functional diversity-area relationship as the 'functional redundancy relationship.' This describes, for a given time bin, the relationship between values of the species area-relationship (on the abscissa) and functional diversity-area relationship at discrete steps of summed MST length. The slope of the functional redundancy relationship reflects an element of functional redundancy, where shallower slopes indicate an increase in functional redundancy.
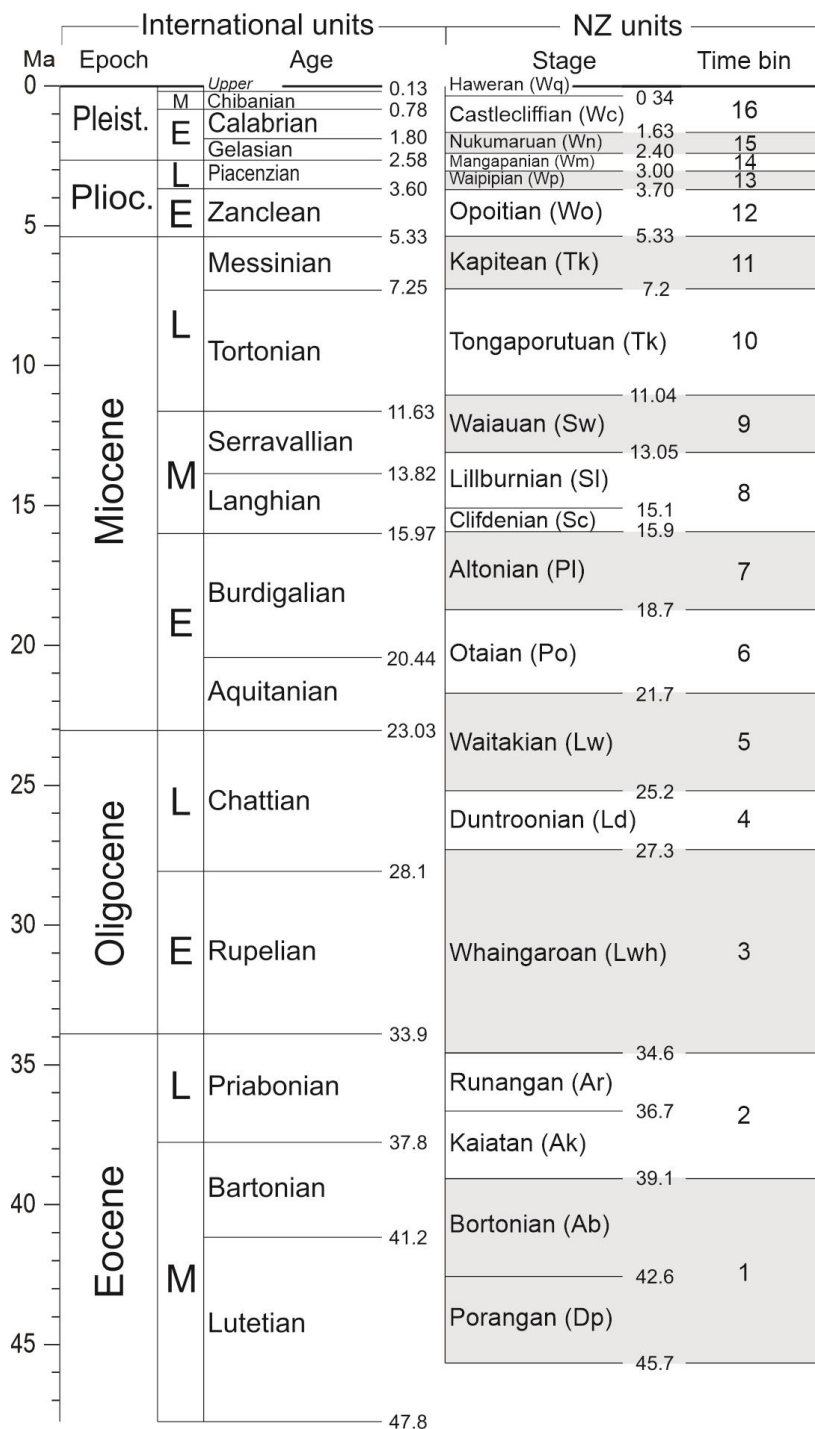
FIGURE 7.2. New Zealand Cenozoic timescale (after Raine et al. 2015). Analyses are undertaken at the resolution of the numerically labelled and highlighted time bins shown here (shaded and unshaded blocks in the right-hand column). Ages are in Ma.

7.2. Methods

7.2.1. Datasets

Occurrence data of Cenozoic shallow-marine shelfal molluscs (bivalves, gastropods, and scaphopods) were obtained from the Fossil Record Electronic Database (FRED) (Clowes et al. 2020) and subsequently cleaned following the protocols in Chapter 3, section 3.2.2. In total, 3,958 collections were analysed, which capture 19,358 taxonomic occurrences. Functional trait data of Cenozoic molluscan species were derived from the New Zealand 'synoptic dataset' and subsequently updated here to include several additional functional traits, assigned based on expert opinion of New Zealand malacologists. Geological ages are given in terms of the New Zealand Cenozoic stages (Raine et al. 2015), and time bins used within this study are based on individual and combined (to even out variation in sampling effort) stages, yielding 16 time bins with an average duration of 2.86 Ma (**Fig.** 7.2). To reduce potentially confounding effects of highly sampled fossil collections, we omitted collections that had raw species counts outside the upper 95% quantile raw collection species counts within each time bin.

7.2.2. Measuring Species Richness

We measured molluscan species richness using coverage-based subsampling (interpolation and extrapolation) of collection-level occurrence data, implemented using the R (R Development Core Team 2021) package *iNEXT* (Hsieh et al. 2016). Extrapolated measures of species richness were limited to twice the observed sample size and estimates that exceeded this limit were omitted, as recommended (Hsieh et al. 2016) (see Chapter 4, section 4.2.2).

7.2.3. Measuring Functional Richness, Evenness and Temporal Turnover

Each molluscan species was assigned to a single functional state following the approach of Edie et al. (2018b), modified to include gastropods and scaphopods. We assigned species to three functional axes: (1) tiering, (2) feeding and (3) motility (locomotion) (Table 7.1) (see Chapter 3, section 3.2.1). Each species was assigned to a functional group based on its unique combination of functional traits (e.g., an individual functional group could represent an epifaunal predator that is actively motile) yielding 49 distinct functional groups (here used as a proxy for ecological niche). We followed the definition of Edie et al. (2018b) and measured functional richness as the number of unique functional groups in a given data partition, and functional evenness as the inverse Simpson's index – which is a measurement of diversity (D) incorporating proportional abundance (in this case derived from the occurrence of species within individual functional groups), expressed as $1/D$ – normalized by the number of functional groups at the spatial scale of observation (explained in more detail below). Functional evenness measured in this way has been shown to identify variation in the distribution of taxa amongst functional groups across the modern latitudinal diversity gradient and in the fossil record (Edie et al. 2018, Schumm et al. 2019).

Functional axes		
1. Tiering	2. Feeding	3. Motility
epifaunal	carnivore	actively motile
semi-infaunal	carnivore & grazer	facultative (unattached)
shallow infaunal	parasite	facultative (attached)
deep infaunal	deposit feeder (detritovore)	non-motile (unattached)
nestler	filter feeder	non-motile (attached)
commensal	filter & deposit feeder	actively motile (swimming)
borer	chemosymbiotic	actively motile (swimming) & byssate
nektonic/pelagic	filter feeder & chemoautotroph	-
-	grazer (herbivore)	-
-	grazer (deposit)	-

TABLE 7.1. Functional traits used for molluscan ecospace. Functional traits are split into three functional axes: (1) tiering, (2) feeding and (3) motility. Molluscan species were assigned to functional groups based on their unique combination of functional traits within the ecospace following the method of Edie et al. (2018b).

We measured functional richness using both coverage-based subsampling and the corrected first-order jackknife (cJ1), implemented using code in Holland (2015) and Alroy (2020). We found that coverage-based subsampling often overestimated functional richness at smaller spatial scales (including instances where extrapolated measures were excluded), as judged by the exceedance of estimates of functional richness at the largest spatial scale, whereas cJ1 was more consistent across a broad range of spatial scales (cf. Figs. 7.3 and 7.4) (and see Chapter 4, sections 4.2.2-4.2.4). Because of this, the results presented here are based solely on cJ1 estimates of functional richness. However, the correlation between oceanic temperature and the slope of the functional redundancy relationship is still observed when functional richness is estimated using coverage-based subsampling ($r_s = -0.741$, $p = 0.014$, see section 7.2.8).
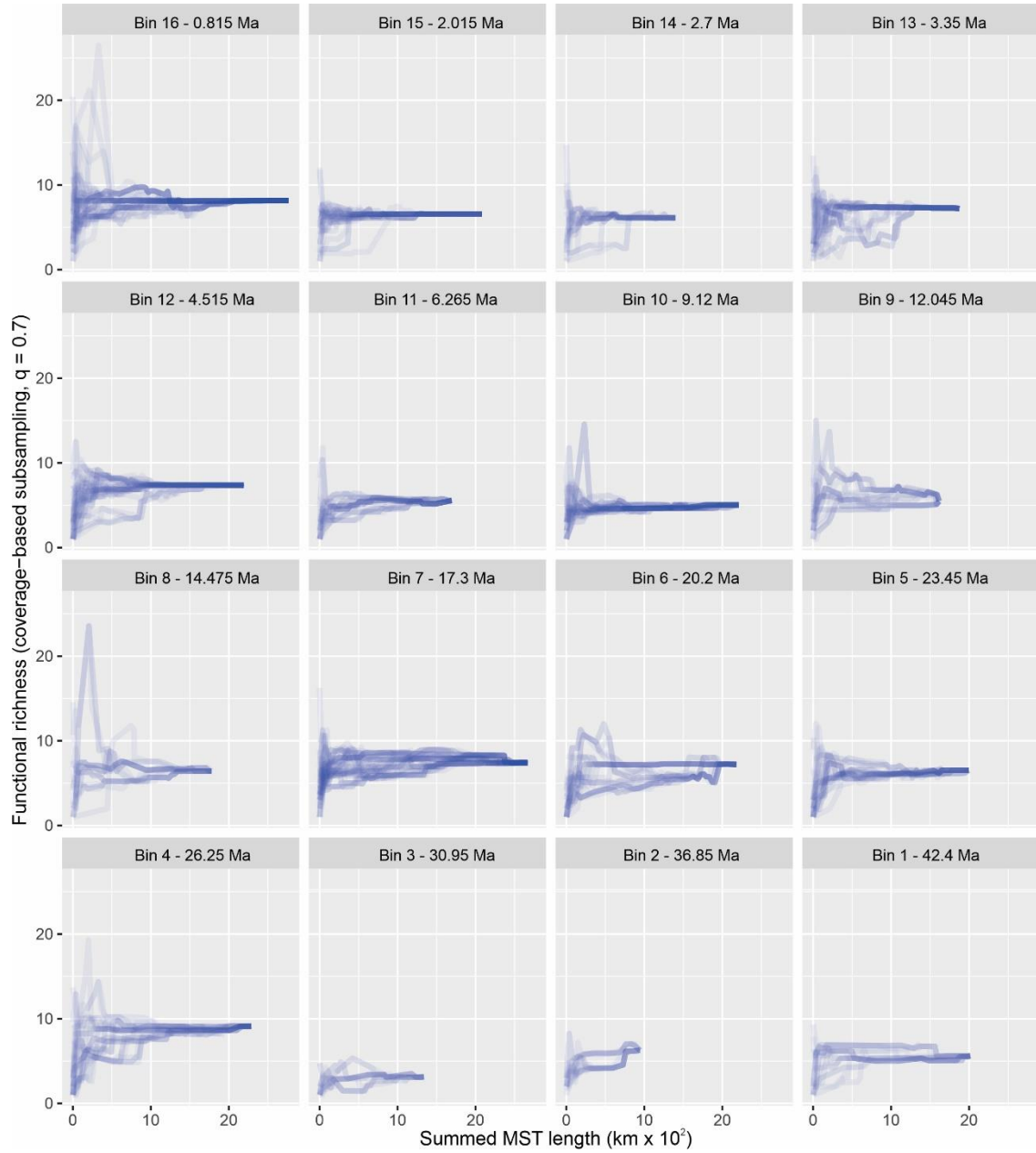
FIGURE 7.3. Functional diversity-area relationship measured using coverage-based subsampling, 45.7 – 0 Ma. The functional diversity-area relationship was plotted for each time bin, labelled by time bin number and mid-point in Ma (see **Fig.** 7.2). Individual lines represent the trajectory of the functional diversity-area relationship starting at each focal cell (see methods). Estimates of functional richness are overestimated at smaller spatial scales compared to the cJ1 estimator (**Fig.** 7.4).
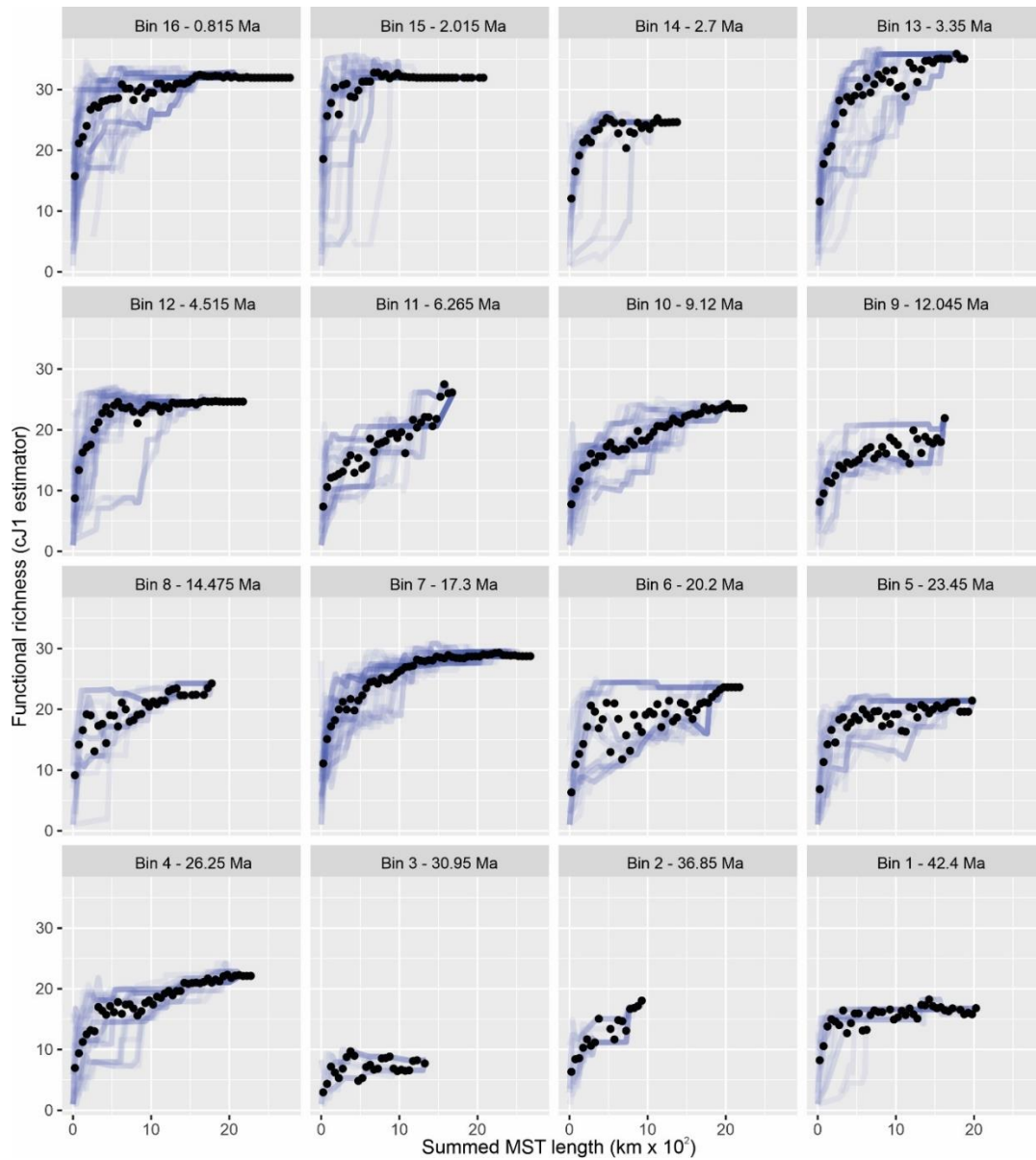
FIGURE 7.4. Functional diversity-area relationship measured using the cJ1 estimator, 45.7 –
0 Ma. The functional diversity-area relationship was plotted for each time bin, labelled by
time bin number and mid-point in Ma (see **Fig.** 7.2). Points represent mean values of species
richness in 50km summed MST bins. Individual lines represent the trajectories of the
functional diversity-area relationship starting at each focal cell (see methods). Bins 2 and 3
were omitted from further analyses based on inadequate sampling and spatial coverage.

We measured functional evenness (a measure of how uniformly species are distributed amongst functional groups) and temporal turnover of functional groups (the change within and between functional groups over time) between successive time bins using frequency distributions derived from coverage-based subsampling trials, based on the code in Holland (2015). Whereas the cJ1 estimator was used to measure functional richness for the function diversity-area relationship, it cannot generate frequency distributions required to estimate functional evenness and temporal turnover, and coverage-based subsampling at the representative spatial scale adopted here performs well (Fig. 7.4). Functional evenness was calculated as the mean evenness from 100 subsampling trials at a representative spatial scale within each time bin (see below) (and see Chapter 4, section 4.6). Results were obtained following the same protocol as outlined for species and functional richness, and the functional evenness-area relationship showed a decreasing trend with increasing area as observed in modern ecological studies (Karadimou et al. 2016).

Temporal turnover of functional groups was measured as the mean pairwise Bray-Curtis dissimilarity between subsampled frequency distributions of successive time-bins, based on 100 resampling trials (in each resampling trial, the metrics were calculated as the mean from 100 subsampling trials) (and see Chapter 4, section 4.6). Pairwise Bray-Curtis dissimilarity was calculated using the R package *betapart* (Baselga and Orme 2012), where total dissimilarity is partitioned into two additive components: balanced-variation (i.e., substitution of species between functional groups) and abundance-gradients (i.e., subsets, or loss of species within functional groups) (Baselga 2017). Whereas temporal turnover is unidirectional (i.e., forward in time) and Bray-Curtis dissimilarity is bidirectional, the metric has been shown to be a good estimate of temporal turnover using abundance data (Magurran and Henderson 2010) (where abundance was here estimated from collection-level incidence frequencies obtained from coverage-based subsampling).

7.2.4. Estimating Area

To estimate spatial area, we generated an equal-area penta-hexagrid (side of 23 km) using the R package *icosa* (Kocsis 2020). We assigned each fossil collection to a grid-cell based on their paleocoordinates, following the protocol in Womack et al. (2020b) (and see Chapter 4, section 4.3). Spatial area was measured as the summed minimum spanning tree (MST) length between the central coordinates of each occupied grid-cell following the protocol outlined below for the species-area relationship and functional diversity-area relationship. Summed MST length has been shown to approximate complex distribution shapes (e.g., surrounding a landmass) and provides a good compromise in summarising the spatial distribution of fossil localities based on their coordinates (Alroy 2010, Close et al. 2020a, Close et al. 2020b, Womack et al. 2020b) (see Chapter 4, section 4.3 and Chapter 6). Whereas summed MST length is not a measure of area, measuring area in this context is highly problematic due to paleogeographic uncertainties and factors that control the spatial extent of preserved and sampled rocks and, importantly, using summed MST length does not violate the basic power-law inference. In addition, summed MST length has been shown to closely follow the perimeter of the paleocoastline in the Cenozoic shallow-marine fossil record of New Zealand (see Chapter 6).

7.2.5. Species-Area Relationship, Functional Diversity-Area Relationship, and the Functional Redundancy Relationship

Species-area relationship and functional diversity-area relationship plots were generated by aggregating collections by grid-cell and calculating summed MST length, and species and functional richness, according to the following protocol. Each grid-cell was treated as the focal cell and measurements were made at progressively greater distances from the focal cell, where distance increments were based on occupied grid-cells ordered

according to great-circle distance from the focal cell. The power function model of the species-area relationship and functional diversity-area relationship were calculated from the mean richness values within 50 km summed MST length bins. As both methods of sampling standardization (coverage-based subsampling and as previously discussed the cJ1 estimator) are calculated from the relative abundance distribution from collection-level occurrence data, richness estimates (particularly functional richness using coverage-based subsampling) can be biased by grid cells that contain large occurrence lists, particularly at smaller spatial scales, and they can exceed estimates of richness at larger spatial scales (see **Figs.** 7.3, 7.4 and 7.5) (and see Chapter 4, section 4.2). As a result, it is possible for the trajectory of the individual species-area relationship and functional diversity-area relationship to decrease with increasing area. The protocol described above was designed to minimize the effect of this bias.
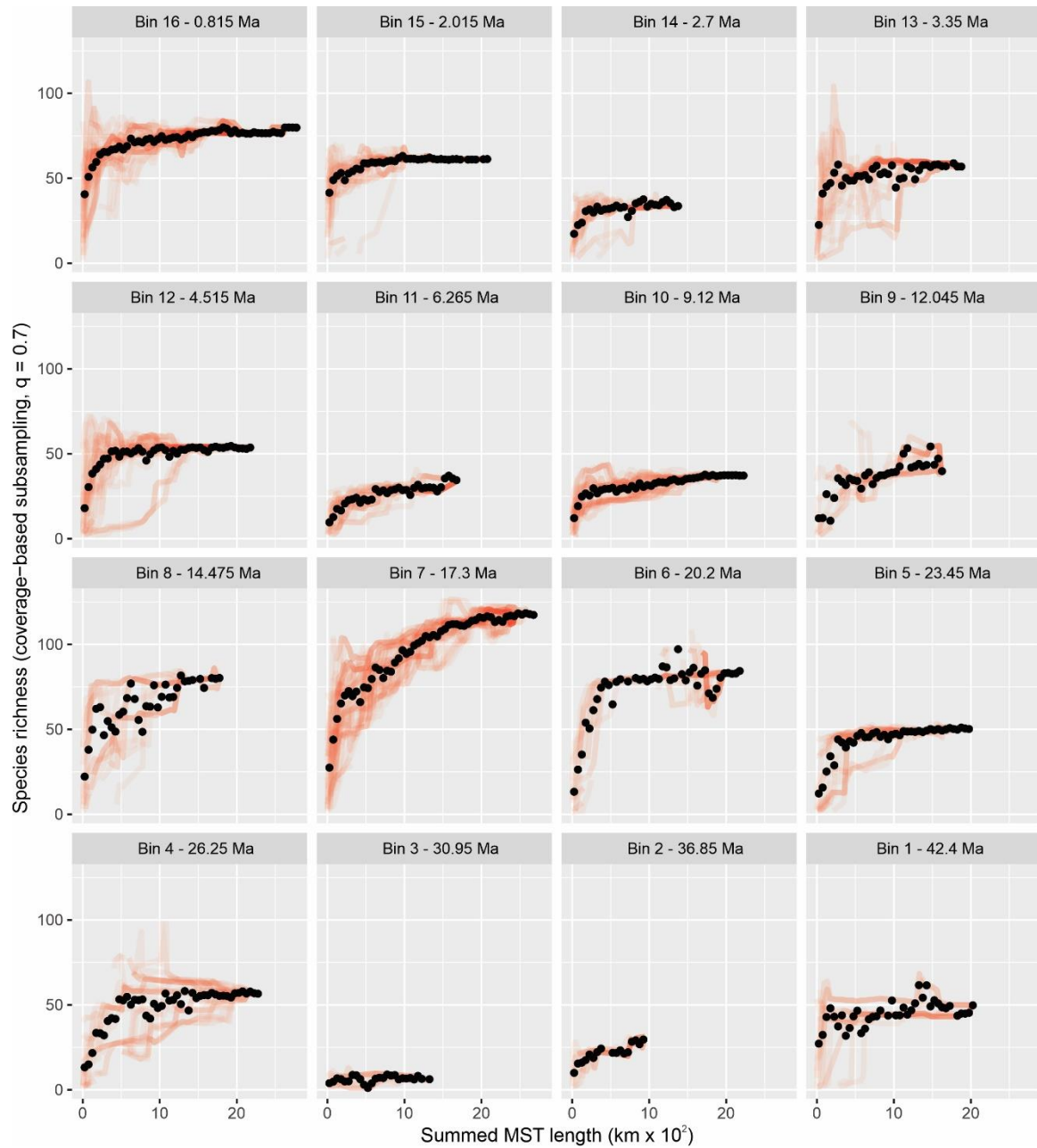
FIGURE 7.5. Species-area relationship, 45.7 – 0 Ma. The species-area relationship was plotted for each time bin, labelled by time bin number and mid-point in Ma (see **Fig.** 7.2). Species richness was measured using coverage-based subsampling (quota = 0.7). Points represent mean values of species richness in 50 km summed MST bins. Individual lines represent the trajectories of the species-area relationship starting at each focal cell (see methods). Bins 2 and 3 were omitted from further analyses because of inadequate sampling and spatial coverage.

214
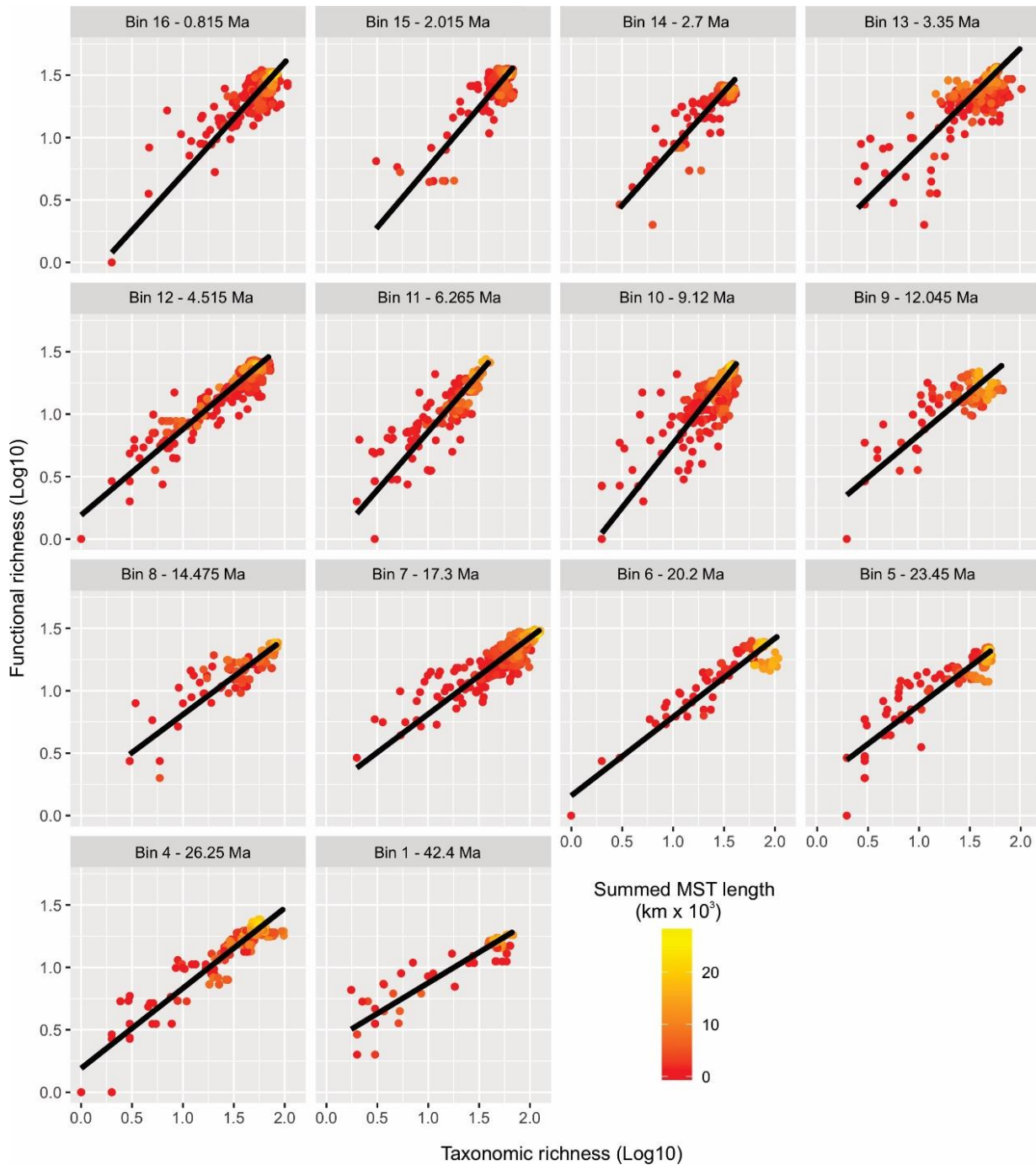
FIGURE 7.6. Power function model fits for the functional redundancy relationship, $45.7 - 0$ Ma. Individual points represent log10 species and functional richness at the same spatial scale (colour coded by summed MST length), plotted for each time bin, labelled by time bin number and mid-point in Ma (see Fig. 7.2). The log-log slope of the functional redundancy relationship (black line) was calculated using reduced major axis regression implemented using the R package *lmodel2* (Legendre 2014).

The functional redundancy relationship for each time bin was generated by plotting species and functional richness recorded at equal summed MST lengths (see **Fig.** 7.6 and **Fig.** 7.7 for graphical workflow). To allow for natural variability in both species and functional richness, the log-log slope of the functional redundancy relationship was calculated using reduced major axis regression implemented using the R package *lmodel2* (Legendre 2014). To determine species and functional richness at a representative spatial scale, we used mean species richness within a single summed MST bin (1300 – 1350 km) and interpolated functional richness at the same spatial scale using the functional redundancy relationship.
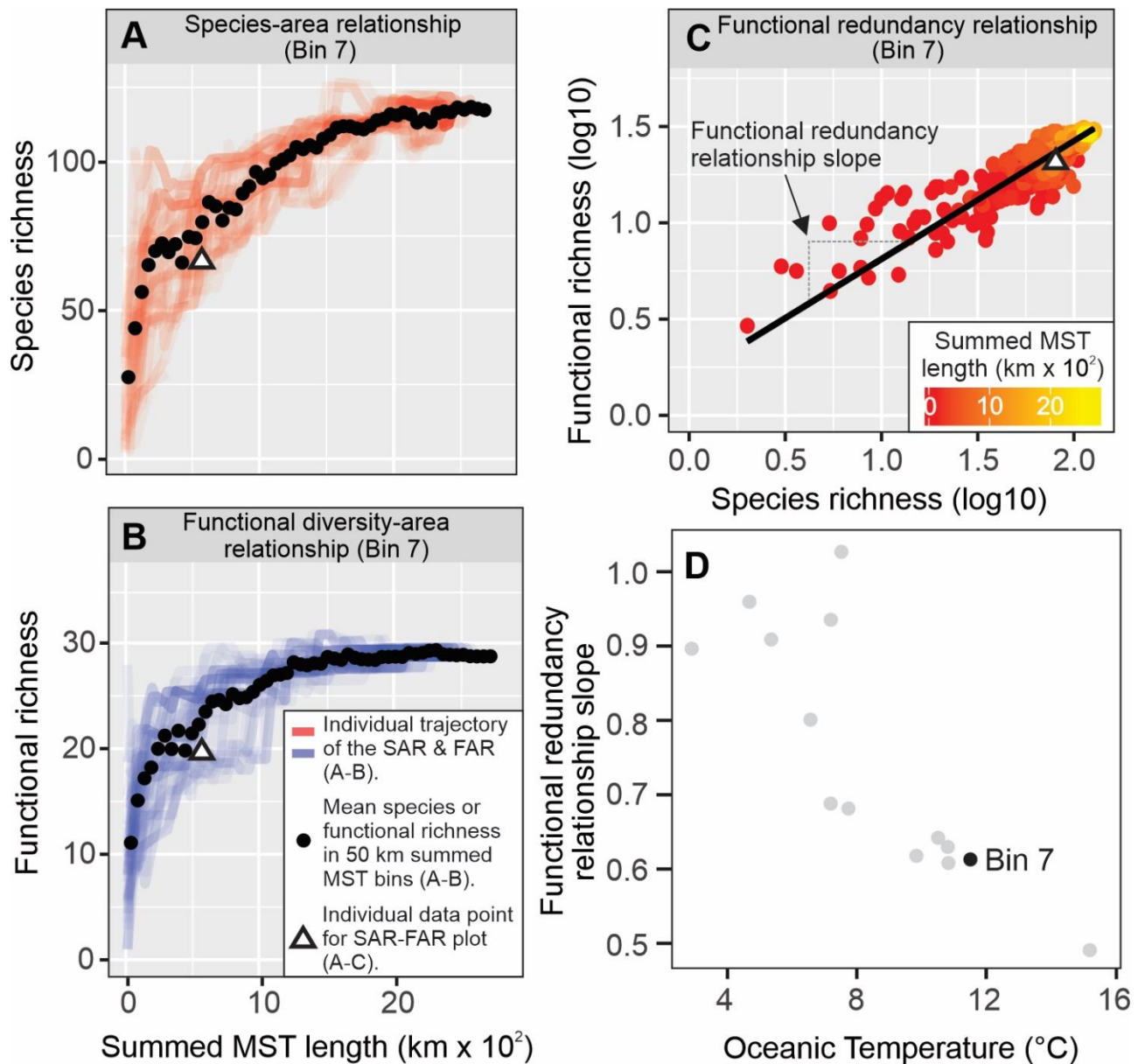
FIGURE 7.7. Graphical workflow illustrating the derivation of the functional redundancy relationship. (A) The species-area relationship and (B) functional diversity-area relationship were plotted against summed MST length for each time bin (using bin 7 as an example). Transparent red and blue lines mark individual spatial trajectories of the species-area relationship and functional diversity-area relationship. (C) We plotted the individual data points of species (A) and functional richness (B) at corresponding steps of spatial scale from individual species-area relationship/functional diversity-area trajectories (individual points are colour coded to the corresponding summed MST length), as illustrated using a single point (triangular symbol) on panels A – C. We measured the reduced major axis (RMA) slope

217

of the functional redundancy relationship for each time bin (slope for bin 7 highlighted). The steepness of this slope reflects an element of functional redundancy, where steeper slopes indicate lower functional redundancy. We plotted the slopes for each time bin against oceanic temperature (D) (same as Fig. 7.8, panel B). Importantly, the variables represented in this plot (D) are not independent. This is corrected for within our analyses.

7.2.6. Measuring oceanic temperature and climatic variability

We estimated the representative oceanic temperature within each time bin using $\delta^{18}O$ records of deep-marine benthic foraminifera from the Pacific Ocean, using the compilation of Cramer et al. (*47*). Oceanic temperature (°C) was calculated for each individual record following the methodology of Veizer and Prokoph (2015), using the transfer function of (Visser et al. 2003). To calculate representative oceanic temperature within each time bin we applied a loess regression and used interpolated values at the mid-point of each time bin.

We calculated climatic variability as the variance of the residuals of the loess curve fit used to interpolate oceanic temperature. To accommodate for the large dispersion of data, particularly during the Miocene climatic optimum, we applied a protocol to remove outliers. However, we acknowledge that ocean temperature and climate variability are intrinsically linked, and dynamics of climate variability during the Miocene climatic optimum "Coolhouse" world and Plio-Pleistocene "Icehouse" were very different (Westerhold et al. 2020); with existing data we cannot fully separate the effects of temperature and climate variability on New Zealand mollusc species richness.

7.2.7. Quantifying the number of sedimentary rock packages

We calculated the total number of sedimentary rock packages within each time bin using the dataset compiled from Macrostrat (Peters et al. 2018) by Crampton et al. (2011).

This provides measurable quantity, analogous to rock volume, to assess the impact of the temporal and spatial structure of the rock record on the resultant patterns of the species-area relationship, functional diversity-area relationship, and functional redundancy relationship.

7.2.8. Time series correlation and autocorrelation

We used two complimentary methods to address the extent of autocorrelation within each time series in order to avoid overestimating the extent of correlation between them. Firstly, we assessed the relationship with oceanic temperature, and several other important explanatory variables including climatic variability, environmental heterogeneity, and sampling effort, by multiple regression using generalized least squares (GLS) autoregressive (AR1) models in the R package *nlme* (Pinheiro et al. 2020). All explanatory variables were standardized (mean centred and scaled to unit variance) prior to running models to allow comparison of variable correlation coefficients and model coefficients. Best-fit models were determined using corrected-Akaike information criterion (AICc) weights using the R package *MuMIn* (Barton 2020). Model significance was calculated by comparison to a null-model and pseudo-R-squared values were calculated following the maximum likelihood approach of Cox and Snell (2018). Secondly, we estimated correlations for pairwise variables using the Spearman rank correlation coefficient, applying first-differences where appropriate. Statistical significance of the correlation coefficients was estimated by Monte Carlo simulations using phase-randomized surrogates, implemented using the R package *astrochron* (Meyers et al. 2014). This method compensates for potential serial correlation of data, particularly for AR1 and AR2 autocorrelation structures (Ebisuzaki 1997). We did not apply first differences to correlations including oceanic temperature. This is because this study focuses on the impact of absolute temperature on a derived slope. Using first-differences would treat a similar decrease in oceanic temperature from warm-warm, warm-cool and cool-cool as equal, when we would not anticipate the interrelationship between the

species-area relationship and functional diversity-area relationship (functional redundancy

relationship) to respond similarly in each of these three scenarios.

7.3. Results

Our key results are, first, that the standardized species-area relationship and functional diversity-area relationship in the Cenozoic shallow-marine molluscan fossil record of New Zealand conform to the power function model generally observed in modern ecological studies (**Figs.** 7.4 and 7.5 and Table 7.2). Two time bins have inadequate sampling and spatial coverage (bins 2 and 3) (**Figs.** 7.4 and 7.5) and are omitted from further analyses, although their inclusion or exclusion does not change the main conclusions of this study. Secondly, the functional redundancy relationship can also be described using a power function model (Table 7.2 and **Fig.** 7.6). Thirdly, we find that the slopes of the functional redundancy relationships significantly correlate to oceanic temperature in the Pacific Ocean (Spearman rank-order correlation coefficient, $r_s = $ -0.829, $p = 0.003$) (See Fig. 7.7 for a graphical workflow and Fig. 7.8), where lower slopes correlate to higher oceanic temperatures. We account for serial autocorrelation using both the Spearman rank-order correlation coefficient with phase-randomized surrogate data (Table 7.3) and generalized least squares (GLS) autoregressive (AR1) models (Table 7.4) (see methods). Models with oceanic temperature as the only explanatory variable outperform models that include a combination of climatic variability, number of sedimentary rock packages, and sampling effort as additive terms (Table 7.4 and see Fig. 7.9 for time series). The individual slopes of the species-area relationship and functional diversity-area relationship show no significant correlation with ocean temperature (Table 7.3). Fourthly, there is evidence that spatially and sampling standardized species richness is positively related to the combined effect of oceanic temperature and climatic variability (Table 7.4). However, we acknowledge that ocean temperature and climate variability are intrinsically linked, and dynamics of climate variability during the Miocene climatic optimum "Coolhouse" world and Plio-Pleistocene

"Icehouse" were very different (Westerhold et al. 2020); with existing data we cannot fully separate the effects of temperature and climate variability on New Zealand mollusc species richness. Lastly, we find no statistically strong evidence that the primary results discussed here result from temporal variability in sampling effort, time bin duration or the quality of the rock record (Tables 7.3 and 7.4).

| Time bins | | Temp. (°C) | Power function model fits | | | | | | | |
| | | | Species-area relationship | | | Functional diversity-area relationship | | | Functional redundancy relationship | |
Bin	Age (Ma)		R^2	coef.	p	R^2	coef.	p	R^2	slope
16	1.63 - 0	2.896	0.913	0.115	<0.001	0.878	0.120	<0.001	0.665	0.896
15	2.4 - 1.63	4.678	0.883	0.083	<0.001	0.702	0.086	<0.001	0.634	0.960
14	3 - 2.4	5.353	0.776	0.153	<0.001	0.751	0.141	<0.001	0.78	0.909
13	3.7 - 3.0	6.567	0.635	0.137	<0.001	0.913	0.226	<0.001	0.609	0.801
12	5.33 - 3.7	7.191	0.731	0.166	<0.001	0.803	0.179	<0.001	0.871	0.688
11	7.2 - 5.33	7.196	0.929	0.284	<0.001	0.887	0.279	<0.001	0.766	0.935
10	11.04 - 7.2	7.535	0.905	0.198	<0.001	0.972	0.247	<0.001	0.647	1.027
9	13.05 -11.04	7.739	0.773	0.356	<0.001	0.882	0.203	<0.001	0.475	0.681
8	15.9 - 13.05	10.831	0.78	0.243	<0.001	0.773	0.181	<0.001	0.792	0.608
7	18.7 - 15.9	11.576	0.966	0.283	<0.001	0.957	0.193	<0.001	0.815	0.614
6	21.7 - 18.7	10.812	0.748	0.309	<0.001	0.651	0.210	<0.001	0.733	0.630
5	25.2 - 21.7	9.842	0.821	0.282	<0.001	0.738	0.178	<0.001	0.828	0.618
4	27.3 - 25.2	10.513	0.831	0.306	<0.001	0.952	0.241	<0.001	0.842	0.642
3	34.6 - 27.3	10.652	0.068	0.118	<0.001	0.419	0.176	<0.001	-	-
2	39.1 - 34.6	14.033	0.887	0.255	<0.001	0.888	0.274	<0.001	-	-
1	45.7 - 39.1	15.203	0.48	0.118	<0.001	0.684	0.122	<0.001	0.827	0.491

TABLE 7.2. Power function model fits and coefficients for the species-area relationship, functional diversity-area relationship, and functional redundancy relationship. The log-log slope of the functional redundancy relationship was obtained using reduced major axis regression, which does not allow the calculation of significance (*p*) values. The log-log slope of the species-area relationship and functional diversity-area relationship was derived using ordinary least squares regression. Temperature values represent the interpolated values of oceanic temperature at the midpoint of each bin (see methods).
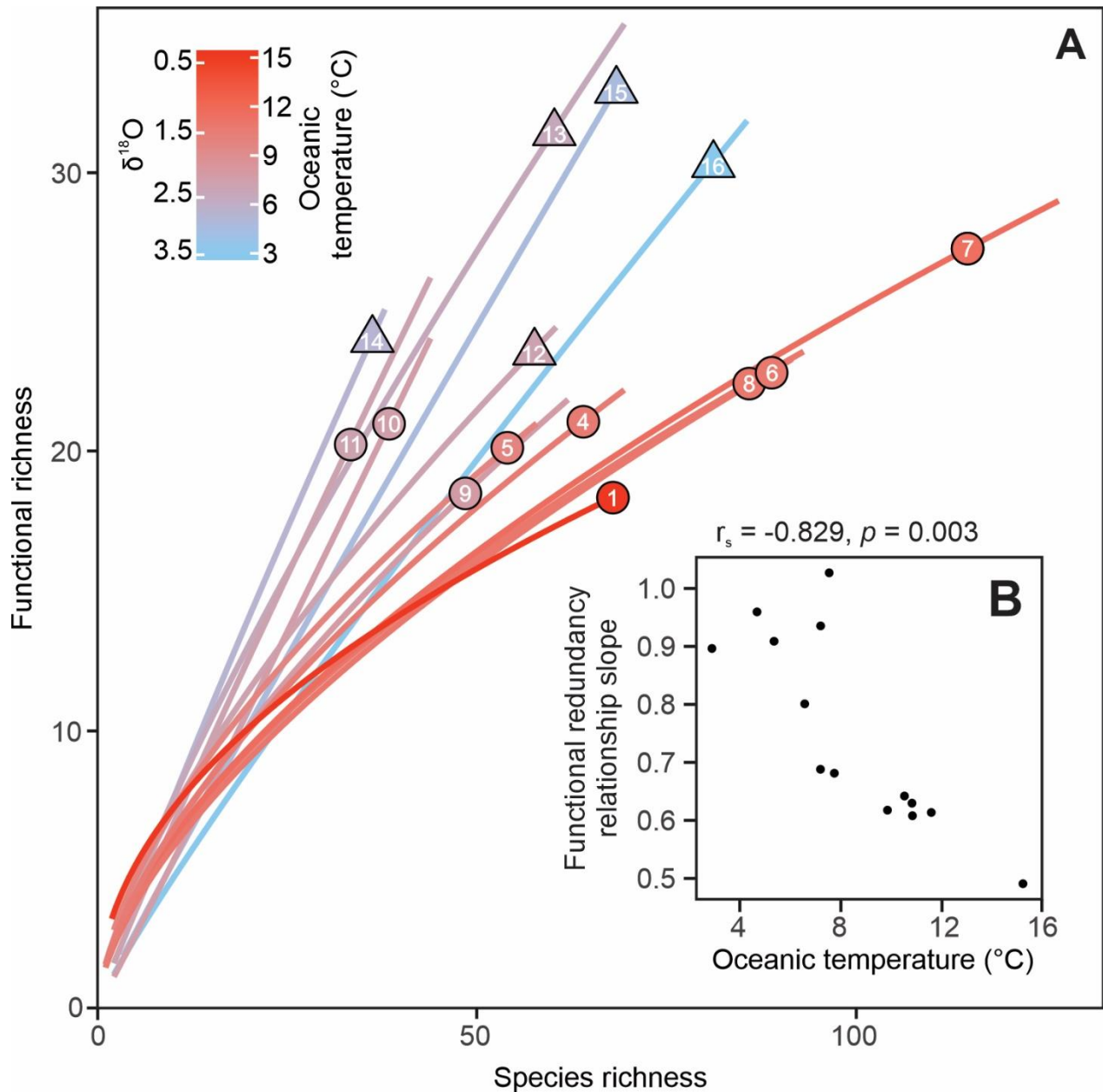
FIGURE 7.8. Relationship between functional redundancy and oceanic temperature, 45.7

0 Ma. (A) Functional redundancy relationship plotted in linear space for each time bin (cf.

Fig. 7.7, panel C, black regression line, but plotted in linear space). Individual trajectories of

the functional redundancy relationship are colour coded to $\delta^{18}O$/oceanic temperature (°C).

Points are labelled by time bin (Fig. 7.2), ordered from oldest (bin 1) to youngest (bin 16),

and represent values of species and functional richness (see methods) at a representative

spatial scale (summed MST length 1300 – 1350 km). Steeper trajectories of individual lines

represent lower functional redundancy. Triangular points represent Plio-Pleistocene time bins and circular points represent time bins older than the Pliocene (> 5.33 Ma). (B) Scatterplot of the slope of the functional redundancy relationship and oceanic temperature (°C) in each time bin. Importantly, the variables represented in this plot (B) are not independent. This is corrected for within our analyses.

Time-series correlation	Spearman-rank correlation coefficient	
	r_s	*p*
Species-area relationship slope v. oceanic temperature	0.394	0.450
Functional diversity-area relationship slope v. oceanic temperature	0.268	0.367
Functional redundancy relationship slope v. oceanic temperature	**-0.829**	**0.003**
(including bins 2 and 3)	**(-0.703)**	**(0.014)**
Δ Species-area relationship slope v. Δ no. of collections	-0.495	0.062
Δ Species-area relationship slope v. Δ time bin duration	-0.566	0.215
Δ Functional diversity-area relationship slope v. Δ no. of collections	-0.319	0.235
Δ Functional diversity-area relationship slope v. Δ time bin duration	-0.341	0.513
Δ Functional redundancy relationship slope v. Δ no. of collections	0.181	0.625
Δ Functional redundancy relationship slope v. Δ time bin duration	0.066	0.759
Δ Species richness v. Δ no. of collections	0.396	0.229
Δ Species richness v. Δ time bin duration	0.082	0.709
Δ Functional richness v. Δ no. of collections	0.500	0.205
Δ Functional richness v. Δ time bin duration	-0.005	0.976
Δ Functional richness v. Δ no. sedimentary rock packages	-0.127	0.704

TABLE 7.3. Correlations between time-series of the species-area relationship, functional diversity-area relationship and functional redundancy relationship slope versus oceanic temperature and number of collections. Where appropriate, time-series are detrended based on first differences, denoted by Δ (see methods). Statistical significance was estimated via Monte Carlo simulations using phase-randomized surrogates using the R package *astrochron* (Meyers 2014).

GLS model	Variables	AICc weight (%)	*Coef.*	Model significance *(p)*	ML-R²	phi	Log likelihood
1. Functional redundancy relationship slope							
null	-	5	-	-	-	-	-
1	Temp. + Temp. var. + Packages + Collections	<1	-	-	-	-	-
2	Temp. + Temp. var. + Collections	<1	-	-	-	-	-
3	Temp. + Temp. var.	12	-	-	-	-	-
4	Temp.	80	-0.04	<0.001	0.50	-0.007	13.55
5	Packages	<1	-	-	-	-	-
6	Temp. var.	<1	-	-	-	-	-
7	Collections	2	-	-	-	-	-
2. Spatially and sampling standardized species richness							
null	-	25	-	-	-	-	-
1	Temp. + Temp. var. + Packages + Collections	<1	-	-	-	-	-
2	Temp. + Temp. var. + Collections	6	-	-	-	-	-
3	Temp. + Temp. var.	40	0.16 0.76	0.007	0.51	-0.184	-12.56
4	Temp.	6	-	-	-	-	-
5	Temp. var.	18	-	-	-	-	-
6	Packages	<1	-	-	-	-	-
7	Collections	5	-	-	-	-	-
3. Spatially and sampling standardized functional richness							
null	-	13	-	-	-	-	-
1	Temp. + Temp. var. + Packages + Collections	<1	-	-	-	-	-
2	Temp. var. + Packages + Collections	<1	-	-	-	-	-
3	Packages + Collections	11	-	-	-	-	-
4	Packages*	28	-0.66	0.018	0.33	-0.067	-15.54
5	Collections*	29	0.63	0.017	0.33	0.061	-15.51
6	Temp. var.	2	-	-	-	-	-
7	Temp.	16	-	-	-	-	-

TABLE 7.4. Summary of generalized least squares (GLS) autoregressive (AR1) models for (1) functional redundancy relationship slopes, and spatially and sampling standardized (2) species and (3) functional richness. Explanatory variables include oceanic temperature, climatic variability and number of sedimentary rock packages and collections per time bin. All explanatory variables were standardized (mean centred and scaled to variance) prior to running models to allow comparison of variable correlation coefficients and model coefficients. Best fit models were determined using corrected-Akaike information criterion (AICc) weights using the R package *MuMIn* (Barton 2020). Pseduo-R^2 values (ML-R^2) for best fit models were calculated following the maximum likelihood approach of Cox and Snell

(1989) and model significance was calculated based on comparison to null models. *There is evidence that the number of sedimentary rock packages is negatively correlated to functional richness and the number of collections per time interval is positively correlated; however, pairwise Spearman rank correlation coefficients are statistically insignificant (Table 7.3).
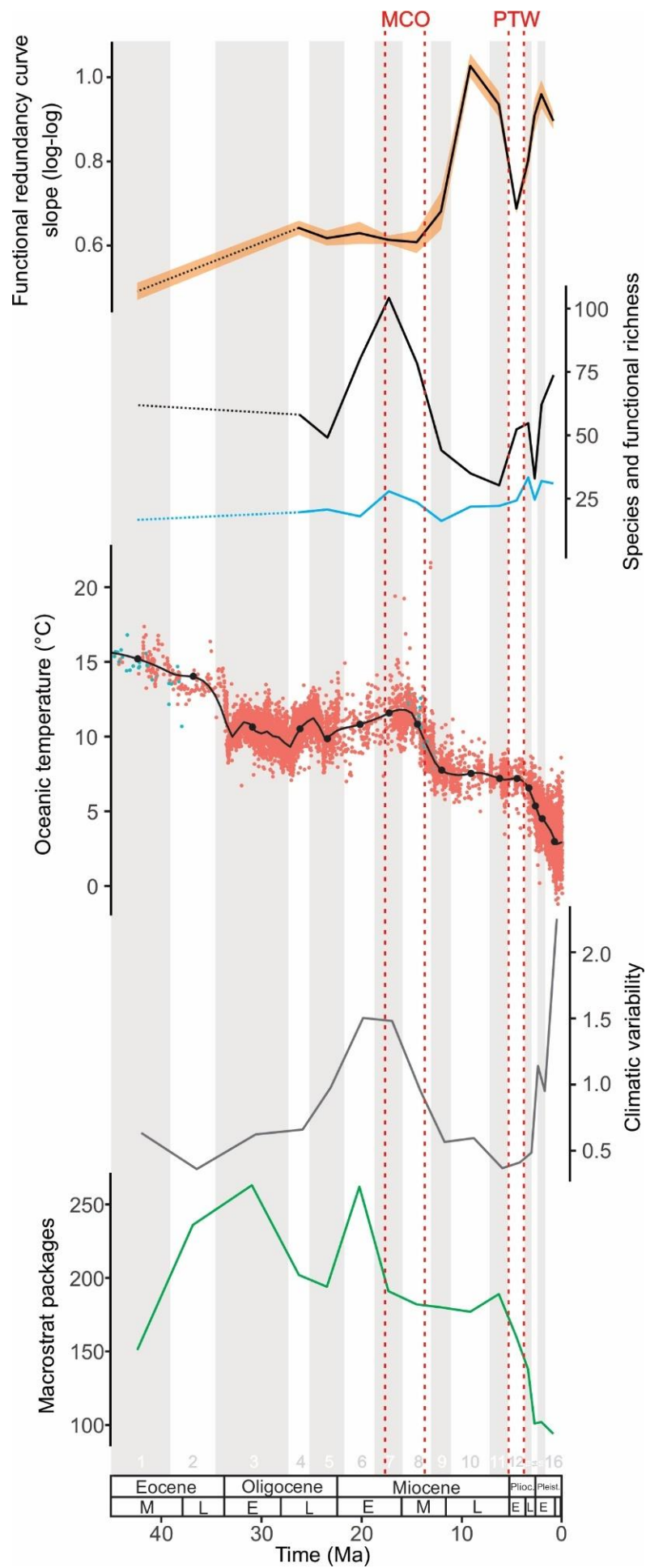
FIGURE 7.9. Time-series of the slope of the functional redundancy relationship, standardized species and functional richness, oceanic temperature, climatic variability, and number of macrostrat packages per time bin, 45.7 – 0 Ma. Analyses were undertaken at the resolution of the time bins indicated by the shaded bars, labelled by time bin number (see Fig. 7.2). Error ribbon of the functional redundancy relationship slopes represent the 2.5 % and 97.5 % confidence interval of the reduced major axis regression calculated using the R package *lmodel2* (Legendre 2014). Species (black line) and functional (blue line) richness were calculated at a representative spatial scale (summed MST length 1300-1350 km) (see methods). Oceanic temperature was calculated within each time bin based on interpolated values of a fitted loess regression at the mid-point of each time bin. Individual red points represent oceanic temperature calculated from a single data point (deep-marine benthic foraminifera). Climatic variability was calculated as the variance of the residuals from the loess regression fitted to oceanic temperature (see methods). Labelled dotted lines represent the approximate timing of the Miocene climatic optimum (labelled MCO) and a period of transient warming in the early Pliocene (PTW).

7.4.1. Temperature Dependency of Functional Redundancy

Our main conclusion is that at warmer oceanic temperatures, species richness accumulates at a faster rate relative to functional richness with increasing spatial scale. At cooler oceanic temperatures, species richness accumulates at a slower rate. This suggests that functional redundancy is relatively elevated at warmer oceanic temperatures, whereby more species occupy the same functional trait space (here used as a proxy for ecological niche). Standardized measures of functional evenness – a measure of how uniformly species are distributed amongst functional groups – suggest that species are generally less evenly distributed between functional groups during global warmth of the Miocene climatic optimum (Fig. 7.10), where functional evenness is low. Functional evenness subsequently increases with global and regional cooling in New Zealand during the middle Miocene climate transition. A similar pattern is observed in the modern latitudinal gradient, where bivalve species and genera are less evenly spread amongst functional groups at mid-latitudes and become more evenly dispersed poleward (Edie et al. 2018, Schumm et al. 2019). The decrease in functional evenness during the Miocene climatic optimum coincides with a significant decrease in the rate and magnitude of temporal turnover of functional groups – the change within and between functional groups over time (see methods). The decrease in temporal turnover is facilitated predominately by a decrease in the replacement of functional groups (Fig. 7.10), indicating the increase in species richness during the Miocene climatic optimum is accommodated largely by existing rather than new functional groups. At a representative fixed spatial scale, functional richness remains relatively static through much of the Cenozoic, except for the Plio-Pleistocene (5.33 – 0 Ma) (Figs. 7.4 and 7.6 and discussed below). Together, these findings suggest that increases in regional molluscan

species richness during the warm Miocene climatic optimum are primarily a result of increased niche packing (i.e., more species occupying the same niche, rather than niche expansion), and the subsequent decrease in species richness during the Miocene climatic transition has minimal impact on functional richness relative to species richness.
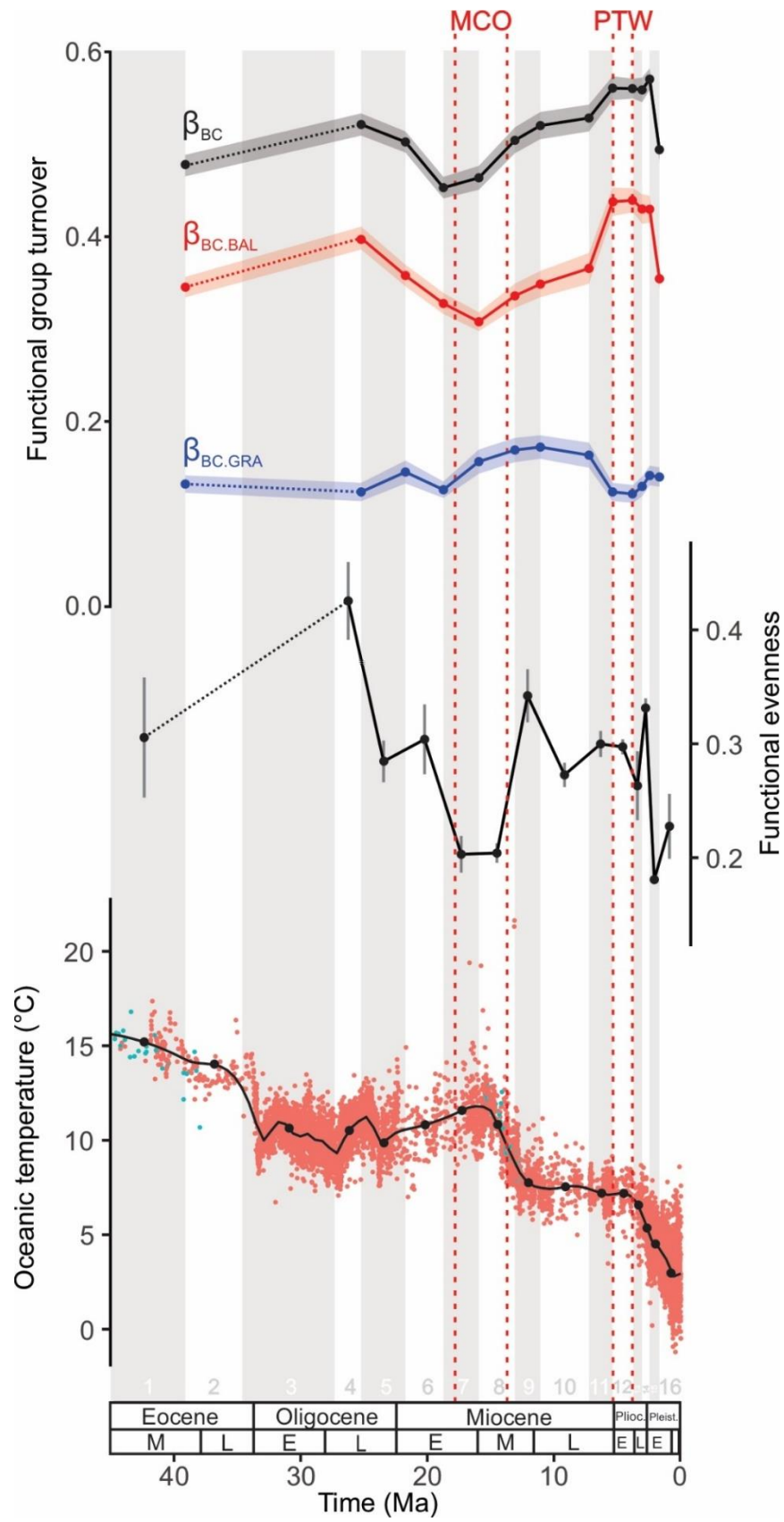
FIGURE 7.10. Time-series of functional group temporal turnover (pairwise Bray-Curtis dissimilarity), functional evenness and oceanic temperature, 45.7 – 0 Ma. Temporal turnover was calculated between successive time bins, plotted at their boundaries, highlighted by the shaded bars, labelled by time bin number (**Fig.** 7.2). β_{BC} (top, black line) represents total dissimilarity, $\beta_{BC.BAL}$ (middle, red line) represents balanced variation in abundance (i.e., substitution, analogous to replacement of functional groups), and $\beta_{BC.AB}$ (bottom, blue line) represents abundance gradients (i.e., subsets, analogous to nestedness) (see methods and Baselga et al. 2017). Error ribbons represents ± 1 standard error. Functional evenness was calculated at a representative spatial scale (summed MST length 1300-1350 km) and error bars represent ± 1 standard error. Oceanic temperature was calculated within each time bin based on interpolated values of a fitted loess regression at the mid-point of each time bin. Individual red points represent oceanic temperature calculated from a single data point (deep-marine benthic foraminifera). Analysis of functional evenness and oceanic temperature were undertaken at the resolution of the highlighted time bins. Labelled dotted lines represent the approximate timing of the Miocene climatic optimum (labelled MCO) and a period of transient warming in the early Pliocene (PTW).

7.4.2. Temperature Dependency of Regional Carrying Capacity

The correlation between temperature and the slope of the functional redundancy relationship implies a relationship between oceanic temperature and regional standing diversity. This is further supported by the positive relationship of species richness with oceanic temperature and climatic variability. Whereas the notion of an ecological carrying-capacity or bound on biodiversity is debated (Harmon and Harrison 2015, Rabosky and Hurlbert 2015, Marshall and Quental 2016, Storch and Okie 2019), our results suggest

that lower oceanic temperatures impose an ecological constraint on regional species richness in New Zealand, particularly before the Plio-Pleistocene (> 5.33 Ma). Notwithstanding this, when we control for spatial scale, we find slightly elevated species richness and elevated functional richness during three Plio-Pleistocene time bins (bins 13, 15 and 16, see **Fig.** 7.6). The Plio-Pleistocene is a period of known climatic variability (Raymo et al. 2006) and, within New Zealand, increased landscape complexity related to accelerated tectonic activity (King et al. 1999), maturation of oceanic fronts and intensification of marine latitudinal temperature gradients (Nelson and Cooke 2001) and is initially marked by a period of transient warming (Beu and Maxwell 1990, Herbert et al. 2016). We find evidence that oceanic temperature and climatic variability contribute positively to species richness, which explains, in part, the early Pliocene and Pleistocene increase in species richness. In contrast, we find no consistent explanation for increases in functional richness. However, we do observe a significant increase in temporal turnover of functional groups at the Plio-Pleistocene boundary, facilitated predominately by an increase in the replacement of functional groups (**Fig.** 7.10). Although it is not possible to determine the exact causes of the increase in functional richness during the Plio-Pleistocene, one plausible cause for the increase in functional group turnover at the Pliocene boundary, and perhaps functional richness, may be the intensification of marine latitudinal gradients around New Zealand. Whereas oceanic temperature is implicated here, the exact mechanisms responsible for the temperature-functional redundancy relationship – the causal cascade – are not clear and could feasibly be related to both abiotic and biotic factors. Overall, our findings suggest New Zealand's species richness fluctuates with oceanic temperature, corroborating recent research that suggests macroevolutionary models should incorporate time-varying (i.e., not static) carrying capacities of species richness (Marshall and Quental 2016).

7.5. Summary

Predicting the ecological response to climate change requires an understanding of how ecosystems have responded to climatic fluctuations in the past, within a spatial context. This is particularly important for temperate marine ecosystems, like those of New Zealand, which span climatic, oceanographic, and biogeochemical boundaries (Reygondeau et al. 2013) where the largest changes in species and functional richness are observed (Edie et al. 2018). We find that despite key differences in both temporal and spatial resolution between paleontological and ecological data, the functional diversity-area relationship in the fossil record conforms to similar patterns observed in modern ecological studies (Whittaker et al. 2014, Karadimou et al. 2016) and, importantly, is consistent over long spans of time at a regional spatial scale. This suggests the functional diversity-area relationship is equally as persistent and predictable as the species-area relationship. Furthermore, the observed correlation between the slope of the functional redundancy relationship and oceanic temperature over the last ~40 Myrs is consistent with observations from the modern latitudinal diversity gradient (Schumm et al. 2019), suggesting a persistent relationship of regional and global ecological importance. Taken at face value, our results suggest that oceanic temperature should increase net species richness and functional redundancy in New Zealand over long-time spans (i.e., multi-centennial to millennial), particularly as we shift to a climate more representative of pre-Pleistocene conditions (Burke et al. 2018). This provides a baseline for what should be expected – at ecological equilibrium – from natural warming in regional, temperate, shallow-marine ecosystems.

Oceanic warming associated with anthropogenic climate change is taking place at a rapid rate and is occurring in tandem with elevated rates of biodiversity loss and habitat destruction (Erwin 2009). The combined effect of these factors on the slope of the functional

redundancy relationship is not clear; however, decreasing habitable area from anthropogenic activity is likely to steepen the slope and decrease functional redundancy. This may alter the future trajectory of the functional redundancy slope predicted from oceanic warming using past analogues and decrease ecosystem resilience (Hooper et al. 2005, Biggs et al. 2020). This is particularly important for New Zealand as it is geographically isolated and has a limited species pool with high levels of endemism.

Chapter 8

Biogeographic Trends and Climatic Change:
Comparisons Through Space and Time

Time - "A big ball of wibbly wobbly, timey wimey stuff"

(*Dr. Who*, 2007)

Preface

Anthropogenic climate change is predicted to have large scale impacts on the biogeography and biodiversity of shallow-marine life (Bindoff et al. 2019, Blowes et al. 2019, Lotze et al. 2019, Antão et al. 2020). Observations and predictions from the distribution of current marine life suggest that tropical and temperate ecosystems will be impacted asymmetrically from oceanic warming. This is based on the prediction of a positive gain in species richness within temperate ecosystems (from the poleward migration of warm water species) and negative impact on tropical ecosystems (due to species loss from elevated heat stress) (Antão et al. 2020, Worm and Lotze 2021). However, the short and long-term impacts of future climatic change on other spatial partitions of taxonomic diversity (i.e., alpha and beta diversity) and on ecosystem function (here measured as functional diversity) are largely unknown. The latter is particularly important as evidence from modern ecology suggests functional diversity and functional redundancy increase ecosystem resilience to environmental change (Biggs et al. 2020).

The distribution of marine life across the modern latitudinal diversity gradient indicates a parallel poleward decrease in taxonomic and functional richness (also termed diversity within this chapter). The poleward decline in taxonomic diversity is attributed, in part, to latitudinal gradients of climate and temperature (Edie et al. 2018b, Schumm et al. 2019, Gagné et al. 2020). However, the fossil record suggests that that taxonomic and functional diversity become decoupled over longer timescales, particularly during intervals of elevated taxonomic loss (Edie et al. 2018b). This spatiotemporal mismatch confounds our understanding of how functional and taxonomic biogeography and biodiversity may respond to future climatic change over long-timescales.

This chapter examines how the Cenozoic shallow-marine fossil record can contribute to our understanding of these contrasting biogeographical spatiotemporal patterns. To do this I quantify the role of climatic change (measured as oceanic temperature and climatic variability), and two other potentially important environmental variables (environmental heterogeneity and habitable area), in controlling the spatial structuring of New Zealand's shallow-marine molluscan biodiversity and macroevolutionary rates (i.e., origination and extinction) through the Cenozoic. I consider: (i) whether taxonomic and functional diversity exhibit congruent temporal and spatial patterns, (ii) whether they respond in the same way to the potential explanatory factors considered here, (iii) how they compare to similar quality regional fossil records from the tropics, and (iv) what the implications are for the future of New Zealand's biodiversity, and similar temperate shallow-marine ecosystems.

In essence, this chapter summarizes much of the work completed within the previous chapters and builds upon the methods and knowledge developed throughout this thesis (particularly Chapter 7). However, as this chapter is formatted as a paper, I also include an additional chapter (Chapter 9) to provide a more detailed summary of the thesis. Chapter 8 is a modified version of a manuscript that will shortly be submitted for publication. Co-authors to this paper include the primary and secondary supervisor of this thesis and D. Strogen and H. Seebeck (both from GNS Science). Input from my primary and secondary supervisor was provided through regular discussion, advice, and reviewing the manuscript. H. Seebeck assisted in the generation of paleocoordinates for fossil collections and stratigraphic data. D. Strogen provided the paleogeographic/paleobathymetric maps (which were created in conjunction with H. Seebeck and the contribution of other scientists at GNS Science) and assisted in quantifying measures of habitable shelf area for New Zealand.

8.1. Introduction

Even under the most conservative estimates, oceanic warming due to anthropogenic climate change over both transient and geological timescales is likely to have a significant impact on the biogeography and biodiversity of shallow-marine life (Mayhew et al. 2008, Blois et al. 2013, Poloczanska et al. 2013, Finnegan et al. 2015, Urban 2015, Bindoff et al. 2019, Blowes et al. 2019, Lotze et al. 2019, Antão et al. 2020). Global studies of taxonomic diversity in the fossil record, and the geographic distribution of species richness along the modern latitudinal gradient, implicate climate, particularly oceanic temperature, as a major positive correlate of marine biodiversity (Mayhew et al. 2008, Mayhew et al. 2012, Edie et al. 2018b, Schumm et al. 2019, Gagné et al. 2020) (see Chapter 2, section 2.4). However, modern conservation efforts are generally focused on spatial scales at, or smaller than, ecological regions, and there is little consensus on the importance of oceanic temperature in the spatial structuring of regional marine ecosystems over multi-centennial to millennial timescales or longer. Here we quantify the roles of oceanic temperature, climatic variability, environmental heterogeneity, and habitable area in controlling the spatial structuring of shallow-marine biodiversity and macroevolutionary rates at the regional scale, using New Zealand's exemplary Cenozoic molluscan fossil record.

Recent research suggests that tropical and temperate marine ecosystems are likely to be impacted asymmetrically by increasing oceanic temperatures over short timescales. Observations and predictive models generally indicate a net increase in species richness in temperate ecosystems in response to warming, largely attributed to poleward migration of warm water species, and concomitant decrease in tropical species richness due to elevated heat stress (Antão et al. 2020, Worm and Lotze 2021). If this holds true over multi-centennial to millennial timescales, we would expect regional shallow-marine biodiversity in temperate

ecosystems (gamma diversity) to be positively correlated with regional oceanic temperature. This could be facilitated by either an increase in origination rate and immigration, or a decrease in extinction rate, or a combination of these factors. However, predictions for the long-term impact of oceanic warming on taxonomic diversity in temperate ecosystems is poorly understood. This is true for diversity at both the local scale (alpha diversity) and in terms of change in community composition along an environmental gradient (beta diversity). Changes in taxonomic diversity at these scales could be overprinted or intensified by other processes such as changes in climatic variability, environmental heterogeneity, or habitable area.

Uncertainty regarding the transient and long-term impact of oceanic temperature on taxonomic diversity is even more marked for ecosystem function, here measured as functional diversity (the number of distinct ecological functional groups occupied by individual species). Functional diversity is becoming increasingly recognised as a fundamental facet of biodiversity in ecological (Petchey and Gaston 2006) and paleontological research (Villéger et al. 2011, Edie et al. 2018a, b, Pimiento et al. 2020). Importantly, observations from modern ecology suggest functional diversity and functional redundancy – the duplication of functional roles by multiple species – increase ecosystem resilience to environmental change (Biggs et al. 2020). This suggests that maintenance of functional diversity is equally as important in biological conservation as the preservation of species richness (Petchey and Gaston 2006, Violle et al. 2014). However, several studies in the fossil record suggest that taxonomic and functional diversity are decoupled through time (Erwin et al. 1987, Foster and Twitchett 2014, Dunhill et al. 2018, Edie et al. 2018b, Alvarez et al. 2019). This contrasts with the congruent spatial patterns observed between functional and taxonomic diversity along the modern marine latitudinal diversity gradient (Edie et al. 2018b, Schumm et al. 2019). Incorporating functional diversity in the same spatial

framework as taxonomic diversity is fundamental to understanding the nature of this decoupling.

Within this study we focus on the role of oceanic temperature, climatic variability, environmental heterogeneity, and habitable area on the taxonomic and functional biogeography of New Zealand's shallow-marine molluscan ecosystems over a time span of ~45 Myrs (Fig. 8.1). We employ several new methods to derive spatially and sampling-standardized measures of partitioned biodiversity (alpha, beta and gamma). In addition, we consider the implications for the future of New Zealand's shallow marine biodiversity and similar temperate regional ecosystems. Our key findings are that oceanic warming seemingly has both a long-term positive effect on regional taxonomic diversity and rates of origination, but a negative correlation with regional functional diversity and its spatial distribution (functional beta diversity).
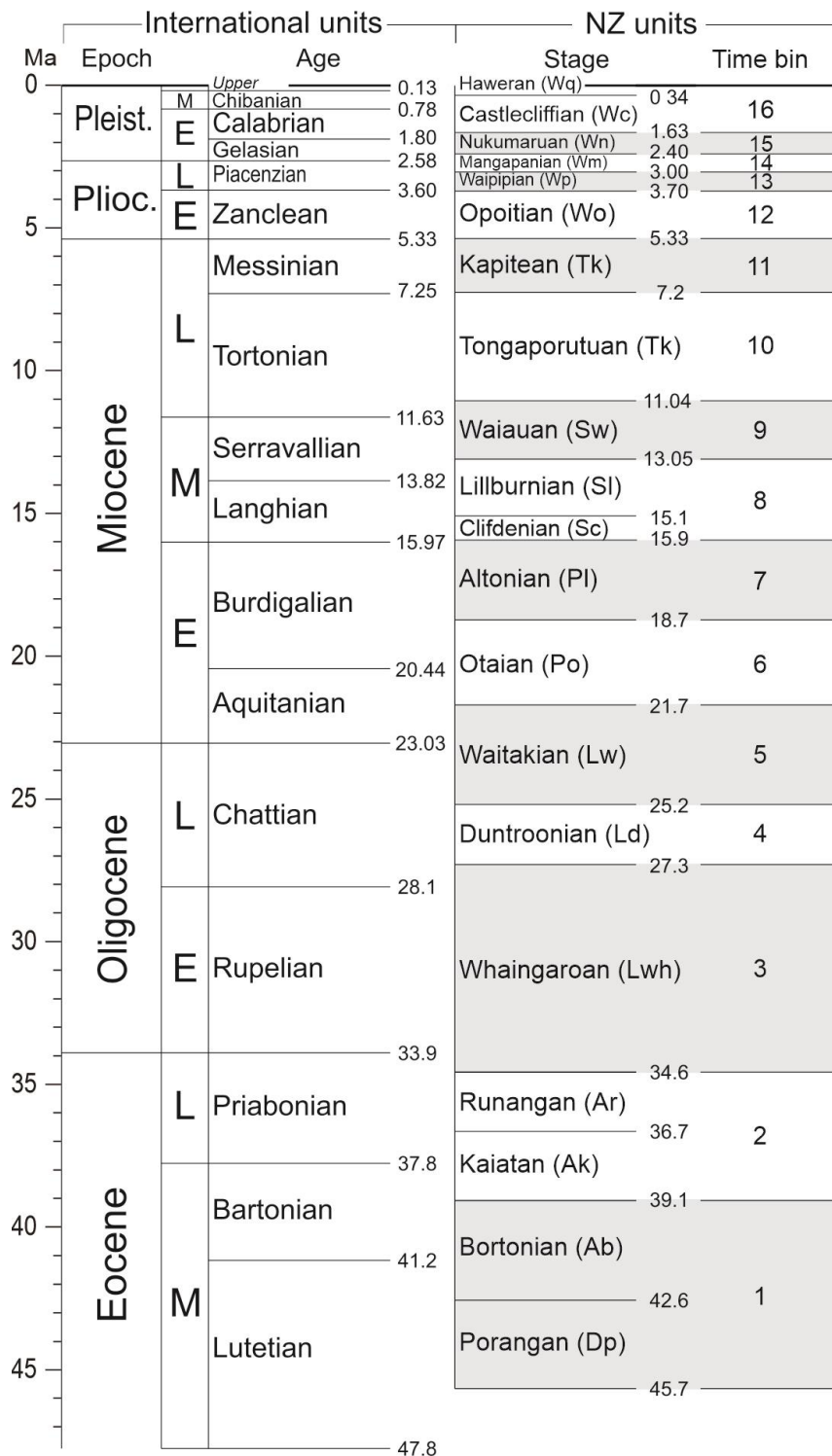
FIGURE 8.1. New Zealand Cenozoic timescale (after Raine et al. 2015). Analyses are undertaken at the resolution of the numerically labelled and highlighted time bins shown here (shaded and unshaded blocks in the right-hand column). Ages are in Ma.

8.2. Methods

8.2.1. Measuring Biodiversity

Biodiversity was quantified in terms of both molluscan taxonomic diversity at the species-level and functional diversity (the number of unique functional groups occupied at the species-level), following the procedures and definitions in Chapter 7, section 7.2. Following Womack et al. (in review), we also omitted two time bins from our biodiversity calculations spanning the mid-late Eocene (Bin 2 - Kaiatan to Runangan Stage: 39.1 – 34.6 Ma, see Fig. 8.1) and the Early Oligocene (Bin 3: Whaingaroan Stage: 34.6 – 27.3 Ma, see Fig. 8.1). These were removed due to the low number and limited spatial distribution of fossil collections from these intervals. Individual fossil collections were assigned to an equal-area penta-hexagrid of varying sizes (side of 23 km, 42 km and 80 km) based on their paleocoordinates using GPlates (Müller et al. 2018), following the methodology outlined in Chapter 6, section 6.2.3 (and see Chapter 4, section 4.3), using the R package *icosa* (Kocsis 2020). Grid sizes were chosen to represent spatial areas at the local (side of 23 km), intermediate (42 km) and basinal (80 km) scales, relative to New Zealand's Cenozoic paleogeography.

Taxonomic and functional diversity were partitioned into gamma, beta and alpha diversity (Whittaker 1960). To accommodate for temporal variation in sampling effort, we applied sampling standardization to each diversity partition. All partitions of taxonomic diversity and functional beta diversity (see below discussion) were standardized using coverage-based subsampling, implemented using the R package *iNEXT* (Hsieh et al. 2016). Gamma and alpha functional diversity were standardized using the corrected first-order jacknife extrapolator, using code provided in Alroy (2020) (see Chapter 4, section 4.2). To accommodate for uneven paleogeographic distribution of fossil collections through time, we

defined gamma diversity at a fixed summed and binned minimum spanning tree (MST) length of 1225 – 1375 km, following the algorithm outlined in Chapter 6 (and see Chapter 7, section 7.2). Beta diversity was measured at the same spatial scale as gamma diversity using the multi-site measure of Sorensen dissimilarity (Baselga 2010), implemented in the R package *betapart* (Baselga and Orme 2012). We extended the multi-site method of beta diversity to include sampling standardization using coverage-based subsampling for both taxonomic and functional beta diversity. Whereas the cJ1 estimator was used to measure functional gamma and alpha diversity, the extrapolator cannot generate frequency distributions required to estimate functional beta diversity (see Chapter 7, section 7.2.3). Hence, we employ coverage-based subsampling in our estimation of functional beta diversity. We defined beta diversity as the average Sorensen dissimilarity between grid cells at the local, intermediate and basinal scales, based on 100 subsampling trials. For each trial, individual grid cells that did not meet a set coverage threshold (here set to 0.4) were omitted. For each grid cell that passed this threshold, a subsampled frequency distribution was created (again at a coverage of 0.4), using the code of Holland (2015). Multi-site beta diversity was then calculated based on the subsampled frequency distributions (converted to binary data) of all available grids. Lastly, alpha diversity was defined as the average diversity, standardized using coverage-based subsampling, across all individual grid cells.

8.2.2. Macroevolutionary Rates

Macroevolutionary rates were estimated using capture-mark-recapture origination and extinction probabilities using the Pradel seniority model (Liow and Nichols 2010, Liow and Finarelli 2014), converted to instantaneous rates (Smiley 2018). Comparisons between the capture-mark recapture rates and other rate estimation methods (calculated using R package *divDyn* (Kocsis et al. 2019) reveal strong similarities (Table 8.1) (and see Chapter 4, section 4.7). Here, we base interpretations on the capture-mark-recapture rates because these

incorporate the full encounter history of each species, information that is wholly or partially lost to the other methods trialled. To compare macroevolutionary rates to our environmental proxies, we created three models to allow for inclusion of environmental heterogeneity at the local, intermediate and basinal scales (as outlined in section 8.2.5).

Rate	Comparison	r_s	p
Origination	Δ Capture-mark recapture v. Δ per-capita rate Foote (2000)	0.84	0.004
	Δ Capture-mark recapture v. Δ gap-filler rate Alroy (2014)	0.77	0.022
Extinction	Δ Capture-mark recapture v. Δ per-capita rate Foote (2000)	0.63	0.019
	Δ Capture-mark recapture v. Δ gap-filler rate Alroy (2014)	0.80	0.009

TABLE 8.1. Comparisons between methods of estimating macroevolutionary rates. Other rates considered include the per-capita rates of Foote (2000) and gap-filler rates of Alroy (2014). Comparisons are based on first-differences (indicated by Δ) using the Spearman rank-order correlation coefficient. In order to account for serial correlation of the data, significance values were estimated using phase-randomized surrogate data, as implemented in the R package *astrochron* (Meyers 2014).

8.2.3.   Oceanic Temperature and Climatic Variability

Oceanic temperature and climatic variability were calculated following the method outline in Chapter 7 (and see Chapter 4, section 4.4). Oceanic temperature was quantified using loess-interpolated temperatures (at individual bin mid-points, with a loess span of 0.1) of Pacific Ocean deep-ocean foraminifera $\delta^{18}O$ values (Cramer et al. 2009) (Fig. 8.2). These values were converted to °C using the linear transfer function following the methodology of Veizer and Prokoph (2015), and using only data from *Cibicidoides* spp. We measured climatic variability as the variance of the residuals of raw data points from the interpolated loess regression used to quantify average Cenozoic Pacific Ocean temperatures (Fig. 8.2). Residual plots suggest several outliers may be present within each time bin. These may

represent true data points but may also represent local climatic conditions at the specific borehole location or erroneous measurements (see Chapter 4, section 4.5). We opted to remove values that fall outside the 95th percentile; including outliers, however, this does not change the main findings of this study.

A

local

intermediate

basinal

Environmental heterogeneity

0.90

0.85

0.80

B

Shelf area (km²)

4e+05

3e+05

2e+05

1e+05

C

Climatic variability

2.0

1.5

1.0

0.5

D

Oceanic temperature (°C)

20

15

10

5

0

1 2 3 4 5 6 7 8 9 10 11 12 16

| Eocene | | Oligocene | | Miocene | | | Plioc. | Pleist. |
| M | L | E | L | E | M | L | E | L | E | I |

Time (Ma)

40   30   20   10   0
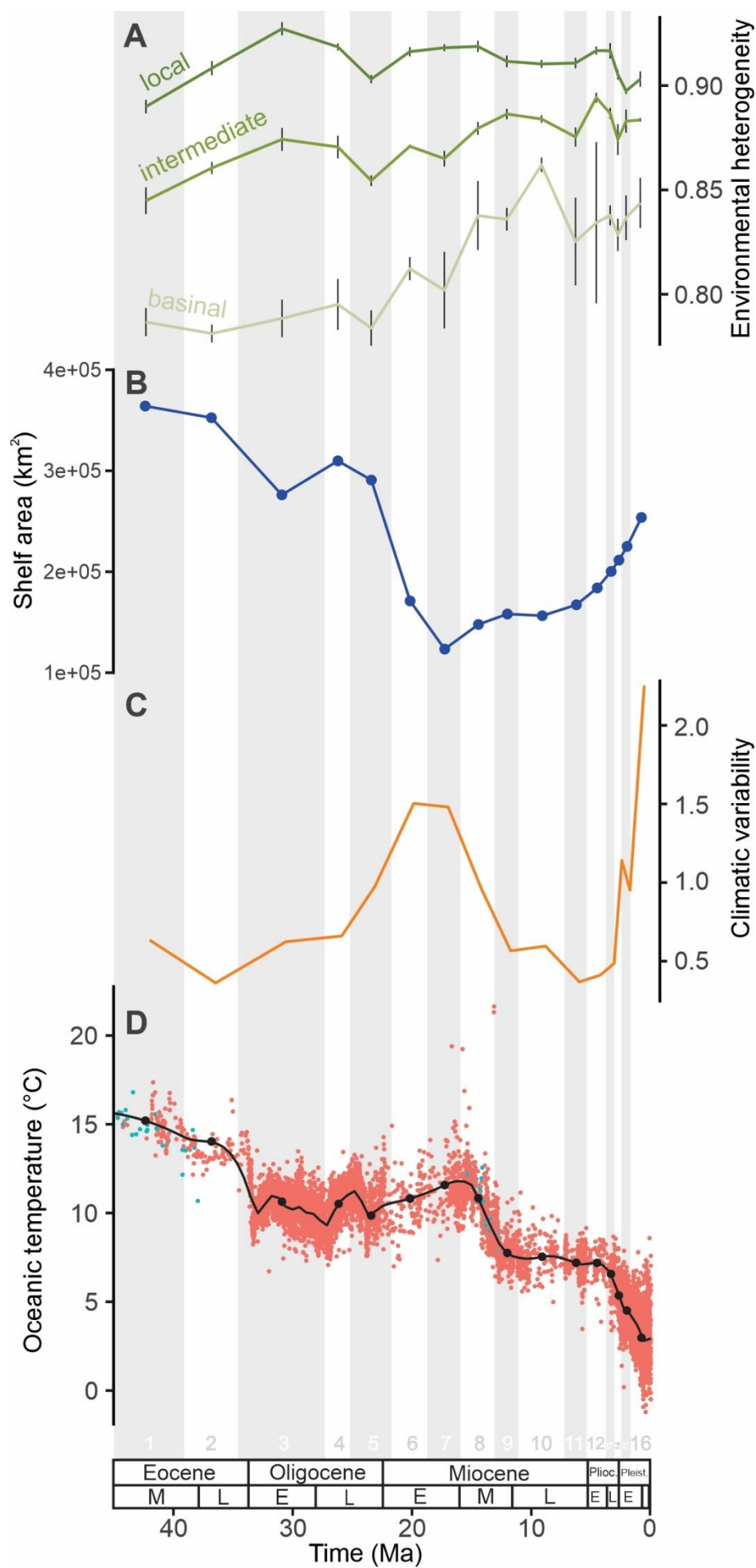
FIGURE 8.2. Time series of key environmental variables through the Cenozoic. (A) Environmental heterogeneity measured as lithological dissimilarity between local, intermediate and basinal grid sizes (see section 8.2.3). Error bars represent ± 1 standard error. (B) Interpolated shelfal area (inner to outer shelf: inferred water depths of 0 – 200 m) derived from spline regression of areas calculated using detailed paleogeographic/paleobathymetric maps of New Zealand. (C) Climatic variability calculated from variance of the residuals of the loess regression used to interpolate oceanic temperature (D). Analyses were undertaken at the resolution of the time bins indicated by the shaded bars, labelled by time bin number (see Fig. 8.1).

8.2.4. Environmental Heterogeneity

Environmental heterogeneity was quantified using an independent dataset compiled from Macrostrat (Peters et al. 2018) by Crampton et al. (2011) This dataset includes a number of stratigraphic sections from across New Zealand that were collated from King et al. (1999), where individual units and formations were aggregated into units of Macrostrat - sedimentary rock packages, herein referred to as packages. Each package includes information on the lithology and depositional environment (the latter encoded as a binary variable: shelfal - considered to have been deposited in a marine environment between 0-200 m water depth, or non-shelfal - terrestrial, or bathyal). To allow meaningful spatial comparisons between measures of biodiversity and environmental heterogeneity, we measured environmental heterogeneity following a similar methodology to gamma diversity. We assigned individual packages to an equal-area penta-hexagrid (using the same three grid sizes noted above) based on their derived paleocoordinates (Fig. 8.2). Where a sedimentary rock package spanned several time bins, we calculated paleocoordinates for that package within each time bin.

Environmental heterogeneity (see Fig. 8.2) was calculated as the average lithological dissimilarity of shelfal packages, and was measured as the multi-site Sorensen dissimilarity of binary lithological states (also implemented in R package *betapart*), at a fixed summed MST length of 1225 – 1375 km. Following Crampton et al (2011), we included unconformities as a coded "lithology" within the analysis on the assumption that intervals of erosion represent an important marine paleoenvironment. Unconformities were only included between packages where at least the preceding or following packages were identified as shelfal.

High environmental heterogeneity

SS MS	SS LM
SS CG UN	SS MS

Low environmental heterogeneity

SS LM	LM
SS LM UN	LM
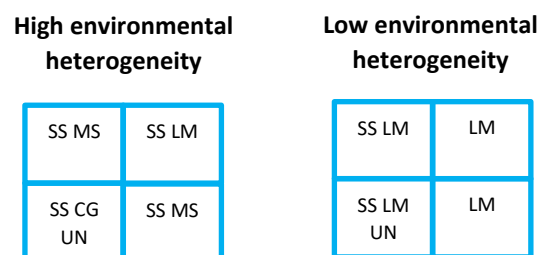
FIGURE. 8.3. Conceptual diagram of a time interval with high versus low environmental heterogeneity based on lithological dissimilarity (SS = Sandstone, MS = Mudstone, CG = Conglomerate, UN = Unconformity). Squares represent individual grid cells (hexagrid in the analysis presented here).

8.2.5. Habitable Area

Habitable area was measured using unpublished detailed Cenozoic paleogeographic maps of New Zealand, compiled by D. Strogen and H. Seebeck (Fig. 8.2, and see Fig. 2.5). The Paleogeographic maps include contours of bathymetric depth based on evidence from fossil and stratigraphic data, compiled for 10 intervals through the Cenozoic. Habitable shelf area was calculated as the area around the main New Zealand landmass between the inferred water depths of 0 – 200 m (see Fig. 8.4). As the maps are not at the same temporal resolution as the fossil and stratigraphic datasets used here, we calculated habitable shelf area using a spline regression interpolated at the mid-points of relevant time bins. Spline regression was used to ensure that the regression intersected the known fixed data points.

KEY

- = Landmass
- = Inner-outer shelf (0-200 m)
- = Upper-mid bathyal (200-1000 m)
- = Lower bathyal (1000-1500 m)

A
4 Ma

B
24 Ma
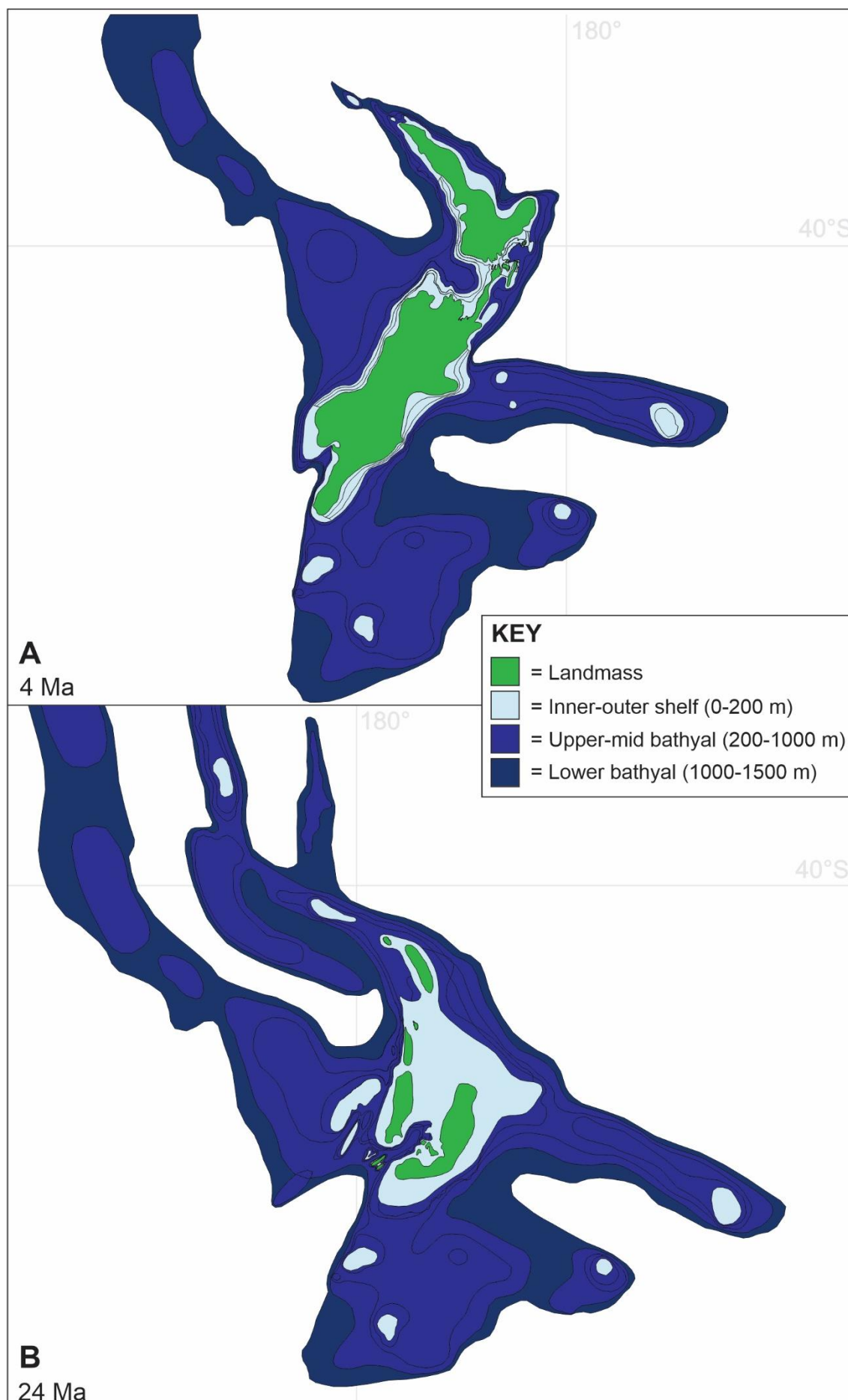
FIGURE 8.4. Paleobathymetric maps of New Zealand from 4 Ma (A) and 24 Ma (B). Habitable area was calculated between the inner to outer shelf contours (0-200 m). The maps show the considerable increase in habitable shelf area (inner to outer shelf) during the Late Oligocene, the time of maximum flooding of the New Zealand subcontinent, relative to Plio-Pleistocene. Maps modified from base maps provided by D. Strogen.

8.2.6. Time Series Comparisons

We assessed the long-term relationships between the biodiversity (taxonomic and functional partitions) of shallow-marine molluscs and their macroevolutionary rates, with the proxies for oceanic temperature, climatic variability, environmental heterogeneity, and habitable area. To assess relationships between these variables we used additive generalized least squares autoregressive (AR1) models using the R package *nlme* (Pinheiro et al. 2020) (Tables 8.2 and 8.4). Measures of biodiversity were standardized for variation in geographic and temporal distribution of sampling (following the procedure outlined in Section 8.2.1 and in more detail in Chapter 7, section 7.2). All explanatory variables were standardized (mean centred and scaled to unit variance) to allow comparison of variable correlation coefficients prior to running models and model coefficients. Best fit models were determined using corrected-Akaike information criterion (AICc) weights using the R package *MuMIn* (Barton 2020). We ran generalized least squares models at three geographic resolutions discussed previously (local, intermediate and basinal).

8.3. Results and Discussion

8.3.1. Diversity dynamics

Our results suggest that New Zealand's regional shallow-marine taxonomic diversity (gamma diversity) is affected positively by oceanic temperature, climatic variability, and within-basin environmental heterogeneity (i.e., environmental heterogeneity at local and intermediate scales) (Table 8.2 and Fig. 8.5 for time series). Furthermore, we find evidence that the relationship between regional taxonomic diversity and environmental heterogeneity may be, in part, related to variation in taxonomic beta diversity at the local scale. This supports the hypothesis that oceanic warming is a driver of regional temperate shallow-marine biodiversity, but also implicates other aspects of climatic change, i.e., climatic variability, and environmental heterogeneity, as approximately equally important factors. However, we find no evidence that any of the environmental variables examined within this study can explain variations in taxonomic alpha diversity (Table 8.2). This implies that biogeographic patterns of taxonomic diversity at the local to basinal scales that result from climatic change are overprinted or intensified by other factors not considered here. Taken together, these results contrast with results from short-term modelling experiments, which suggest that the most important environmental determinant of living marine biodiversity, other than water depth, is average water temperature (Gagné et al. 2020).

FIGURE 8.5. Time series of taxonomic diversity through the Cenozoic, partitioned into (A) alpha, (B) beta and (C) gamma diversity. For clarity error bars are only shown for partitions of taxonomic diversity at the local scale and represent ± 1 standard error. Analyses were undertaken at the resolution of the time bins indicated by the shaded bars, labelled by time bin number (see Fig. 8.1).

	Diversity partition	Grid size (hexagrid side km)	Explanatory variable(s)	Coefficient(s)	Pseudo-R2	phi	AICc weight (%)	p	Log-likelihood
Taxonomic	Gamma	23	Temp. var. + Env. Het.	14.2, 21.7	0.63	-0.74	48	0.001	-53.51
		42	Temp. + Temp. var. + Env. Het.	18.1, 21.8, 12.7	0.71	-0.53	29	<0.001	-51.51
		80	Temp. var.	11.7	0.29	0.27	30	0.030	-56.89
	Beta	23	Env.Het.	0.01	0.27	0.23	31	0.037	37.49
		42	*Area*	-0.02	0.32	0.46	23	0.019	35.67
		80	-	-	-	-	35	-	-
	Alpha	23	-	-	-	-	33	-	-
		42	-	-	-	-	34	-	-
		80	-	-	-	-	40	-	-
Functional	Gamma	23	*Temp.*	-3.82	0.42	-0.02	39	0.006	-35.46
		42	*Temp.*	-3.31	0.32	0.12	33	0.019	-37.23
		80	*Temp.*	-3.91	0.37	-0.2	47	0.011	-39.76
	Beta	23	*Temp.*	-0.02	0.5	-0.35	45	0.002	33.43
		42	*Temp.*	-0.02	0.39	0.11	50	0.020	33.28
		80	-	-	-	-	36	-	-
	Alpha	23	-	-	-	-	62	-	-
		42	-	-	-	-	32	-	-
		80	-	-	-	-	32	-	-

TABLE 8.2. Summary of generalized least squares (GLS) autoregressive (AR1) models. Results are presented for the best-fit models for taxonomic and functional diversity, both split into alpha, beta and gamma diversity and subdivided into the three grid sizes employed within this study. Model comparison was undertaken using the weighted corrected-Akaike information criterion (AICc). Explanatory variables in italics represent negative relationships. All explanatory variables were standardized (mean centred and scaled to unit variance) prior to running models to allow comparison of variable correlation coefficients and model coefficients. Pseduo-R^2 values (ML-R^2) for best fit models were calculated following the maximum likelihood approach of Cox and Snell (1989).

Functional gamma and beta diversities are correlated negatively with oceanic temperature as the sole explanatory variable (Table 8.2, see Fig. 8.6 for time series). Previous studies on New Zealand's Cenozoic shallow-marine molluscan fossil record observed a positive correlation between functional redundancy and oceanic temperature through the Cenozoic (see Chapter 7), where functional redundancy is defined as the duplication of functional roles by multiple species. The findings of the present study are complementary and suggest that the relationship between oceanic temperature and functional redundancy may, in part, be mediated by functional beta diversity, i.e., increased functional compositional similarity between grid cells at warmer temperatures. Importantly, the long-term negative correlation between oceanic temperature and functional gamma and beta diversities contrasts with the observed spatial pattern of poleward decreasing functional diversity along the modern latitudinal diversity gradient (Edie et al. 2018b, Schumm et al. 2019). Furthermore, we observe that major losses in regional taxonomic diversity between time bins, particularly the Miocene climatic transition (~14 Ma), are not associated with a similar loss in functional groups.
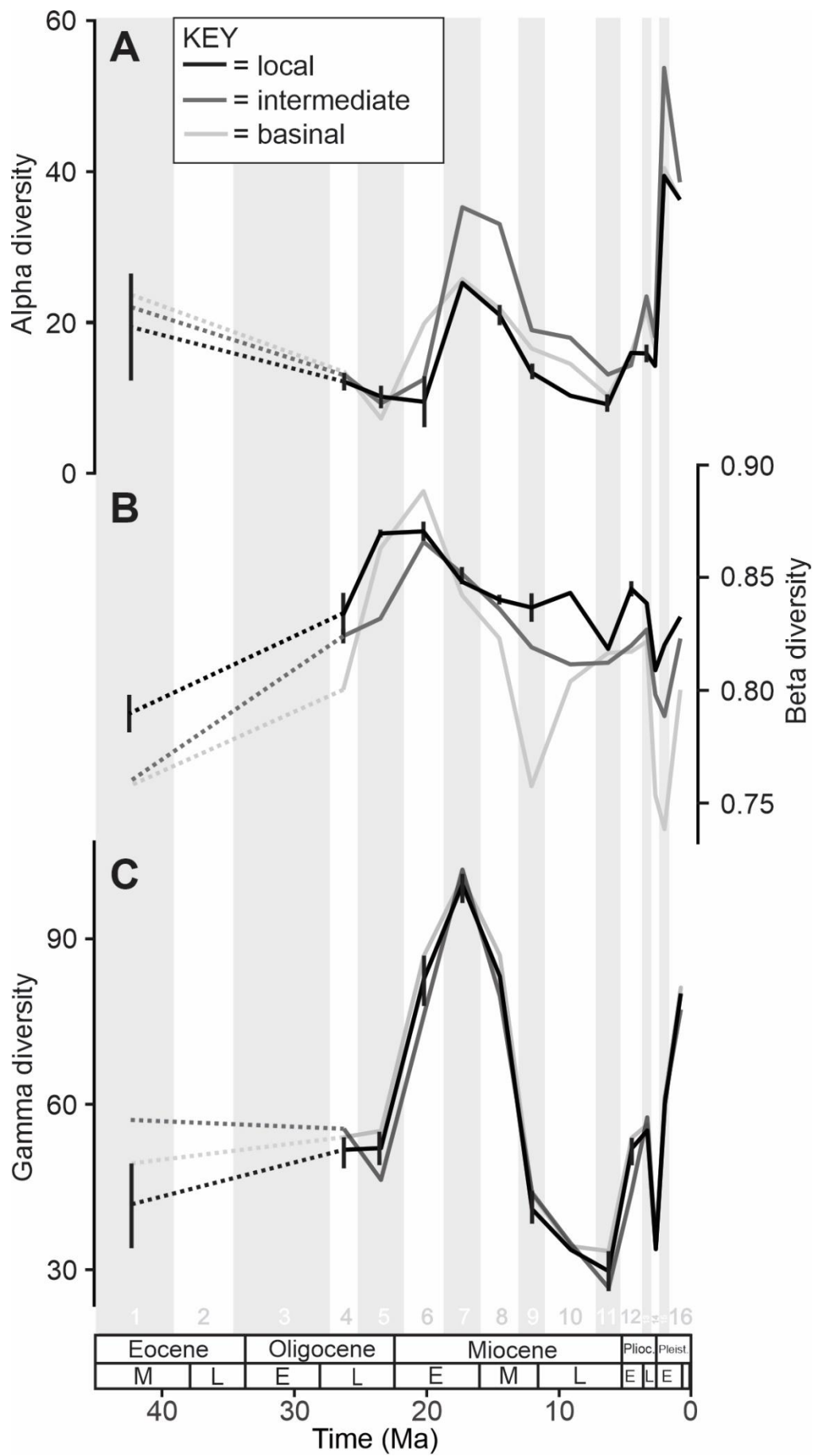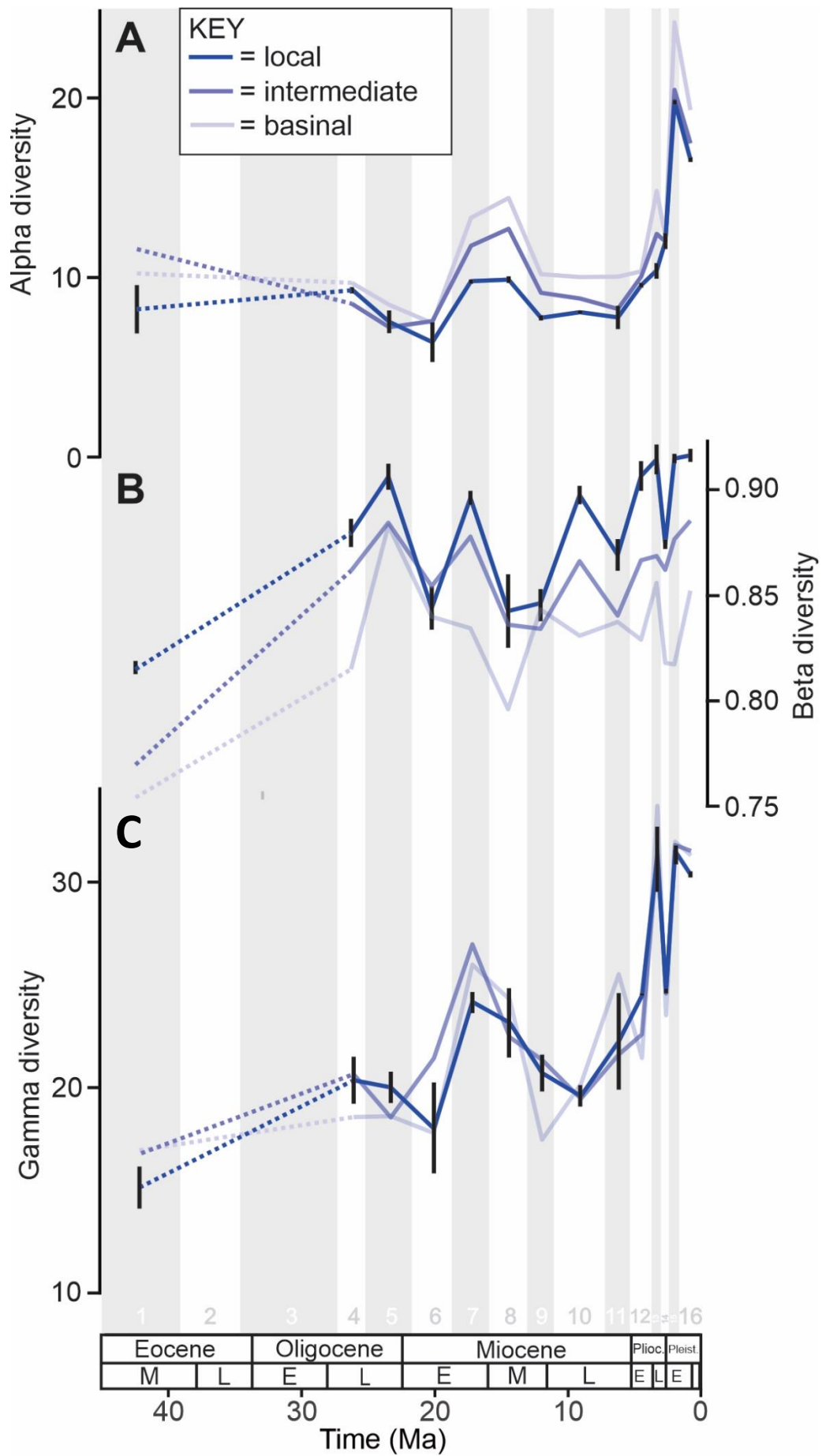
FIGURE 8.6. Time series of functional diversity through the Cenozoic, partitioned into (A) alpha, (B) beta and (C) gamma diversity. For clarity error bars are only shown for partitions of taxonomic diversity at the local scale and represent ± 1 standard error. Analyses were undertaken at the resolution of the time bins indicated by the shaded bars, labelled by time bin number (see Fig. 8.1).

The decoupling of taxonomic and functional diversity described above has been recorded in the global marine molluscan fossil record during mass extinction events (Foster and Twitchett 2014, Aberhan and Kiessling 2015, Edie et al. 2018b), and associated with moderate extinction events in the Caribbean Neogene molluscan fossil record (Pimiento et al. 2020). Our findings are consistent with these studies, but they also imply that in the absence of mass extinction events, New Zealand's temperate benthic marine ecosystems appear to be functionally resilient in the face of moderate taxonomic extinction (i.e., not mass extinction events). On this basis, the negative relationship between regional functional diversity and oceanic temperature through the Cenozoic likely reflects, in part, the accumulation of functional groups with minimal loss during intervals of decreasing taxonomic diversity. The maintenance of functional groups during warmer intervals may have been facilitated by increased functional redundancy within existing functional groups (see Chapter 7) (Fig. 8.7).

Time 1   Time 2   Time 3

A

Functional group 1

Functional group 2

Functional group 3

warming   cooling

B

= Δ taxonomic diversity
= Δ functional diversity

C

= functional redundancy
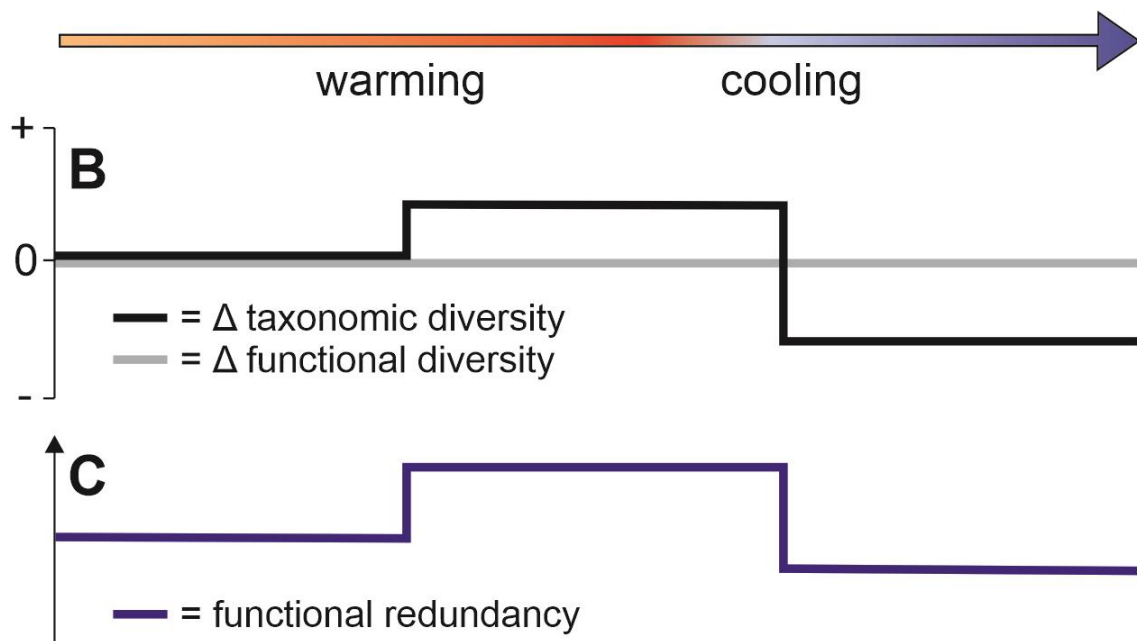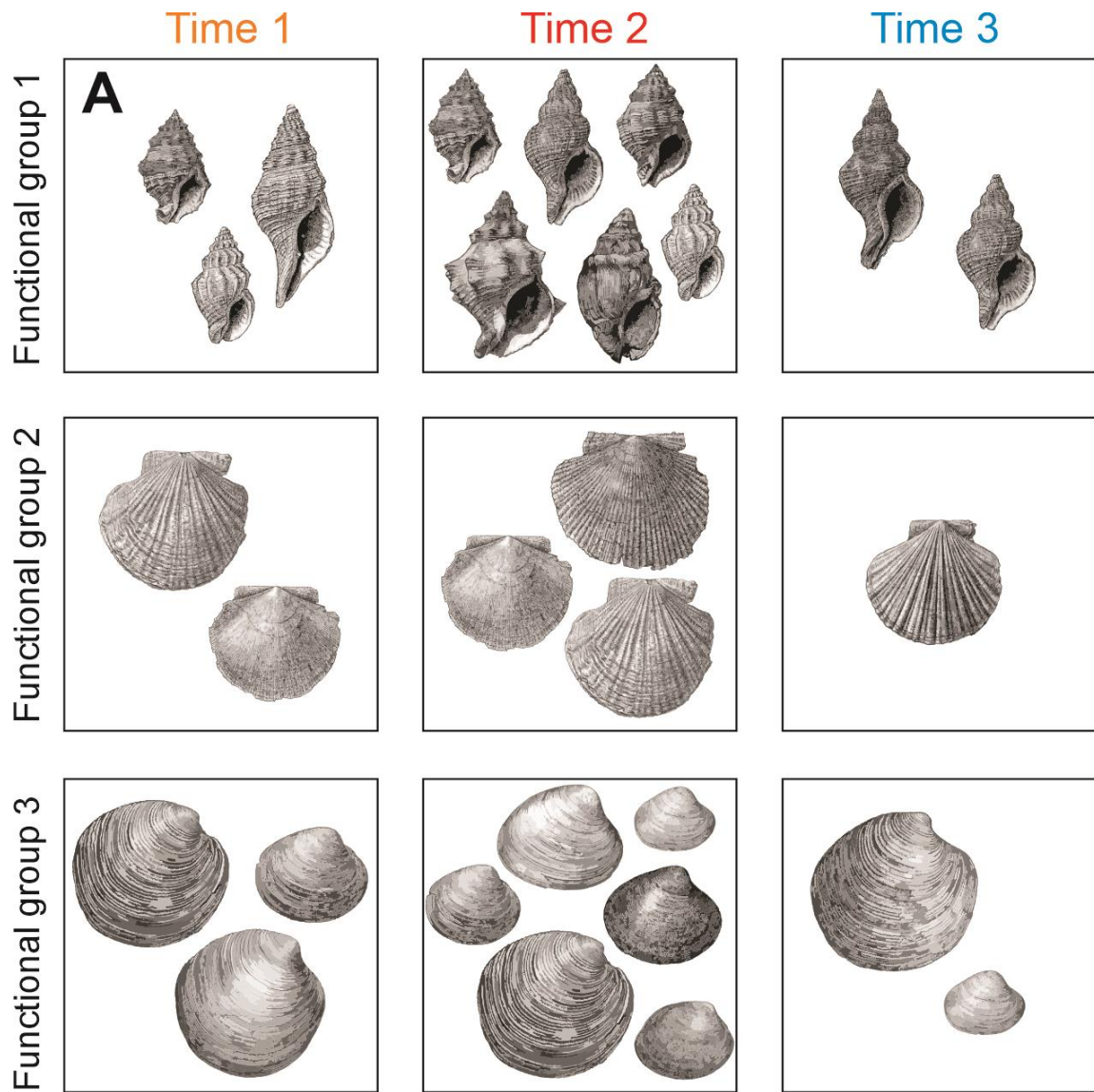
FIGURE 8.7. Conceptual figure depicting the relationships between taxonomic and functional diversity, and functional redundancy with oceanic warming and cooling for New Zealand's temperate shallow-marine ecosystems. (A) Simplified schematic showing species-level diversity within three functional groups during three time intervals that span a phase of warming (Time 1 to 2) and cooling (Time 2 to 3). Oceanic warming is predicted to increase relative (indicated by Δ) species-level diversity (Time 2) due to the arrival of warm-water genera, and subsequent in-situ speciation within existing functional groups, without changing relative functional diversity (B). Functional redundancy is likewise predicted to increase during warming (C). Drawings from Beu and Maxwell (1990, plates 11, 13, 17, 19, 21 and 28, used with permission of GNS Science, Lower Hutt).

8.3.2. Macroevolutionary rates

We find a consistently strong signal for oceanic temperature as the primary positive correlate of origination rates, and to a lesser extent climatic variability as a negative correlate (Table 8.3 and see Fig. 8.8 for time series). This implies that origination rates, on average, reflect a predictable time-continuous process over long time spans. Conversely, extinction rates reveal contrasting results, with no clear relationship between extinction and the environmental proxies examined here. Notwithstanding this, previous studies on New Zealand's Cenozoic shallow-marine molluscan fossil record observed a correlation between short-term (i.e., between interval) variation in extinction rate and sampling probability, and indirectly, with the spatial structuring of the environment (similar to the measure of environmental heterogeneity derived here) (Crampton et al. 2011). The indirect relationship between the spatial structuring of the environment and extinction, along with evidence from previous studies (Crampton et al. 2006b), was attributed to a pulsed model of extinction,

hypothesizing that extinction was more pulsed in nature than origination in New Zealand's shallow-marine ecosystems during the Cenozoic (Crampton et al. 2011). This hypothesis is supported by several studies at the global scale which similarly suggest origination rates in the marine realm are less pulsed than extinction rates (Foote 2005, Heim and Peters 2011). Despite differences in the methods and metrics employed within this study, we also find evidence that environmental heterogeneity correlates to extinction rate. However, we find the correlation between environmental heterogeneity to be both positive and negative, at the local and basinal scale, respectively (Table 8.3 and see Fig. 8.2 for time series). The positive correlation between environmental heterogeneity and extinction at the local scale also includes habitable shelf area as an additive positive correlate. This spatial contrast could be related to the different underlying drivers of environmental heterogeneity at the different spatial scales. Variation measured at the local scale likely reflects widespread heterogeneity in regional environmental conditions as higher values of lithological dissimilarity between grid cells requires widespread variation in lithology. This could feasibly be driven by a number of factors, such as tectonic cycles of uplift and erosion or changes in relative sea level. Similarly, increases in shelf area in New Zealand are considered to reflect tectonic phases, rather than eustacy (see Chapter 2, section 2.2.2). On this basis, environmental heterogeneity measured at the local scale and habitable shelf area may be driven by similar factors related to regional tectonism. Conversely, basinal scale environmental heterogeneity captures variation in the environmental conditions between depositional sedimentary basins. Whereas this could also be driven by a number of factors, higher values of lithological dissimilarity suggest that basins are more environmentally differentiated. Thus, it is plausible that large-scale environmental change (as measured by environmental heterogeneity at the local scale) could contribute to increased extinction rates in conjunction with increases in shelf area, and between basin heterogeneity could reduce or buffer extinction (e.g., by

providing refugia and limiting biotic interactions). However, at the resolution of our data, it is not possible to discern the exact nature of this correlation and further study is required. Whereas our findings do not provide any direct support to a model of pulsed extinction, they do corroborate some of the findings of Crampton et al. (2011). Regardless of the mode of extinction the decoupled loss of taxonomic diversity relative to functional diversity suggests a non-random element in extinction selectivity, i.e., species extinction is evenly spread across functional groups, rather than concentrated in selected groups.
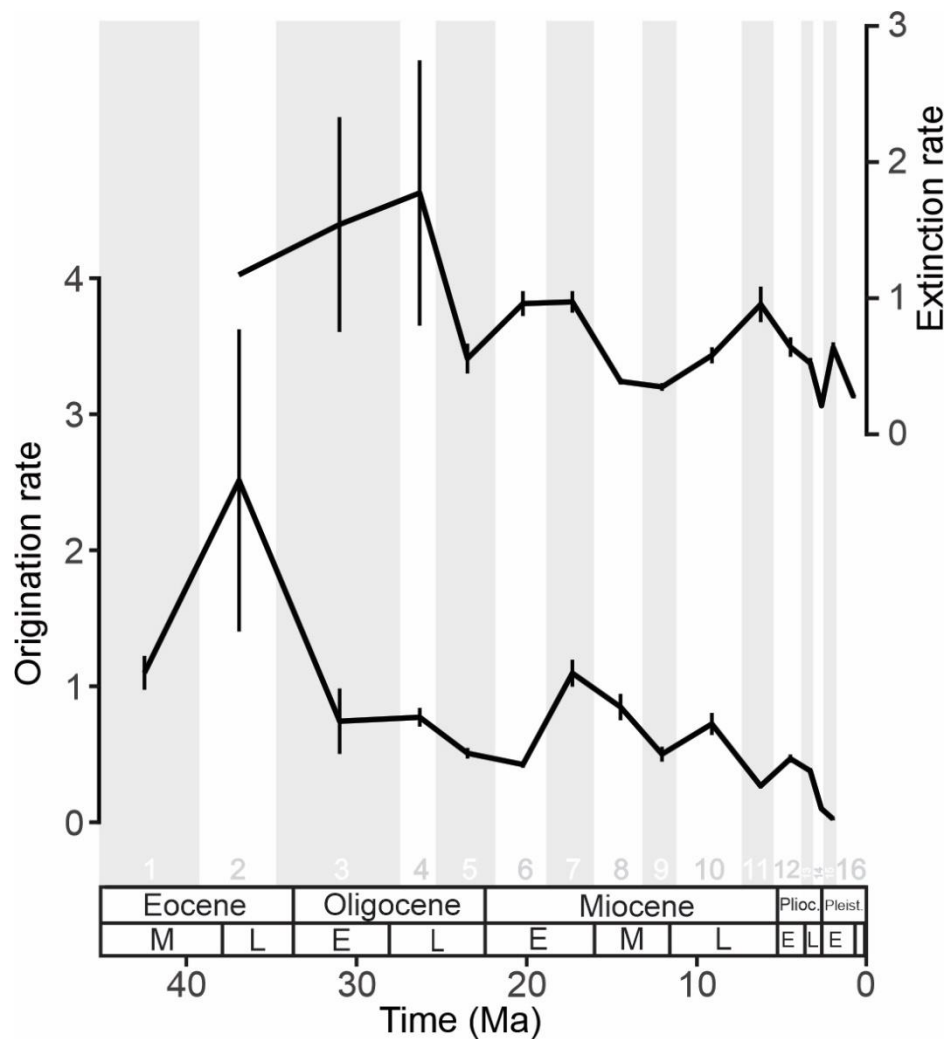
FIGURE 8.8. Time series of origination and extinction rates through the Cenozoic. Macroevolutionary rates were calculated using probabilities derived from the capture-mark recapture Pradel seniority model converted to instantaneous rates. Error bars represent ± 1 standard error. Analyses were undertaken at the resolution of the time bins indicated by the shaded bars, labelled by time bin number (see Fig. 8.1).

	Grid size (hexagrid side km)	Explanatory variable(s)	Coefficient(s)	Pseudo-R2	phi	AICc weight (%)	_p_	Log-likelihood
	23	Temp. + _Temp. var._	0.50, -0.15	0.7	-0.54	45	<0.001	-3.16
Origination	42	Temp. + _Temp. var._	0.50, -0.15	0.7	-0.54	41	<0.001	-3.16
	80	Temp. + _Temp. var._	0.50, -0.15	0.7	-0.54	44	<0.001	-3.16
	23	Area + Env. Het.	0.27, 0.36	0.51	-0.19	41	0.005	-1.97
Extinction	42	-	-	-	-	38	-	-
	80	_Env. Het._	-0.35	0.34	-0.2	46	0.012	-4.15

TABLE 8.3. Summary of generalized least squares (GLS) autoregressive (AR1) models. Results are presented for the best-fit models for rates of

origination and extinction derived using probabilities from the capture-mark recapture Pradel seniority model converted to instantaneous rates.

Model comparison was undertaken using the weighted corrected-Akaike information criterion (AICc). Explanatory variables in italics represent

negative relationships. Pseduo-R^2 values (ML-R^2) for best fit models where calculated following the maximum likelihood approach of Cox and

Snell (1989)

The exact mechanism by which oceanic temperature may increase rates of origination, and how this relationship manifests biogeographically, is unclear. Nevertheless, New Zealand's geographic isolation during the Cenozoic precludes the wholesale role of species immigration and emigration (see Chapter 2, sections 2.2 and 2.4). As a result, origination at the species-level is largely attributed to in-situ evolution (Crampton et al. 2006a). Previous research suggests that warm intervals, particularly the Miocene climatic optimum, are associated with the arrival of warm-water molluscan genera from Australia and the Pacific (Beu and Maxwell 1990). Both genera that arrived during the Miocene climatic optimum and exiting genera show evidence of in-situ adaptive radiation at the species-level (i.e., cladogenesis, *sensu* Allmon 2017). This is based on the clear increase of within genus species richness in genera during the Miocene climatic optimum (Beu and Maxwell 1990). Conversely, during cooler intervals, particularly the latest Miocene and Pliocene-early Pleistocene, there is taxonomic evidence that anagenesis (*sensu* Allmon 2017) was a significant mode of speciation (Beu 2012). This is based on the observations of taxonomic experts, rather than evidence from biodiversity patterns. The mid- to late Pleistocene (< 1.63 Ma) marks a shift in evolutionary mode from the Pliocene and early to mid- Pleistocene, with speciation entirely attributed to cladogenesis (in particular, allopatric speciation) (Beu 2012). This is perhaps a result of increasing environmental variability related to increased amplitude of Milankovitch-driven global climate cycles (Beu 2012). On this basis, we speculate that part of the long-term relationship between oceanic temperature and origination rates may relate to different modes and paces of speciation, whereby warmer intervals are associated with higher cladogenetic origination rates and cooler intervals are associated with lower origination rates related to anagenesis (Fig. 8.9). Within this model, assuming pulsed extinction, we would expect regional biodiversity to be positively correlated with oceanic temperature, which is consistent with our observations. We suggest that the relationship

between evolutionary mode and oceanic temperature may weaken or diminish once a threshold of environmental variability is crossed (i.e., during the mid- to late Pleistocene), as suggested by Beu (2012). However, origination rates remain low during the mid- to late Pleistocene, suggesting oceanic temperature has an impact on the pace of speciation regardless of the prevalent evolutionary mode. In addition, it is acknowledged that it is possible the evidence of anagenesis is not preserved in warmer time intervals, and that the decrease in species richness observed in cooler intervals could also be related to a decrease in the cladogenesis. Further work is required to fully understand these relationships.
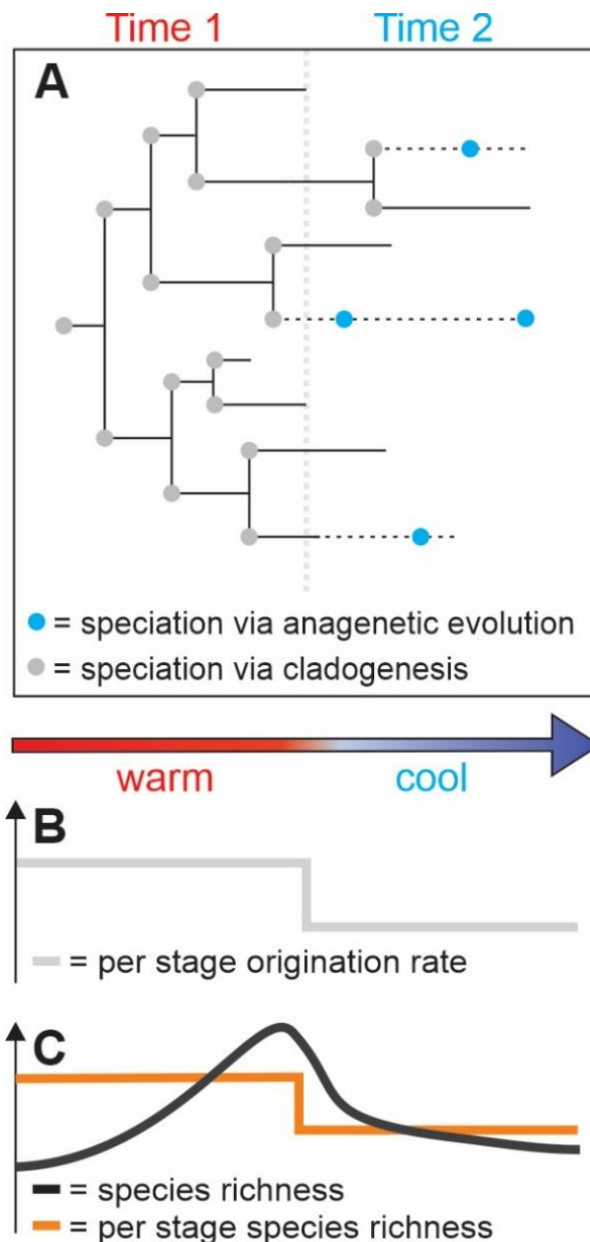
FIGURE 8.9. Conceptual figure depicting an hypothesized model for the relationship between speciation and oceanic temperature in New Zealand's shallow-marine ecosystems. (A) Illustrative simplified evolutionary tree showing a transition from warmer oceanic temperatures (Time 1), where cladogenesis is the dominant mode of speciation, to cooler (Time 2) oceanic temperatures, where anagenesis is the dominant mode of speciation. Within this model, per-stage origination and species richness is expected to decrease with cooling oceanic temperatures (B, C).

8.3.3. Temperate-tropical comparisons

Comparisons between the temperate molluscan shallow-marine record of New Zealand with similar quality records of Neogene tropical shallow-marine molluscs of the Caribbean reveals several interesting similarities and differences. Firstly, in New Zealand, peak regional taxonomic diversity occurred in the warm interval during the Miocene climatic optimum (see Chapter 2, sections 2.3 and 2.4), and subsequently declined during cooling of the Miocene climatic transition. In contrast, regional taxonomic diversity of the Caribbean shows minimal change in species richness during the Miocene climatic optimum and an increase following the Miocene climatic transition, with diversity peaking in the late Pliocene (Pimiento et al. 2020). Second, we find peak functional diversity in New Zealand during Plio-Pleistocene global cooling, coinciding with peak functional beta diversity. This broadly corresponds to the patterns of regional functional diversity observed in Caribbean molluscs, which also peaked during the Plio-Pleistocene (Pimiento et al. 2020). Lastly, previous research on New Zealand's shallow-marine molluscan fossil record suggests that functional redundancy generally decreased during the Neogene (see Chapter 7). Conversely, functional redundancy in the Caribbean mollsucan fossil record increased throughout the Neogene and peaked in the late Pliocene (Pimiento et al. 2020). Although the impact of oceanic temperature has not been quantified in the Caribbean molluscan fauna during the Neogene, the comparison between temperate and tropical regional biodiversity suggests asymmetrical patterns of taxonomic diversity and functional redundancy, but congruent patterns of functional diversity. This contrasts to observations along the modern latitudinal diversity gradient that indicate a poleward decline in both taxonomic and functional diversity (Edie et al. 2018b, Schumm et al. 2019).

8.3.4. Implications for Modern Regional Temperate Marine Ecosystems

Our results have two important implications for the future of New Zealand's shallow marine biodiversity and similar temperate regional ecosystems. Firstly, functional diversity in temperate marine ecosystems appears to be resilient in the face of modest taxonomic extinction through the Cenozoic, potentially as a result of increased functional redundancy during warm intervals. Modern conservation generally focuses on biodiversity loss measured as species richness, with less emphasis on ecosystem function. However, if the current global and regional loss of functional diversity and redundancy (Petchey and Gaston 2002, Rosenfeld 2002, Barnosky et al. 2011, Ceballos et al. 2015, Valiente-Banuet et al. 2015, McWilliam et al. 2018) goes unabated, the unfolding sixth mass extinction may be qualitatively different (i.e., higher levels of functional extinction) from previous mass extinction events, and have correspondingly unforeseeable consequences (Payne et al. 2016). Second, under natural warming scenarios, we may expect increased speciation rates in New Zealand's shallow-marine ecosystems through the arrival and radiation of immigrant genera. The expanding list of threatened New Zealand marine invertebrates (Freeman et al. 2010, Gordon et al. 2010), which includes a large number of endemics with narrow biogeographic ranges, may be vulnerable to extinction from biotic interactions from warm-water arrivals, e.g., competition and predation. This is particularly true for the highly endemic subantarctic fauna of southern New Zealand, with no major landmasses between New Zealand and Antarctica to act as refugia. For New Zealand, our study highlights the need to incorporate long-term predictions into future conservation strategy and the need for more temporally resolved regional studies in temperate ecosystems focusing on conservation paleobiology.

8.4. Summary

Quantifying the impact of anthropogenic climate change on biodiversity requires an understanding of the mechanisms that drive biodiversity change, origination, and extinction, and the ways that these mechanisms are expressed both spatially and temporally. Here we show contrasting responses of biodiversity and macroevolutionary rates to climatic change, both spatially and over geological timescales within a regional, temperate, shallow-marine fauna, and between temperate and tropical ecosystems. Overall, our findings suggest the driving mechanisms of New Zealand's shallow-marine biodiversity change through the Cenozoic are complex and related to several factors. Notwithstanding this, oceanic temperature is clearly the single most important factor of those examined here. We find evidence that oceanic warming positively impacts origination rates and contributes to regional taxonomic richness, but negatively impacts regional functional and beta diversities. This negative long-term relationship between regional functional diversity and oceanic temperature is likely driven by functional resilience to extinction during cool intervals rather than a negative response to warming per se. Although we find some support for an asymmetrical response in regional taxonomic diversity to oceanic warming between tropical and temperate shallow-marine ecosystems through the Neogene, we find symmetrical patterns of regional functional diversity. These findings are not consistent with biogeographic patterns observed along the modern latitudinal diversity gradient, which display congruent relationships between temperature and taxonomic and functional diversity (Edie et al. 2018b, Schumm et al. 2019), and require further exploration and explanation.

Chapter 9

Conclusion

"*It's the job that's never started as takes longest to finish.*"

(*Samwise Gamgee*, The Lord of the Rings, J.R.R Tolkien, 1954)

9.1. Introduction

Within this final Chapter, I revisit the three research questions outlined in the introduction (Chapter 1). In addition, I consider the wider context of the results, in terms of: (1) implications for the future of New Zealand's ecosystems, and (2) scope for future work. Lastly, I provide a brief summary of the thesis in its entirety.

9.2. Research Questions

9.2.1. Research Question 1 - *What is the structure of apparent diversity patterns in New Zealand's shallow-marine fossil record?*

This research question is further subdivided into two related questions: (1) is the structure of the apparent diversity pattern in New Zealand's shallow-marine fossil record consistent with the nature of the global diversity history that records a putative rise in taxonomic diversity towards the recent, and (2) are the resultant patterns impacted by biases that are known to affect the global diversity trajectory, in particular the Pull of the Recent.

In Chapter 5 I show that taxonomic diversity in New Zealand, measured using biostratigraphic ranges (range-through) at the species and generic level, increases through the Cenozoic, but peaks during the Miocene climatic optimum (Otaian and Altonian stages: 21.7 – 15.9 Ma). This contrasts to the global trajectory of marine diversity (similarly based on range-through diversity), which peaks in the recent (Jablonski et al. 2003). Despite this contrast, estimates of the impact of the Pull of the Recent in New Zealand are in line with the global estimate of the Pull of the Recent for marine molluscs reported for a similar time interval (Jablonski et al. 2003). The results indicate that the Pull of the Recent effects < 2% of genera and < 4 % of species when considering taxa missing from the youngest ~5 Myrs of the New Zealand fossil record. The taxonomic composition of molluscs missing from the youngest part of the fossil record cannot easily be explained by effects related to shell mineralogical composition, body size, habitat, or taxonomic class. This implies that the apparent rise in New Zealand's marine diversity and peak in species richness during Miocene climatic optimum either represents a true biological signal, or results from other confounding effects not considered here, or reflects a combination of factors.

9.2.2. Research Question 2 - *Are patterns of biodiversity at the regional scale confounded by the uneven distribution of fossil collections?*

Measures of biodiversity in New Zealand are confounded by the spatial distribution of fossil collections through time. In Chapter 6 I show that beta diversity is spatially dependent at local to regional spatial scales and uneven spatial sampling can influence recovered temporal trends in beta diversity. In Chapter 7 I also show that this issue extends to measures of regional taxonomic and functional diversity (defined as gamma diversity in this thesis). The findings of Chapter 7 imply that even sampling standardized measures of regional taxonomic and functional diversity taken at face value (i.e., at the maximum preserved spatial distribution of fossil collections) can alter the recovered temporal trends. The magnitude of spatial bias shown here is much smaller than recent estimates for the global fossil record (Close et al. 2020). Nevertheless, the results of Chapters 6 and 7 highlights that temporal variation in the spatial distribution of fossil collections at the regional spatial scale still needs to be considered in studies of regional biodiversity dynamics.

Within research question 2 I also query whether spatial patterns of taxonomic and functional diversity in New Zealand's shallow-marine molluscan fossil record, i.e., the species-area relationship and functional diversity-area relationship, adhere to patterns observed in modern ecology. In Chapter 7 I show that both the species-area relationship and functional diversity-area relationship conform to the power function model typically observed in modern ecological studies (Whittaker et al. 2014, Karadimou et al. 2016). This suggests that the functional diversity-area relationship, which has not been previously quantified in the fossil record, is equally persistent and predictable as the species-area relationship.

In Chapter 7 I extend the findings of research question 2 to consider how these spatial patterns of biodiversity can be used to investigate fundamental biogeographic and

macroecological patterns. The results of Chapter 7 indicate that the relationship between the number of functional groups (functional richness/diversity) and the number of species at increasing spatial scales, referred to as the functional redundancy relationship within this thesis, conforms to a similar asymptotic relationship observed within modern latitudinal climatic zones (temperate and tropical) (Schumm et al. 2019). The asymptotic relationship in modern ecosystems is steeper in temperate climatic zones relative to tropics, implying: (1) functional redundancy is elevated in the tropics, and (2) this relationship may be driven by latitudinal gradients of climate and oceanic temperature. In Chapter 7 I show that the slope of the functional redundancy relationship is strongly correlated to Pacific Ocean temperature over the last ~45 Myrs. This result is consistent with observations from the modern latitudinal diversity gradient (Schumm et al. 2019). This suggests that functional redundancy in temperate ecosystems is temperature dependent and the functional redundancy relationship is a persistent biogeographic and macroecological feature of regional and global importance.

9.2.3.   Research Question 3 - *What are the key drivers of long-term biodiversity change and macroevolution in New Zealand's shallow-marine ecosystems through the Cenozoic?*

In Chapter 7 I find evidence for a long-term correlation between regional (gamma) species-level taxonomic diversity, and oceanic temperature and climatic variability. In Chapter 8 I extend these analyses to other spatial partitions of biodiversity and additional explanatory factors (environmental heterogeneity and habitable area). The results imply that the long-term driving mechanisms of New Zealand's shallow-marine biodiversity change through the Cenozoic are complex and related to several factors. Notwithstanding this, oceanic temperature is clearly the single most important factor of those examined here. I find evidence that oceanic warming positively impacts origination rates and contributes positively to regional taxonomic diversity at the species-level, but negatively impacts regional functional and beta diversities. This negative long-term relationship between regional

functional diversity and oceanic temperature is likely driven by resilience to functional extinction during cool intervals rather than a negative response to warming per se. I suggest that the maintenance of functional groups during warmer intervals may have been facilitated by increased functional redundancy within existing functional groups, as inferred in Chapter 7. In addition, I speculate that part of the long-term relationship between oceanic temperature and origination rates may relate to different modes and paces of speciation, whereby warmer intervals are associated with higher cladogenetic origination rates and cooler intervals are associated with lower origination rates related to anagenesis (Fig. 8.8).

Within research question 3, I also query whether the long-term drivers of biodiversity change are consistent with predictions and observations from modern ecology. Observations and predictions from the distribution of current marine life suggest that tropical and temperate ecosystems will be impacted asymmetrically from oceanic warming. This is based on the prediction of a positive gain in species richness within temperate ecosystems (from the poleward migration of warm water species) and negative impact on tropical ecosystems (due to species loss from elevated heat stress) (Antão et al. 2020, Worm and Lotze 2021). Within Chapter 8 I find some support for an asymmetrical response in regional taxonomic diversity to oceanic warming between tropical and temperate shallow-marine ecosystems through the Neogene. However, the results indicate symmetrical patterns of regional functional diversity. These findings are not consistent with biogeographic patterns observed along the modern latitudinal diversity gradient, which display congruent relationships between temperature and both taxonomic and functional diversity (Edie et al. 2018, Schumm et al. 2019), and require further exploration and explanation.

9.3. Implications for the Future

Within this section I focus on how this thesis aligns with the growing field of conservation paleobiology. In addition, I outline how the results of this thesis: (1) contributes to our understanding of New Zealand's biodiversity in relation to anthropogenic climate change and the modern biodiversity crisis, and (2) consider the broader implications for our understanding of modern macroecology.

9.3.1. Conservation Paleobiology

Over the last few decades, the field of conservation paleobiology has become a major sub-discipline within paleontology and ecology. The integration of ecological and paleontological data at different spatial and temporal scales is now recognised as a principal element in understanding the current biodiversity crisis and future impacts of anthropogenic climate change (Price and Schmitz 2016). However, conservation paleobiology still faces several challenges. Perhaps the most important of these is creating a common framework and currency of biodiversity that transcends the traditional units of space and time typically employed in isolated ecological and paleontological studies. Within this thesis I place emphasis on: (1) using New Zealand's shallow-marine fossil record to test whether observations and predictions from the distribution of modern life are consistent over longer (geological) time frames. Quantifying the spatiotemporal nature of these predictions and observations improves our understanding of macroecological and macroevolutionary processes, which ultimately underpins our ability to model how diversity may respond to future climatic and environmental change. (2) I also focus on quantifying biodiversity in equivalent currencies and spatial scales to modern regional/bioregional ecological studies (e.g., Antão et al. 2020).

9.3.2. Implications for New Zealand

The results presented within this thesis provide an insight into the long-term drivers of New Zealand's shallow-marine biogeography and biodiversity change, providing a baseline for what should be expected from natural warming in New Zealand. Based on the data presented here, we may expect the following changes in New Zealand's shallow-marine ecosystems (and similar regional, temperate, shallow-marine ecosystems elsewhere on the globe) with increasing oceanic temperatures over long-timescales (multi-centennial to millennial):

- *A net gain in species richness* of shallow-marine ecosystems in New Zealand as a whole. This is likely to be facilitated by the poleward migration of species from subtropical regions to the north of New Zealand (i.e., Australia and Pacific Islands). In the past the migration of molluscs and other benthic marine biota to New Zealand was largely restricted to species with free swimming larval stages (Beu et al. 1997). However, anthropogenic activity is likely to increase the rate of the translocation of "alien" species (i.e., non-native), particularly within ballast water and hull fouling of transport ships. There are already at least 177 naturalized alien species in New Zealand's coastal waters, mostly in ports and harbours (Gordon et al. 2010).

- *An increase in functional redundancy*, meaning that there will be more species performing particular ecological roles within shelfal marine ecosystems. If correct, this might be expected to improve ecosystem resilience to anthropogenic or natural perturbation.

- *Limited change in the regional functional diversity*, the number of ecological roles performed within shelfal marine ecosystems in New Zealand, but an increase in the similarity of functional roles performed between local regions

278

(i.e., *reduced functional beta diversity*). As with functional redundancy, this may be expected to improve ecosystem resilience to anthropogenic or natural perturbation and suggests that oceanic warming, in isolation, is unlikely to cause regional scale functional extinction in New Zealand.

- *Variable response of species richness within local regions*. Although species richness is predicted to increase in New Zealand as a whole, the net loss and gain of species at smaller spatial scales, measured by alpha and beta diversity (as defined in this thesis), is expected to vary across the latitudinal range of New Zealand. Other variables such as habitable shelf area, which may increase in tandem with global sea-level rise, and environmental heterogeneity may interact to modify temperature-driven changes in local species richness. Thus, changes in species richness related to oceanic warming may not manifest in local regions.

- *Increased speciation rates* through the arrival and radiation of immigrant genera. As noted earlier, this is likely to facilitate an increase in species richness in New Zealand as a whole.

Oceanic warming associated with anthropogenic climate change is taking place at a rapid rate (Kemp et al. 2015) and is occurring in tandem with elevated rates of biodiversity loss and habitat destruction (Erwin 2009). If the current rate of climate change and unprecedented loss of biodiversity and disruption of marine habitats is maintained, it is likely that patterns of biodiversity in the future will deviate from the predictions outlined above. However, as noted, these long-term predictions provide a baseline. This baseline implies oceanic warming should lead to a net increase in New Zealand's regional biodiversity, and functional diversity should be largely unaffected. On this basis, even small losses in biodiversity, particularly anthropogenically forced extinction of functional groups and

279

endemic species, may significantly alter the future biodiversity and function of New Zealand's shallow-marine ecosystems. This is particularly important for New Zealand as it is geographically isolated and has a limited species pool with high levels of endemism (Beu and Maxwell 1990, Crampton et al. 2006, Spencer et al. 2009, Gordon et al. 2010). Furthermore, the expanding list of threatened New Zealand marine invertebrates (Freeman et al. 2010, Gordon et al. 2010), which includes a large number of endemics with narrow biogeographic ranges, may be particularly vulnerable to extinction from biotic interactions with warm-water arrivals, e.g., competition and predation. The most vulnerable ecosystems are likely to be the highly endemic subantarctic fauna of southern New Zealand, with no major landmasses between New Zealand and Antarctica to act as refugia.

9.3.3. Implications for modern ecology

Beta Diversity.— In Chapter 5 I demonstrate that measurements of beta diversity in the fossil record are spatially dependent and apparent temporal patterns are confounded by the uneven spatial distribution of fossil collections. Although the spatial pattern of beta diversity in modern ecosystems may differ from those identified in the fossil record (e.g. Barton et al. 2015), it is important to realise that measurement of modern beta diversity is still spatially dependent. This is particularly important as beta diversity is a key measure of biodiversity in ecological studies and conservation planning (Bush et al. 2016, Socolar et al. 2016); it is possible that the results of such studies may be influenced by spatial bias and misleading. This is particularly true for studies utilizing large ecological databases, which are becoming increasingly prevalent. Similar to studies utilizing fossil databases, these studies have no control over the spatial sampling regime.

Functional Redundancy Relationship.— In Chapter 7 I define a new biogeographic phenomenon, the functional redundancy relationship, which describes the relationship

between the number of species and number of functional groups at increasing spatial scales. The asymptotic form of the functional redundancy relationship is consistent with patterns observed within the modern latitudinal diversity gradient (Schumm et al. 2019), which implies that it is a persistent macroecological pattern of regional and global ecological importance. Evidence from modern ecology suggests functional redundancy increases ecosystem resilience to environmental change (Biggs et al. 2020). The results presented herein corroborate this conclusion, suggesting that elevated functional redundancy during warm periods may buffer functional loss during moderate taxonomic extinction over long timescales. However, measures of functional redundancy are rarely quantified in conservation research (Biggs et al. 2020). The functional redundancy relationship highlights the importance of functional redundancy, but the slope of this relationship also provides a spatially explicit method of quantifying the loss (or gain) of multiple dimensions of biodiversity in modern ecosystems, whilst simultaneously capturing aspects of functional redundancy. Notwithstanding this, further work is required to explore the application of the functional redundancy relationship in modern ecological settings.

Global Biogeography and Climatic Change. — As noted earlier (section 9.2.3), recent research suggests that tropical and temperate marine ecosystems are likely to be impacted asymmetrically by increasing oceanic temperatures over short timescales. This asymmetrical relationship is mirrored in the distribution of marine life across the modern latitudinal diversity gradient, which indicates a parallel poleward decrease in taxonomic and functional richness. Both patterns are attributed, in part, to current or predicted short-term changes in latitudinal gradients of climate and temperature (Edie et al. 2018b, Blowes et al. 2019, Schumm et al. 2019, Antão et al. 2020, Gagné et al. 2020). However, the fossil record suggests that taxonomic and functional diversity become decoupled over longer timescales, particularly during intervals of elevated taxonomic loss (Edie et al. 2018a, Edie et al. 2018b).

The results provide some support for an asymmetrical response in regional taxonomic diversity to oceanic warming between tropical (Pimiento et al. 2020) and temperate shallow-marine ecosystems through the Neogene. However, the results imply symmetrical patterns of regional functional diversity. These findings are not consistent with biogeographic patterns observed along the modern latitudinal diversity gradient (Edie et al. 2018b, Schumm et al. 2019), but the evidence for asymmetry in temperate-tropical taxonomic diversity is in line with short-term predictions and observations from modern ecology (Blowes et al. 2019, Antão et al. 2020). Taken at face value, these results suggest that over geological timescales and in the face of global warming, we expect a net increase in New Zealand marine species richness and an increase in functional redundancy.

Lastly, and importantly, the results of this thesis imply that functional diversity may respond differently to anthropogenic climate change than past episodes of natural climate change. In particular, I suggest New Zealand's shallow-marine ecosystems are functionally resilient in the face of moderate taxonomic extinction (i.e., not mass extinction events), consistent with interpretations from tropical ecosystems during the Neogene (Pimiento et al. 2020) and evidence from the global marine fossil record (Edie et al. 2018b). However, if the current global and regional loss of functional diversity and redundancy (Petchey and Gaston 2002, Rosenfeld 2002, Barnosky et al. 2011, Ceballos et al. 2015, Valiente-Banuet et al. 2015, McWilliam et al. 2018) goes unabated, the unfolding sixth mass extinction may be qualitatively different (i.e., higher levels of functional extinction) from previous mass extinction events, and have correspondingly unforeseeable consequences (Payne et al. 2016).

9.4. Summary and future work

This thesis confirms the high quality of the New Zealand shallow-marine fossil record and identifies oceanic temperature as a major positive correlate of regional taxonomic diversity, origination, and functional redundancy. This correlation is not consistent across spatial scales, a reflection of other abiotic and biotic processes that are likely important. These findings are relevant to our understanding of macroevolutionary and macroecological processes, particularly how temperate shallow-marine ecosystems may respond to future warming. For New Zealand, this research suggests, as a baseline, that natural warming may increase equilibrium taxonomic diversity and functional redundancy in New Zealand over geological time scales. However, anthropogenic climate change is associated with rates of biodiversity loss and habitat destruction that have few analogues in the geological record. This is likely to alter the future trajectory of New Zealand's biodiversity change in ways that cannot be anticipated from past natural warmings, which may have far-reaching impacts for the short- and long-term future of New Zealand's unique and highly endemic marine fauna.

This highlights the need for further study in shallow-marine temperate ecosystems, particularly at finer time scales than examined here, to better understand diversity dynamics, and importantly, extinction risk. There is minimal work on either the short- or long-term impact of anthropogenic climate and environmental change on New Zealand's biodiversity that utilizes information from the fossil record. This thesis only covers the long-term relationships of New Zealand's shallow-marine molluscan biogeography and biodiversity to global change, and there is much to gain from examining other taxonomic groups and extending analyses to finer temporal and spatial scales.

Several areas of potential future work have been identified during the course of this thesis:

1. New Zealand specific research into the extinction risk of shallow-marine ecosystems over both short and long timeframes based on oceanic warming scenarios outlined by the IPCC (Bindoff et al. 2019).

2. Research establishing a local historical baseline for diversity change in response to climate in New Zealand's shallow-marine ecosystems. The Whangnaui Basin (see Fig 2.3) holds a unique record of shallow-marine ecosystems (predominately during interglacials) through the Plio-Pleistocene to recent at a high temporal resolution (40 – 100 Kyrs) and is estimated to be up to 88 % complete (Johnson and Curry 2001); this record would represent at excellent case study. The integration of fossil data from the basin with modern ecological data could directly inform modern conservation, management and policy (e.g., Dietl et al. 2015, Barnosky et al. 2017) .

3. Research into the long-term drivers of biodiversity patterns in other taxonomic groups in New Zealand, and their comparison. New Zealand hosts high quality fossil records of foraminifera (Hayward et al. 2009), bony fish (Teleostei) (Schwarzhans 2019) and terrestrial plants (pollen and spores) (Prebble et al. 2021) through the Cenozoic. Recent research suggests that many different phyla share biogeographic boundaries and that environmental conditions, particularly temperature, explained some of the observed variation (Holman et al. 2021). Whilst Holman et al. (2021) attribute these patterns to long-term ecological and evolutionary mechanisms, it is not known whether these biogeographic synergies are maintained over longer timescales.

4. Further investigation into the global spatiotemporal mismatch between taxonomic and functional diversity, and the role of regional ecosystems in the evolution and future trends of the latitudinal diversity gradient.

References

Aberhan, M., and W. Kiessling. 2012. Phanerozoic marine biodiversity: a fresh look at data, methods, patterns and processes. Pp. 3-22. Earth and life. Springer.

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 112(23):7207-7212.

Agresti, A., and M. Kateri. 2011. Categorical data analysis. Springer.

Allmon, W. D. 2017. Species, lineages, splitting, and divergence: why we still need 'anagenesis' and 'cladogenesis'. Biological Journal of the Linnean Society 120(2):474-479.

Alroy, J. 2008. Dynamics of origination and extinction in the marine fossil record. Proceedings of the National Academy of Sciences 105(Supplement 1):11536-11542.

Alroy, J. 2010a. Fair sampling of taxonomic richness and unbiased estimation of origination and extinction rates. The Paleontological Society Papers 16:55-80.

Alroy, J. 2010b. Geographical, environmental and intrinsic biotic controls on Phanerozoic marine diversification. Palaeontology 53(6):1211-1235.

Alroy, J. 2010c. The shifting balance of diversity among major marine animal groups. Science 329(5996):1191-1194.

Alroy, J. 2014. Accurate and precise estimates of origination and extinction rates. Paleobiology 40(3):374-397.

Alroy, J. 2017. Effects of habitat disturbance on tropical forest biodiversity. Proceedings of the National Academy of Sciences 114(23):6056-6061.

Alroy, J. 2020. On four measures of taxonomic richness. Paleobiology 46(2):158-175.

Alroy, J., M. Aberhan, D. J. Bottjer, M. Foote, F. T. Fürsich, P. J. Harries, A. J. W. Hendy, S. M. Holland, L. C. Ivany, W. Kiessling, M. A. Kosnik, C. R. Marshall, A. J. McGowan, A. I. Miller, T. D. Olszewski, M. E. Patzkowsky, S. E. Peters, L. Villier, P. J. Wagner, N. Bonuso, P. S. Borkow, B. Brenneis, M. E. Clapham, L. M. Fall, C. A. Ferguson, V. L. Hanson, A. Z. Krug, K. M. Layou, E. H. Leckey, S. Nürnberg, C. M. Powers, J. A. Sessa, C. Simpson, A. Tomašových, and C. C. Visaggi. 2008. Phanerozoic Trends in the Global Diversity of Marine Invertebrates. Science 321(5885):97-100.

Alroy, J., C. Marshall, R. Bambach, K. Bezusko, M. Foote, F. Fürsich, T. Hansen, S. Holland, L. Ivany, and D. Jablonski. 2001. Effects of sampling standardization on estimates of Phanerozoic marine diversification. Proceedings of the National Academy of Sciences 98(11):6261-6266.

Alvarez, S. A., S. J. Gibbs, P. R. Bown, H. Kim, R. M. Sheward, and A. Ridgwell. 2019. Diversity decoupled from ecosystem function and resilience during mass extinction recovery. Nature 574(7777):242-245.

Anderson, M. J., T. O. Crist, J. M. Chase, M. Vellend, B. D. Inouye, A. L. Freestone, N. J. Sanders, H. V. Cornell, L. S. Comita, and K. F. Davies. 2011. Navigating the multiple meanings of β diversity: a roadmap for the practicing ecologist. Ecology letters 14(1):19-28.

Antão, L. H., A. E. Bates, S. A. Blowes, C. Waldock, S. R. Supp, A. E. Magurran, M. Dornelas, and A. M. Schipper. 2020. Temperature-related biodiversity change across temperate marine and terrestrial systems. Nature Ecology & Evolution:1-7.

Antão, L. H., B. McGill, A. E. Magurran, A. M. Soares, and M. Dornelas. 2019. β-diversity scaling patterns are consistent across metrics and taxa. Ecography 42(5):1012-1023.

Antell, G. S., W. Kiessling, M. Aberhan, and E. E. Saupe. 2020. Marine biodiversity and geographic distributions are independent on large scales. Current Biology 30(1):115-121. e5.

Archibald, S. B., W. H. Bossert, D. R. Greenwood, and B. D. Farrell. 2010. Seasonality, the latitudinal gradient of diversity, and Eocene insects. Paleobiology 36(3):374-398.

Bambach, R. K. 1977. Species richness in marine benthic habitats through the Phanerozoic. Paleobiology:152-167.

Bambach, R. K. 1999. Energetics in the global marinefauna: A connection between terrestrial diversification and change in the marine biosphere. Geobios 32(2):131-144.

Barker, S., I. Cacho, H. Benway, and K. Tachikawa. 2005. Planktonic foraminiferal Mg/Ca as a proxy for past oceanic temperatures: a methodological overview and data compilation for the Last Glacial Maximum. Quaternary Science Reviews 24(7-9):821-834.

Barnosky, A. D., M. A. Carrasco, and E. B. Davis. 2005. The impact of the species–area relationship on estimates of paleodiversity. PLoS Biol 3(8):e266.

Barnosky, A. D., M. A. Carrasco, and R. W. Graham. 2011a. Collateral mammal diversity loss associated with late Quaternary megafaunal extinctions and implications for the future. Geological Society, London, Special Publications 358(1):179-189.

Barnosky, A. D., N. Matzke, S. Tomiya, G. O. Wogan, B. Swartz, T. B. Quental, C. Marshall, J. L. McGuire, E. L. Lindsey, and K. C. Maguire. 2011b. Has the Earth's sixth mass extinction already arrived? Nature 471(7336):51-57.

Barton, P. S., S. A. Cunningham, A. D. Manning, H. Gibb, D. B. Lindenmayer, and R. K. Didham. 2013. The spatial scaling of beta diversity. Global Ecology and Biogeography 22(6):639-647.

Barton, K. 2020. MuMIn: Multi-Model Inference. R package version 1.43.17. Available at: http://CRAN. R-project. org/package= MuMIn.

Baselga, A. 2010. Partitioning the turnover and nestedness components of beta diversity. Global Ecology and Biogeography 19(1):134-143.

Baselga, A. 2017. Partitioning abundance-based multiple-site dissimilarity into components: Balanced variation in abundance and abundance gradients. Methods in Ecology and Evolution 8(7):799-808.

Baselga, A., and F. Leprieur. 2015. Comparing methods to separate components of beta diversity. Methods in Ecology and Evolution 6(9):1069-1079.

Baselga, A., and C. D. L. Orme. 2012. betapart: an R package for the study of beta diversity. Methods in Ecology and Evolution 3(5):808-812.

Benton, M. J. 1995. Diversification and extinction in the history of life. Science 268(5207):52-58.

Beu, A. 1990. Molluscan generic diversity of New Zealand Neogene stages: extinction and biostratigraphic events. Palaeogeography, Palaeoclimatology, Palaeoecology 77(3-4):279-288.

Beu, A. 2006. Marine Mollusca of oxygen isotope stages of the last 2 million years in New Zealand. Part 2. Biostratigraphically useful and new Pliocene to Recent bivalves. Journal of the Royal Society of New Zealand 36(4):151-338.

Beu, A. 2011. Marine Mollusca of isotope stages of the last 2 million years in New Zealand. Part 4. Gastropoda (Ptenoglossa, Neogastropoda, Heterobranchia). Journal of the Royal Society of New Zealand 41(1):1-153.

Beu, A. 2012. Marine Mollusca of the last 2 million years in New Zealand. Part 5. Summary. Journal of the Royal Society of New Zealand 42(1):1-47.

Beu, A., B. Alloway, B. Pillans, T. Naish, and J. Westgate. 2004. Marine Mollusca of oxygen isotope stages of the last 2 million years in New Zealand. Part 1: Revised generic positions and recognition of warm-water and cool-water migrants. Journal of the Royal Society of New Zealand 34(2):111-265.

Beu, A., M. Griffin, and P. Maxwell. 1997. Opening of Drake Passage gateway and Late Miocene to Pleistocene cooling reflected in Southern Ocean molluscan dispersal: evidence from New Zealand and Argentina. Tectonophysics 281(1-2):83-97.

Beu, A. G., and P. A. Maxwell. 1990. Cenozoic Mollusca of New Zealand. New Zealand Geological Survey Paleolontological Bulletin 58:1-518.

Beu, A. G., and J. Raine. 2009. Revised descriptions of New Zealand Cenozoic Mollusca from Beu and Maxwell (1990). GNS Science miscellaneous series 27:5-16.

Biggs, C. R., L. A. Yeager, D. G. Bolser, C. Bonsell, A. M. Dichiera, Z. Hou, S. R. Keyser, A. J. Khursigara, K. Lu, and A. F. Muth. 2020. Does functional redundancy affect ecological stability and resilience? A review and meta-analysis. Ecosphere 11(7):e03184.

Bindoff, N., W. W. Cheung, J. Kairo, J. Arstegui, V. Guinder, R. Hallberg, N. Hilmi, N. Jiao, M. Karim, and L. Levin. 2019. Changing ocean, marine ecosystems, and dependent communities. IPCC special report on the ocean and cryosphere in a changing climate.

Blois, J. L., P. L. Zarnetske, M. C. Fitzpatrick, and S. Finnegan. 2013. Climate change and the past, present, and future of biotic interactions. Science 341(6145):499-504.

Blowes, S. A., S. R. Supp, L. H. Antão, A. Bates, H. Bruelheide, J. M. Chase, F. Moyes, A. Magurran, B. McGill, and I. H. Myers-Smith. 2019. The geography of biodiversity change in marine and terrestrial assemblages. Science 366(6463):339-345.

Bostock, H. C., B. W. Hayward, H. L. Neil, A. T. Sabaa, and G. H. Scott. 2015. Changes in the position of the Subtropical Front south of New Zealand since the last glacial period. Paleoceanography 30(7):824-844.

Brocklehurst, N., M. O. Day, and J. Fröbisch. 2018. Accounting for differences in species frequency distributions when calculating beta diversity in the fossil record. Methods in Ecology and Evolution 9(6):1409-1420.

Burgman, M. A., and J. C. Fox. 2003. Bias in species range estimates from minimum convex polygons: implications for conservation and options for improved planning. Pp. 19-28. Animal Conservation forum. Cambridge University Press.

Burke, K., J. Williams, M. Chandler, A. Haywood, D. Lunt, and B. Otto-Bliesner. 2018. Pliocene and Eocene provide best analogs for near-future climates. Proceedings of the National Academy of Sciences 115(52):13288-13293.

Carrasco, M. A., A. D. Barnosky, and R. W. Graham. 2009. Quantifying the extent of North American mammal extinction relative to the pre-anthropogenic baseline. PloS one 4(12):e8331.

Carter, J. G. 1990. Skeletal Biomineralization: Patterns, Processes and Evolutionary Trends: Volume I. Springer.

Ceballos, G., P. R. Ehrlich, A. D. Barnosky, A. García, R. M. Pringle, and T. M. Palmer. 2015. Accelerated modern human–induced species losses: Entering the sixth mass extinction. Science advances 1(5):e1400253.

Chao, A., and L. Jost. 2012. Coverage-based rarefaction and extrapolation: standardizing samples by completeness rather than size. Ecology 93(12):2533-2547.

Chao, A., and T. Shen. 2010. User's guide for program SPADE (Species prediction and diversity estimation). Taiwan: National Tsing Hua University.

Chen, H., P. Cohen, and S. Chen. 2010. How big is a big odds ratio? Interpreting the magnitudes of odds ratios in epidemiological studies. Communications in Statistics—Simulation and Computation® 39(4):860-864.

Cleary, T. J., B. C. Moon, A. M. Dunhill, and M. J. Benton. 2015. The fossil record of ichthyosaurs, completeness metrics and sampling biases. Palaeontology 58(3):521-536.

Close, R., R. Benson, E. Saupe, M. Clapham, and R. Butler. 2020a. The spatial structure of Phanerozoic marine animal diversity. Science 368(6489):420-424.

Close, R. A., R. B. Benson, J. Alroy, M. T. Carrano, T. J. Cleary, E. M. Dunne, P. D. Mannion, M. D. Uhen, and R. J. Butler. 2020b. The apparent exponential radiation of Phanerozoic land vertebrates is an artefact of spatial sampling biases. Proceedings of the Royal Society B 287(1924):20200372.

Close, R. A., R. B. Benson, P. Upchurch, and R. J. Butler. 2017. Controlling for the species-area effect supports constrained long-term Mesozoic terrestrial vertebrate diversification. Nature Communications 8(1):1-11.

Close, R. A., S. W. Evers, J. Alroy, and R. J. Butler. 2018. How should we estimate diversity in the fossil record? Testing richness estimators using sampling-standardised discovery curves. Methods in Ecology and Evolution 9(6):1386-1400.

Clowes, C. D., J. S. Crampton, K. J. Bland, K. S. Collins, J. G. Prebble, J. I. Raine, D. P. Strogen, M. G. Terezow, and T. M. Womack. 2020. The New Zealand Fossil Record File: a unique database of biological history. New Zealand Journal of Geology and Geophysics:1-10.

Collins, K., S. Edie, T. Gao, R. Bieler, and D. Jablonski. 2019. Spatial filters of function and phylogeny determine morphological disparity with latitude. PloS one 14(8):e0221490.

Collins, K., S. Edie, G. Hunt, K. Roy, and D. Jablonski. 2018. Extinction risk in extant marine species integrating palaeontological and biodistributional data. Proceedings of the Royal Society B: Biological Sciences 285(1887):20181698.

Connolly, S. R., and A. I. Miller. 2001. Joint estimation of sampling and turnover rates from fossil databases: capture-mark-recapture methods revisited. Paleobiology:751-767.

Cooper, A., and R. A. Cooper. 1995. The Oligocene bottleneck and New Zealand biota: genetic record of a past environmental crisis. Proceedings of the Royal Society of London. Series B: Biological Sciences 261(1362):293-302.

Cooper, R. A., P. A. Maxwell, J. S. Crampton, A. G. Beu, C. M. Jones, and B. A. Marshall. 2006. Completeness of the fossil record: estimating losses due to small body size. Geology 34(4):241-244.

Cornell, H. V. 2013. Is regional species diversity bounded or unbounded? Biological Reviews 88(1):140-165.

Cox, D. R., and E. J. Snell. 2018. Analysis of binary data. Routledge.

Crame, J. A. 2000. Evolution of taxonomic diversity gradients in the marine realm: evidence from the composition of Recent bivalve faunas. Paleobiology 26(2):188-214.

Crame, J. A. 2020. Early Cenozoic evolution of the latitudinal diversity gradient. Earth-Science Reviews 202:103090.

Cramer, B., K. Miller, P. Barrett, and J. Wright. 2011. Late Cretaceous–Neogene trends in deep ocean temperature and continental ice volume: Reconciling records of benthic foraminiferal geochemistry ($\delta 18O$ and Mg/Ca) with sea level history. Journal of Geophysical Research: Oceans 116(C12).

Cramer, B., J. Toggweiler, J. Wright, M. Katz, and K. Miller. 2009. Ocean overturning since the Late Cretaceous: Inferences from a new benthic foraminiferal isotope compilation. Paleoceanography 24(4).

Crampton, J. S., A. G. Beu, R. A. Cooper, C. M. Jones, B. Marshall, and P. A. Maxwell. 2003. Estimating the rock volume bias in paleobiodiversity studies. Science 301(5631):358-360.

Crampton, J. S., A. G. Beu, R. A. Cooper, C. M. Jones, I. Matcham, M. Foote, B. A. Marshall, and P. A. Maxwell. 2006a. The ark was full! Constant to declining Cenozoic shallow marine biodiversity on an isolated midlatitude continent. Paleobiology 32(4):509-532.

Crampton, J. S., R. Cooper, A. G. Beu, M. Foote, and B. A. Marshall. 2010. Biotic influences on species duration: interactions between traits in marine molluscs. Paleobiology 36(2):204-223.

Crampton, J. S., M. Foote, A. G. Beu, R. A. Cooper, I. Matcham, C. M. Jones, P. A. Maxwell, and B. A. Marshall. 2006b. Second-order sequence stratigraphic controls on the quality of the fossil record at an active margin: New Zealand Eocene to Recent shelf molluscs. Palaios 21(1):86-105.

Crampton, J. S., M. Foote, R. A. Cooper, A. G. Beu, and S. E. Peters. 2011. The fossil record and spatial structuring of environments and biodiversity in the Cenozoic of New Zealand. Geological Society, London, Special Publications 358(1):105-122.

Curran-Everett, D. 2000. Multiple comparisons: philosophies and illustrations. American Journal of Physiology-Regulatory, Integrative and Comparative Physiology 279(1):R1-R8.

Darroch, S. A., and E. E. Saupe. 2018. Reconstructing geographic range-size dynamics from fossil data. Paleobiology 44(1):25-39.

Drakare, S., J. J. Lennon, and H. Hillebrand. 2006. The imprint of the geographical, evolutionary and ecological context on species–area relationships. Ecology letters 9(2):215-227.

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(1):133-148.

Dunhill, A. M., and M. A. Wills. 2015. Geographic range did not confer resilience to extinction in terrestrial vertebrates at the end-Triassic crisis. Nature Communications 6(1):1-8.

Dunne, E. M., R. A. Close, D. J. Button, N. Brocklehurst, D. D. Cashmore, G. T. Lloyd, and R. J. Butler. 2018. Diversity change during the rise of tetrapods and the impact of the 'Carboniferous rainforest collapse'. Proceedings of the Royal Society B: Biological Sciences 285(1872):20172730.

Edie, S. M., S. Huang, K. S. Collins, K. Roy, and D. Jablonski. 2018a. Loss of biodiversity dimensions through shifting climates and ancient mass extinctions. Integrative and comparative biology 58(6):1179-1190.

Edie, S. M., D. Jablonski, and J. W. Valentine. 2018b. Contrasting responses of functional diversity to major losses in taxonomic diversity. Proceedings of the National Academy of Sciences 115(4):732-737.

Ennyu, A., and M. A. Arthur. 2004. 13. Data Report: Oxygen and Carbon Stable Isotope Records of the Miocene Calcareous Microfossils from ODP Leg 189 Sites 1170 (South Tasman Rise) and 1172 (East Tasman Plateau).

Erwin, D. H. 2009. Climate as a driver of evolutionary change. Current Biology 19(14):R575-R583.

Erwin, D. H., J. W. Valentine, and J. J. Sepkoski Jr. 1987. A comparative study of diversification events: the early Paleozoic versus the Mesozoic. Evolution 41(6):1177-1186.

Ebisuzaki, W. 1997. A method to estimate the statistical significance of a correlation when the data are serially correlated. *Journal of Climate*, *10*(9), 2147-2153.

Field, R., B. A. Hawkins, H. V. Cornell, D. J. Currie, J. A. F. Diniz-Filho, J. F. Guégan, D. M. Kaufman, J. T. Kerr, G. G. Mittelbach, and T. Oberdorff. 2009. Spatial species-richness gradients across scales: a meta-analysis. Journal of Biogeography 36(1):132-147.

Fine, P. V. 2015. Ecological and evolutionary drivers of geographic variation in species diversity. Annual Review of Ecology, Evolution, and Systematics 46:369-392.

Finnegan, S., S. C. Anderson, P. G. Harnik, C. Simpson, D. P. Tittensor, J. E. Byrnes, Z. V. Finkel, D. R. Lindberg, L. H. Liow, and R. Lockwood. 2015. Paleontological baselines for evaluating extinction risk in the modern oceans. Science 348(6234):567-570.

Fleming, C. A. 1975. The geological history of New Zealand and its biota. Pp. 1-86. Biogeography and ecology in New Zealand. Springer.

Fleming, C. H., and J. M. Calabrese. 2017. A new kernel density estimator for accurate home-range and species-range area estimation. Methods in Ecology and Evolution 8(5):571-579.

Foote, M. 1997. Estimating taxonomic durations and preservation probability. Paleobiology:278-300.

Foote, M. 2000. Origination and extinction components of taxonomic diversity: general problems. Paleobiology 26(sp4):74-102.

Foote, M. 2003. Origination and extinction through the Phanerozoic: a new approach. The Journal of Geology 111(2):125-148.

Foote, M. 2005. Pulsed origination and extinction in the marine realm. Paleobiology 31(1):6-20.

Foote, M., J. S. Crampton, A. G. Beu, and R. A. Cooper. 2008. On the bidirectional relationship between geographic range and taxonomic duration. Paleobiology 34(4):421-433.

Foote, M., J. S. Crampton, A. G. Beu, B. A. Marshall, R. A. Cooper, P. A. Maxwell, and I. Matcham. 2007. Rise and fall of species occupancy in Cenozoic fossil mollusks. Science 318(5853):1131-1134.

Foote, M., J. S. Crampton, A. G. Beu, and C. S. Nelson. 2015. Aragonite bias, and lack of bias, in the fossil record: lithological, environmental, and ecological controls. Paleobiology 41(2):245-265.

Foote, M., and D. M. Raup. 1996. Fossil preservation and the stratigraphic ranges of taxa. Paleobiology:121-140.

Foster, W. J., and R. J. Twitchett. 2014. Functional diversity of marine ecosystems after the Late Permian mass extinction event. Nature Geoscience 7(3):233-238.

Fox, J., and S. Weisberg. 2018. An R companion to applied regression. Sage publications.

Freeman, D., B. Marshall, S. Ahyong, S. Wing, and R. Hitchmough. 2010. Conservation status of New Zealand marine invertebrates, 2009. New Zealand Journal of Marine and Freshwater Research 44(3):129-148.

Gagné, T. O., G. Reygondeau, C. N. Jenkins, J. O. Sexton, S. J. Bograd, E. L. Hazen, and K. S. Van Houtan. 2020. Towards a global understanding of the drivers of marine and terrestrial biodiversity. PloS one 15(2):e0228065.

Gibbard, P. L., Head, M. J., Walker, M. J., & Subcommission on Quaternary Stratigraphy. 2010. Formal ratification of the Quaternary System/Period and the Pleistocene Series/Epoch with a base at 2.58 Ma. *Journal of Quaternary Science*, *25*(2), 96-102.

Good, I. J. 1953. The population frequencies of species and the estimation of population parameters. Biometrika 40(3-4):237-264.

Gordon, D. P., J. Beaumont, A. MacDiarmid, D. A. Robertson, and S. T. Ahyong. 2010. Marine biodiversity of aotearoa New Zealand. PloS one 5(8):e10905.

Gradstein, F. M., J. G. Ogg, and A. G. Smith. 2004. A geologic time scale 2004. Cambridge University Press.

Hall, R., M. A. Cottam, and M. E. Wilson. 2011. The SE Asian gateway: history and tectonics of the Australia–Asia collision. Geological Society, London, Special Publications 355(1):1-6.

Halley, J. M., V. Sgardeli, and K. A. Triantis. 2014. Extinction debt and the species–area relationship: a neutral perspective. Global Ecology and Biogeography 23(1):113-123.

Hannisdal, B., J. Henderiks, and L. H. Liow. 2012. Long-term evolutionary and ecological responses of calcifying phytoplankton to changes in atmospheric CO 2. Global Change Biology 18(12):3504-3516.

Haq, B. U., J. Hardenbol, and P. R. Vail. 1987. Chronology of fluctuating sea levels since the Triassic. Science 235(4793):1156-1167.

Harmon, L. J., and S. Harrison. 2015. Species diversity is dynamic and unbounded at local and continental scales. The American Naturalist 185(5):584-593.

Hawkins, B. A., R. Field, H. V. Cornell, D. J. Currie, J.-F. Guégan, D. M. Kaufman, J. T. Kerr, G. G. Mittelbach, T. Oberdorff, and E. M. O'Brien. 2003. Energy, water, and broad-scale geographic patterns of species richness. Ecology 84(12):3105-3117.

Hayward, B. 1977. Lower Miocene corals from the Waitakere Ranges, North Auckland, New Zealand. Journal of the Royal Society of New Zealand 7(1):99-111.

Hayward, B. W., A. T. Sabaa, A. Kolodziej, M. P. Crundwell, S. Steph, G. H. Scott, H. L. Neil, H. C. Bostock, L. Carter, and H. R. Grenfell. 2012. Planktic foraminifera-based sea-surface temperature record in the Tasman Sea and history of the Subtropical Front around New Zealand, over the last one million years. Marine Micropaleontology 82:13-27.

Hayward, B. W., O. Tendal, R. Carter, H. Grenfell, H. Morgans, G. H. Scott, C. Strong, and J. Hayward. 2009. Phylum Foraminifera. *In* D. Gordon, ed. New Zealand Inventory of Biodiversity. Canterbury University Press, Canterbury, New Zealand.

He, J., H. Kreft, S. Lin, Y. Xu, and H. Jiang. 2018. Cenozoic evolution of beta diversity and a Pleistocene emergence for modern mammal faunas in China. Global Ecology and Biogeography 27(11):1326-1338.

Heim, N. A., and S. E. Peters. 2011. Covariation in macrostratigraphic and macroevolutionary patterns in the marine record of North America. GSA Bulletin 123(3/4):620-630.

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(4):511-528.

Hendy, A. J. 2005. Lithification and the measurement of biodiversity—is missing alpha stuck between a rock and a hard place. P. 117. Geological Society of America Abstracts with Programs.

Hendy, A. J. 2009. The influence of lithification on Cenozoic marine biodiversity trends. Paleobiology:51-62.

Herbert, T. D., K. T. Lawrence, A. Tzanova, L. C. Peterson, R. Caballero-Gill, and C. S. Kelly. 2016. Late Miocene global cooling and the rise of modern ecosystems. Nature Geoscience 9(11):843-847.

Hillebrand, H. 2004. On the generality of the latitudinal diversity gradient. The American Naturalist 163(2):192-211.

Hofmann, R., M. Tietje, and M. Aberhan. 2019. Diversity partitioning in Phanerozoic benthic marine communities. Proceedings of the National Academy of Sciences 116(1):79-83.

Holland, S. M. 2010. Additive diversity partitioning in palaeobiology: revisiting Sepkoski's question. Palaeontology 53(6):1237-1254.

Holland, S. M. 2015. Estimating diversity with SQS.

Holland, S. M. 2017. Structure, not Bias. Journal of Paleontology 91:1315-1317.

Holland, S. M., and J. A. Sclafani. 2015. Phanerozoic diversity and neutral theory. Paleobiology 41(3):369-376.

Hooper, D. U., F. Chapin Iii, J. Ewel, A. Hector, P. Inchausti, S. Lavorel, J. H. Lawton, D. Lodge, M. Loreau, and S. Naeem. 2005. Effects of biodiversity on ecosystem functioning: a consensus of current knowledge. Ecological monographs 75(1):3-35.

Hornibrook, N. d. B. 1992. New Zealand Cenozoic marine paleoclimates: a review based on the distribution of some shallow water and terrestrial biota. Pacific Neogene: environment, evolution, and events. University of Tokyo Press, Tokyo:83-106.

Hsieh, T., K. Ma, and A. Chao. 2016. iNEXT: an R package for rarefaction and extrapolation of species diversity (H ill numbers). Methods in Ecology and Evolution 7(12):1451-1456.

Huber, M., H. Brinkhuis, C. E. Stickley, K. Döös, A. Sluijs, J. Warnaar, S. A. Schellenberg, and G. L. Williams. 2004. Eocene circulation of the Southern Ocean: Was Antarctica kept warm by subtropical waters? Paleoceanography 19(4).

Huber, M., and R. Caballero. 2003. Eocene El Nino: Evidence for robust tropical dynamics in the" hothouse". Science 299(5608):877-881.

Hunt, G., and K. Roy. 2006. Climate change, body size evolution, and Cope's Rule in deep-sea ostracodes. Proceedings of the National Academy of Sciences 103(5):1347-1352.

Jablonski, D. 2005. Evolutionary innovations in the fossil record: the intersection of ecology, development, and macroevolution. Journal of Experimental Zoology Part B: Molecular and Developmental Evolution 304(6):504-519.

Jablonski, D., C. L. Belanger, S. K. Berke, S. Huang, A. Z. Krug, K. Roy, A. Tomasovych, and J. W. Valentine. 2013. Out of the tropics, but how? Fossils, bridge species, and thermal ranges in the dynamics of the marine latitudinal diversity gradient. Proceedings of the National Academy of Sciences 110(26):10487-10494.

Jablonski, D., S. Huang, K. Roy, and J. W. Valentine. 2017. Shaping the latitudinal diversity gradient: new perspectives from a synthesis of paleobiology and biogeography. The American Naturalist 189(1):1-12.

Jablonski, D., K. Roy, J. W. Valentine, R. M. Price, and P. S. Anderson. 2003. The impact of the pull of the recent on the history of marine diversity. Science 300(5622):1133-1135.

Johnson, K. G., and G. B. Curry. 2001. Regional biotic turnover dynamics in the Plio-Pleistocene molluscan fauna of the Wanganui Basin, New Zealand. Palaeogeography, Palaeoclimatology, Palaeoecology 172(1-2):39-51.

Hyndman, R., Athanasopoulos, G., Bergmeir, C., Caceres, G., Chhay, L., O'Hara-Wild, M., Petropoulos, F., Razbash, S., Wang, E., Yasmeen, F. 2020. forecast: Forecasting functions for time series and linear models. R package version 8.13. Available at: http://CRAN. R-project. org/package= forecast.

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 6:35420.

Kennett, J. P. 1986. Miocene to early Pliocene oxygen and carbon isotope stratigraphy in the southwest Pacific, Deep Sea Drilling Project Leg 90. Initial reports of the deep sea drilling project 90:1383-1411.

King, P. 2000. New Zealand's changing configuration in the last 100 million years: plate tectonics, basin development, and depositional setting. 2000 New Zealand Petroleum Conference Proceedings. Crown Minerals, Ministry of Commerce Wellington, New Zealand.

King, P. R., T. R. Naish, G. H. Brown, B. D. Field, and S. W. Erdbrooke. 1999. Cretaceous to recent sedimentary patterns in New Zealand. Lower Hutt, New Zealand: Institue of Geological & Nuclear Sciences Limited.

Kocsis, A. T., Reddin, C. J., Alroy, J., & Kiessling, W. 2019. The R package divDyn for quantifying diversity dynamics using fossil sampling data. *Methods in Ecology and Evolution*, *10*(5), 735-743.

Kocsis, A. T. 2020. Triangular and Penta-Hexagonal Grids Based on Tessellated Icosahedra. R package version 0.10. 0. https://CRAN.R-project.org/package=icosa.

Koleff, P., K. J. Gaston, and J. J. Lennon. 2003. Measuring beta diversity for presence–absence data. Journal of Animal Ecology 72(3):367-382.

Laake, J. L. 2013. RMark: an R interface for analysis of capture-recapture data with MARK.

Lagomarcino, A. J., and A. I. Miller. 2012. The relationship between genus richness and geographic area in Late Cretaceous marine biotas: epicontinental sea versus open-ocean-facing settings. PloS one 7(8):e40472.

Lamb, S. 2011. Cenozoic tectonic evolution of the New Zealand plate-boundary zone: A paleomagnetic perspective. Tectonophysics 509(3-4):135-164.

Lande, R. 1996. Statistics and partitioning of species diversity, and similarity among multiple communities. Oikos:5-13.

Lane, A., and M. J. Benton. 2003. Taxonomic level as a determinant of the shape of the Phanerozoic marine biodiversity curve. The American Naturalist 162(3):265-276.

Lear, C. H., H. Elderfield, and P. Wilson. 2000. Cenozoic deep-sea temperatures and global ice volumes from Mg/Ca in benthic foraminiferal calcite. Science 287(5451):269-272.

Legendre, P. 2014. lmodel2: Model II Regression. R package version 1.7-2. Available at: http://CRAN. R-project. org/package= lmodel2

Lennon, J. J., P. Koleff, J. Greenwood, and K. J. Gaston. 2001. The geographical structure of British bird distributions: diversity, spatial turnover and scale. Journal of Animal Ecology 70(6):966-979.

Lewis, O. T. 2006. Climate change, species–area curves and the extinction crisis. Philosophical Transactions of the Royal Society B: Biological Sciences 361(1465):163-171.

Liow, L. H., and J. A. Finarelli. 2014. A dynamic global equilibrium in carnivoran diversification over 20 million years. Proceedings of the Royal Society B: Biological Sciences 281(1778):20132312.

Liow, L. H., and J. D. Nichols. 2010. Estimating rates and probabilities of origination and extinction using taxonomic occurrence data: capture-mark-recapture (CMR) approaches. Quantitative methods in paleobiology. Paleontological Society Papers 16:81-94.

Livermore, R., A. Nankivell, G. Eagles, and P. Morris. 2005. Paleogene opening of Drake passage. Earth and Planetary Science Letters 236(1-2):459-470.

Lotze, H. K., D. P. Tittensor, A. Bryndum-Buchholz, T. D. Eddy, W. W. Cheung, E. D. Galbraith, M. Barange, N. Barrier, D. Bianchi, and J. L. Blanchard. 2019. Global ensemble projections reveal trophic amplification of ocean biomass declines with climate change. Proceedings of the National Academy of Sciences 116(26):12907-12912.

Lyle, M., S. Gibbs, T. C. Moore, and D. K. Rea. 2007. Late Oligocene initiation of the Antarctic circumpolar current: evidence from the South Pacific. Geology 35(8):691-694.

MacArthur, R. H., and E. O. Wilson. 2001. The theory of island biogeography. Princeton university press.

Magurran, A. E., and P. A. Henderson. 2010. Temporal turnover and the maintenance of diversity in ecological assemblages. Philosophical Transactions of the Royal Society B: Biological Sciences 365(1558):3611-3620.

Mannion, P. D., R. B. Benson, M. T. Carrano, J. P. Tennant, J. Judd, and R. J. Butler. 2015. Climate constrains the evolutionary history and biodiversity of crocodylians. Nature Communications 6(1):1-9.

Mannion, P. D., P. Upchurch, R. B. Benson, and A. Goswami. 2014. The latitudinal biodiversity gradient through deep time. Trends in ecology & evolution 29(1):42-50.

Marchitto, T., W. Curry, J. Lynch-Stieglitz, S. Bryan, K. Cobb, and D. Lund. 2014. Improved oxygen isotope temperature calibrations for cosmopolitan benthic foraminifera. Geochimica et Cosmochimica Acta 130:1-11.

Markov, A. V., and A. V. Korotayev. 2007. Phanerozoic marine biodiversity follows a hyperbolic trend. Palaeoworld 16(4):311-318.

Marshall, C. R., and T. B. Quental. 2016. The uncertain role of diversity dependence in species diversification and the need to incorporate time-varying carrying capacities. Philosophical Transactions of the Royal Society B: Biological Sciences 371(1691):20150217.

Matthews, K. J., K. T. Maloney, S. Zahirovic, S. E. Williams, M. Seton, and R. D. Mueller. 2016. Global plate boundary evolution and kinematics since the late Paleozoic. Global and Planetary Change 146:226-250.

Mayhew, P. J., M. A. Bell, T. G. Benton, and A. J. McGowan. 2012. Biodiversity tracks temperature over time. Proceedings of the National Academy of Sciences 109(38):15141-15145.

Mayhew, P. J., G. B. Jenkins, and T. G. Benton. 2008. A long-term association between global temperature and biodiversity, origination and extinction in the fossil record. Proceedings of the Royal Society B: Biological Sciences 275(1630):47-53.

McKnight, M. W., P. S. White, R. I. McDonald, J. F. Lamoreux, W. Sechrest, R. S. Ridgely, and S. N. Stuart. 2007. Putting beta-diversity on the map: broad-scale congruence and coincidence in the extremes. PLoS Biol 5(10):e272.

McWilliam, M., M. O. Hoogenboom, A. H. Baird, C.-Y. Kuo, J. S. Madin, and T. P. Hughes. 2018. Biogeographical disparity in the functional diversity and redundancy of corals. Proceedings of the National Academy of Sciences 115(12):3084-3089.

Meyers, S. R.. 2014. astrochron:An R Package for Astrochronology. R package version 1.0. Available at: http://CRAN. R-project. org/package=astrochron.

Miller, A. I. 2000. Conversations about Phanerozoic global diversity. Paleobiology 26(S4):53-73.

Miller, K. G., J. V. Browning, W. J. Schmelz, R. E. Kopp, G. S. Mountain, and J. D. Wright. 2020. Cenozoic sea-level and cryospheric evolution from deep-sea geochemical and continental margin records. Science advances 6(20):eaaz1346.

Mitchell, J. S. 2015. Preservation is predictable: quantifying the effect of taphonomic biases on ecological disparity in birds. Paleobiology 41(2):353-367.

Mittelbach, G. G., D. W. Schemske, H. V. Cornell, A. P. Allen, J. M. Brown, M. B. Bush, S. P. Harrison, A. H. Hurlbert, N. Knowlton, and H. A. Lessios. 2007. Evolution and the latitudinal diversity gradient: speciation, extinction and biogeography. Ecology letters 10(4):315-331.

Mittermeier, R. A., W. R. Turner, F. W. Larsen, T. M. Brooks, and C. Gascon. 2011. Global biodiversity conservation: the critical role of hotspots. Pp. 3-22. Biodiversity hotspots. Springer.

Mortimer, N., and H. Campbell. 2014. Zealandia: Our continent revealed. Institute of Geological & Nuclear Sciences Ltd 271.

Mortimer, N., H. J. Campbell, A. J. Tulloch, P. R. King, V. M. Stagpoole, R. A. Wood, M. S. Rattenbury, R. Sutherland, C. J. Adams, and J. Collot. 2017. Zealandia: Earth's hidden continent. GSA today 27(3):27-35.

Mudelsee, M., T. Bickert, C. H. Lear, and G. Lohmann. 2014. Cenozoic climate changes: A review based on time series analysis of marine benthic δ18O records. Reviews of Geophysics 52(3):333-374.

Müller, R. D., J. Cannon, X. Qin, R. J. Watson, M. Gurnis, S. Williams, T. Pfaffelmoser, M. Seton, S. H. Russell, and S. Zahirovic. 2018. GPlates: building a virtual Earth through deep time. Geochemistry, Geophysics, Geosystems 19(7):2243-2261.

Murphy, M., and J. P. Kennett. 1985. Development of latitudinal thermal gradients during the Oligocene: oxygen-isotope evidence from the southwest Pacific. Initial Reports of Deep Sea Drilling Project, XC.

Myers, N., R. A. Mittermeier, C. G. Mittermeier, G. A. Da Fonseca, and J. Kent. 2000. Biodiversity hotspots for conservation priorities. Nature 403(6772):853-858.

Na, L., and W. Kiessling. 2015. Diversity partitioning during the Cambrian radiation. Proceedings of the National Academy of Sciences 112(15):4702-4706.

Nelson, C., C. Hendy, A. Cuthbertson, and G. Jarrett. 1986. Late Quarternary carbonate and isotope stratigraphy, sub-antarctic site 594, southwest pacific. Initial reports of the deep sea drilling project 90:1425-1436.

Nelson, C. S., and P. J. Cooke. 2001. History of oceanic front development in the New Zealand sector of the Southern Ocean during the Cenozoic—a synthesis. New Zealand Journal of Geology and Geophysics 44(4):535-553.

Nichols, J. D., and K. H. Pollock. 1983. Estimating taxonomic diversity, extinction rates, and speciation rates from fossil data using capture-recapture models. Paleobiology:150-163.

Oksanen, J., R. Kindt, P. Legendre, B. O'Hara, M. H. H. Stevens, M. J. Oksanen, and M. Suggests. 2007. The vegan package. Community ecology package 10:631-637.

Patzkowsky, M. E., and S. M. Holland. 2012. Stratigraphic paleobiology: understanding the distribution of fossil taxa in time and space. University of Chicago Press.

Pateiro-Lopz, B., Rodriguez-Casal, A. 2019. alphahull: Generalization of the convex hull of a sample of point in the plane. R package version 2.2. Available at: http://CRAN. R-project. org/package=alphahull.

Paul, C. R., and S. Donovan. 1998. Adequacy, completeness and the fossil record. The Adequacy of the Fossil Record. Wiley, Chichester 1:28.

Payne, J. L., Bush, A. M., Heim, N. A., Knope, M. L., & McCauley, D. J. 2016. Ecological selectivity of the emerging mass extinction in the oceans. *Science*, *353*(6305), 1284-1286.

Penny, A., and B. Kröger. 2019. Impacts of spatial and environmental differentiation on early Palaeozoic marine biodiversity. Nature Ecology & Evolution 3(12):1655-1660.

Petchey, O. L., and K. J. Gaston. 2002. Extinction and the loss of functional diversity. Proceedings of the Royal Society of London. Series B: Biological Sciences 269(1501):1721-1727.

Petchey, O. L., and K. J. Gaston. 2006. Functional diversity: back to basics and looking forward. Ecology letters 9(6):741-758.

Peters, S. E. 2008. Macrostratigraphy and its promise for paleobiology. The Paleontological Society Papers 14:205-231.

Peters, S. E., and N. A. Heim. 2011. Macrostratigraphy and macroevolution in marine environments: testing the common-cause hypothesis. Geological Society, London, Special Publications 358(1):95-104.

Peters, S. E., J. M. Husson, and J. Czaplewski. 2018. Macrostrat: a platform for geological data integration and deep-time Earth crust research. Geochemistry, Geophysics, Geosystems 19(4):1393-1409.

Peters, S. E., and M. McClennen. 2016. The Paleobiology Database application programming interface. Paleobiology 42(1):1-7.

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 287(1931):20201162.

Pinheiro, J. B., Bates, D., DebRoy, S., Sarkar, D., R Core Development Team. 2020. nlme:Linear and Nonlinear Mixed Effects Models. R package version 3.1-149. Available at: http://CRAN. R-project. org/package=nlme.

Poloczanska, E. S., C. J. Brown, W. J. Sydeman, W. Kiessling, D. S. Schoeman, P. J. Moore, K. Brander, J. F. Bruno, L. B. Buckley, and M. T. Burrows. 2013. Global imprint of climate change on marine life. Nature Climate Change 3(10):919-925.

Powell, A. W. B. 1979. New Zealand Mollusca: marine, land, and freshwater shells. Collins.

Powell, M. G., B. R. Moore, and T. J. Smith. 2015. Origination, extinction, invasion, and extirpation components of the brachiopod latitudinal biodiversity gradient through the Phanerozoic Eon. Paleobiology 41(2):330-341.

Pradel, R. 1996. Utilization of capture-mark-recapture for the study of recruitment and population growth rate. Biometrics:703-709.

Prebble, J. G., E. M. Kennedy, T. Reichgelt, C. Clowes, T. Womack, D. C. Mildenhall, J. I. Raine, and E. M. Crouch. 2021. A 100 million year composite pollen record from New Zealand shows maximum angiosperm abundance delayed until Eocene. Palaeogeography, Palaeoclimatology, Palaeoecology 566:110207.

R Development Core Team. 2021. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Rabosky, D. L., and A. H. Hurlbert. 2015. Species richness at continental scales is dominated by ecological limits. The American Naturalist 185(5):572-583.

Raine, J., A. Beu, A. Boyes, H. Campbell, R. Cooper, J. Crampton, M. Crundwell, C. Hollis, H. Morgans, and N. Mortimer. 2015. New Zealand geological timescale NZGT 2015/1. New Zealand Journal of Geology and Geophysics 58(4):398-403.

Raup, D. M. 1972. Taxonomic diversity during the Phanerozoic. Science 177(4054):1065-1071.

Raup, D. M. 1979. Biases in the fossil recrod of species and genera. Bulletin of the Carnegie Museum of Natural History 13:85-91.

Raup, D. M., & Sepkoski, J. J. 1982. Mass extinctions in the marine fossil record. *Science*, *215*(4539), 1501-1503.

Raymo, M. E., L. Lisiecki, and K. H. Nisancioglu. 2006. Plio-Pleistocene ice volume, Antarctic climate, and the global δ18O record. Science 313(5786):492-495.

Rex, M. A., R. J. Etter, A. J. Clain, and M. S. Hill. 1999. Bathymetric patterns of body size in deep-sea gastropods. Evolution 53(4):1298-1301.

Reygondeau, G., A. Longhurst, E. Martinez, G. Beaugrand, D. Antoine, and O. Maury. 2013. Dynamic biogeochemical provinces in the global ocean. Global Biogeochemical Cycles 27(4):1046-1058.

Ricklefs, R. E. 1987. Community diversity: relative roles of local and regional processes. Science 235(4785):167-171.

Ricklefs, R. E. 2004. A comprehensive framework for global patterns in biodiversity. Ecology letters 7(1):1-15.

Rosenfeld, J. S. 2002. Functional redundancy in ecology and conservation. Oikos 98(1):156-162.

Rosenzweig, M. L. 1995. Species diversity in space and time. Cambridge University Press.

Sabaa, A. T., E. L. Sikes, B. W. Hayward, and W. R. Howard. 2004. Pliocene sea surface temperature changes in ODP Site 1125, Chatham Rise, east of New Zealand. Marine Geology 205(1-4):113-125.

Sahney, S., and M. J. Benton. 2017. The impact of the Pull of the Recent on the fossil record of tetrapods. Evolutionary Ecology Research 18(1):7-23.

Sahney, S., M. J. Benton, and P. A. Ferry. 2010. Links between global taxonomic diversity, ecological diversity and the expansion of vertebrates on land. Biology letters 6(4):544-547.

Scheiner, S. M. 2003. Six types of species-area curves. Global Ecology and Biogeography 12(6):441-447.

Scher, H. D., and E. E. Martin. 2006. Timing and climatic consequences of the opening of Drake Passage. Science 312(5772):428-430.

Schumm, M., S. Edie, K. Collins, V. Gómez-Bahamón, K. Supriya, A. White, T. 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 286(1908):20190745.

Schwarzhans, W. 2019. Reconstruction of the fossil marine bony fish (Teleostei) from the Eocene to Pleistocene of New Zealand by means of otoliths. Museo di Storia Naturale di Milano.

Sclafani, J. A., and S. M. Holland. 2013. The species-area relationship in the Late Ordovician: a test using neutral theory. Diversity 5(2):240-262.

Sepkoski, D. 2012. Rereading the fossil record: the growth of paleobiology as an evolutionary discipline. University of Chicago Press.

Sepkoski, J. J. 1976. Species diversity in the Phanerozoic: species-area effects. Paleobiology 2(4):298-303.

Sepkoski, J. J., R. K. Bambach, D. M. Raup, and J. W. Valentine. 1981. Phanerozoic marine diversity and the fossil record. Nature 293(5832):435-437.

Sepkoski Jr, J. J. 1988. Alpha, beta, or gamma: where does all the diversity go? Paleobiology:221-234.

Shackleton, N. 1974. Attainment of isotopic equilibrium between ocean water and the benthonic foraminifera genus Uvigerina: isotopic changes in the ocean during the last glacial.

Shackleton, N. J., and J. P. Kennett. 1975. Paleotemperature history of the Cenozoic and the initiation of Antarctic glaciation: oxygen and carbon isotope analyses in DSDP Sites 277,279, and 281. Initial Reports of Deep Sea Drilling Project 29:743-756.

Silvestro, D., N. Salamin, A. Antonelli, and X. Meyer. 2019. Improved estimation of macroevolutionary rates from fossil data using a Bayesian framework. Paleobiology 45(4):546-570.

Silvestro, D., N. Salamin, and J. Schnitzler. 2014. PyRate: a new program to estimate speciation and extinction rates from incomplete fossil data. Methods in Ecology and Evolution 5(10):1126-1131.

Smiley, T. M. 2018. Detecting diversification rates in relation to preservation and tectonic history from simulated fossil records. Paleobiology 44(1):1-24.

Smith, A. B. 2001. Large–scale heterogeneity of the fossil record: implications for Phanerozoic biodiversity studies. Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences 356(1407):351-367.

Smith, A. B. 2007. Marine diversity through the Phanerozoic: problems and prospects. Journal of the Geological Society 164(4):731-745.

Song, H., S. Huang, E. Jia, X. Dai, P. B. Wignall, and A. M. Dunhill. 2020. Flat latitudinal diversity gradient caused by the Permian–Triassic mass extinction. Proceedings of the National Academy of Sciences 117(30):17578-17583.

Spencer, H., B. Marshall, and R. Willan. 2009. Chapter 10; Phylum Mollusca - chitons, clams, tusk shells, snails, squids and kin. *In* D. P. Gordon, ed. The New Zealand Inventory of Biodiversity. Canterbury University Press, Christchurch.

Spencer, H., R. Willan, B. Marshall, and T. Murray. 2017. Checklist of the living Mollusca recorded from the New Zealand Exclusive Economic Zone. accessed November 2019 10:2017.

Squires, D. F. 1958. The Cretaceous and Tertiary corals of New Zealand: NZ Geol. Surv., Pal. Bull 29.

Stanley, S. M. 2007. An analysis of the history of marine animal diversity. Paleobiology 33(4_Suppl):1-55.

Starrfelt, J., and L. H. Liow. 2016. How many dinosaur species were there? Fossil bias and true richness estimated using a Poisson sampling model. Philosophical Transactions of the Royal Society B: Biological Sciences 371(1691):20150219.

Storch, D., and J. G. Okie. 2019. The carrying capacity for species richness. Global Ecology and Biogeography 28(10):1519-1532.

Thomas, C. D., A. Cameron, R. E. Green, M. Bakkenes, L. J. Beaumont, Y. C. Collingham, B. F. Erasmus, M. F. De Siqueira, A. Grainger, and L. Hannah. 2004. Extinction risk from climate change. Nature 427(6970):145-148.

Torsvik, T. H., R. Van der Voo, U. Preeden, C. Mac Niocaill, B. Steinberger, P. V. Doubrovine, D. J. Van Hinsbergen, M. Domeier, C. Gaina, and E. Tohver. 2012. Phanerozoic polar wander, palaeogeography and dynamics. Earth-Science Reviews 114(3-4):325-368.

Toumoulin, A., Y. Donnadieu, J. B. Ladant, S. Batenburg, F. Poblete, and G. Dupont-Nivet. 2020. Quantifying the effect of the Drake Passage opening on the Eocene Ocean. Paleoceanography and Paleoclimatology 35(8):e2020PA003889.

Tuomisto, H. 2010a. A diversity of beta diversities: straightening up a concept gone awry. Part 1. Defining beta diversity as a function of alpha and gamma diversity. Ecography 33(1):2-22.

Tuomisto, H. 2010b. A diversity of beta diversities: straightening up a concept gone awry. Part 2. Quantifying beta diversity and related phenomena. Ecography 33(1):23-45.

Ulrich, W., and M. Almeida-Neto. 2012. On the meanings of nestedness: back to the basics. Ecography 35(10):865-871.

Urban, M. C. 2015. Accelerating extinction risk from climate change. Science 348(6234):571-573.

Valentine, J. W. 1969. Patterns of taxonomic and ecological structure of the shelf benthos during Phanerozoic time. Palaeontology 12(4):684-709.

Valentine, J. W. 1973. Evolutionary paleoecology of the marine biosphere.

Valentine, J. W., and D. Jablonski. 2015. A twofold role for global energy gradients in marine biodiversity trends. Journal of Biogeography 42(6):997-1005.

Valiente-Banuet, A., M. A. Aizen, J. M. Alcántara, J. Arroyo, A. Cocucci, M. Galetti, M. B. García, D. García, J. M. Gómez, and P. Jordano. 2015. Beyond species loss: the extinction of ecological interactions in a changing world. Functional Ecology 29(3):299-307.

Van Valen, L. M. 1984. A resetting of Phanerozoic community evolution. *Nature*, *307*(5946), 50-52.

Veech, J. A., and T. O. Crist. 2007. Habitat and climate heterogeneity maintain beta-diversity of birds among landscapes within ecoregions. Global Ecology and Biogeography 16(5):650-656.

Veizer, J., and A. Prokoph. 2015. Temperatures and oxygen isotopic composition of Phanerozoic oceans. Earth-Science Reviews 146:92-104.

Vellend, M. 2001. Do commonly used indices of β-diversity measure species turnover? Journal of Vegetation Science 12(4):545-552.

Venerables, W. N., Ripley, B. D. 2002. MASS: Modern Applied Statistics with S. R package version 7.3-54. Available at: http://CRAN. R-project. org/package= MASS.

Vergés, A., P. D. Steinberg, M. E. Hay, A. G. Poore, A. H. Campbell, E. Ballesteros, K. L. Heck Jr, D. J. Booth, M. A. Coleman, and D. A. Feary. 2014. The tropicalization of temperate marine ecosystems: climate-mediated changes in herbivory and community phase shifts. Proceedings of the Royal Society B: Biological Sciences 281(1789):20140846.

Vermeij, G. J., and L. R. Leighton. 2003. Does global diversity mean anything? Paleobiology 29(1):3-7.

Vilhena, D. A., and A. B. Smith. 2013. Spatial bias in the marine fossil record. PloS one 8(10):e74470.

Villéger, S., P. M. Novack-Gottshall, and D. Mouillot. 2011. The multidimensionality of the niche reveals functional diversity changes in benthic marine biotas across geological time. Ecology letters 14(6):561-568.

Violle, C., P. B. Reich, S. W. Pacala, B. J. Enquist, and J. Kattge. 2014. The emergence and promise of functional biogeography. Proceedings of the National Academy of Sciences 111(38):13690-13696.

Visser, K., R. Thunell, and L. Stott. 2003. Magnitude and timing of temperature change in the Indo-Pacific warm pool during deglaciation. Nature 421(6919):152-155.

Wernberg, T., S. Bennett, R. C. Babcock, T. De Bettignies, K. Cure, M. Depczynski, F. Dufois, J. Fromont, C. J. Fulton, and R. K. Hovey. 2016. Climate-driven regime shift of a temperate marine ecosystem. Science 353(6295):169-172.

Westerhold, T., N. Marwan, A. J. Drury, D. Liebrand, C. Agnini, E. Anagnostou, J. S. Barnet, S. M. Bohaty, D. De Vleeschouwer, and F. Florindo. 2020. An astronomically dated record of Earth's climate and its predictability over the last 66 million years. Science 369(6509):1383-1387.

White, G. C. MARK. 2021. http://www.phidot.org/software/mark/index.html

Whittaker, R. H. 1960. Vegetation of the Siskiyou mountains, Oregon and California. Ecological monographs 30(3):279-338.

Whittaker, R. J., F. Rigal, P. A. Borges, P. Cardoso, S. Terzopoulou, F. Casanoves, L. Pla, F. Guilhaumon, R. J. Ladle, and K. A. Triantis. 2014. Functional biogeography of oceanic islands and the scaling of functional diversity in the Azores. Proceedings of the National Academy of Sciences 111(38):13709-13714.

Wiese, R., J. Renaudie, and D. B. Lazarus. 2016. Testing the accuracy of genus-level data to predict species diversity in Cenozoic marine diatoms. Geology 44(12):1051-1054.

Willig, M. R., D. M. Kaufman, and R. D. Stevens. 2003. Latitudinal gradients of biodiversity: pattern, process, scale, and synthesis. Annual Review of Ecology, Evolution, and Systematics 34(1):273-309.

Womack, T. M., J. S. Crampton, and M. J. Hannah. 2020a. The Pull of the Recent revisited: negligible species-level effect in a regional marine fossil record. Paleobiology:1-8.

Womack, T. M., J. S. Crampton, and M. J. Hannah. 2020b. Spatial scaling of beta diversity in the shallow-marine fossil record. Paleobiology:1-15.

Worm, B., and H. K. Lotze. 2021. Marine biodiversity and climate change. Pp. 445-464. Climate change. Elsevier.

Worton, B. J. 1989. Kernel methods for estimating the utilization distribution in home-range studies. Ecology 70(1):164-168.

Wright, D. H., B. D. Patterson, G. M. Mikkelson, A. Cutler, and W. Atmar. 1997. A comparative analysis of nested subset patterns of species composition. Oecologia 113(1):1-20.

Xing, Y., R. E. Onstein, R. J. Carter, T. Stadler, and H. Peter Linder. 2014. Fossils and a large molecular phylogeny show that the evolution of species richness, generic diversity, and turnover rates are disconnected. Evolution 68(10):2821-2832.

Zachos, J., M. Pagani, L. Sloan, E. Thomas, and K. Billups. 2001. Trends, rhythms, and aberrations in global climate 65 Ma to present. Science 292(5517):686-693.

Appendices

Appendix 1.

Simplified step-by-step breakdown of the original implementation of Shareholder Quorum Subsampling (SQS) (Alroy 2010a). Note this is based on sampling individual taxa, not collections.

1. Estimate Good's u (see Eq. 4.1). If above 0 (i.e. not all singletons) continue.

2. Calculate initial taxon frequency*.

$$int.freq_i = n_i - \frac{(s_1 + s_2)/2}{S}$$

3. Calculate taxon frequency*.

$$freq_i = \frac{int.freq_i}{N - \max(n_i)}$$

4. For a single subsampling trial:

   a. Randomly sample taxon, one at a time, from the sample abundance distribution.

   b. For each randomly sampled taxon, calculate the cumulative taxon frequency (coverage) of all sampled taxon. Importantly, the frequency for an individual species can only be included once.

   c. Stop sampling once the cumulative taxon frequency (coverage) is greater than the set coverage level (quota).

   d. Calculate richness from the subsampled frequency distribution.

5. Calculate the geometric mean of richness from all subsampling trials.

Steps 2 and 3 and the equations given adjust the frequency of individual taxa to accommodate for the completeness of the sample. Using the raw frequencies of each taxa would overestimate their actual coverage.

Appendix 2.

Simplified step-by-step breakdown of the updated Shareholder Quorum Subsampling (SQS) (Alroy 2014), following the outline provided in Close et al. (2018).

1. If required - randomly draw a reference (without replacement)

2. Randomly draw a collection from the reference.

3. Draw individual occurrences from the collection, calculating Good's u after each occurrence is added.

4. Record richness each time sample coverage meets the target coverage (just below or above).

5. Repeat the process drawing new collections from the reference (until the maximum number of set collections is drawn – n collections (recommended to be 3).

6. Once n collections are drawn from the reference, sample another reference and repeat until all references have been evaluated.

7. For each trial, the richness is recorded as the median richness at all crossing points.