OPEN ACCESS

University of BRISTOL

This electronic thesis or dissertation has been downloaded from Explore Bristol Research, http://research-information.bristol.ac.uk

Author:
Flannery-Sutherland, Joe T

Title:
Revolution on land and sea

bias, biodiversity and biogeography through the Triassic

General rights
Access to the thesis is subject to the Creative Commons Attribution - NonCommercial-No Derivatives 4.0 International Public License. A copy of this may be found at https://creativecommons.org/licenses/by-nc-nd/4.0/legalcode This license sets out your rights and the restrictions that apply to your access to the thesis so it is important you read this before proceeding.

Take down policy
Some pages of this thesis may have been removed for copyright restrictions prior to having it been deposited in Explore Bristol Research. However, if you have discovered material within the thesis that you consider to be unlawful e.g. breaches of copyright (either yours or that of a third party) or any other law, including but not limited to those relating to patent, trademark, confidentiality, data protection, obscenity, defamation, libel, then please contact collections-metadata@bristol.ac.uk and include the following information in your message:

•Your contact details
•Bibliographic details for the item, including a URL
•An outline nature of the complaint

Your claim will be investigated and, where appropriate, the item in question will be removed from public view as soon as possible.

Revolution on Land and Sea: Bias, Biodiversity and Biogeography through the Triassic

Joseph T. Flannery Sutherland

Supervised by Professor Michael J. Benton.

A dissertation submitted to the University of Bristol in accordance with the requirements for award of the degree of Doctor of Philosophy in the Faculty of Science, School of Earth Sciences.

August 2023

55511 words

Abstract

Novel analytical methods emerging over the last decade, combined with a wealth of fossil occurrence data accrued over the last half century, have transformed our understanding of past biodiversity and its drivers on the grandest macroevolutionary scales. This has been accompanied, however, by ever increasing awareness of two pervasive biases which distort the deep history of the biosphere as portrayed by its fragmentary geological record. One of these biases, the uneven distribution of fossils through geological time, has received substantial conceptual and methodological attention. The other bias, the uneven distribution of fossils across geographic space has only recently come to the attention of palaeontologists. Its effects on the inextricably linked facets of biodiversity and biogeography are simultaneously a profound cause for concern but also a largely unexplored frontier within palaeobiology. In this thesis I draw upon advances in quantifying diversification dynamics from fossil data and Earth-system modelling in deep time to conduct spatially explicit analyses of biodiversity and biogeography and explore ways of circumventing the spatial biases posed by the geological record. I use the Triassic period as a testing ground for these approaches to unpick the dynamic interplay of environmental and biological events that set the foundations for our modern biosphere during this interval, focusing on marine biodiversity through times of biotic crisis, the rise of scleractinian coral reefs, and the radiation of archosauromorph reptiles. I demonstrate that while geographically sensitive analytical methods augmented by the spatially complete views of climate and palaeogeography from Earth-system models can provide powerful insight into past biodiversity and biogeography, substantial challenges remain before the biases inherent in the fossil record can be mitigated.

Author’s Declaration

I declare that the work in this dissertation was carried out in accordance with the requirements of the University's Regulations and Code of Practice for Research Degree Programmes and that it has not been submitted for any other academic award. Except where indicated by specific reference in the text, the work is the candidate's own work. Work done in collaboration with, or with the assistance of, others, is indicated as such. Any views expressed in the dissertation are those of the author.

Joseph T. Flannery-Sutherland

20/08/2023

Statement of Collaboration

The chapters in this thesis have involved varying degrees of collaboration ranging from data access and provision to co-author support during project development and publication. Chapters one and six are entirely the work of the author. Chapters two and three are the published results of research collaborations led by the author. Chapters four and five are unpublished research efforts conducted by the author, with project-critical data provided by several different collaborators. All individuals and their contributions are disclosed at the beginnings of each chapter. All collaborations were carefully managed to ensure that the vast majority of the work in each chapter was solely the effort of the author.

COVID-19 Statement

The themes and topics of this thesis changed substantially due to the impacts of COVID-19. Originally, my research was to focus on identifying the effects of an enigmatic mass extinction event taking place in the Late Triassic, the Carnian Pluvial Episode (CPE), on marine and terrestrial biotas. I had planned two major phases of fieldwork to collect new, high-resolution data on the spatiotemporal distributions of fossils across the event to resolve its timing and ecological impacts. The summer of 2020 had been earmarked for data collection from shallow marine strata in the Italian Dolomites, with plans to investigate terrestrial strata in South America the following year.

I was planning the first phase of fieldwork in March 2020, in the run-up to the first lockdown. Subsequent travel restrictions meant that I was unable to collect any novel data to probe the impacts of the CPE during this critical early phase of my PhD. This led me to target digitally accessible data sources, chiefly community-led fossil occurrence databases and reconstructions of palaeogeography and global climate simulations available through the University of Bristol's School of Geographical Sciences. The absence of novel data on the CPE also necessitated a shift in the scope of the research, leading me to encompass the Triassic more broadly so I could investigate the emergence of key groups of organisms whose success was potentially associated with the CPE. To achieve this, I capitalised on emerging methods for analysing diversification dynamics and biogeography using these data sources.

Reworking the project on the fly also exposed my work to developing methodological concerns surrounding the use of fossil occurrence data. This necessitated further adaptation of my research strategy to address these issues, revisions which rapidly became major components of the final thesis. By the end of the first lockdown, I had identified the rise of scleractinian corals and archosauromorph reptiles during the Triassic as suitable candidates for investigation, alongside development of a more general framework for spatially explicit analysis of diversification dynamics. My supervisor and I considered the potential for fieldwork after the first lockdown ended but concluded that the situation was not certain enough to permit viable plans to be made. This would also have necessitated waiting to the summer of 2021 to conduct fieldwork due to the restrictions imposed by winter weather and daylight hours, leading us to pivot to the set of fully computational approaches used herein.

Acknowledgements

This thesis caps four years of work which would not have been possible without the veritable hoard of people who have supported me in so many ways. Firstly, I must thank my supervisor Michael Benton for developing the original research proposal with me and securing funding for the project, taking me on as a PhD student and allowing me to freely pursue a rambling course of scientific enquiries that have eventually clung together into something coherent, and most recently for advising and guiding me in working out what to do next. Philip Donoghue and Emily Rayfield played indispensable roles as advisors in my annual progress meetings, providing healthy doses of critique and encouragement and much needed nudges to clarify the structure of the thesis. I must also thank the NERC GW4+ DTP for their generous financial support of my studentship and I'd like to pick out Sara Berrow and Emeliana Palk for all their help with the mountain of underlying administration, particularly their efforts in securing funding extensions in the aftermath of the COVID 19 pandemic. Finally, I must again thank Philip Donoghue in conjunction with Alex Dunhill for their examination of this thesis, with their comments and corrections helping to greatly improve the finished article.

My collaborators have played integral parts in facilitating my research, again particularly considering the practical limitations imposed by the pandemic. The data cleaning challenge posed by the University of Erlangen emerged as an unexpected but immensely rewarding opportunity during this time, so I am grateful to Nussaïbah Raja, Adàm Kocsis and Wolfgang Kiessling for issuing the challenge in the first place and for their subsequent support in developing my submission into a fully-fledged piece of research. I must thank Daniele Silvestro for fielding my barrages of emails regarding PyRate and its associated methods, and for his superlative contributions as both a collaborator and more recently as an academic reference. Alex Farnsworth and Dan Lunt were kind enough to give me unfettered access to their stunning reconstructions of global climate and palaeogeography, which are a leitmotif of this thesis, and their expertise in the intricacies of Earth-system modelling have been invaluable. Armin Elsler is also thanked for providing me with access to his supertree of early tetrapods and Jens Landwehrs for his assistance with additional climate model data. Invaluable academic and technical support has come from several places besides my project collaborators. I would like to acknowledge Ciara O'Donovan and Andrew Meade for their patient tutelage and assistance with BayesTraits, the high-performance computing team at the University of Bristol for all

their help in navigating the dark world of the Blue Pebble terminal, Roger Close and Lewis Jones whose comments during peer review were integral to improving the quality of my research, and January Weiner, Sara Varela, and Damaris Zurell for their instruction during the two excellent Physalia training courses I attended during my PhD.

It would be an understatement to say that the thesis has been stressful at times, so I am especially grateful to all the people who have kept my spirits buoyant and my wellbeing in check, within academia and without (although plenty of you have fallen into both camps at times). To avoid this page looking like the credits of a film and the risk of missing anyone out, I shall refer to groups as well as individuals, but know that each one of you has been so, so important in making my PhD experience a wonderful four years, even with its ups and downs. Within academia, I must thank the 'Tea at Three' crew who inhabit the first floor of the Life Sciences Building for many hours of morale boosting chitchat in the breakout spaces, followed by many hours more at the pub, and the camaraderie and support given by Ilja Kogan, Tamas Kapitany and Meerim Derbisheva during my adventures in Kyrgyzstan. Beyond academia, I want to give shoutouts to the 'Kickboxing Lot', who I've always been able to rely on for a good laugh, an indispensable antidote to work, and by dint of the sport a good scrap; and everyone from the 'Palaeo Group' who have been there for me, whether you started the Bristol adventure with me eight years ago or were unwittingly roped into our collective at a later stage – cheers buds.

Finally, those closest to me, my parents and Hollie. You reminded me of my best times during the worst. You have my gratitude always.

Table of Contents

Front Matter

Chapter One

Chapter Two

Chapter Three

Chapter Four

Chapter Five

Chapter Six

Back Matter

List of Figures and Tables

Figures

Tables

Chapter One

General Introduction

In this introductory chapter, I outline the conceptual scope of the thesis and the structure of its succeeding research chapters then survey the corresponding suite of methods upon which it draws, chiefly techniques used to investigate patterns of biodiversity and biogeography in deep time, with comments on their limitations and prospects. I review the issues faced by approaches to quantifying deep time diversification dynamics based on phylogenetic versus fossil data. I then examine how integrated models of diversification and fossil sampling bridge the gaps between these methodological poles, forming the basis of the Bayesian techniques used in this thesis for inferring these processes. Finally, I consider the issue of spatial sampling bias in the geological record, outlining its consequences for inference of past biodiversity, the diversification process, and its drivers, along with the some of the possible solutions to this major hurdle facing palaeobiologists.

Author contributions: This chapter was written and developed by J. T. Flannery-Sutherland. Comments were provided on a draft by M. J. Benton. The author contributed all of the work in this review chapter.

1.1 Scope of the thesis

A detailed picture of the distribution of biodiversity across space and time forms the basis for investigating its underlying drivers (Close et al., 2020a; Benson et al., 2021). Fossils provide the only empirical evidence of past biodiversity and so the only direct means of assessing the factors responsible for its spatiotemporal structure (Marshall 2017). In turn, fossil occurrence databases, which record the taxonomic affinities, present-day locations, depositional settings, and stratigraphic ages of millions of fossil observations, have enabled the palaeobiological community to capitalise on decades of accrued knowledge of the fossil record (Uhen et al., 2013; Flannery Sutherland et al., 2022) to query its patterns and genetic processes.

Successive compilations of the fossil record, even since the early days of palaeobiological research, have identified broad trends in Phanerozoic diversity which have remained robust to repeated scrutiny (Fig. 1.1), leading palaeobiologists to seek general models for these phenomena (Raup 1972; Sepkoski et al., 1981; Sepkoski 2001; Alroy et al., 2008; Cermeño et al., 2022). Palaeobiologists have questioned whether life has self-regulated its own abundance through geological time by processes of competition and predation, or whether taxonomic richness across a wide variety of clades has been controlled and curtailed by Earth-system regulators, in particular the interlinked drivers of global tectonics and climate (Raup 1972; Sepkoski 1981; Benton 1997; Barnosky 2001; Cermeño et al., 2022). Critical to this debate was the recognition that mass extinction events appeared to separate successive biotic regimes in Earth history, leading to the notion that catastrophic diversity loss and biological innovation during subsequent phases of recovery have progressively promoted Earth's capacity to support biodiversity (Sepkoski 1981; Jablonski 2008; Krug and Jablonski 2012; Muscente et al., 2018; Rojas et al., 2021). Consequently, high quality fossil occurrence data is critical to investigate the ramifications of these pivotal events through geological time. Advancements in radiometric and stratigraphic dating (e.g., Fan et al., 2020) have continually refined estimates of the tempo of waxing and waning fossil biodiversity, permitting precise identification of the coupled episodes of mass extinction and collapse, and recovery and radiation responsible for transforming the biosphere through geological time. Concurrently, increasingly sophisticated palaeogeographic models have enabled the richness of the fossil record to be placed within its original spatial context and facilitated numerical modelling of Earth-system processes (Wright et al., 2013), complementing the wealth of information on past environments available from the broader geological record.

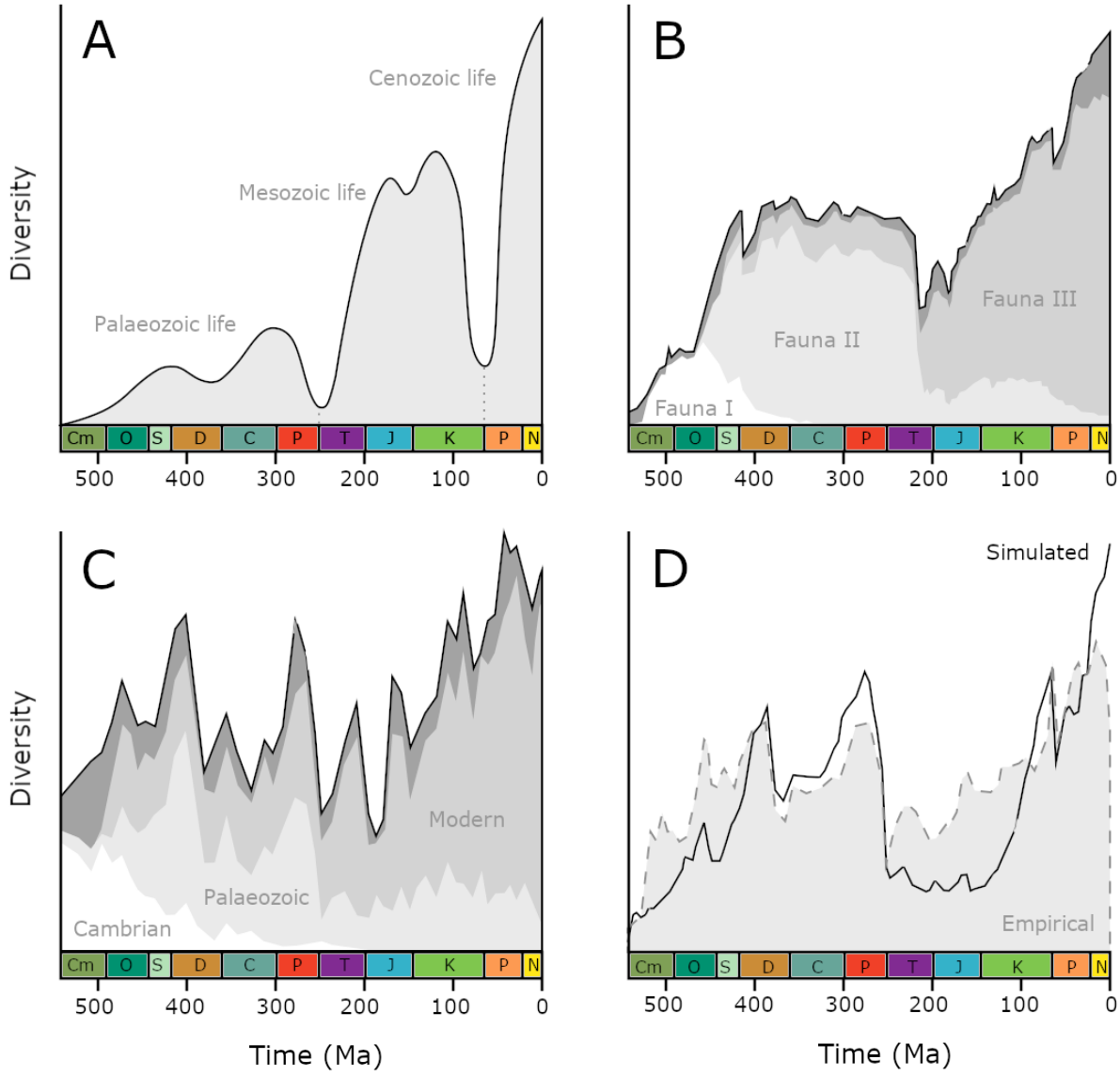


Fig. 1.1. A view of Phanerozoic diversity through palaeontological history. (A). Diversity in the British fossil record from Phillips (1860), the earliest comprehensive curve in the palaeontological literature. **(B).** Consensus Phanerozoic diversity from Sepkoski et al. (1981), the product of decades of accrued data and debate. **(C).** Sampling standardised Phanerozoic diversity from Alroy (2010), using the Paleobiology Database and modern methods for mitigating sampling bias through geological time. **(D).** Simulated Phanerozoic diversity from Cermeño et al. (2022), using Earth-system modelling and palaeogeographic constraints. All curves were redrawn from their sources.

Despite advances in the availability of fossil occurrence data, the precision of stratigraphic dating and the detail of palaeogeographic models, determining past patterns of biodiversity and its underlying drivers have not been simple endeavours. Growing knowledge of the spatiotemporal distribution of the fossil record has been accompanied by the understanding that this distribution does not readily reflect original patterns of biodiversity but

is instead heavily distorted by a multitude of sampling biases (Raup 1972; Alroy et al., 2001, 2008; Close et al., 2020a, 2020b, Benson et al., 2021). The fossil record is an incomplete, imperfect sample of original biodiversity and this imperfection distorts our view of the biosphere through time and across space. More recently, palaeontologists have considered the importance of biogeography when interpreting patterns of past species richness. Biodiversity is expected to differ between different parts of the globe due to corresponding spatial variation in its underlying drivers, yet this fact has historically been overlooked by palaeontologists who have instead sought to derive views of changing taxonomic richness on global scales (e.g., Valentine 1969; Sepkoski et al., 1981; Alroy et al., 2008; Zaffos et al., 2017). Not only do analyses which treat the fossil record as singular, homogenous dataset fall prey to the myriad of sampling biases, they have the undesirable effect of glossing over the spatial nuances of how individual clades, biotas and ecosystems has been shaped by mass extinctions, evolutionary radiations, and the equally spatially varied influences of palaeogeography and climate (Benson et al., 2021).

While growth of fossil occurrence databases will naturally enhance our understanding of some parts of the fossil record, more data will not necessarily remedy existing sampling biases and may even enhance the magnitude of the discrepancies between different geological intervals and parts of the globe (Raja et al., 2022). Continued database growth creates further issues which palaeobiologists must also overcome, particularly as the scale of fossil occurrence databases tips them ever further into the realm of 'Big Data' (Allmon et al., 2018; Marshall et al., 2018). Firstly, database growth will naturally promote accumulation of errors, whether present in the original literature, arising during human imputation of their records (or more recently during data extraction by machine learning methods), or due to updated scientific knowledge or convention rendering older data erroneous. Palaeontological studies often devote significant manual effort to eliminating poor quality records prior to analysis, but such efforts will become increasingly unfeasible as the pool of available data grows (Flannery Sutherland et al., 2018). Secondly, tools and methods must be developed or adapted to accommodate the wealth of novel data available for analysis, including its uncertainties and biases, and the increased computational burden their implementations may impose.

No matter how much data palaeontologists accumulate, or how sophisticated their analytical methods become, the fossil record will always remain an incomplete sample of past diversity. The accessible extent of the sedimentary rock record sets a hard limit for where

fossils can theoretically be found (Husson and Peters 2018; Ye and Peters 2023); for parts of the globe occupied by igneous or metamorphic formations, or those where the bedrock is obscured by ice, sand, or soil, we will never be able to collect empirical fossil data on their past ecosystems. Overcoming this fundamental limitation of the fossil record is crucial to permit inference of biodiversity and biogeographic history in geological intervals and areas of the globe where direct fossil evidence is unavailable. This thesis tackles elements of the issues outlined above, utilising methodologies developed herein, or adopted from advances in modelling diversification dynamics, biogeography, and Earth-system processes. For the remainder of this first chapter, I highlight how fossil occurrence data remain indispensable to macroevolutionary enquiry despite the limitations of the fossil record, and review key methods used to infer diversification dynamics and biogeography in deep time. I consider the major hurdles these methods have faced, the outstanding challenges posed by spatial heterogeneity in the fossil record, and how the integration of diversification dynamics, biogeography and Earth-system modelling may help to overcome these respective issues. This review thus provides a technical context for my subsequent chapters where I apply a suite of these techniques to three empirical analyses of the Triassic fossil record.

The Triassic is a critical period in geological history during which the roots of our modern biosphere were established, bounded by two of the most catastrophic mass extinctions in geological history, the end-Permian and end-Triassic mass extinctions (Chen and Benton 2012; Wignall and Atkinson 2020). While this period was highly dissimilar in multiple ecological, climatological, and geographic respects to the present, critical relationships between the biosphere and the Earth-system were formed during this interval which permanently transformed the capacity of ecosystems to support biodiversity, namely stabilisation of long-term carbon cycling by the evolution of calcareous nannoplankton and increased delivery of energy to the base of marine ecosystems by the evolution of large-celled phytoplankton (Preto et al., 2012; Knoll and Follows 2016). Furthermore, phylogenetic and fossil evidence indicate the origins of a diverse range of clades in the Triassic which would continue to characterise terrestrial and marine ecosystems through the remainder of the Mesozoic and to the present day, including mammals, dinosaurs, lizards, turtles, lissamphibians, scleractinian corals and teleost fish (Dal Corso et al., 2020; Wu and Benton 2022). Cultivation of fossil occurrence databases and improvements to the stratigraphic resolutions of fossil, lithological and isotopic records through the Triassic have enabled the construction of a detailed narrative for the tempo

and mode of life's recovery in the wake of the end-Permian mass extinction (EPME) and its proliferation during subsequent phases of environmental stability and upheaval. This narrative shows that taxonomic components and ecological structuring within the Triassic biosphere that we would recognise as biologically modern did not originate simultaneously during the initial recovery of marine and terrestrial ecosystems, but in a more episodic pattern set against a dramatic backdrop of broader climatic, tectonic, and ecological events (Chen and Benton 2012; Song et al., 2018). Study of the episodic nature of recovery and proliferation during the Triassic is ongoing, but its spatial components remain understudied by comparison, with our view of these processes often being limited to the geographically sporadic snapshots provided by the geological record. Consequently, the Triassic provides an ideal testing ground for new approaches to diversification dynamics and biogeography, from which we can understand the processes and events that established the course of the biosphere for the last 250 million years.

1.2 Inferring past diversity dynamics

The fossil record, displaying taxa with a clear lack of living representatives led early palaeontologists to develop the concept of species demise, extinction, while the formulation of evolutionary theory provided the corresponding mechanism for their origination, speciation (Rudwick 1997; Gould 2002). The balance between speciation and extinction rates, diversification, then determines changing diversity through time. An array of methods for estimating these dynamics in the geological past exist and can be broadly divided into a spectrum ranging from those that consider the diversification process responsible for a phylogenetically resolved set of extant taxa (the neontological pole; see 1.2.1) to those that analyse diversification from sets of stratigraphically dated fossil occurrence data without any notion of their phylogenetic relationships (the palaeontological pole; see 1.2.2), or in the parlance of Sanderson and Donoghue (1996) 'tree-based' versus 'time-based'. Despite using data ultimately derived from the same diversification process, these poles have historically operated separately, frequently producing strongly discrepant estimates, and encountering their own sets of caveats and limitations (see 1.2.3). This has led evolutionary biologists to seek solutions that reconcile empirical diversification signals recovered from the fossil record with phylogenetic signals inferred from extant taxa and their evolutionary relationships (see 1.3).

1.2.1 Neontological diversification estimation

Diversity dynamics may be modelled as a birth-death process (Kendall 1948), where lineages arise via speciation (births) and terminate via extinction (deaths). The topology of a phylogenetic tree contains information about speciation rate variation within a clade; when branch lengths are also available, a phylogeny of extant taxa can subsequently be used to quantify speciation and extinction rates through time by maximum likelihood estimation of the birth-death process (Stadler 2013). Recognition of this property enabled evolutionary biologists to capitalise on the emergence of time-calibrated molecular phylogenies of extant taxa to infer diversification processes through geological time (Hey 1992; Nee et al., 1992a, 1992b, Harvey et al., 1994), information that was previously the purview of palaeobiologists studying the fossil record.

Phylogenetic estimation methods rapidly evolved to account for the common scenario of incomplete sampling of a clade's true richness inducing systematic errors in inferred rates (Yang and Rannala 1997; Stadler 2009), to model rate variation through time and between lineages (reviewed by Morlon 2014), and to incorporate uncertainties in phylogenetic divergence times using Bayesian approaches (BayesRate; Silvestro et al., 2011). Under the assumption of constant rates within the birth-death process, the diversification rate for a clade can alternatively be estimated from its root age and its extant diversity, the method-of-moments estimator (Magallon and Sanderson, 2001). This estimator has proven immensely popular amongst evolutionary biologists as it eliminates the need for a phylogenetic hypothesis beyond monophyly and is computationally lightweight compared to more complex methods. Method-of-moments has frequently been criticised given frequent empirical violations of its core assumption, spurring the development of estimators that can account for temporal and lineage-wise rate variation, and divergence time uncertainties (BAMM; Rabosky 2014), but simulations have demonstrated its robustness and superior performance even when birth-death rates are heterogenous (Meyer and Wiens 2017).

Two linked characteristics of all methods within the neontological pole is that diversity through time is a fixed property arising from the total number of sampled lineages and their branch lengths, and that the methodological goal is to infer the rates responsible for that diversity using the birth-death process. These methods consequently share the general advantage that a phylogenetically constrained history of extant taxa enables birth-death models of their diversification process to extend to areas of biotic history that fall beyond the bounds

of the fossil record. The multi-billion-year span of unicellular eukaryotic and prokaryotic evolutionary history through the Archaean and Proterozoic, or even multicellular clades with sparse fossil records in the otherwise well-sampled eon of 'evident life', the Phanerozoic, can still be accessed through the phylogenetic relationships of their extant descendants (e.g., Betts et al., 2018).

1.2.2 Palaeontological diversification estimation

The palaeontological pole of methods for estimating diversity uses the fossil record as a sample of taxon richness through geological time. Phylogenetic hypotheses are not required, nor do the sampled taxa need to form a monophyletic clade or represent extant lineages, enabling estimation of diversity dynamics for entire biotas or through any range of geological intervals provided fossil data are available. This flexibility is advantageous given that phylogenetic hypotheses are unavailable for most fossil taxa. Early palaeobiologists recognised that the fossil record, consisting of an ultimately monophyletic set of lineages with discrete stratigraphic durations bounded by the ages of their oldest and youngest fossils, could be described using the birth-death process (Raup et al., 1973; Raup 1985). Incomplete stratigraphic sampling, however, means that the oldest and youngest fossils of a lineage likely do not record its true duration, producing tardy times of speciation and premature times of extinction – the Signor Lipps effect (Signor and Lipps 1982). Further, the fossil record is not a random sample of past diversity but instead an uneven, incomplete sample of its distribution through geological time. Consequently, both diversity and its underlying diversification rate are unknown quantities, leading palaeontologists to develop separate methods for their estimation which correct for the inadequacies of the fossil record in contrast to the neontological pole where only rates are considered.

One class of methods for diversity estimation utilises a geological proxy for sampling intensity to correct its biasing effect on fossil diversity. Phillips (1860) normalised fossil diversity by the volume of rock through geological time, leading to discovery of broad trends in Phanerozoic fossil diversity that are still recognised today. Progressively larger compilations of fossil taxon ranges repeatedly recovered these broad trends (Newell 1956; Valentine 1969; Sepkoski et al. 1981; Sepkoski 1993), but palaeobiologists noted that diversity still showed a strong correspondence with the age-wise distribution of exposed lithologies globally – the rock record bias hypothesis (Raup 1972; Smith 2001; McGowan and Smith 2011). Palaeobiologists subsequently explored the use of geological sampling proxies to model true diversity, utilising

either the residuals from a linear regression directly between the proxy and diversity (Raup 1976) or by independently ordering both variables prior to model fitting so that the proxy perfectly predicts sampled diversity, and the residuals predict true diversity (Smith and McGowan 2007; Lloyd 2012). In both cases, the model residuals cannot be attributed to sampling effort and instead are taken to record the biological signal of changing diversity. The proxy-based approach has been used to study diversification histories in a wide range of clades with a variety of sampling proxies, including the number of named geological formations, rock volume or outcrop area. Later authors, however, have demonstrated that such modelling residuals is statistically unjustified, producing misleading diversity trends (Brocklehurst 2015; Sakamoto et al. 2016), and geological proxies are typically not independent, reliable measures of sampling effort (Dunhill et al., 2017). Instead, variation in the intensity of fossil sampling and an apparent geological proxy can instead be simultaneously driven by a third factor like sea level (the common cause hypothesis; Peters 2005, 2006; Hannisdal and Peters 2011; Peters and Heim 2012) or drive one another through mutual discovery (the redundancy hypothesis; Benton et al., 2011, 2013; Dunhill et al., 2014). Consequently, proxy-based methods have been abandoned.

The other class of approaches for estimating diversity from fossil data utilise sample size standardisation. These take inspiration from analogous problems encountered by ecologists examining modern taxa, where differential human sampling effort may distort patterns of relative richness between different areas or environments. By exploiting the relationship between the rate at which richness changes with sample size (the collector curve) unevenly sampled populations can be standardised to equal degrees of sampling effort, yielding an unbiased estimate of their relative diversity. Raup (1975) was again the first to apply this method of correction to the fossil record, although his approach utilised the changing discovery rate between higher taxonomic groups and their constituent species or genera. The focus on compilation of taxon stratigraphic ranges limited standardisation methods to Raup's approach until the emergence of the first fossil occurrence databases, most notably the Paleobiology Database. This effort spurred sustained interest in identifying and eliminating sampling biases in the fossil record, including proxy-based approaches, but also the application of standardisation methods directly to fossil occurrences, the fundamental palaeontological unit of sampling (Alroy et al., 2001, 2008). Standardisation uses either interpolation to move 'down' the collector curve, or extrapolation to move 'up' or even beyond the limit of the collector

curve to achieve even sampling between populations or geological intervals. Interpolative and extrapolative approaches were extensively reviewed and tested by Close et al. (2018), who demonstrated that shareholder quorum subsampling (SQS; Alroy 2010a, 2010b) is the most robust method for diversity estimation from fossil occurrence data. Supported by the abundance of fossil occurrences within the Paleobiology Database, SQS was adopted as the gold standard for diversity estimation, even before the abandonment of proxy-based approaches.

Diversification rate estimation methods are more numerous and have enjoyed a greater pedigree than diversity estimators due to the longstanding availability of the requisite taxon stratigraphic range data and interest in quantifying the balance of speciation and extinction processes rather than simply their net effect of changing taxon richness. Older methods based on tallying taxon originations and extinctions in discrete time intervals suffer from a series of confounding methodological issues relating to the sizes of the stratigraphic windows across which lineage durations are observed and show sensitivity to incomplete sampling and the Signor Lipps effect (Foote 1994, 2000a, 2000b, 2005). Palaeobiologists recognised that modelling taxonomic turnover using exponential decay was a more realistic representation of diversification under the birth-death process, enabling calculation of rates using time-continuous decay equations (Raup 1985; Alroy 2000). The subsequent emergence of the Paleobiology Database permitted the development of decay equations that utilise occurrences rather than ranges to describe the exponential decay and turnover of cohorts of taxa, rendering the estimated rates more robust to incomplete sampling (Alroy 2008). Occurrence-based rate estimators have been further refined to account for cases where turnover is exceptionally high or sampling extremely poor (Alroy 2014a, 2015). In general, two rate estimators have seen regular use by palaeontologists, the boundary crosser method of Foote (2001) and the three-timer metrics of Alroy (2008), These two methods have received a degree of comparison (Allen et al., 2023) although there is currently no study that has simultaneously compared the efficacies of all traditional approaches using simulations across a wide range of diversification and sampling scenarios.

1.2.3 Rocks versus clocks and richness versus rates

While neontological methods do not use fossil data themselves, they all require time-scaled phylogenies or an estimate of the root age, necessitating fossil data for empirical constraint of absolute clade age or for node or tip calibration (Stadler, 2013; Heath et al., 2014). Consequently, the availability and quality of the data is of utmost importance due to the high

sensitivity of calibration methods to the chosen node priors (Marshall 2019; Barido-Sotani et al., 2020). The fossil record further demonstrates that assumptions of rate constancy through time and between lineages in a phylogeny are often unrealistic, with palaeontological estimates of diversification rates frequently diverging from neontological estimates. Without fossil taxa, extant phylogenies will always display positive diversification rates and increasing diversity, with the corresponding effect of inducing unrealistically low extinction rates (Quental and Marshall 2010). This is in stark contrast to exceptionally high extinction rates incurred during mass extinction events and intervals when diversification rates are genuinely negative, producing the waxing and waning pattern of diversity displayed by the empirical fossil record.

Some authors have gone as far as to claim that extinction rates cannot be reliably estimated in the absence of fossil data (Rabosky 2008). While attempts have been made to model periods of negative diversification from extant phylogenies with some success (Morlon et al., 2011), this remains a challenging endeavour (Quental and Marshall 2009; Liow et al., 2010; Morlon 2014). More recently it has been demonstrated that inference of zero extinction rate is mathematically unsupported, that multiple combinations of speciation and extinction rates are congruent with the same time-calibrated phylogeny, and that relative support for different combinations within this class of congruent solutions cannot be distinguished using the phylogeny alone (i.e., the true solution within the congruence class is non-identifiable, an issue extending to the method-of-moments estimator; Louca and Pennell 2020, 2021; Rabosky and Benson 2021). In some cases, qualitative inferences about rate variation may still be drawn if all solutions within the congruence class indicate the same result (Kopperud et al., 2023), but otherwise fossil data remain an indispensable source of information regarding diversification processes in deep time despite the advances achieved using phylogenies of extant taxa (Quental and Marshall 2010; Morlon et al., 2022; Wright et al., 2022).

While the palaeontological pole provides insight into past diversification processes that the neontological pole fails to recover, its methods also bear limitations and caveats. SQS will underestimate diversity for species-rich intervals when sampling is poor and is restricted to a range of coverage determined by the taxon frequencies within the dataset to which it is applied (Close et al., 2018), potentially undermining estimation of sampling-corrected diversity through time if insufficient occurrences are unavailable for all intervals through the study duration. This is alleviated somewhat by the implementation of Chao and Jost (2012) which combines interpolation and extrapolation in an analytical rather than an algorithmic solution,

but the diversity estimates may still be unreliable when sampling is poor and cannot be achieved if occurrences are entirely absent within an interval. Similarly, simulation-based exploration of rate metric accuracy has shown that the popular boundary-crosser and three timer metrics may be highly unreliable under conditions of low turnover and poor sampling (Warnock et al., 2020).

More alarmingly, the palaeontological pole suffers from a methodological disconnect between estimation of richness and rates which have received very little discussion in the literature, despite longstanding recognition of fossil taxa as manifestations of a birth-death process. In a face value reading of fossil data, achieved by discarding any notion of incomplete sampling of true taxon richness and their stratigraphic durations, tallying the overlapping stratigraphic ranges of fossil taxa through time yields a range-through-time (RTT) diversity trajectory with extinction and speciation rates that can be readily calculated from the set of lineage durations. The empirical reality of incomplete sampling, however, necessitates a formal concept of sampling rate, the number of fossils recovered per lineage per time unit, to quantify the quality of sampling. Any sampling rate naturally induces a non-random sampling bias against shorter stratigraphic durations, whether those durations are divisions of the geological timescale or the lifespans of lineages, and increased risk of range truncation (Silvestro et al., 2014). Consequently, as sampling rate decreases, leading to more and more lineages failing to be sampled in all the geological intervals they originally spanned (including cases where lineages fail to be sampled altogether), the RTT diversity curve in the fossil record progressively decouples from true diversity (Fig. 1.2).

The most popular rate metrics and diversity estimators indirectly consider sampling rate by examining fossil occurrences rather than whole taxon ranges. By using the relationship between the sampling intensity in a bin and sampled-in-bin (SIB) diversity (the number of taxonomically unique fossil occurrences observed within it; Fig. 1.2), sampling-corrected diversity can be estimated. Three-timer rates and their relatives also use occurrences, although they consider cohorts of SIB taxa across several adjacent time bins rather than just one. A SIB diversity curve, however, will often be incongruent with its already distorted parent RTT curve due to incomplete sampling throughout lineage durations, and so not map onto the speciation and extinction rates implied by the latter. Further, sampling-corrected diversity curves or rate estimates will not necessarily be compatible with either their uncorrected parent RTT curve or with each other, nor can this latter issue simply be addressed by using one component to infer

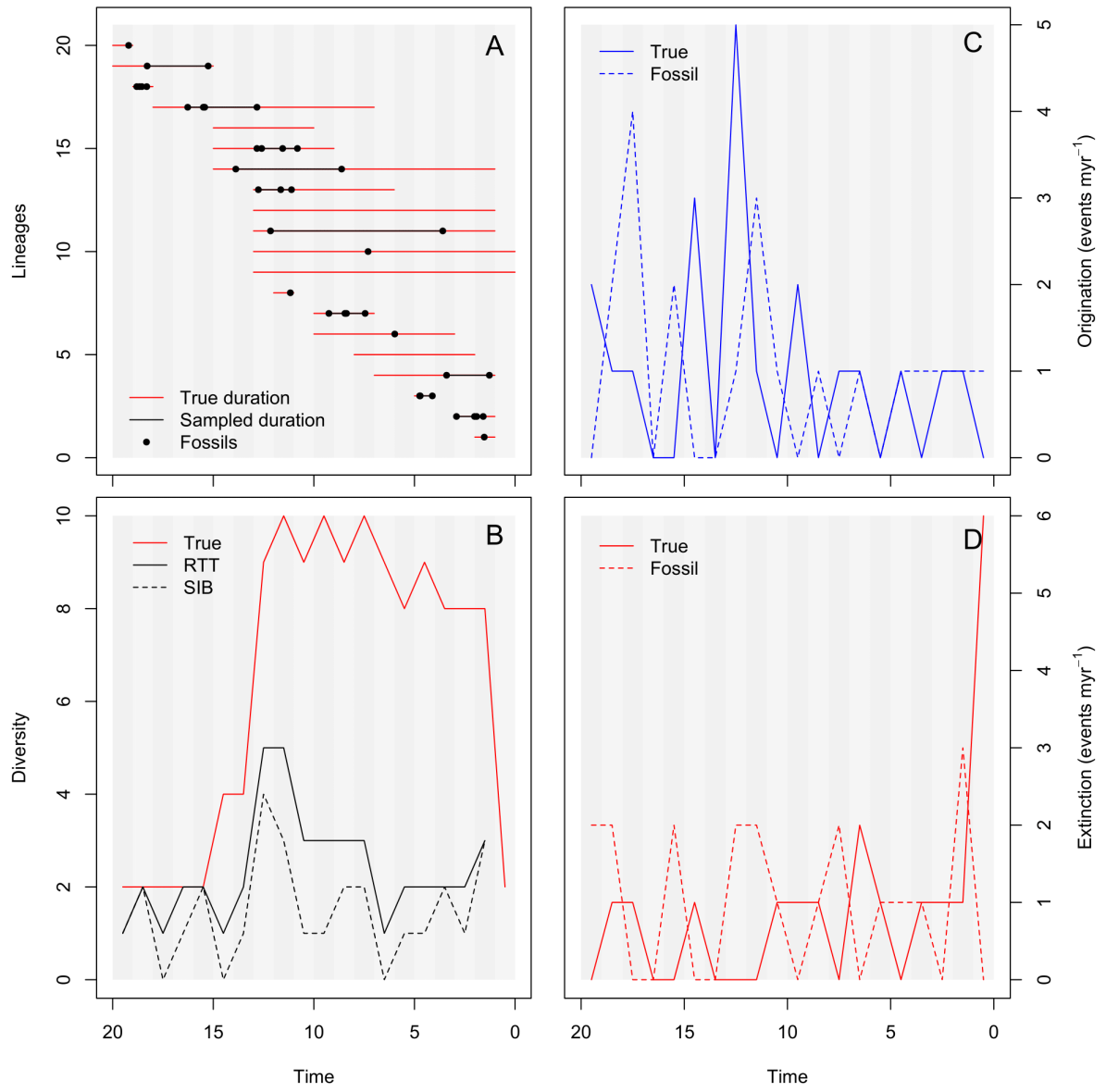


Fig. 1.2. The incongruence of palaeontological estimates of richness and rates. (A). Hypothetical set of stratigraphic ranges with varying durations and degrees of fossil sampling. **(B).** True diversity resulting from the stratigraphic ranges in part A, along with the range-though-time (RTT) and sampled-in-bin (SIB) diversities implied by the fossil samples in part A. **(C).** True versus fossil lineage origination rates based on part A. **(D).** True versus fossil lineage extinction rates based on part A. All times are in arbitrary units.

the other. It is trivial to calculate the diversification rate responsible for any diversity curve, but there may be multiple compatible combinations of extinction and speciation rates, analogous to the problem of non-identifiability faced by neontological methods. Conversely, net diversification resulting from the balance of speciation and extinction rates cannot be used to calculate a diversity trajectory in isolation, requiring at least one absolute measurement of

past diversity, whether supplied as a raw observation or a corrected estimate. Even though simulation-based studies have confirmed that the separate sets of methods can perform well individually (Close et al., 2018; Warnock et al., 2020), the continued incongruence between palaeontological estimates of rates and richness is conceptually problematic given that they reflect aspects of the same birth-death process. Not only should estimates of rates and richness be congruent, but it is also preferable to have methods that make full use of the available data on the diversification process, yet available methods fundamentally fail to so do by their use of SIB rather than RTT data.

1.3 Uniting fossils and phylogenies in diversity dynamics

The palaeontological pole is plagued by methods which do not necessarily produce estimates of diversity and diversification rate that are compatible within the birth-death process they aim to estimate (richness versus rates). Conversely, the neontological pole has suffered from its inability to reproduce the realities of diversification histories recorded in the fossil record despite predication of its methods on the mathematics of the birth-death process (rocks versus clocks). The solution to these problems is the development of methods which treat the generation of biodiversity and its incomplete sample of fossils as components of the same underlying process, offering the potential for fully congruent estimation of sampling-corrected diversity, speciation rates and extinction rates from a set of fossil occurrence data.

Part of the discrepancy between rate estimates from both poles arises from fundamental differences in how evolutionary biologists have treated speciation within birth-death models versus how palaeontologists have treated speciation amongst fossil morphospecies (Foote 1996; Ezard et al., 2012; Bapst 2013a; Silvestro et al., 2018). In the former, lineages give birth to one additional lineage (cladogenesis via budding), then lineage deaths occur independently of any births. However, palaeontologists recognise that an ancestral lineage may give birth to two descendant lineages and go extinct by this replacement (cladogenesis via bifurcation), or the continuous replacement of one lineage by another, leading to the extinction of the former by the birth of the latter (anagenesis). Extension of the standard birth-death process to incorporate these scenarios (the birth-death-chronospecies model) achieves more concordant rate estimates between the phylogeny and the fossil record of a given clade, subject to the quality of its fossil record and the present-day sampling of its diversity (Silvestro et al., 2018).

Consequently, continued discrepancies between fossil and phylogenetic data can be informative on a clade's speciation mode (Silvestro et al., 2018).

While comparing fossil and phylogenetic data within a coherent framework is useful, greater advances have come from integrating their signals into a unified picture of the diversification and fossilisation processes. These birth-death-sampling models are not just helping to overcome the individual issues faced by the palaeontological and neontological poles, but also to bridge the gap between the two, aided by the development of Bayesian methods to incorporate various sources of uncertainty in phylogenetic and fossil data, and by the increases in computing power required for application of these complex models to ever larger data sets.

1.3.1 The fossilised birth-death model

The mathematical formulation of serial sampling through lineages generated by a birth-death process (independently derived by Stadler 2010 and Didier et al., 2012) constitutes the fossilised birth-death (FBD) model (Fig. 1.3; Heath et al., 2014). In a palaeobiological context (given its alternative applications in epidemiology), the FBD model combines a birth-death process describing the speciation and extinction rates driving the diversification of lineages within an underlying phylogeny, a Poisson process describing the sampling rate of fossils from each lineage through their stratigraphic durations, and the probability of sampling of its extant tips in the present. This yields a coherent description of tree shape and the placement of fossils along its branches, allowing the FBD model to be used as a prior for phylogenies with any combination of extant and extinct taxa, incorporating purely neontological (zero preservation rate) and purely palaeontological phylogenies (total extinction before the present) as special cases.

Since its formulation, the FBD model has evolved to incorporate non-uniform, incomplete sampling in the present, and speciation, extinction, and sampling rate variation through time (the skyline FBD model; Stadler et al., 2013; Gavryushkina et al., 2014, 2017; Zhang et al., 2016) and between lineages (multistate FBD, fossil BAMM; Kuhnert et al., 2016; Mitchell et al., 2019). This capacity for flexible, realistic modelling of complex diversification and sampling scenarios in a single statistical framework has played a pivotal role in bridging the gap between palaeontological and neontological data. Integration of modern and fossil taxa within phylogenies by combining molecular and morphological data (termed the

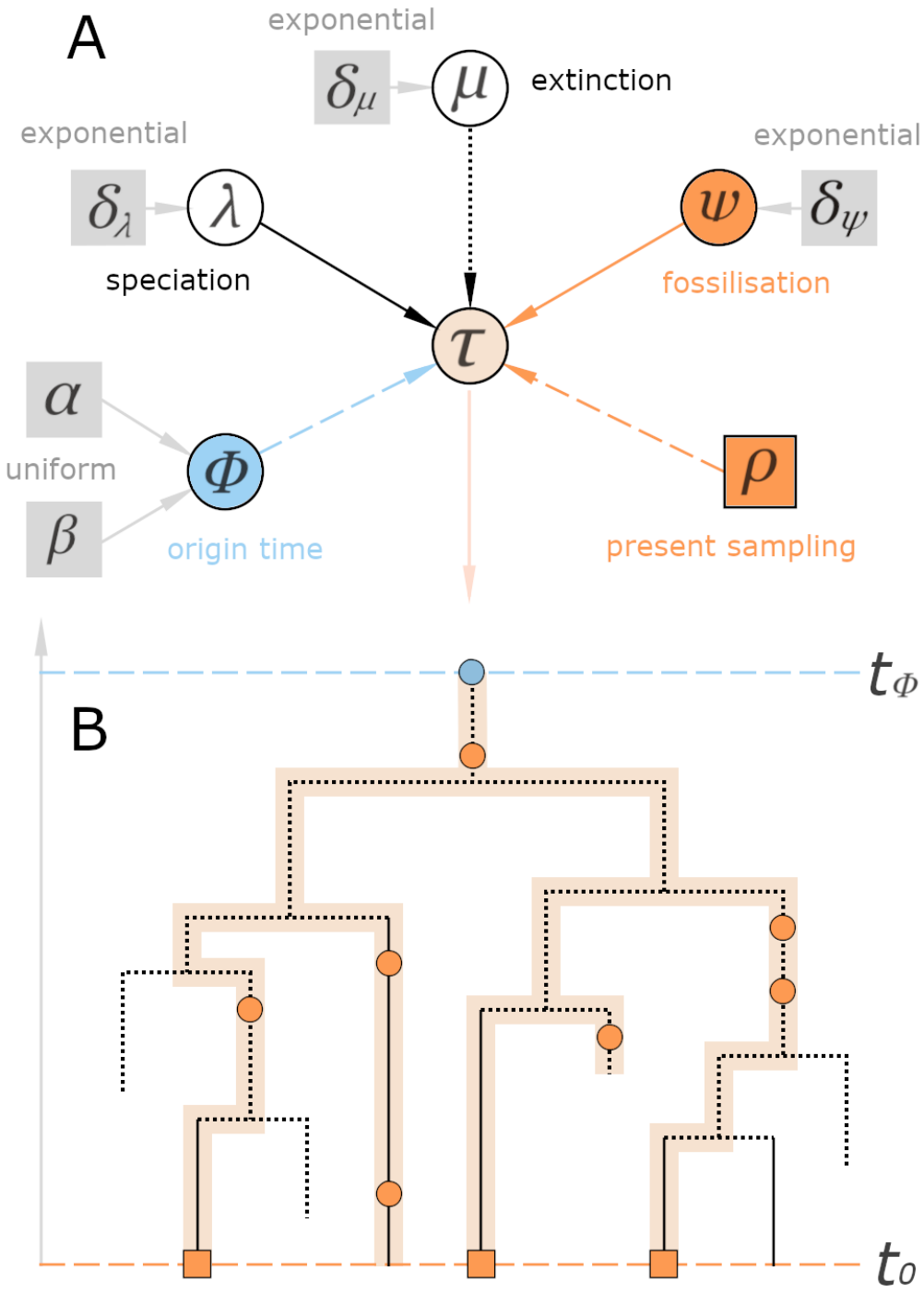


Fig. 1.3. Illustration of the fossilised birth-death model. The FBD model provides a unified description of the processes responsible for the observed topology of a phylogeny, along with its sampled record (fossils in the palaeontological case). **(A).** Components of the FBD model and their typical priors. **(B).** Schematic realisation of the FBD model, highlighting its relationships to the components in A. After Gavryushkina and Zhang (2016) and Heath et al. (2019).

'simultaneous', or more recently the 'total evidence' approach to phylogenetic inference; Kluge 1989; Nixon and Carpenter 1996), has been of interest to evolutionary biologists for several decades due to the potential for fossil lineages to clarify the evolutionary relationships between extant lineages (Donoghue et al., 1989; Wiens 2009; Legg et al., 2013; Miyashita et al., 2019). The importance of this endeavour was further emphasised by the need to incorporate fossil lineages into extant phylogenies to enable accurate inference of extinction and speciation rates under the birth-death process (Quental and Marshall 2010). Estimating reliable divergence times (and so branch lengths) in these trees proved challenging, however, as the *a priori* requirement by node calibration methods of fixed positions for fossil tips meant their phylogenetic and stratigraphic uncertainties could not be accounted for within the final topology (O'Reilly et al., 2015). The total-evidence approach was consequently extended to permit phylogenetic analysis using both molecular and morphological data whilst simultaneously allowing the ages of the fossil tips to calibrate divergence times of their associated nodes (hence the associated terms 'tip-dating' or 'tip-calibration'; Pyron 2011; Ronquist et al., 2012). These early efforts, however, were limited by the use of simple tree priors that did not account for the impacts of fossilisation and sampling of extant taxa on tree topology, problems consequently addressed by deployment of the FBD model as a prior within this framework (Heath et al., 2014; Zhang et al., 2016).

By enabling integration of fossil and extant taxa in a total evidence framework, use of the FBD model and its extensions (reviewed by MacPherson et al., 2021) has flourished amongst evolutionary biologists and palaeobiologists alike for inference of phylogenies, diversification histories and their drivers (Lopez-Antonanzas et al., 2022; Wright et al., 2022). While this integration has been largely driven by desire to improve the utility of phylogenies and birth-death models by practitioners of the neontological pole, the advances presented by the FBD model have also benefitted the palaeontological pole. The FBD model has been used for *post-hoc* time calibration of phylogenies consisting of purely fossil taxa, the cal3 method of Bapst (2013b, 2014). Later authors have advocated for total-evidence analysis of topology and divergence times using the original morphological and stratigraphic data underlying such phylogenies (Bapst and Hopkins 2017). Divergence times can still be optimised under the FBD model alone, however, by using blank morphological matrices in lieu of a true character matrix (Zhang et al., 2016) and this clockless tip-dating method is particularly useful for calibration of fossil supertrees where morphological matrices are unavailable (Lloyd and Slater 2021; Wright et al., 2022).

1.3.2 Birth-death-sampling models for fossil occurrence data

More pertinently to the palaeontological pole, birth-death-sampling approaches are more frequently being applied to fossil stratigraphic ranges and occurrences, offering a new lease of life for the decades of data collated within fossil occurrence databases. Some of these methods still incorporate a phylogeny, restricting their application to clades with available trees but consequently maximising the use of all palaeontological and neontological data. Others are phylogeny-free, or more accurately treat the phylogeny as an unknown parameter, retaining their flexible application to wherever sufficient fossil occurrence data is available. The earliest and most widely used of these methods, which emerged prior to widespread adoption of the FBD model, is PyRate (Silvestro et al., 2014a, 2014b, 2019) on which this thesis draws substantially.

PyRate uses the Markov Chain Monte Carlo (MCMC) method to jointly estimate fossil sampling rate, the sampling-corrected times of origination and extinction for each fossil lineage, and the corresponding rates of origination and extinction in a hierarchical Bayesian framework. Sampling rate is modelled using a Poisson process that describes the expected number of sampled fossils through time, with the potential to model complex, realistic sampling rates being one of PyRate's major benefits. The process can be a homogeneous Poisson process (HPP), a non-homogeneous, bell-shaped trajectory (NHPP), or vary through time as a piecewise series of constant rates separated by pre-defined rate shifts (TPP) akin to the FBD skyline model. Sampling rates are expected to vary between lineages as well as through time, so each Poisson process can be coupled with a gamma distribution which discretises this lineage-wise heterogeneity into a user-defined number of sampling rate multipliers. Origination and extinction rates are modelled using a stochastic birth-death process (Kendall 1948) with the assumption that these lineages form an underlying, complete phylogeny, thus using the Keiding (1975) maximum likelihood estimate of the process. The tree topology, however, is regarded as unknown and does not affect the estimated rates as these are based purely on the times of speciation and extinction of each lineage (Nee et al., 1994; Silvestro et al., 2014). Even though the assumption of complete lineage sampling will be violated by most fossil datasets, particularly when strongly polyphyletic assemblages are analysed (e.g., Flannery-Sutherland et al., 2022), the maximum likelihood point estimates of speciation and extinction rates are again unaffected if incomplete lineage sampling is random

(Silvestro et al., 2014). Rates of origination and extinction may also be expected to vary through time and instances of these rate shifts can be fixed prior to analysis, although this risks subjectivity in the choice of timing. Instead, reverse jump (rj) MCMC can be used to propose the timing of rate shifts in the birth-death model, with the frequency of accepted rate shifts providing a *post hoc* estimate of their statistical support. Finally, the stratigraphic uncertainty associated with empirical fossil occurrences can be incorporated into a PyRate analysis by randomly fixing stratigraphic ages within a dataset, then analysing and combining the results from multiple randomised replicates, accounting for this critical property of fossil occurrence data.

The birth-death-sampling models used by the palaeontological and neontological poles (PyRate versus FBD) developed in parallel (Table 1.1), but the latter have since been adapted for analysis of fossil lineage ranges from their first and last occurrences, the FBD range model (Stadler et al., 2018; Warnock et al., 2020), making it highly comparable to PyRate. Comparison of the performance of both models, plus boundary crosser and three timer rates, has shown that FBD approach is generally most accurate, particularly when fossilisation rates are low, as it explicitly accounts for incomplete lineage sampling (Warnock et al., 2020). PyRate has since received an extension that accounts for this effect, using the FBD range likelihood rather than Keiding likelihood to estimate the divergence time of each sampled lineage within the underlying, unknown phylogeny (Warnock et al., 2020), although its significant computational burden has limited its empirical application. Conversely, the FBD range model makes the unrealistic assumptions of constant rates through time and between lineages, in contrast to PyRate and the FBD skyline model (Warnock et al., 2020). PyRate has also been compared to fossil BAMM, with both methods showing variable performance depending on the underlying diversification dynamics and quality of sampling, although this may only be an issue for relatively small datasets (Cerny et al., 2021).

In some cases, a sampled fossil will bear enough character data that its phylogenetic position can be estimated. In other cases, a fossil will be taxonomically distinguishable, but will not provide any useable phylogenetic data. Nonetheless any fossil taxonomic observation, whether phylogenetically constrained or otherwise, constitutes a sample from a birth-death-sampling process, including occurrences from within the duration of a lineage. Consequently, it is possible to distinguish samples with character data, enabling them to be assigned to specific

Table 1.1. Comparison of birth-death-sampling models

	Fossil BAMM	PyRate	FBD
Phylogeny required	Yes	No	Yes, aside for range
Lineage sampling	Incomplete	Complete*	Incomplete
Taxonomic sampling rates	Yes	Yes	Yes, with multistate
Temporal sampling rates	Yes	Yes	Yes, with skyline
Taxonomic birth-death rates	Yes	Yes	No
Temporal birth-death rates	Yes	Yes	Yes, with skyline
Use of fossil occurrences	All	All	FADs (ordinary, skyline, multistate) + LADs (range) + all (OBDS)

* Can account for incomplete sampling using FBD range likelihood, but currently computationally prohibitive

lineages within a phylogenetic tree, and samples with occurrence data, which can still be assigned to a specific lineage even if their phylogenetic affinities are unknown. Only the former class of samples can inform the phylogenetic topology of a birth-death process, but both classes of samples provide useful information on the total number of lineages produced by that process. PyRate uses all occurrences through the durations of lineages to inform sampling, speciation, and extinction rate estimates, maximising its use of all available fossil occurrence data, while the standard FBD model considers just the first occurrence of a lineage, and the FBD range model its first and last occurrences. The 'occurrence' property of fossil samples, however, has led to extensions of the FBD model which consider sampling of lineages throughout their lifetimes (Vaughan et al., 2019; Gupta et al., 2020; Manceau et al., 2021), along with variation in birth, death and sampling rates through time - the occurrence birth-death-sampling model (OBDS; Andreoletti et al., 2022). The OBDS model therefore offers the potential for fully coherent estimates of diversity, speciation rate and extinction rate that utilise all available data, account for temporal and taxonomic rate variation, and accommodate incomplete sampling. Its

main limitation that it still requires a phylogenetic component while FBD range and PyRate can instead be applied to any set of fossil occurrence data, with the former offering potentially greater accuracy at the expense of incomplete use of the available data and the assumption of constant rates.

1.3.3. Beyond birth-death-sampling models

The birth-death-sampling framework has spurred development of a radiation of methods for estimating aspects of the diversification process that bridge the palaeontological and neontological poles. Workers in both camps are nonetheless continuing to strive for new ways to tackle inadequacies of the fossil record and their impacts on current estimation techniques. One new approach to estimating sampling-corrected diversity is presented in this thesis (Flannery Sutherland et al., 2022), with chapters three and four providing a description of the method, validation using simulated data, results from its empirical applications, and discussions of its benefits and limitations. Two other recent approaches are worth mentioning here, however.

The first of these recent developments is the use of deep-learning methods to link the birth-death-sampling process to an empirical fossil dataset (Cooper, Flannery-Sutherland, Silvestro, IN REVIEW). Briefly, this approach uses a mechanistic simulation framework that combines a spatially explicit birth-death-sampling process with a dispersal process, along with their associated spatial, temporal, and taxonomic heterogeneities, to generate realistic manifestations of diversity trajectories and their incomplete fossil records. A deep-learning model is trained on the outputs of this framework to enable the parameters of the same process to be estimated from an empirical fossil occurrence dataset. As with PyRate a phylogeny is not required, facilitating its application to a wide variety of fossil datasets. The other recent development represents an intriguing departure from the long-established tradition of modelling diversity as a birth-death process with the addition of sampling when considering fossil data: the Bayesian Brownian Bridge (BBB; Silvestro et al., 2021; Carlisle et al., 2023). Rather than model changing diversity as the outcome of a birth-death-sampling process, it is instead treated as a random walk between the origin of a clade (diversity = 1) at some point in the geological past to its diversity in the present. The individual speciation and extinction parameters responsible for present-day and fossil diversity are treated as unknown and the BBB instead uses random walks to integrate across a diverse range of diversification and sampling scenarios. The BBB is consequently an attractive development as it is not prone to the

potentially restrictive or unrealistic assumptions required by a formal birth-death-sampling model and is phylogeny-free, permitting its ready application to the fossil record and avoiding the additional uncertainties posed by molecular or morphological data (Silvestro et al., 2021; Carlisle et al., 2023).

Ultimately, palaeobiologists have a range of sophisticated tools for flexible analysis of fossil occurrence data depending on the quality of the data and the availability of a phylogeny. These tools, however, are encountering new challenges stemming from an emerging desire to understand how biodiversity has varied spatially in the fossil record, along with how sampling biases may themselves be spatially structured.

1.4 Towards a spatially explicit view of diversity dynamics

The species-area effect, where species richness is positively correlated with geographic area, is a well-established biogeographic phenomenon (Arrhenius 1921; Gleason 1922; Connor and McCoy 1979; Lomolino 2000), but the theory of island biogeography also states that finite geographic space ultimately limits biodiversity due to eventual saturation of available ecological niches by a diversifying pool of species (MacArthur and Wilson 1967; Brown and Lomolino 2001; Warren et al., 2015). Palaeobiologists have therefore spent substantial time debating whether Earth's capacity to support new species is similarly limited, leading them to consider the total fossil record as global sample of biodiversity terms to test whether it shows expansionist patterns characteristic of an absent or unrealised limit on fecundity (Valentine 1979; Benton 1995, 1997) versus equilibrium patterns indicative of a ceiling on global biodiversity (Raup 1972, 1976; Sepkoski 1976, 1978, 1979, 1981, 1984). This debate has co-evolved with methodological advances in quantifying sampling-corrected patterns of past diversity, notably the subsampling approaches outlined above (e.g., Alroy 2010) and more recently simulation-based approaches to model diversity from its abiotic drivers (Hagen et al., 2021a, 2021b; Cermeño et al., 2022).

Set against these global analyses, palaeobiologists have historically examined spatial patterns within the fossil record with respect to individual clades to infer their biogeographic histories (e.g., Cracraft 1975; Smith 1988; Barrett et al., 2003; Lieberman 2003) or used spatial similarities in faunal composition to delimit past bioregions or continental configurations (e.g., Forety and Cocks 1998; Brocklehurst and Frobish 2018; Kocsis et al., 2018; Penn-Clarke and

Harper 2021). Renewed recognition of the heterogeneity of past biodiversity and its fossil remnants through space and time is now drawing palaeobiological interest away from viewing the fossil record as a global sample of diversity in favour of a mosaic of regional signals where the nuances underpinning these heterogeneities can be more precisely understood (Close et al., 2020a, 2020b; Benson et al., 2021; Flannery-Sutherland et al., 2022; Allen et al., 2023). Spatial variation in biodiversity is a critical aspect of many biogeographic phenomena. As different sides of the same coin, analyses of biodiversity and biogeography can help strengthen understanding of the other. These remain challenging research endeavours, however, due to the spatial inconsistency of the fossil record through geological time.

1.4.1 Spatial heterogeneity in richness and rocks through time

Major challenges have recently arisen to the established paradigm of treating the total fossil record as a global sample of biodiversity. Surveys of rock availability and sampling effort have repeatedly demonstrated that the fossil record is biased towards the global north, particularly North America and Western Europe, leading to suspicions that 'global' diversity curves are instead disproportionately driven by fossil data from Europe and North America (McGowan and Smith 2008; Raja et al., 2020; Ye and Peters 2023). Palaeobiologists have also repeatedly recognised that the fossil record is not just unevenly distributed with respect to stratigraphic age and present-day location, but also that its spatial distribution also shows substantial variation between geological intervals (Vilhenna and Smith 2013; Close et al., 2020a, 2020b; Ye and Peters, 2023). Variation in the geographic extent of the observable fossil record through geological time consequently induces a bias in our view of global biodiversity through the species-area effect, nor is this bias readily predictable as the richness does not increase smoothly with area, but instead shows substantial heterogeneity between habitats, latitude, and continents (Vermeij 1987; Allison and Briggs 1993; Bush et al., 2004; Barnosky et al., 2005; Vilhena and Smith 2013). More alarmingly, heterogeneous sampling of environments and latitudes exacerbates the impact of spatial sampling bias in the fossil record as spatial biases continue to distort curves of 'global' diversity through geological time even after correction for uneven sampling using the palaeontological gold standard of SQS (Fig. 1.4; Close et al., 2017, 2020a, 2020b), nor do more recently developed methods for estimating diversity dynamics like PyRate, FBD-type models or the BBB account for these issues.

Not only are past diversity signals affected by spatially heterogeneous sampling, but inference of their drivers may be skewed in the same way. Biotic drivers, for example diversity

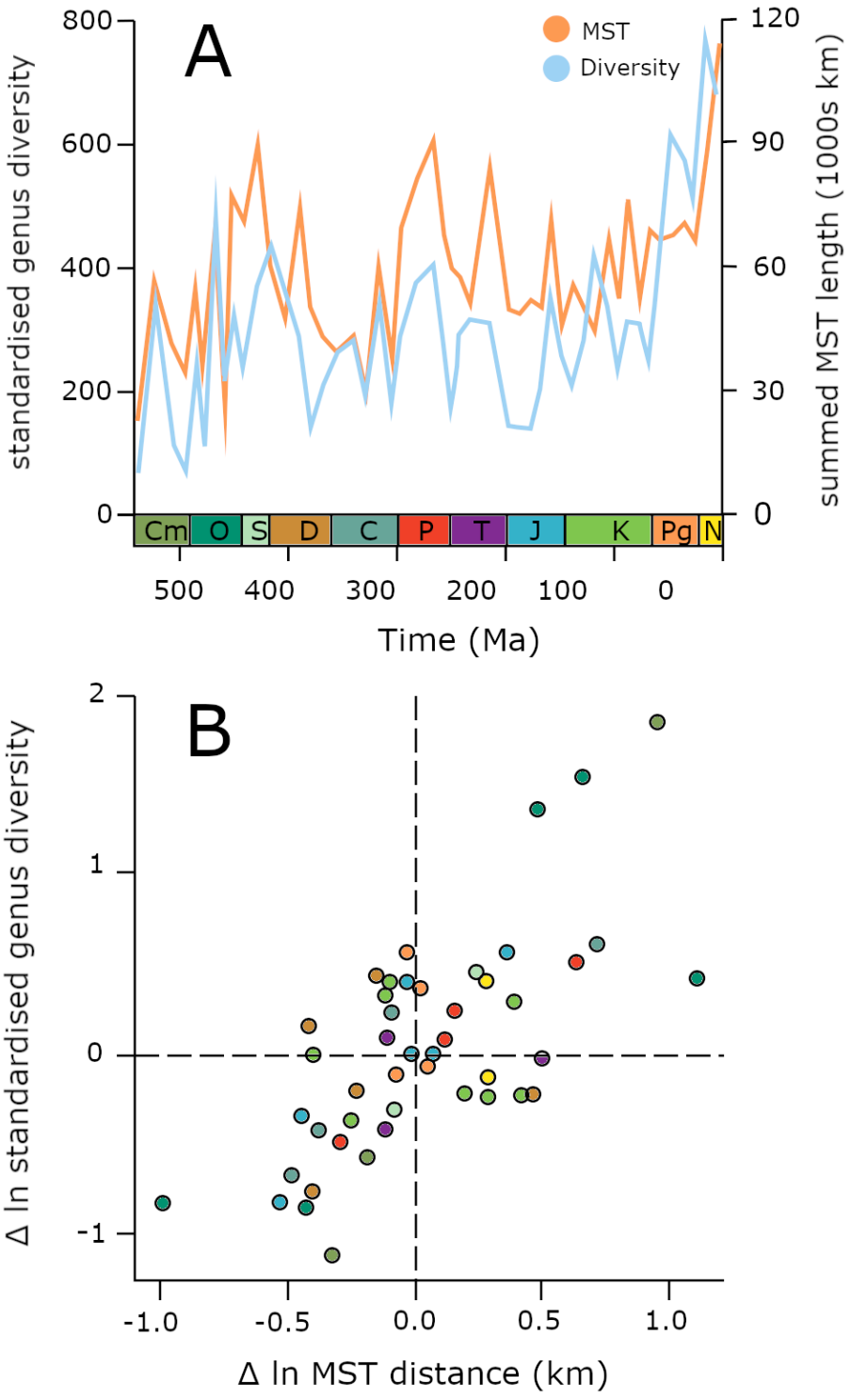


Fig. 1.4. The pervasive influence of spatial sampling extent on Phanerozoic diversity. Even after correction for uneven sampling through geological time, Phanerozoic diversity still displays a severe bias arising from uneven sampling across geographic space. **(A).** Phanerozoic genus-level marine diversity in the PBDB versus the spatial extent of the occurrence data, measured using minimum spanning tree length (MST). **(B).** Epoch-wise relationship between diversity and spatial extent. Point colours correspond to the ribbon in part A. Modified from Close et al. (2020b).

dependent effects or predation pressure, are measured from the same pool of biased fossil occurrence data. Spatially heterogenous sampling through geological time also biases compilations of climatologically sensitive proxies that are often used to infer the long-term state of abiotic Earth-system variables, for example sea surface temperature (Jones and Eichenseer 2022). This leads to complications when comparing supposedly global diversity curves to supposedly global curves of their potential drivers; instead, both are misleading and the effects of the latter on the former cannot be confidently identified. Spatially heterogenous sampling will also impact the completeness of taxon geographic distributions, complicating inference of their abiotic and biotic constraints and limiting our view of their contribution to ecosystem compositions and biogeographic patterns in the geological past (Brocklehurst et al., 2018; Gardner et al., 2019). Finally, the integration of heterogeneously spatially sampled fossil taxa into total-evidence trees may conceivably affect phylogenetic analyses of diversification dynamics due to non-random, incomplete lineage sampling. The consequent impacts on rates of lineage originations and terminations remain largely unexplored in a palaeontological context, however, compared to a neontological or even epidemiological context (Gardner et al., 2019; Guidon and De Maio 2021; Hancock et al., 2022).

Deep-learning methods provide a potentially comprehensive solution to the spatiotemporal sampling issues posed by the fossil record (Cooper, Flannery-Sutherland, Silvestro, IN REVIEW), but this comes with substantial methodological assumptions that have yet to receive critique from the wider palaeobiological community, nor can these methods currently return spatially resolved components of the single global diversity estimate (Cooper pers. comm.). Instead, it may be potentially simpler and more informative in some respects to examine fossil biodiversity at regional scales, where spatial sampling bias can be more properly mitigated and the rich mosaic of spatial trends in biodiversity can emerge (Close et al., 2020a, 2020b; Flannery-Sutherland et al., 2022; Allen et al., 2023). This paradigm shift away from global biodiversity estimation recognises that while spatially heterogeneous sampling confounds estimates of past diversity and its biogeographic structure at multiple geographic scales, biodiversity is also expected to vary globally due to corresponding spatial heterogeneity in its underlying drivers (Benson et al., 2021).

Regional analyses will in turn necessitate regional compilations of the drivers of diversity, although the spatial extent of geological sampling may limit the availability of such data separately to the availability of fossil data on regional diversity. While these endeavours

are challenging, they remain essential as the signal of geographical variation in past diversity will conceivably remain within the purview of the geological record. Phylogenies, even when augmented by fossil tips, will never be able scry polyphyletic, ecosystem-level taxonomic diversity in the pelagic zone of the Western Interior Seaway, the shelf of Iapetus, or the swamp forests of Carboniferous Pangaea, particularly when considering the entirely non-phylogenetic distribution of biodiversity hotspots like reefs or rainforests. Instead, the fossil record remains the only empirical signal of these patterns across space and through time, requiring that fossil occurrence data is of the highest stratigraphic and taxonomic quality, along with analytical methods that can account for its spatiotemporal sampling biases.

1.4.2 Earth-system models can go beyond the bounds of the geological record

A spatially explicit view of past biodiversity necessitates placing the fossil record in its original palaeogeographic context. This goal was first made possible on a global scale in the 1970s with the emergence of plate tectonic models which describe the architecture of Earth's lithospheric blocks and their kinematics (Dewey and Bird 1970; Smith et al., 1973; Scotese and Baker 1975; Scotese 1976). A variety of models have accrued in subsequent decades from academic to commercial research efforts, with newer reconstructions often basing their tectonic and kinematic configurations after older models whilst integrating new sources of data (see reviews in Verard 2019; Torsvik and Cocks 2019). These efforts can be broadly divided into 'single control' and 'dual control' approaches (Verard et al., 2015). The former (typically used by older models) focus on the positions of tectonic elements based on all relevant data for a given interval, describing those positions in terms of the rotation of rigid plates around a Euler pole without necessarily resolving discrepancies between the underlying data and resulting model, or within the model itself (Verard 2019). Dual control approaches instead focus on deriving key assemblies of geological data which precisely describe plate tectonic structure, deformation and kinematics for a specific reconstruction age, then smoothly interpolating structural and geometric changes between these key assemblies. Despite being substantially more complex to create, dual control models yield globally complete plate tectonic reconstructions which can be iteratively refined to achieve geological congruence with their underlying data, geometrical consistency with respect to their component plate boundaries, and kinematic and geodynamic plausibility with the mechanisms and forces proposed for their movements (Verard et al., 2015; Verard 2019).

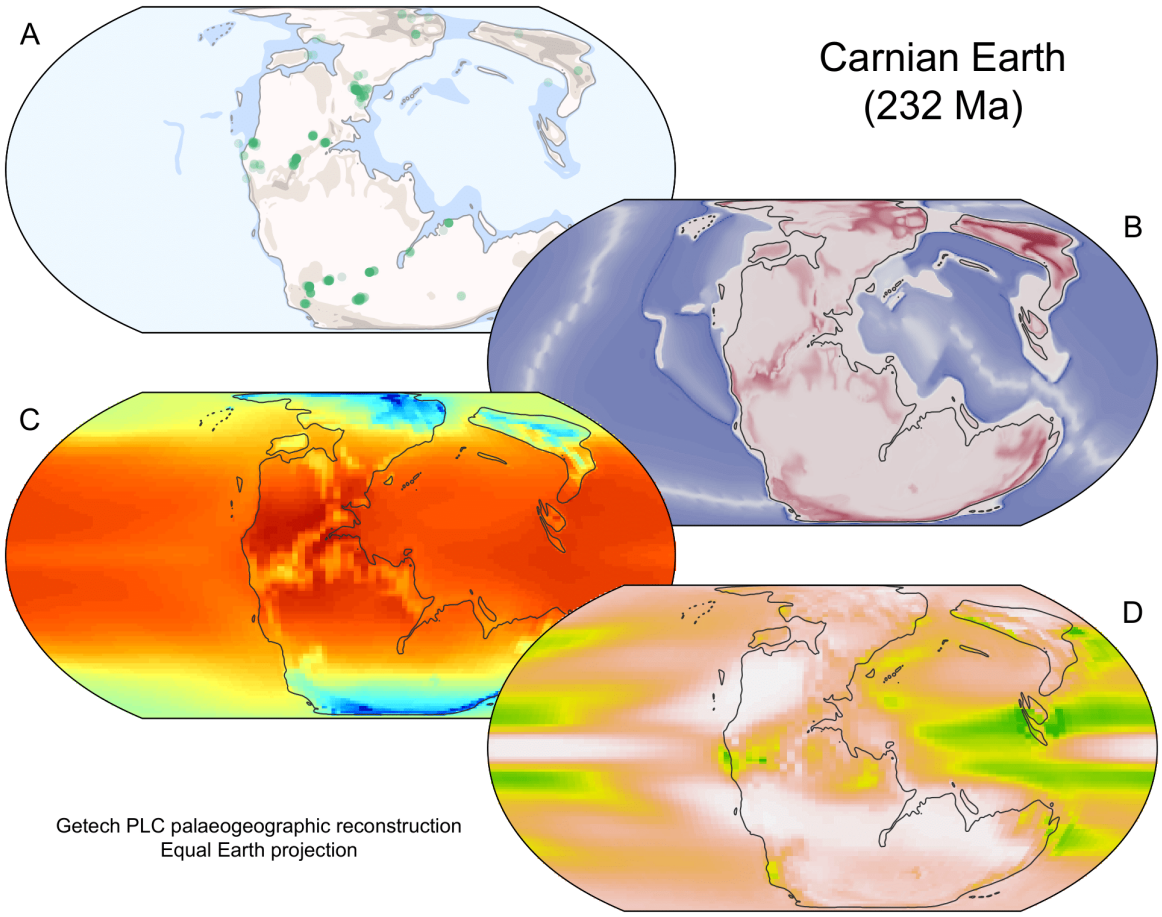


Fig. 1.5. Overview of the Carnian Earth. (**A).** Getech PLC palaeogeographic reconstruction of the highlands, coastlines, and continental shelf during the Carnian (~232 Ma), highlighting terrestrial fossil occurrences from the Paleobiology Database. All maps are figured using Equal Earth projection. (**B).** Digital elevation model forming the basis for part A. (**C).** HadCM3L climate simulation for mean annual temperature, with boundary conditions taken from part B. **(D).** Annual precipitation from the same climate simulation.

Plate tectonic models provide a basis for reconstructions of continental block and oceanic basin configurations and their associated topographies and bathymetries. Such reconstructions predate plate tectonic theory, but the merging of palaeogeography with plate tectonics has enabled production of increasingly detailed global maps through the Phanerozoic and beyond (see review in Markwick 2019). Palaeogeographic maps have historically been qualitative representations of the distributions of different palaeoenvironments which can be semi-quantitatively mapped to bathymetries or topographies, but the desire to apply Earth-system models (ESMs) to the geological past has driven the development of formal digital elevation models (DEMs) for paleogeographic reconstructions (Markwick and Valdes 2004; Markwick

2019). DEMs provide the 2D and 3D boundary conditions required by spatially resolved ESMs, enabling Earth-system processes to be simulated for a given input landscape (Fig. 1.5). A full review of applications of ESMs to palaeogeographic DEMs is beyond the scope of this thesis, but major applications include climate modelling (Lunt et al., 2016; Valdes et al., 2021), biogeochemical cycling (Ridgwell et al., 2007; Mills et al., 2021), and erosion, sedimentation, and landscape evolution (Lyster et al., 2020; Salles et al., 2023).

Plate tectonic models enable the fossil record to be placed in its original spatial context, whilst also forming the basis for the DEMs and ESMs that can illuminate their environmental context. While some models may be more reliable than others given the differential quantities and qualities of data on which they are built, all such models recover broadly concordant patterns of tectonic and continental configurations through their respective timespans (Seton et al., 2023). Nonetheless, precise positions of otherwise geographically homologous features become increasingly discrepant with increasing geological age of reconstructions from the paucity of oceanic crust older than the Cretaceous (Muller et al., 1997; Seton et al., 2020). In turn, discrepant plate rotation models introduce a degree of uncertainty into any palaeogeographic reconstruction of fossil data, in combination with uncertainty within the chosen model itself, and compounded by stratigraphic uncertainty in fossil ages during pointwise rotation. The potential impact of these uncertainties on analyses of diversity and biogeography in the fossil record have long been recognised by palaeontologists but are only just beginning to be formally investigated (Boddy et al., 2021; Jones et al., 2022; Buffan et al., 2023). Uncertainties in rotation models also propagate to DEMs and in turn to ESMs, as these latter models are demonstrably sensitive to variations in their boundary conditions (Lunt et al., 2016).

While ESMs are by no means infallible, they offer an independent framework against which geological proxies for past environmental conditions can be compared, and which can fill in the blanks beyond the spatial bounds of the geological record. Consequently, they can provide the environmental data required for regional analyses of the drivers of biodiversity, for mechanistic modelling of patterns of past diversity on global scales, or for determining the constraints on clade distributions (see 1.4.3). This is not necessarily a new approach as ESMs have been applied to the geological past for nearly three decades, aiding in investigation and interpretation of the fossil record (e.g., Parrish 1993; Woods 2005). More recent advances in their accuracy (Lunt et al., 2016; Valdes et al., 2021) and the computing power needed for their

implementation, however, provide a wellspring of potential that is only just being recognised by palaeontologists.

1.4.2 Biogeography can go beyond the bounds of the fossil record

The incompleteness of the fossil record limits our view of the distributions of clades and biodiversity through time and across space, but fossils still provide a signal of the broader biogeographic patterns to which they belong. Consequently, methods for biogeographic inference of ancestral geographic ranges and origination points can provide an indirect view of past trends in the spatial biodiversity of clades and ecosystems that may extend beyond the bounds of the observable fossil record. These techniques can be broadly divided into a diverse array of methods that consider various combinations of anagenetic and cladogenetic mechanisms for range modification in combination with discrete, pre-defined geographic areas (see Matzke 2013 for a summary of commonly used models), and a second category of models which consider range evolution as a spatially continuous random walk in a planar or spherical domain (Lemmon and Lemmon 2008; Bouckaert 2016; O'Donovan et al., 2018). Mirroring diversity estimation by the neontological pole, these methods were originally developed using extant phylogenies but estimates of ancestral geographic ranges are substantially improved by inclusion of fossil tips (Crisp et al., 2011; Matzke 2013; Wood et al., 2013; Meseguer et al. 2015; Wisniewski et al., 2022). While these methods still suffer from the incompleteness of the fossil record both temporally and spatially (Silvestro et al., 2016), and are limited to cases where phylogenies are available, they yield a picture of a clade's potential geographic distribution through time and in turn its contribution to spatiotemporal patterns of biodiversity. A smaller number of phylogeny-free models are available, which consider processes of dispersal, extinction and sampling in fossil occurrence data, but these are less common by comparison (Silvestro et al., 2016).

Alternatively, biogeographic patterns can be investigated using taxon-environment relationships, pertinently to this thesis by prediction of taxon distributions from their environmental constraints using ecological niche modelling (ENM; also referred to as species distribution modelling or habitat suitability modelling). This heterogeneous set of techniques emerged from a desire for statistically robust, predictive approaches to taxon-environment relationships in modern ecology (Peters 1991; Guisan and Zimmerman 2000). ENM subsequently benefitted from the rise of digitally accessible neontological databases of species occurrences, and geographic information system (GIS) databases of climatological and

biogeochemical conditions from ESMs or satellite and weather station measurements (Morales-Barbero and Vega-Alvarez 2018; Ball-Damerow et al., 2019). Corresponding growth of spatiotemporally resolved fossil occurrence databases and climate simulations for palaeogeographic reconstructions has in turn enabled the application of ENM to fossil taxon distributions (e.g., Chiarenza et al., 2018, 2019; Jones et al., 2019; 2021; Saupe et al., 2019).

Projection of a niche model to the full spatial extent of its environmental calibration data provides an indication of potential clade distributions and a powerful means of circumventing the geographical limits of the fossil record. The same estimate of habitable area may also be used as an indirect measure of past diversity using the assertion that under the species area effect, a larger habitable area will naturally engender greater diversity (Chiarenza et al., 2018), although this latter use is questionable given that a clade's fundamental abiotic niche often differs to its realised geographic extent (Pulliam 2002). ENM may even be combined with geographic ancestral state estimation to estimate geographic distribution through a clade's phylogeny based on their changing climatic tolerances, the emerging field of phylogenetic niche modelling (also termed phyloclimatic modelling; Yesson and Culham 2006; Evans et al., 2009; Meseguer et al., 2015; Guillon and Brown, 2021). Estimation of geographic ranges through geological time provides insight into the changing spatial diversity of clades and their contributions to broader biogeographic patterns in the fossil record. As with ancestral state estimation, however, ENM is sensitive to missing data, and the challenges of applying such techniques in deep time are still being explored (Varela et al., 2011). The diversity of available methods come with a corresponding set of technical caveats and limitations (Peterson et al., 2011; Peterson and Anamza 2015), and the spatial uncertainties introduced by imprecise stratigraphic ages of fossil occurrences and the climatic uncertainties arising from differences between palaeogeographic boundary conditions and ESM architecture pose further issues for deep time ENM. An innovative solution to at least the former issue, and one which has been extended to other biogeographic problems in the geological past, is to quantify taxon-environment relationships in the present day where the spatial quality of the data may be more rigorously controlled, then project those relationships backwards in time using the outputs of deep time ESMs (Kocsis et al., 2018a; Jones et al., 2022). While this requires strong assumptions regarding the stability of the present-day relationship in the geological past, such approaches highlight how spatially complete deep time ESMs can provide a powerful means of circumventing the limits of the fossil record.

More recently, methods have emerged for mechanistic prediction of spatiotemporal diversity patterns from their abiotic drivers based on lithological proxies (e.g., Hagen et al., 2021a, 2021b) or ESM simulations (Rangel et al., 2018; Saupe et al., 2020; Cermeño et al., 2022). Broadly, these methods operate by simulating the acquisition of diversity via cladogenesis, dispersal and extinction processes, with the chosen abiotic drivers influencing these parameters and the carrying capacities of the geographic regions under consideration (Hagen 2021; Pilowsky et al., 2022). These approaches provide a powerful alternative to calibrating models using spatially incomplete fossil data or risking flawed assumptions of constant taxon-environment relationships when projecting models calibrated in the present. As they are constructed in the absence of empirical biological data regarding taxon-environment relationships, however, their outputs require scrutiny to ensure that they accurately recover past or present patterns of biodiversity. Mechanistic models also incur the issue of equifinality, analogous to the problems of non-identifiability faced by phylogenetic models of diversification: varied parameters across a given set of mechanisms may produce the same trends in diversity without a clear way to distinguish the best-fitting scenario. The high degree of control offered by this modelling framework, however, correspondingly comes with strong potential for hypothesis testing by comparison of different diversification scenarios, or even unrealised geological histories. For example, Hagen et al. (2021b) experimentally removed Andean uplift during simulations of continental biodiversity through the Cenozoic, demonstrating that tectonic evolution was a major control on the formation of the Amazonian biodiversity hotspot, along with the high spatial heterogeneity of biodiversity across South America more generally.

1.5 Current outlook

Fossils will always be our only empirical source of information on past diversification dynamics. A rich ecosystem of methods for estimating these dynamics has developed based on phylogenies and fossil occurrence data, but the greatest advances have undoubtedly come from unifying these once disparate approaches using birth-death-sampling models. While phylogenies provide critical biogeographic insight into the distribution and dispersal of clades through geological time, fossil occurrence data are again irreplaceable for examination of patterns of biodiversity through space and time, particularly at the ecosystem level and greater

when the assemblages under examination are strongly polyphyletic. Perhaps the greatest challenge facing palaeobiologists currently, however, is the spatial limitations of the fossil record. Overcoming this challenge is not straightforward, but a solution lies with regionalised approaches to diversification dynamics, where bias may be separated from biodiversity and the spatial nuances underpinning the biosphere's original biological patterns revealed, in combination with ecological and Earth-system modelling which can circumvent spatial limitation altogether. These latter approaches nonetheless bear their own set of caveats stemming from uncertainties in palaeogeographic reconstructions, necessitating substantial investigation to fully understand their effects on our inference of Earth-system processes and taxon-environment relationships in the geological past.

The technical considerations comprising this first chapter underpin the remainder of this thesis. It is critical that methods can take full advantage of all available fossil occurrence data, that the data are of high quality, and that such methods are complemented by approaches that can tackle the spatial inadequacies and uncertainties of the geological record. In the second chapter, I develop automated solutions to identify erroneous records in fossil occurrence databases and improve their stratigraphic precision and consistency. In the third chapter, I combine state-of-the-art Bayesian techniques for inferring diversification dynamics from fossil occurrence data (PyRate and associated methods) whilst accounting for heterogenous sampling through geological time with a new method for eliminating the effects of uneven sampling across geographic space. I apply this combination of methods to analyse regional trends in marine biodiversity through the Triassic to document their spatiotemporal heterogeneity, particularly across the two mass extinction events bounding this geological interval. In the fourth chapter, I then draw upon biogeographic and Earth-system modelling methods to examine the rise of scleractinian corals during the Triassic, the clade which forms the basis for the most diverse marine ecosystems today. I utilise the Bayesian methods adopted earlier in the thesis to analyse their diversity, but use ecological niche models, palaeogeographic reconstructions and deep time climate simulations to infer their global distributions in the face of the spatial limitations of their fossil record. In the fifth chapter, I continue to probe the origins of the modern biosphere from a spatial perspective by examining the biogeographic origins and rise of archosauromorph reptiles during the Triassic. Archosauromorph reptiles would continue to dominate terrestrial ecosystems, in the guise of non-avian dinosaurs during the Mesozoic, then birds and crocodiles during the Cenozoic. I use phylogenetic comparative methods to infer

their geographic origins and patterns of dispersal, then combine these inferences with landscape connectivity analysis to investigate how palaeogeography and climate shaped their biogeographic history, leveraging the availability of high-resolution palaeogeographic DEMs and climate simulations.

Chapter Two

Improving the Quality of Fossil Occurrence Datasets

Fossil occurrence databases are indispensable resources to the palaeontological community, yet present unique data cleaning challenges. Many studies devote significant attention to cleaning fossil occurrence data prior to analysis, but such efforts are typically bespoke and difficult to reproduce. There are also no standardised methods to detect and resolve errors despite the development of an ecosystem of cleaning tools fuelled by the concurrent growth of neontological occurrence databases. As fossil occurrence databases continue to increase in size, the demand for standardised, automated, and reproducible methods to improve data quality will only grow. Here, I present semi-automated cleaning solutions to address these issues with a new R package *fossilbrush*. I apply these cleaning protocols to the Paleobiology Database to assess their efficacy and impact on the prevalence of anomalous entries. Anomalies may be effectively resolved by comparison against a published compendium of stratigraphic ranges, improving the stratigraphic quality of the data, and through methods which detect outliers in taxon-wise occurrence stratigraphic distributions. Despite this, anomalous entries remain prevalent throughout major clades, with often more than 30% of genera in major fossil groups (e.g., bivalves, echinoderms) displaying stratigraphically suspect occurrence records. These methods provide a way to flag and resolve anomalous taxonomic data before downstream palaeobiological analysis and may also aid in the automation and targeting of future cleaning efforts. It should be stressed, however, that these methods are semi-automated and are primarily for the detection of potential anomalies for further scrutiny, as full automation should not be a substitute for expert vetting. It is also worth noting that some of these methods do not rely on external databases for anomaly resolution and so are also applicable to occurrences in neontological databases, expanding the utility of the *fossilbrush* R package.

Author contributions: This chapter is based on Flannery-Sutherland et al. (2022b) in Methods in Ecology and Evolution, 13:2404–2418. The work was developed in response to an open research challenge posed by co-authors N. Raja, A. Kocsis and W. Kiessling, who provided on a draft of the paper. The author contributed almost all the work in this research chapter.

2.1 Introduction

Biological occurrence databases record the present and historical spatial distributions of organisms across the tree of life, enabling assessment of the drivers of biodiversity across a wide variety of geographic, taxonomic, and temporal scales (Zizka et al., 2019). These databases are critical to predicting the impacts of habitat change on species distributions and future ecosystems, particularly to inform conservation efforts (Bell-Damerow et al., 2019; Stephenson & Stengel 2020). Such efforts are strengthened by incorporating organismal distribution records in the geological past from fossil occurrence databases (Dietl. 2019; Kiessling et al., 2019), including Triton (Fenton et al., 2021), FAUNMAP (Graham & Lundelius, 2010), PaleoReefs (Kiessling and Krause, 2022), Neotoma (Williams et al., 2018), the PBDB (Alroy et al., 2008), FRED (Clowes et al., 2021) and NOW (The NOW Community, 2020). Alongside geographic and systematic data, fossil occurrence databases record the geological ages of their entries, permitting connection of palaeontological and neontological information in the recent past (Eduardo et al., 2018) and analysis of diversity dynamics, macroecological structuring and biogeography in deep time (Peters and McClennen, 2016).

The proliferation of occurrence databases has promoted development of a series of tools for database access and cleansing, typically as packages for the R programming environment (R Core Team, 2023). Errors in taxonomy and spatial coordinates inevitably accrue due to improper data imputation and changing taxonomic convention in the primary literature (Grenie et al., 2022). Tools for coordinate cleaning cover both fossil and modern occurrence databases (Zizka et al., 2019), but available tools for spell checking and harmonisation of taxonomic names are usually restricted to extant species (Norman et al., 2020; Ribeiro et al., 2022). Consequently, these are unsuited for resolution of fossil taxonomy and fossil data cleaning is typically done in isolation without support from standardised tools. As well as taxonomic and coordinate errors, fossil occurrence databases may contain erroneous geological ages arising from increasing refinements to chronostratigraphic timescales, improved lithostratigraphic constraints for fossil-bearing formations, or typographical errors during data imputation. Finally, fossil occurrences present the unique problem where misidentification of an occurrence may result in an incorrect stratigraphic range for the taxon to which it is assigned.

The PBDB is one of the most well-known fossil occurrence databases, with over 1.5 million occurrences for >45,000 taxa. Its data underpin seminal works on the quality of the fossil record and the broadest patterns of biotic change throughout the Phanerozoic (Alroy et al., 2008), yet the presence of taxonomic, spatial and stratigraphic errors leads some authors to condemn the PBDB altogether (Prothero, 2015). Manual cleaning of the entire database is rendered infeasible by its scale, and such an effort would quickly become outdated with addition of new data and changing taxonomic and stratigraphic conventions in the primary literature. New databases may also emerge in the future, for example through application of automated data extraction and machine-learning methods to digitised publications (Peters et al., 2014; Kopperud et al., 2018) or from dark data in museum collections (Allmon et al., 2018; Marshall et al., 2018), and these may also contain inconsistencies in need of revision. As such, cleaning of fossil occurrence data is constant challenge requiring automated, standardised and reproducible solutions capable of scaling to large datasets, and which tackle the unique challenges posed by the stratigraphic component of the data.

2.2 Materials and methods

Here, I present several data-cleaning considerations and tools to resolve misspellings and identify inconsistencies in taxonomic schemes. These operate independently of external databases and can be applied to any occurrence dataset that contains higher taxonomic information. Tools to address spatial coordinate errors already exist (Zizka et al., 2019), but accurate geological ages are required for calculation of accurate palaeocoordinates and I provide methods to update chronostratigraphic ages for fossil occurrence datasets. Resolution of taxonomic and stratigraphic errors in isolation is predicated on expert knowledge of clade diagnostic characters and spatiotemporal distributions, and regional rock unit relationships. In cases where taxonomic misidentification results in stratigraphic range anomalies, however, error detection can be predicated on outlier detection, permitting development of parameterised, automated, and reproducible statistical approaches. I develop three methods to flag and resolve outliers in taxon stratigraphic ranges which have been made available within a new R package, *fossilbrush*. I apply these cleaning protocols to the PBDB to demonstrate how they can be applied in a standardised, reproducible fashion, then examine the efficacy of anomaly detection and resolution, along with their impact on the database.

2.2.1 Resolving database inconsistency

Occurrence chronostratigraphy and analytical strategy should conform to the same chronostratigraphic timescale. The Paleobiology Database currently uses the International Chronostratigraphic Time Scale 2013 (ICS 2013; Cohen et al., 2013), yet studies often use more recent chronostratigraphic schemes for data analysis, which may lead to erroneous division of occurrences into successive, temporally discrete partitions. For example, the base of the Triassic sits at 252.17 Ma in ICS 2013 compared to 251.9 Ma in more recent timescales. This is a pertinent example as it coincides with the end-Permian mass extinction where well-resolved and correctly binned data is critical to determining extinction magnitude and timing. A Geologic Time Scale 2020 (Gradstein et al., 2020) is the most recently published comprehensive chronostratigraphic standard that covers both regional and international geological intervals and is being adopted both analytically (e.g., Marshall et al., 2021; Metcalfe and Crowley, 2021) in databasing efforts (e.g., Fenton et al., 2021), and as the basis for still newer timescales (e.g., ICS 2023). I provide a lookup table and function `chrono_scale()`, to enable quick revision of chronostratigraphic dates in existing fossil occurrence datasets. While the lookup table will inevitably accrue errors as chronostratigraphic timescales receive refinement, the majority of the dates are expected to remain valid, and a user can either update the lookup table in R or supply a new one entirely to `chrono_scale()` to ensure that their data and analytical strategy conform to the latest chronostratigraphic standard.

Spelling variations and inconsistent higher classifications risk incomplete representation of a taxon by its occurrences, leading to underestimation of spatiotemporal ranges and overestimation of diversity. I provide a modular taxonomic harmonisation workflow through the function `check_taxonomy()`, with options to report flagged issues or resolve them automatically. This workflow is applicable to any occurrence dataset with hierarchically organised taxonomic information, including composite datasets compiled from multiple databases, and ensures that all classifications are internally consistent so that taxa are fully represented by all their occurrences at any given taxonomic level. Following previous recommendations (Grenie et al., 2022), this taxonomic harmonisation function initially standardises taxon name formatting as a prerequisite to improve detection of inconsistent spelling using fuzzy string methods (where characters are inexactly matched between text strings to detect expected spelling error or variation). Instances of inconsistent higher

classifications can be displayed graphically for user inspection using the `plot_taxa()` function and automatically resolved within `check_taxonomy()` if desired.

The workflow in `check_taxonomy()` consists of four steps. Firstly, taxonomic name formatting is scanned. The default function behaviour is to flag instances of non-letter characters and extraneous words in taxon names, which is preferable as it forces the user to manually investigate these irregularities. Name formatting should be standardised prior to applying the other modular steps, but this formatting can optionally be automated within the function with the following routing. The first letter of each name is capitalised while all other letters are set to lowercase. Next, abbreviations are removed so that the quality of fuzzy string matching of names is not reduced by these elements. This includes taxonomic qualifiers such as sp., spp., cf., and can be supplemented with user-defined deletion terms to aid in tailoring the formatting routine to their dataset. It should be noted, however, that qualifiers have their justification in taxonomic names and the general convention is that occurrences with such entries should be excluded from a dataset prior to cleaning and analysis. Removal of qualifiers and abbreviations, however, is necessary to produce consistently formatted names which can then be effectively scanned for spelling irregularities. Any remaining non-letter characters and Roman numerals are replaced with whitespaces (avoiding accidental concatenation of other parts of taxon names, e.g., *Tyto_alba* becoming *Tytoalba*), then the first word of each name taken (e.g., *Felix (Linnaeus)* becomes *Felix)*, aside from species-level assignments where the first two words are taken (e.g., *Panthera tigris (subsp.) sondaica* becomes *Panthera tigris*).

In the second part of the routine, fuzzy string methods are used to check for potential spelling variations of a taxon name. As above, the default behaviour is to flag and report these instances rather than resolve them automatically, but automatic resolution is possible. To prevent flagging of similar names from disparate clades, comparisons are conducted in groups defined at the next highest taxonomic level, for example all genus names within the family Pteriidae (class Bivalvia). As there is no single best method for fuzzy string matching, two similarity metrics are used to compare potential matches and their similarity scores are reported to the user for inspection, subsequently allowing for automatic filtering and resolution using a given threshold similarity. The two default methods were chosen with respect to their complementary characteristics: Jaro-Winkler distance and q-gram distance. Jaro-Winkler distance measures the positional differences in shared characters between a pair of character

strings (i.e., the number of edits needed to make the positions match) rescaled to the range [0,1], such that strings with no common characters will have distance = 1). String distance is reduced when the first few letters are shared and so is heuristic rather than truly metric (as in a Euclidean distance), but this means that Jaro-Winkler distance is well suited to short, human-typed strings under the assumption that the starting letters are more likely to be correct, compared to the rest of the string (Winkler 1990). q-gram distance is then used as it is a non-edit-based, Euclidean metric where distance is directly interpretable as percentage dissimilarity, unlike Jaro-Winkler distance (van der Loo 2014).

The third part of the cleaning routine checks for the reuse of taxonomic names at different levels. This can occur when a higher taxon is used as a placeholder name in databases at a lower level, or from rare cases of homonymy between higher taxa. These cases are flagged and reported to the user by default, and in the case of the automated cleaning option numeric prefixes are added to distinguish between the names at different levels. This cleaning method is necessary for the fourth step to correctly function as it requires that names are unique to their given taxonomic level. To detect inconsistent higher classifications, columns of taxonomic classifications are treated as a hierarchical edge list and converted to a directed graph, such that edge directions flow from lower to higher taxonomic ranks. Graph node topology then provides a simple way to detect inconsistent higher assignments of the same taxonomic name, supported by functions from the *igraph* R package (Csardi and Nepusz). Homonyms (accidental re-use of the same name for disparate taxa) are distinguished based on the number of higher taxonomic levels for which they remain divergent (Fig. S1). For example, a genus with divergent classifications is likely a homonym if it only shares the same phylum level classification, for example the trilobite *Panderia* (order Corynexochida) and the ostracod *Panderia* (order Palaeocopida) within the common phylum Arthropoda. Resolution can be then automated if desired by infilling missing classifications where possible, application of the more frequent or more complete classification scheme for inconsistent higher taxonomy, and by appending suffixes to distinguish between homonyms at the same taxonomic rank. For automatic resolution of inconsistently classified taxa, the choice of lumping or splitting is set in `check_taxonomy()` based on the number of levels for which they are divergent. This parameter must therefore be adjusted based on the number of taxonomic levels present in the dataset being cleaned and the desired degree of taxonomic stringency. Any flagged synonyms are split through the addition of suffixes to give unique names.

2.2.2 Stratigraphic outlier detection against an independent database

Taxon and occurrence age ranges may be flagged against a reference database with similar basic data. The following notation is used (Fig. 2.1) where 0 and 1 denote error versus validity for the FAD (left) and LAD (right) of a range (R): `1R1` = valid, `0R1` = FAD outside of reference, `1R0` = LAD outside of reference, `0R0` = FAD and LAD both exceed reference, `00R` = older than reference, `R00` = younger than reference, `000` = unrecorded in the reference database. Occurrence ages can be checked individually against the reference database, with the magnitude of age discrepancies helping to inform whether they are potentially anomalous. Stage to substage-level discrepancies may mark discoveries whose FADs and LADs genuinely supersede their reference range while those which overlap with their reference range (`0R1`, `0R0`, `1R0`) have a partially plausible stratigraphic range and may reflect cases where dating

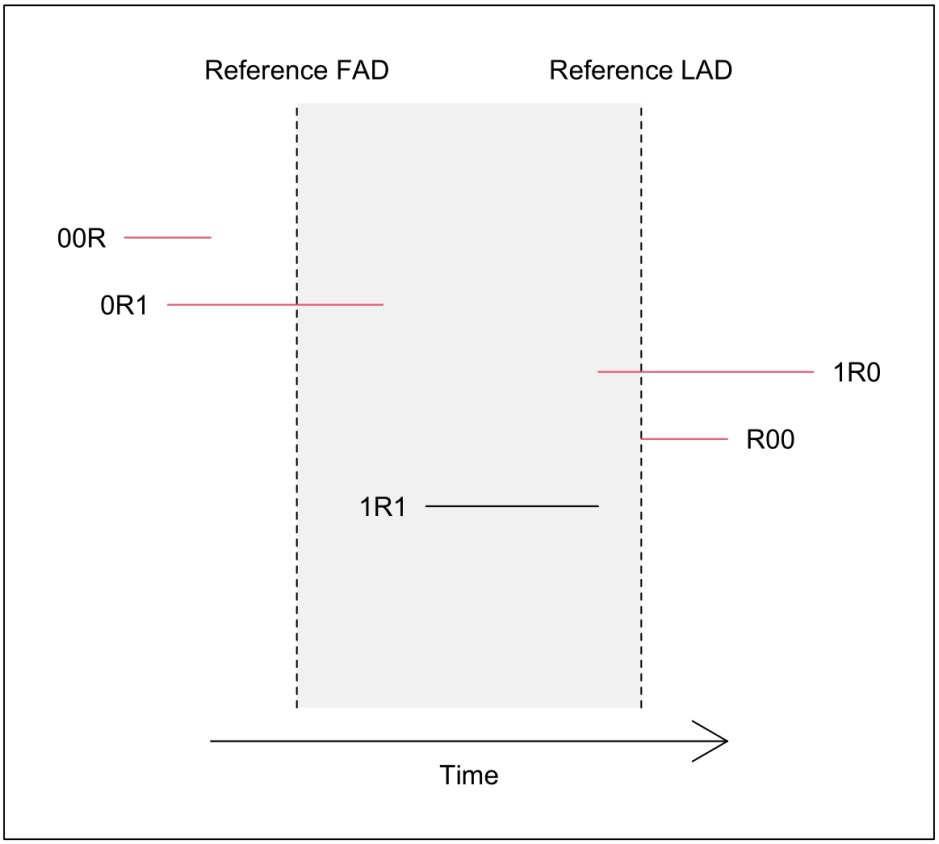


Fig. 2.1. Schematic representation of stratigraphic error flagging codes. Dotted lines demarcate the reference range of a hypothetical taxon and solid lines the age range of occurrences assigned to that taxon. Solid red lines indicate occurrences with stratigraphic age anomalies relative to the reference range, while solid black lines are occurrences with valid stratigraphic ranges. Definitions of the error codes are in the main text (see 2.2.2).

imprecision for the parent stratigraphic unit results in erroneous range extension. On the other hand, occurrences which fall outside the reference range (`00R`, `R00`) are more likely to reflect taxonomic error, particularly if that error is substantial (epoch-level or greater). This is a common issue for wastebasket taxa where superficial taxonomic work can result in backward range extension of relatively recent taxa by hundreds of millions of years, for example the brachiopod *Lingula* or bivalve mollusk *Ostrea* (Plotnick and Wagner, 2006).

Comparison against a reference database is implemented in the `flag_ranges()` function, with the option to tag occurrence age discrepancies greater than a user-defined threshold to quickly distinguish between cases of stratigraphic versus taxonomic error. This is a useful exploratory procedure but is limited to taxa common to both databases. Additionally, occurrences in the PBDB are organised into collections, in idealised terms a discrete community from a single bedding plane (i.e., a precise point in space and time), but in reality a time-averaged assemblage bearing the stratigraphic uncertainty of its parent lithological unit. Fossil occurrences in the same collection must possess the same stratigraphic age range, but resolving occurrence ages individually may violate this property, so the flagging procedure instead uses the reference database to determine if a collection range age is plausible. The procedure retrieves the available reference ranges for any taxa in the collection, then determines if a threshold proportion of those ranges overlap. If the overlap falls below the threshold, then the entire collection is treated as bearing a high prevalence of stratigraphic and taxonomic anomalies. If a common interval exceeding the threshold can be identified, the collection age should fall within this interval to be considered stratigraphically plausible and can be modified according to this robust consensus of calibrations. This procedure is implemented in the function `revise_ranges()` with a default, arbitrary threshold of 75% overlap denoting stratigraphic plausibility. While a more stringent threshold could be used, a more relaxed threshold better accommodates erroneous occurrences in some collections when searching for a plausible interval. As above, any discrepancies between original and revised ages of occurrences within a collection can then be examined to determine whether they represent plausible range extension, taxonomic error, or stratigraphic error.

2.2.3 Occurrence density distribution

As reference databases may still contain errors and are unavailable for many groups, the density of fossil occurrences through time can alternatively be used to detect and resolve stratigraphic

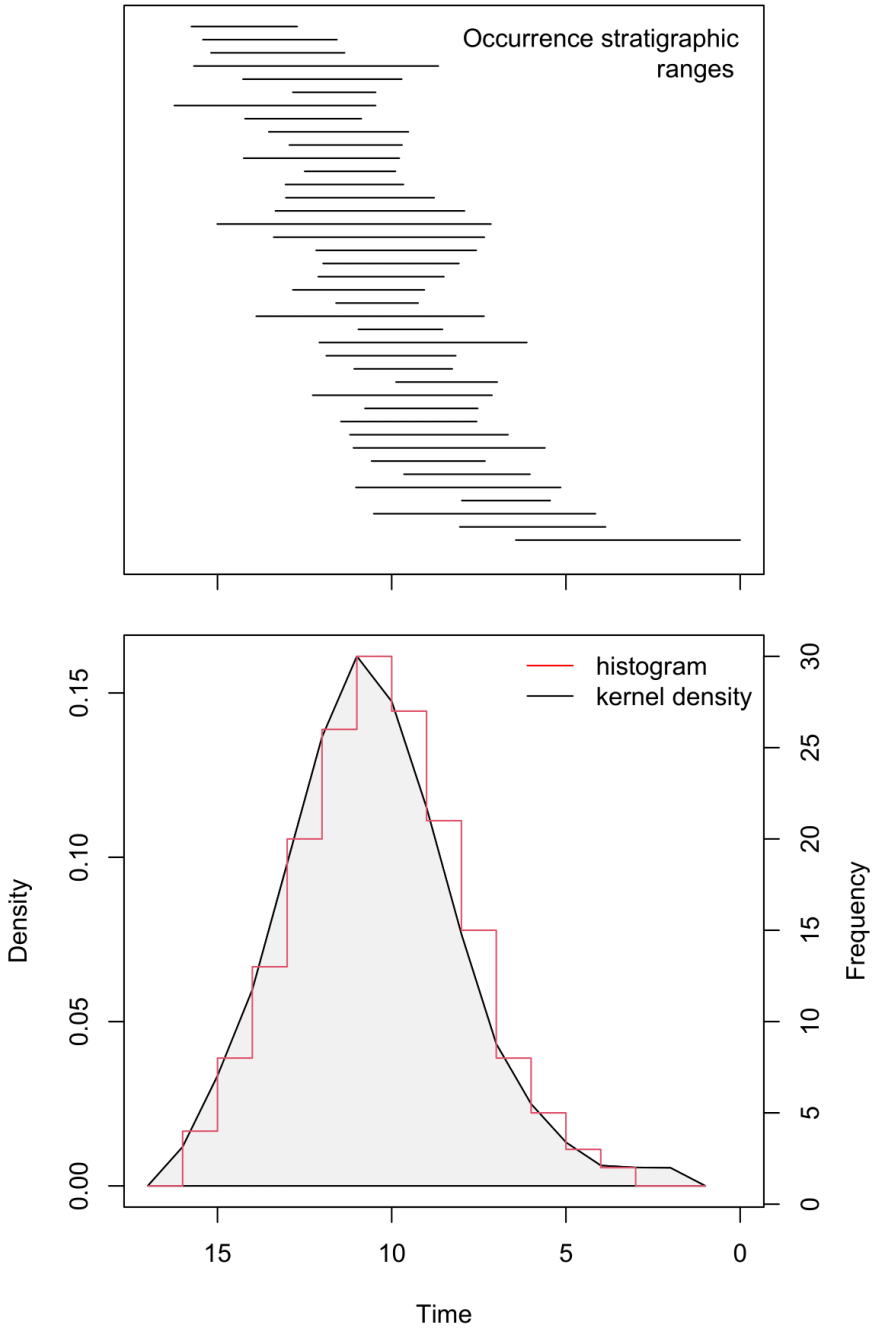


Fig. 2.2. **An idealised occurrence stratigraphic density distribution.** Using simulated occurrences (top), this theoretical taxon displays the idealised bell-curve (bottom) expected under the biological sequence of origination, population expansion, population decline and extinction. Time is in arbitrary units.

anomalies. Fossil occurrence age uncertainties may be treated as uniform probability density distributions on their point-wise ages (Silvestro et al., 2014a; Zhang et al., 2016). Combining the uniform density distributions for a set of taxon occurrences gives a composite density distribution of all possible pointwise observations of that taxon through time (Fig. 2.2). At face

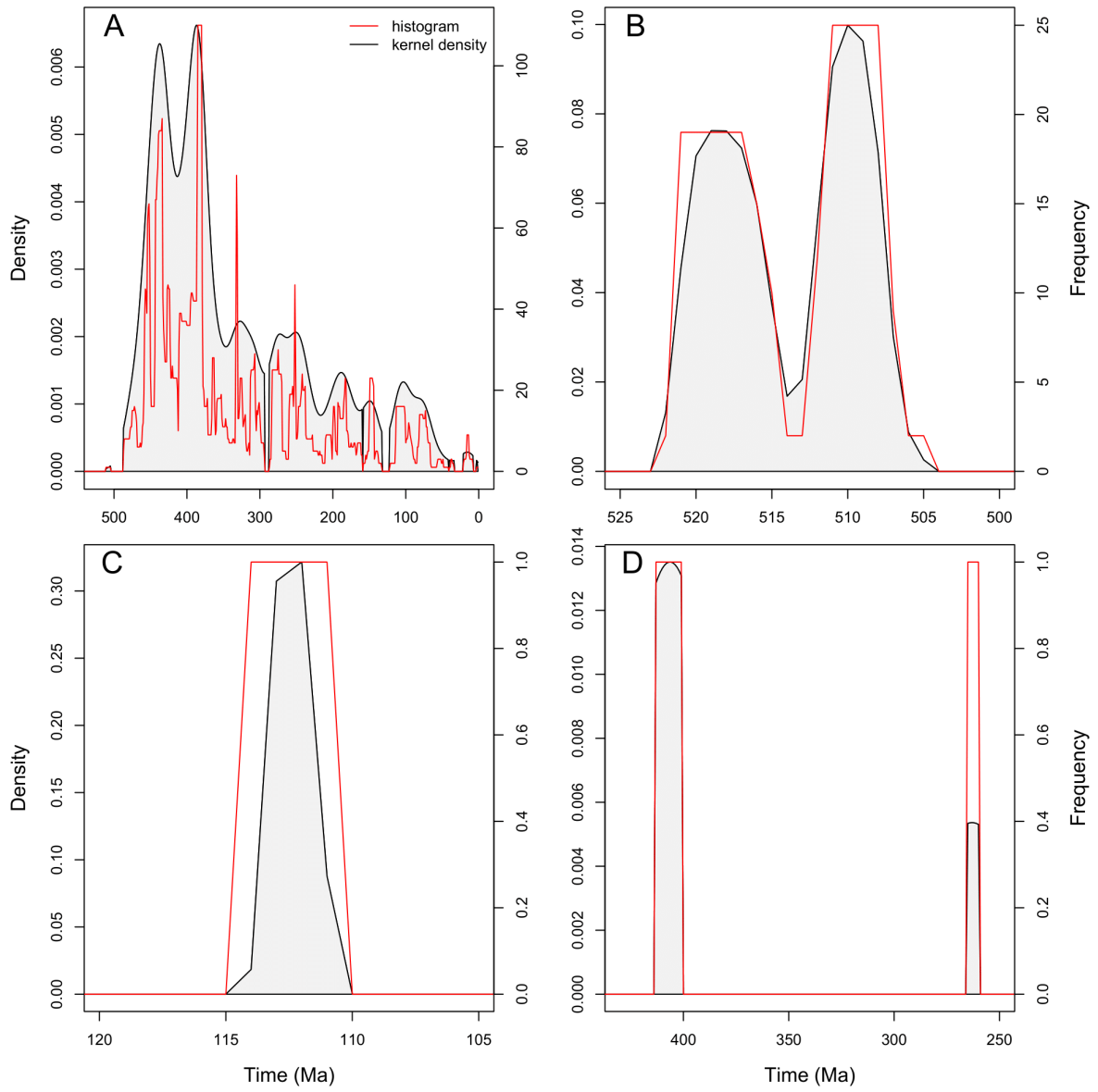


Fig. 2.3. Empirical stratigraphic density distributions of PBDB taxa. (A). A complex, multipeak distribution for the brachiopod *Lingula*, reflecting its frequent misuse throughout the fossil record for various linguliform taxa. **(B).** A bimodal distribution resulting from patchy sampling of the valid Cambrian arthropod taxon *Waptia*. **(C).** The expected density distribution for a stratigraphically plausible taxon, here the Cretaceous turtle *Caririemys*. **(D).** A bimodal density distribution which is stratigraphically suspect given the gap between the peaks and lack of intervening sampling, resulting from re-use of the name *Nanochilina* for unrelated taxa.

value, a density distribution represents how palaeontologists have assigned stratigraphically varied fossil observations with associated age uncertainties to a single morphospecies. The probability of fossil observations through time and so the idealised shape of their density distribution is expected to follow biological principles. A taxon originates in a discrete area, expand its range and population size, then declines through local extirpation until its total

extinction, producing a bell curve (Fig. 2.2) with symmetry dependent on the rapidity of expansion, extirpation, and extinction (Foote, 2007; Foote et al., 2007; Liow et al., 2010b; Silvestro et al., 2014). This unimodal distribution may conceivably become multimodal in several ways. Firstly, boom-bust cycles in population size may be detectable through intensive bed-wise sampling in finely resolved or recent sedimentary sections (Kidwell and Flessa, 1995; Kidwell, 2015), but it is assumed here that such signals are unlikely to manifest in temporally coarser macrofossil occurrence databases. Secondly, heterogeneous sampling through time may introduce peaks and troughs into an idealised observation record (Barido-Sottani et al., 2019). Thirdly, density distributions for stratigraphically distinct taxa may be conflated into a single multimodal density record due to their superficially similar morphologies (an invalid morphospecies or wastebasket taxon; Fig. 2.3A) or through taxonomic misclassification of individual occurrences (Fig. 2.3D). Multimodality arising from sampling effects versus taxonomic error is distinguished based on the assumption that, as sampling probability increases exponentially with time (Stadler, 2010), short durations between peaks are more likely to represent sampling variation (Fig. 2.3B) while longer durations are more likely to represent conflation of distinct taxonomic records (Fig. 2.3D), with average genus longevities within a taxon's parent clade helping to distinguish between these cases. Taxonomic misidentifications resulting in stratigraphic anomalies are expected to be concentrated towards the ends of a taxon's stratigraphic range, leading to long tails in its density distribution (Lazarus et al., 2012). Alternatively, long tails may arise from high stratigraphic uncertainty in a handful of occurrences. As such, the density distribution of macrofossil observations provides information on the stratigraphic plausibility and consistency of a morphospecies and so a means of detecting outliers.

2.2.4 Stratigraphic outlier detection using occurrence densities

Using the above treatment of a taxon's set of stratigraphically uncertain fossil occurrences as a density distribution of all their possible observations, I develop two methods for detecting stratigraphically anomalous occurrences: interpeak thresholding and Pacmacro. Interpeak thresholding ascertains whether interpeak durations in a multimodal occurrence density distribution more likely reflect incomplete sampling versus taxonomic error. As short overlaps in occurrence ranges or minor fluctuations in sampling may produce small peaks in density (Fig. 2.4A), a density distribution is first smoothed within a local window to reduce noise, then the local peak (i.e., a density value in excess of its immediate neighbours) taken as significant

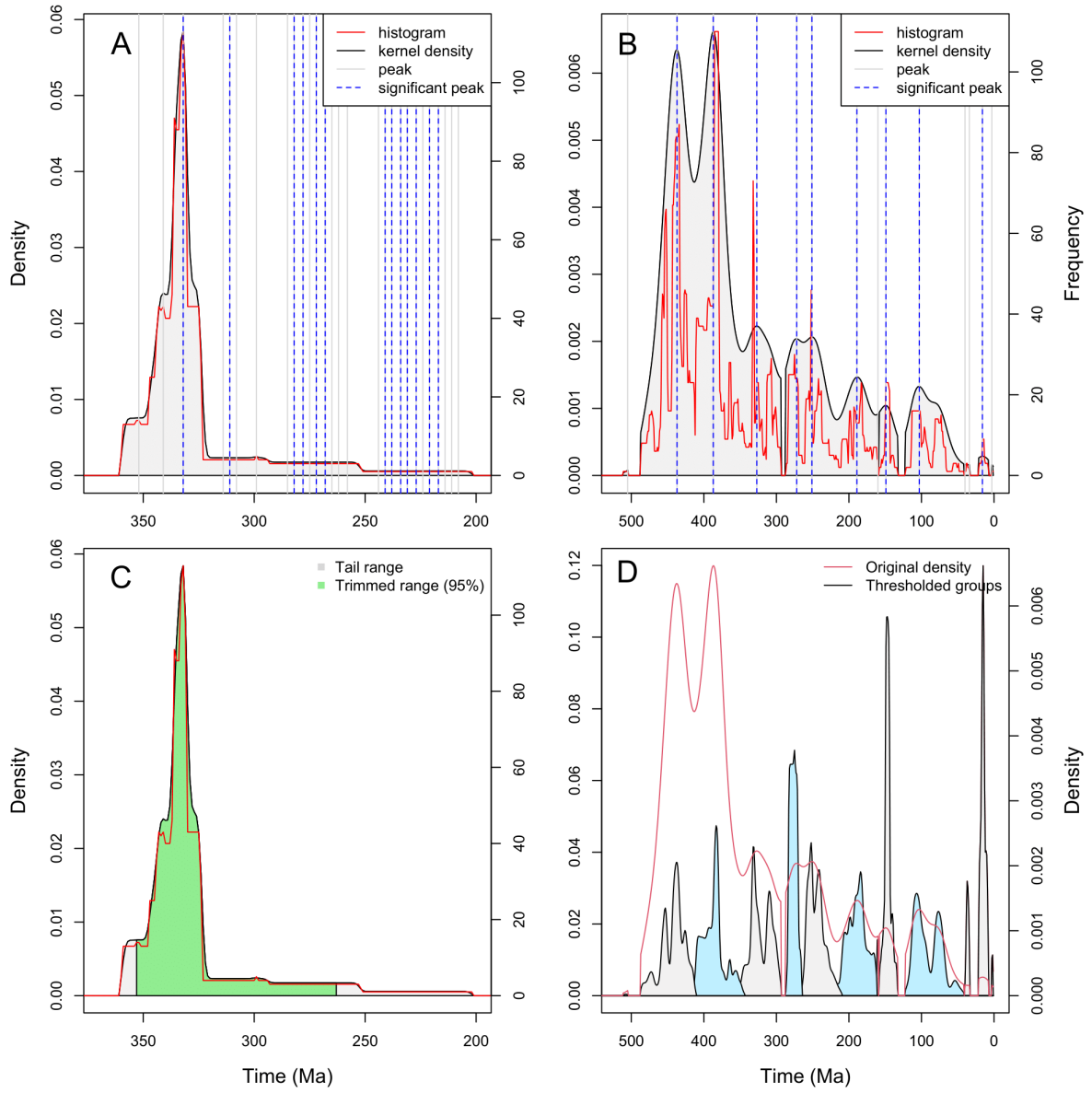


Fig. 2.4. Application of detection methods to empirical occurrence density distributions. (A). Density distribution, local peaks and significant peaks for the Carboniferous coral *Lithostrotion*. **(B).** Density distribution, local peaks and significant peaks for the brachiopod *Lingula*. **(C).** Tail detection and trimming by Pacmacro for *Lithostrotion*, with the trimmed range highlighted in green. **(D).** Interpeak thresholding for *Lingula*, dividing its occurrences into nine stratigraphically coherent groups of linguliform morphologies.

if its value exceeds the mean plus standard deviation of the local window (Fig. 2.4B). If one significant peak is detected, the distribution conforms to the unimodal expectation and all occurrences are considered valid. For multimodal distributions, interpeak durations are sequentially examined forwards in time with an interpeak duration below a threshold gap length taken to indicate a peak split by poor sampling. Otherwise, the peaks are considered to represent

separate taxa and their separation point in time defined as the interpeak nadir. This equates to dividing a stratigraphically suspect morphospecies into a successive set of stratospecies. While earlier palaeontologists have advocated for abandonment of stratospecies and chronospecies concepts in favour of character-based definitions (Forey et al., 2004), this is within the content of formal phylogenetic analysis and such concepts are nonetheless useful to elucidate different speciation modes within a phylogeny (Pardo et al., 2008; Silvestro et al., 2018).

As outlined above, the gap threshold value may be chosen based on the expected stratigraphic duration for the taxon in question. Careful choice is critical here: too broad and anomalies will not be detected, too short and valid morphospecies will be erroneously subdivided, including in the likely rare cases where population-level dynamics manifest as peaks and troughs in occurrence density distributions. As interpeak thresholds should relate to the durations of entire taxa on macroevolutionary timescales, however, these are expected to encompass peaks and troughs present at population-level timescales. In turn, average taxon durations provide one potential means of selecting a suitable interpeak threshold (see 2.2.5), although this will then depend upon the taxonomic composition of an occurrence dataset and how well the sampled taxon durations reflect true taxon durations. Similarly, the interpeak threshold may also vary through time with coarseness of the different chronostratigraphic intervals on which occurrence ages are based.

The other method, Pacmacro, is analogous to Pacman profiling, used for detecting stratigraphic errors in microfossil occurrence data from sediment cores (Lazarus et al., 2012). Pacman profiling scrutinises the density of microfossil occurrences versus core depth (i.e., time) to minimise errors at the tails of a range, for example due to reworking of fossils in the sediment or through conflation of ancestors and descendants within anagenetic lineages. As Pacman profiling is linked to well-constrained age models, microfossil occurrences can be treated as point-wise observations rather than bearing stratigraphic uncertainties and a simple percentage trim is applied to the tails of a taxon density distribution in the core. Under the assumption that errors will be primarily distributed towards the ends of the distribution, this procedure preferentially targets the noisy tail portions of a set of occurrences, naturally improving the signal to noise ratio (Lazarus et al., 2012). Using a percentage trim has the further benefits of locally adapting to the density distribution at the tail ends of data and freedom from any assumptions about the shape of distribution itself (Lazarus et al., 2012). A non-parametric approach is preferable as macrofossil density distributions show varied shapes

(Fig. 2.3) but applying a percentage trim directly may be inappropriate as macrofossil occurrences are not pointwise observations. Instead, Pacmacro flags anomalously long tails in a stratigraphic density distribution based on the proportion of the total stratigraphic duration they represent, then trims the tails from the taxon range (Fig. 2.4C). This procedure is implemented in the function `pacmacro_ranges()`, with testing of Pacmacro with tail proportions of 1% to 90% followed by breakpoint analysis on the relationship between tail proportion and the numbers of flagged taxa using the the *segmented* R package (Vito and Muggeo, 2008) highlighting 5% as a suitably conservative tail density proportion (Fig. S2). Under this conservative tail density proportion, stratigraphic proportions of 40 – 60% appear suitable for detecting anomalously long tails. Ranges trimmed by Pacmacro can then be supplied to `flag_ranges()` to identify anomalous occurrences, and suspicious density distributions flagged by either Pacmacro or interpeak thresholding plotted for further inspection using the function `plot_dprofile()`.

2.2.5 Application to the PBDB

Here, I apply my data harmonisation and stratigraphic flagging procedures to the PBDB to demonstrate how these protocols can be applied and documented in a transparent, reproducible, and standardised manner, and to examine the structure and prevalence of anomalies in the database. The entire PBDB was downloaded on 15/06/21 and occurrence chronostratigraphy updated to GTS 2020 using `chrono_scale()`. Next, formatting irregularities were scanned for using `check_taxonomy()`. Suprageneric synonyms were identified and resolved manually due to the higher frequency of false positives, genus-level synonyms below a conservative q-gram threshold of 0.2 automatically resolved as the more frequent name, and inconsistent higher taxonomic schemes automatically resolved using default settings. The Sepkoski Compendium, a database of stratigraphic ranges for >36,000 Phanerozoic marine genera (Sepkoski, 2002), was used for anomaly flagging. While this precluded checking of terrestrial occurrences and plants, marine animals comprise the majority of PBDB entries, and the Sepkoski Compendium remains taxonomically useful despite some accumulated errors as it was based on secondary literature such as the Treatise of Invertebrate Paleontology, The Compendium also uses interval-based epoch to substage-level dating which could be to a modern chronostratigraphic standard, although this comes with the possibility that some of its age ranges could be invalid compared to the modern literature. The Sepkoski Compendium

was downloaded using the `fetch()` function of the *chronosphere* R package (Kocsis and Raja, 2019). Minor typographical errors in interval notations were corrected manually without any alteration to the interval ranges themselves, chronostratigraphy updated to GTS 2020 using `chrono_scale()`, then spelling errors and higher taxonomy checked as above. Finally, the Sepkoski Compendium was appended to the cleaned PBDB dataset and `check_taxonomy()` reapplied to align the higher taxonomy between both databases.

Stratigraphic anomalies were flagged against the Sepkoski Compendium using `flag_ranges()`, then collection ages revised by `resolve_ranges()` with the default consensus threshold of 75%. Pacmacro was applied to detect extended stratigraphic ranges using `pacmacro_ranges()`, with the default tail density proportions of 5% and a tail stratigraphic proportion of 40%. Further anomalous occurrences were then identified using `flag_ranges()` and the Pacmacro-trimmed stratigraphic ranges. Finally, the stratigraphic consistency of taxonomic names was queried using `threshold_ranges()`. Mean and median class-wise genus durations were calculated from the PBDB and the Sepkoski Compendium, with their distributions (Fig. S3) indicating 15 million years as a suitably relaxed default for cases where class-specific durations were unavailable. To determine the potential impact of these data cleaning procedures on empirical palaeobiological analysis, range-through diversity, and second-for-third speciation and extinction rates were calculated at 5 Ma intervals through the Phanerozoic using the `divDyn()` function from the *divDyn* R package (Kocsis et al., 2019).

2.3 Results

2.3.1 Data imputation errors

The `check_taxonomy()` function flagged 67,629 name formatting irregularities in the 1,526,026 database entries (Table 2.1). Flagged irregularities became increasingly common as taxonomic level decreased, arising at higher levels from PBDB-specific formatting for missing higher taxonomy (e.g., NO_CLASS_SPECIFIED), and at genus level from frequent inclusion of bracketed subgenus names, rather than due to the presence of taxonomic modifiers. The majority of suprageneric synonyms were taxonomically distinct (e.g., Homocrinidae –

Table 2.1. Counts of data imputation anomalies in the PBDB

	Phylum	Class	Order	Family	Genus
Non-letter characters	2747	15,413	20,601	13,276	7172
Incorrect word count	0	0	0	0	7020
Potential synonyms	2	2	38	210	885
Cross-rank homonyms	3	5	6	3	11
Inconsistently classified	0	1	3	32	199

n = 1,526,026

Holocrinidae), with q-gram distances typically > 0.2 (i.e., < 80%) similarity. Otherwise, synonyms mostly reflected inconsistent spelling or use of subclades or superclades at formal taxonomic ranks (e.g., Pyrotheridae – Pyrotheriidae, Bothriolepididae – Bothriolepidae). While these latter cases are not true synonyms, it is still inconsistent to use, for example, a subclass classification at the class level when the class level classification is also present (e.g., Actinopteryi – Actinopterygii). The number of genuine synonyms rose substantially at the genus level, arising primarily from spelling errors and assonance (e.g., *Allonnia – Allonia*, *Sichuanolenus – Szechuanolenus*, *Drepanochilus – Drepanocheilus*). The PBDB dynamically generates a coherent taxonomic scheme, so cases of inconsistent higher classifications were unsurprisingly infrequent at just 235 instances, the majority of which were homonyms between distant clades or instances of missing higher classifications for some occurrences.

2.3.2 Anomaly resolution using the Sepkoski Compendium

Prior to resolving occurrence ages collection-wise, 24.6% of marine PBDB taxa with entries in the Sepkoski Compendium had stratigraphic distributions concordant with their reference ranges. `0R0` anomalies were most prevalent (Table 2.2), indicating frequent range overextension by their occurrence records. Taxa fully outside of their ranges (`00R`, `R00`) were roughly equal while `0R1` anomalies were more common than `1R0` anomalies. Conversely the proportion of valid occurrences was substantially greater than the proportion of valid taxa, and occurrence-wise `0R0` anomalies were least common (Table 2.2) which is expected given that occurrence age uncertainties should typically be smaller than taxon age ranges. Resolution using the Sepkoski Compendium substantially reduced taxon-wise and occurrence-wise error prevalence

Table 2.2. Prevalence of taxon-wise and occurrence-wise anomalies in the PBDB

	Error type prevalence (%)					
	00R	0R1	0R0	1R0	R00	1R1
Sepkoski						
Taxa (pre)	13.7	17.2	23.6	9.1	11.9	24.6
Taxa (post)	12.9	6.6	9.9	4.7	12.9	52.9
Occs (pre)	5.6	6.3	1.1	6.0	7.5	73.5
Occs (post)	3.6	1.3	0.1	1.3	5.2	88.5
Pacmacro						
Taxa	12.1	10.7	70.0	2.0	5.5	-
Occs	24.5	36.1	2.0	38.1	25.5	-

(Fig. 2.5A, 2.5B). The increase in the proportion of valid occurrences was modest, but the proportion of valid taxa more than doubled and `0R0` anomalies indicative of high age imprecision were virtually eliminated (Table 2.2).

Unexpectedly, taxon-wise `00R` anomalies decreased only slightly while `R00` anomalies increased slightly, despite substantial reductions in the prevalence of all other error types for both taxa and occurrences and the high stratigraphic plausibility of collections after revision. Of the 216,568 collections in the download, 68.4% went unchecked as they did not contain any taxa present in the Sepkoski Compendium, highlighting a major limitation of the reference database approach. Of those which could be checked, 10.8% did not meet the consensus threshold for revision and 2.4% retained their original age, while 86.5% had their age revised with the vast majority displaying 100% consensus in the range overlaps of constituent taxa present in the Sepkoski Compendium (Fig. S4). Taxon-wise FAD and LAD anomalies showed similar distributions (Fig. 2.5B) in contrast to occurrence-wise anomalies, where LAD anomalies were more prevalent overall but FAD anomalies in the range of a few million years were disproportionately more common (Fig. 2.5D).

While the consensus revisions returned less precise ages during the Carboniferous and Permian, they substantially improved median collection age precision throughout the Cambrian

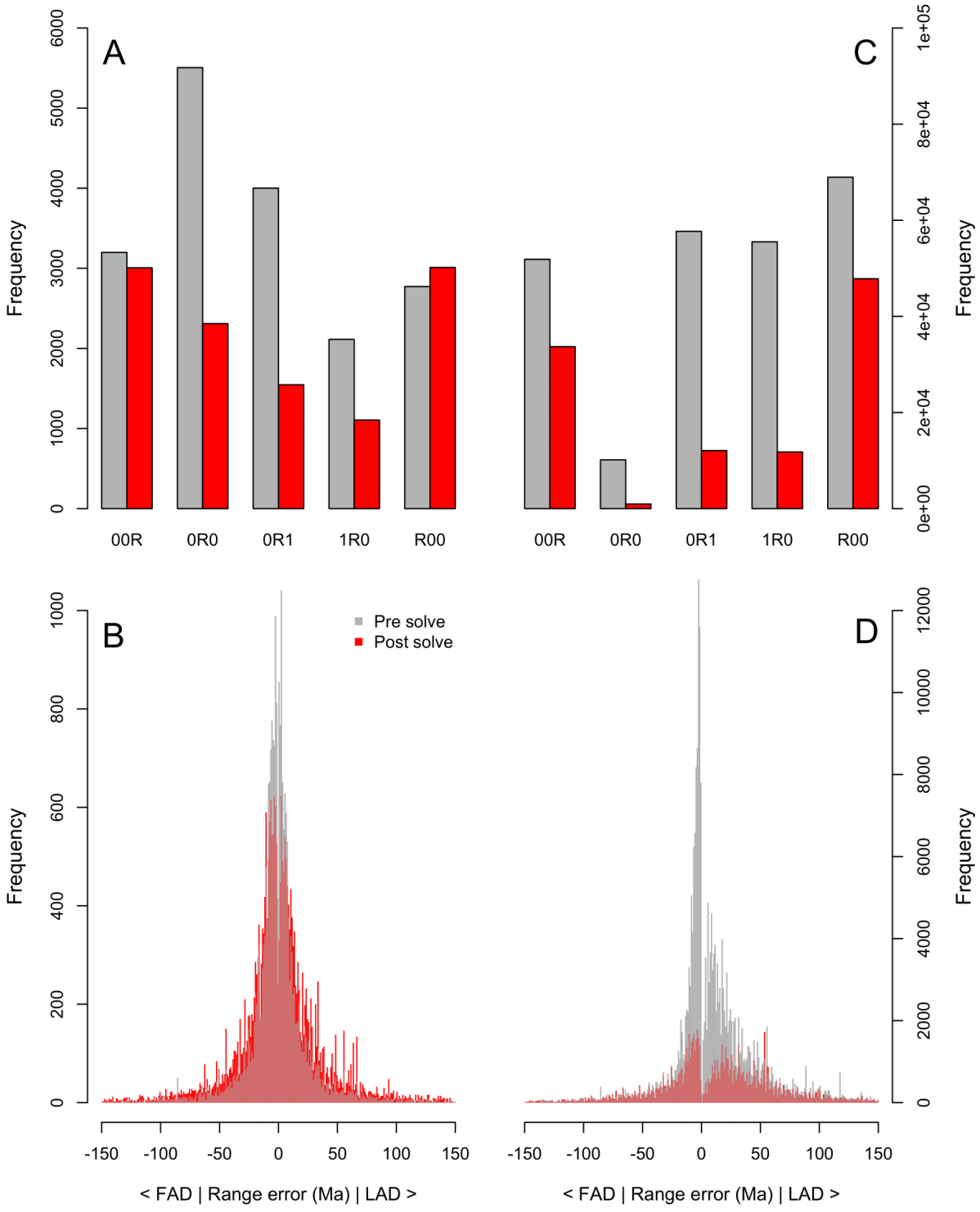


Fig. 2.5. Impact of collection age revision using the Sepkoski Compendium. (A). Taxon-wise anomaly prevalence. **(B).** Taxon-wise distribution of FAD and LAD anomalies. **(C).** Occurrence-wise anomaly prevalence. **(D).** Occurrence-wise distribution of FAD and LAD anomalies. Definitions of the error codes are in the main text (see 2.2.2).

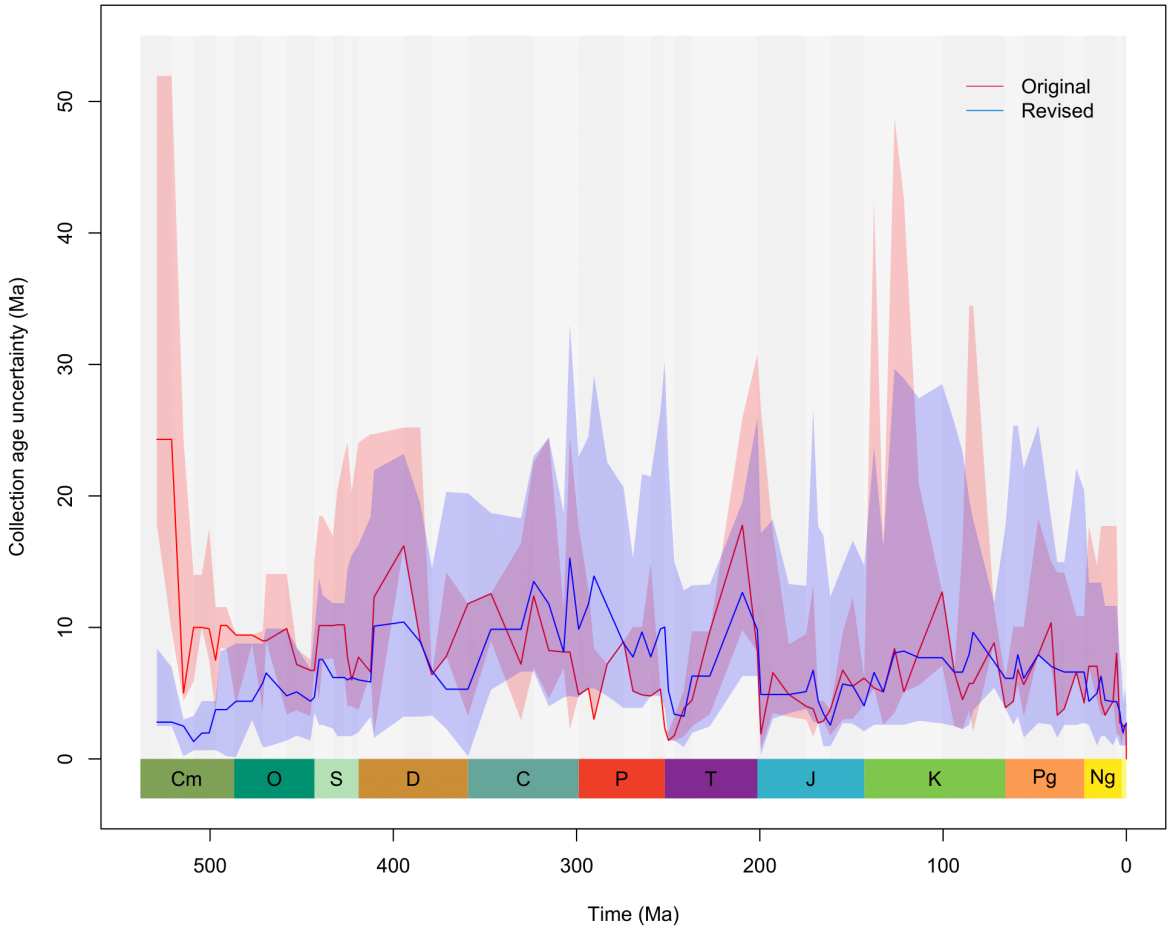


Fig. 2.6. Pre- and post-revision collection age uncertainty through the Phanerozoic. Collection age revision was conducted using the Sepkoski Compendium. Solid lines indicate the median and the shaded regions the interquartile range. Grey bars demarcate the epochs of the Phanerozoic.

to Devonian and during the Triassic and provided median age precision comparable to the unrevised ages throughout the remainder of the Phanerozoic in the absence of any lithostratigraphic information (Fig. 2.6). Age precisions showed a greater interquartile range after revision, indicating that the method performs variably between collections. Nonetheless, the 25% quartile on precision fell below that for unrevised age precision in virtually each time bin (Fig. 2.6), demonstrating that improved stratigraphic resolution was always achieved for at least some collections.

2.3.3 Density methods

2487 genera (3.7%) showed long tails in their occurrence density distributions, with distinct differences in taxon-wise and occurrence-wise anomalies when flagged against Pacmacro-

trimmed ranges. Occurrence-wise 0R0 anomalies were the least prevalent, with relatively even balances of 00R to R00 and 0R1 to 1R0 anomalies (Table 2.2). Consequently, 0R0 anomalies indicative of two-tailed stratigraphic density distributions were the most prevalent amongst taxa, reflecting the relatively even distribution of occurrence-wise FAD and LAD anomalies, but anomalies on taxon FADs (00R, 0R1) were disproportionately more common than anomalies on taxon LADs (Table 2.2). As an example, *Lithostrotion*, a colonial rugose coral abundant during the Visean (Early Carboniferous), has a density distribution contains multiple local peaks but just one true peak (Fig. 2.4A), marking it as stratigraphically coherent morphospecies. Rugose corals went extinct during the Permo-Triassic mass extinction (Wang et al., 2018), but records of *Lithostrotion* persist to the end of the Triassic (Fig. 2.4A). Anomalously long tails were successfully flagged for the genus and its range truncated to within the Palaeozoic by the 5% tail trim (Fig. 2.4C). While its range in the Sepkoski Compendium is more conservative (346 – 326 Ma), Pacmacro still removed highly erroneous portions of the occurrence record in this case, returning a range that is concordant for rugose corals and the family Lithostrotionidae (Carboniferous – Late Permian; Wang et al., 2018).

Using class-specific interpeak thresholds, 10,733 genera were split stratigraphically (16.1%). The number of splits showed a positive relationship with original genus duration (Fig. S5), demonstrating that longer stratigraphic durations are less likely to show plausible occurrence density distributions, and more likely to be split into greater numbers of groups. Several examples are highlighted here, chiefly *Lingula*, a brachiopod reported throughout the Phanerozoic and a so-called 'living fossil' given its supposed antiquity. In reality *Lingula* is a recent genus, with older occurrences representing multiple different genera which convergently evolved shell morphologies adapted for burrowing (Emig, 2003). This is reflected by its occurrence density record which contains multiple significant peaks (Fig. 2.4B), split by the thresholding method into nine stratigraphically distinct groups of linguliform occurrences (Fig. 2.4D). Similarly, the two peaks for *Nanochilina* are separated by nearly 250 Ma (Fig. 2.4D) and demonstrates a case where a homonym has resulted in a stratigraphically implausible density record. Conversely, the Cambrian arthropod *Waptia* bears two peaks in its density record (Fig. 2.4B) which fall within the interpeak threshold, indicating a stratigraphically coherent, taxonomically valid record split by sampling artefacts. This is supported by its limited fossil distribution but distinctive morphology and apomorphies, with identifiable specimens

derived primarily from the Burgess Shale of British Columbia (Taylor, 2002), and from older occurrences in the Marjum Shale and Wheeler Formation of Utah (Briggs et al., 2008).

2.4 Discussion

2.4.1 Outliers and anomalies in PBDB occurrences

The even balance of taxon-wise FAD and LAD anomalies (Fig. 2.5B) shows that there is no particular bias in the PBDB towards either mode of range overestimation at the genus level. Instead, the disproportionate frequency of small magnitude occurrence-wise FAD anomalies versus the greater total prevalence of LAD anomalies over a much greater range of magnitudes (Fig. 2.5D) suggests overestimation of FADs relates more strongly to stratigraphic imprecision, and overestimation of LADs to taxonomic misidentification. This is supported by independent flagging of outliers against Pacmacro-trimmed ranges. The even distributions of occurrence-wise FAD and LAD anomalies flagged against Pacmacro-trimmed ranges and the resulting prevalence of 0R0 anomalies and two-tailed anomalies on taxon stratigraphic density distributions is congruent with the lack of bias towards predominant FAD or LAD overestimation. Similarly, the greater prevalence of occurrence-wise FAD anomalies relative to the Sepkoski Compendium is matched by the disproportionate frequency of FAD anomalies relative to Pacmacro-trimmed ranges.

Taxonomic representation in the PBDB is highly disproportionate, so class-wise numbers of flagged genera were normalised by class-wise genus diversity to investigate proportional error rates under each detection method. The distribution of error proportions versus clade size indicates that comparison against the Sepkoski Compendium is the most stringent detection procedure and Pacmacro profiling the most conservative in its approach to identifying anomalies (Fig. S6, Table S1, S2). Under each method, the highest proportional error rates typically occur in classes with relatively small numbers of genera, reflecting the patchiness of their fossil records. Plants in particular show stratigraphically suspect occurrence densities with a high prevalence of long stratigraphic tails and splitting by interpeak thresholding (Table S2), corroborating previous concerns over their poor representation in the PBDB (Cleal and Thomas, 2010; Silvestro et al., 2015). Nonetheless, these relatively poorly sampled clades contribute only a small amount of the overall anomalies in the database. More concerning is

the repeated flagging within well-sampled, taxonomically diverse clades comprising the bulk of PBDB occurrences which, despite their lower proportion of anomalies, contribute the most to the overall error pool. Major clades containing thousands of genera, including gastropods, chondrichthyans, ostracods, brachiopods, bivalves, cephalopods, osteichthyans and tetrapods show range error proportions >50% prior to resolution against the Sepkoski Compendium. Collection age resolution had a positive impact on class-wise error proportions, but these remain in the range of 20-50% for major clades. Density methods also reveal moderate class-wise error proportions, with division of stratigraphically suspect occurrence density distributions by interpeak thresholding of 20-30%, and anomalous stratigraphic tails in the range of 2-9% (Table S1, S2). Long tails and range anomalies call into question the validity of individual occurrence identifications while the impact of interpeak thresholding calls into question the validity of entire genera with stratigraphically suspect occurrence density distributions. Conversely, the discrepancy between the number of taxon-wise and occurrence-wise anomalies flagged against the Sepkoski Compendium suggests that a relatively small number of occurrence-wise anomalies are responsible for the alarming prevalence of taxon-wise range anomalies in the PBDB, while most occurrences are otherwise valid.

2.4.2 Phanerozoic diversity dynamics and error structure

Temporal structuring in discrepancies between original and revised FADs and LADs (Fig. 2.7) highlights forward-smearing of a number of Cambrian and Ordovician-aged collections and of variably aged Palaeozoic FADs and LADs to revised ages coincident with the Permo-Triassic boundary, back-smearing of FADs at approximately 66 Ma coincident with the end-Cretaceous mass extinction, back-smearing of LADs around 35 Ma at the end of the Eocene, and back-smearing of present-day LADs. The magnitudes of some of these revisions are on the order of hundreds of millions of years, suggesting that frequent misidentification of age-diagnostic taxa in younger assemblages may produce ‘relict’ collections of superficially greater antiquity. Bands of back-smearing appear to coincide at least partially with major phases of taxonomic turnover, highlighting their effect as watersheds between ecologically discrete assemblages on broader temporal scales (Muscente et al., 2018; Rojas et al., 2021). For example, there is virtually no revision of collection ages across the Permo-Triassic boundary, highlighting how the event creates a clear division between Palaeozoic and Mesozoic morphospecies and assemblages. This watershed effect may be responsible for the marked grid-wise pattern of age revision in the post-Palaeozoic portion of the dataset, with the increasing taxonomic

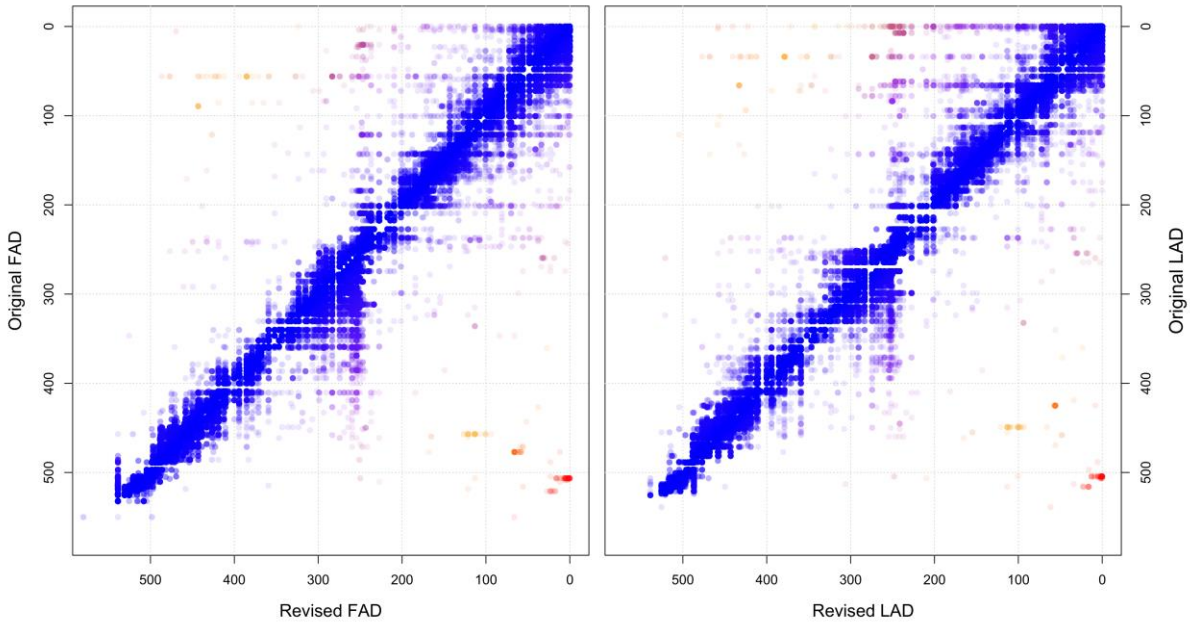


Fig. 2.7. Pre- and post-revision collection FADs and LADs through the Phanerozoic. Warmer colours indicate increasing discrepancy between original and revised ages.

misidentification of increasingly prevalent 'modern' morphospecies driving the greater degree of age restructuring. This could also conceivably be tied to variation in stage and substage length throughout the Phanerozoic, with coarser bin lengths inducing greater age revisions. Consequently, suitable thresholds for the interpeak method and tail proportions for Pacmacro may also range through time as well as between clades due to varying precision of Phanerozoic stages.

There are clear differences in the tempo of diversification throughout the Phanerozoic due to taxonomic and stratigraphic revision (Fig. 2.8). Most notably, diversity accrued more rapidly in the wake of the Late Devonian mass extinction before a trend of Late Permian decline up to the end-Permian mass extinction in contrast to the pre-cleaning trend of continued increase, although the magnitude of the latter extinction is muted by the use of 5 Ma bins instead of more finely resolved chronostratigraphic stages where the difference in pre-and post-extinction diversity are clearer. While the results themselves are inaccurate due to the known impact of spatial sampling bias and regional heterogeneity on global estimates of diversity, speciation and extinction rates through geological time (Close et al., 2020; Flannery Sutherland et al., 2022b), they nonetheless demonstrate the impact of taxonomic and stratigraphic revision on estimates of richness and rates. Set against the context of the entire PBDB, the anomalies

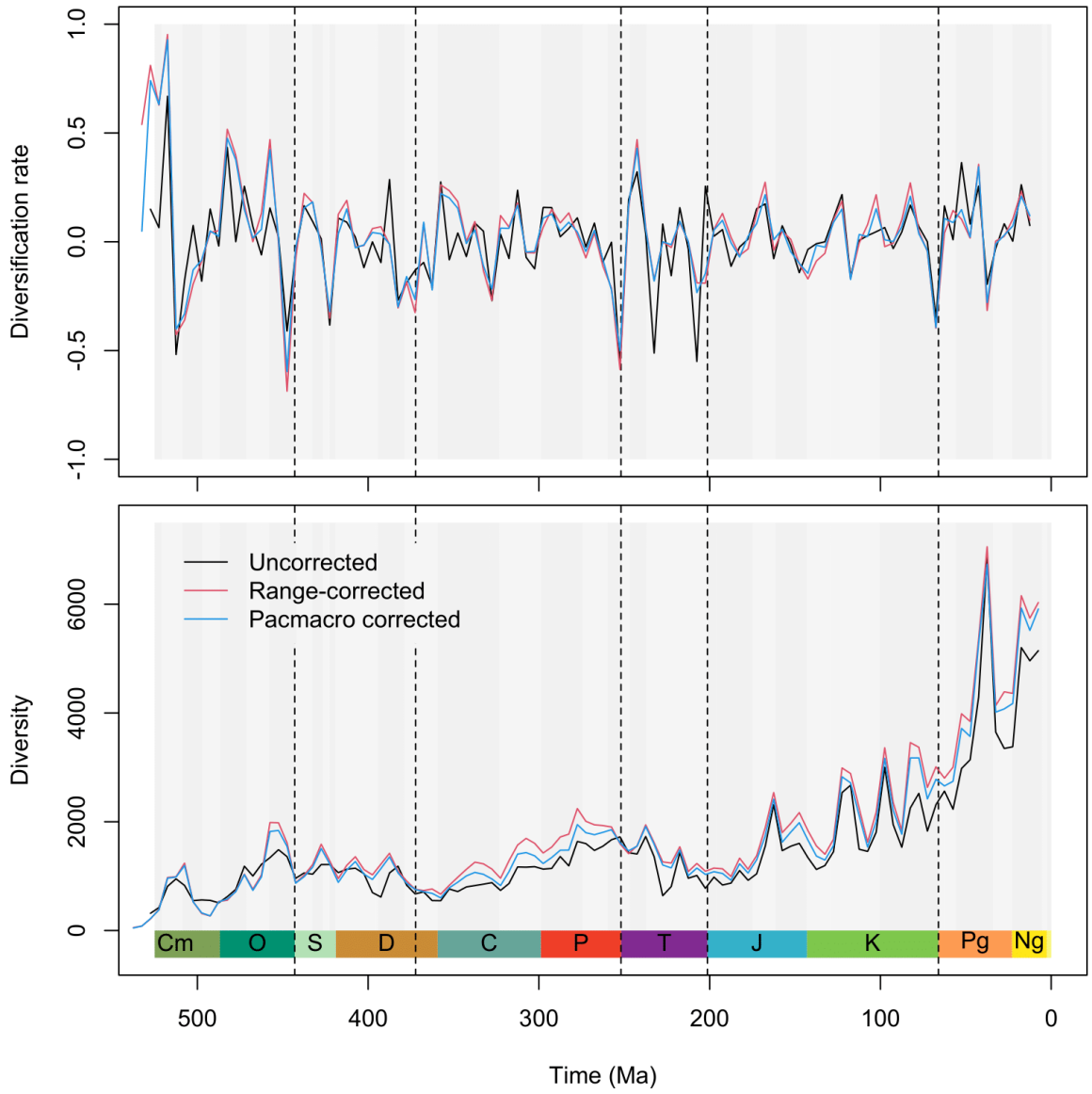


Fig. 2.8. Pre- and post-correction diversification rate and diversity through the Phanerozoic. Black lines are uncorrected, red lines from data with stratigraphic ages corrected using the Sepkoski Compendium, and green lines from data additionally split using interpeak thresholding. Dashed lines mark the Big Five mass extinctions. Grey bars demarcate the epochs of the Phanerozoic. All measurements were conducted at the genus level.

detected and resolved here using purely statistical methods have only a modest impact on patterns of diversity (Fig. 2.8). This corroborates previous analysis of the impact of taxonomic revision on PBDB data (Wagner et al., 2007) and comparison of broad patterns of diversity before and after taxonomic revision of the Sepkoski Compendium additionally found that higher level trends remained stable despite a high prevalence of errors, as their distribution throughout the entire database was random (Adrain and Westrop, 2000). Recovered error prevalence within individual clades may differ if detection were followed by manual vetting

by taxonomic specialists, however, and so the impact of error resolution on downstream analysis is expected to be greater for smaller taxonomic subsets of fossil occurrence data.

2.4.3 Utility of gaps versus observations in stratigraphic range analysis

While the methods here focus on instances where taxon stratigraphic ranges are overextended by imprecisely dated or incorrectly identified occurrences, palaeontologists have traditionally focused on the more common problem of taxon stratigraphic ranges truncated by the Signor Lipps effect. Early methods for estimating complete taxon stratigraphic durations considered fossil occurrences at the section level, where occurrence horizons could be assigned pointwise ages. Under assumption that the probability of fossil recovery is uniform through its stratigraphic range, the subsequent distribution of fossil observations can be considered random with respect to their taxon's true stratigraphic range and gaps between observations modelled as exponentially distributed waiting times between observations events in a Poisson process, enabling calculation of point estimates and confidence intervals for taxon FADs and LADs (Strauss and Sadler 1987, 1989). The assumption of non-randomness is frequently violated in the geological record, leading to development of methods which relax or dispense with this assumption (Marshall 1994, 1997; Solow 2003; Solow and Roberts 2003).

With the rise of fossil occurrence databases, stratigraphic correction methods have been applied to the total fossil records of taxa to estimate their true stratigraphic durations (e.g., Powell 2007; Aguirre 2016; Calede 2020) and many have implementations available in R (e.g., Zaffos 2017; Claramunt 2022). As fossil occurrences bear stratigraphic age uncertainties, rather than displaying the precise pointwise ages needed to strictly define the gaps between their observation, this has prompted more recent development of Bayesian techniques which can incorporate these uncertainties, along with prior information on variable fossil recovery potential if available (Alroy 2014b; Silvestro et al., 2014a, 2014b; Wang et al., 2015). Their performance, however, is still predicated on the quality of the fossil occurrence data for a given taxon. Suspect fossil occurrences responsible for anomalously long stratigraphic durations will subsequently cause these correction methods to estimate still more erroneous FADs and LADs. As such the methods developed here for truncation of stratigraphic ranges can complement methods for the estimation of true stratigraphic durations by ensuring that they are not misled by poor quality data.

Methods developed to correct for the Signor Lipps effect may also provide inspiration for improved methods for detecting anomalous stratigraphic occurrences. Analogous to using the gaps between fossil observations, the interpeak method uses the gaps between occurrence sampling peaks in conjunction with a simple threshold to determine whether a lineage is a stratigraphically cohesive entity or otherwise. Explicit sampling models based on the Poisson process may allow the method to better adapt to the data comprising an occurrence density distribution and permit probabilistic tests between different thresholds, making the procedure more quantitatively robust compared to arbitrary selection of a threshold, or selection based on average taxon durations. While Pacmacro ignores the gaps in taxon stratigraphic records and instead focuses on their occurrences, incorporating their age uncertainties into the procedure as uniform distributions, those gaps nonetheless contain information on potential lineage durations and occurrences may not show uniformly distributed stratigraphic age uncertainties. Consequently, a Bayesian, model-based approach to Pacmacro may make better use of all available information in the underlying data.

2.5 Conclusions

Robust palaeobiological research is predicated on high quality data, yet techniques which address errors in fossil occurrences databases are underdeveloped. I add to existing tools for resolving errors in geographic coordinate data with methods for standardisation and cleaning of taxonomic names, and for the unique challenges presented by the stratigraphic component of fossil occurrence data within the new R package *fossilbrush*. The `check_taxonomy()` function does not rely on any external databases for checking. As such it can be applied to any dataset with taxonomic information present, and with any number of levels in the recorded taxonomic hierarchy and so its utility extends beyond application solely to fossil occurrence datasets.

The most stringent error detection method utilises a reference database (here the Sepkoski Compendium) to assess entire assemblages of fossils, querying the validity of occurrences, taxa and collections. Collection ages can be effectively revised using plausible consensus ages of their taxa, often improving their stratigraphic precision in the process, but anomalies in taxon ranges relative to the Sepkoski Compendium are frequent in the PBDB. FAD anomalies may

relate to stratigraphic imprecision, while LAD anomalies may arise more from taxonomic misidentification, informing where future cleaning efforts based on expert knowledge should be targeted. I also provide conceptual advances on how occurrences records may be treated as observation densities which incorporate stratigraphic uncertainty, along with how the properties of these density distributions may be used to flag anomalous occurrences and stratigraphically suspect taxa.

Finally, it must be stressed that while these methods appear to function effectively and scale well to the challenges presented by a large occurrence database, the occurrence density methods rely on outlier detection, a statistical solution, rather than drawing upon expert knowledge where a definitive solution can be achieved. As such, they are best applied in concert, so that occurrences which are repeatedly flagged by each method may be confidently assessed as erroneous.

Chapter Three

Mitigating Spatial Sampling Bias during Fossil Biodiversity Analysis

Global diversity patterns in the fossil record comprise a mosaic of regional trends, underpinned by spatially non-random drivers and distorted by variation in sampling intensity through time and across space. Sampling-corrected diversity estimates from spatially standardised fossil datasets retain their regional biogeographic nuances and avoid these biases, yet diversity-through-time arises from the interplay of origination and extinction, the processes that ultimately shape macroevolutionary history. Here I present a subsampling algorithm to eliminate spatial sampling bias in a manner that produces fossil occurrence datasets suitable for investigating these diversification dynamics. I then re-examine the Late Permian to Early Jurassic marine fossil record, an interval spanning several global biotic upheavals that shaped the origins of the modern marine biosphere. I use PyRate and a new Bayesian method, mcmcDivE, to estimate origination and extinction rates and diversity trajectories which are corrected for temporal sampling bias, from fossil occurrence datasets corrected for spatial sampling bias. I find that diversification rates are regionally heterogenous even during biotic events and upheavals that manifested globally, highlighting the need for spatially explicit views of macroevolutionary processes through geological time.

Author contributions: This chapter is based on Flannery-Sutherland et al. (2022a) in Nature Communications, 13:2751. The work in this chapter was developed in collaboration with co-authors D. Silvestro and M. J. Benton. D.S. created the mcmcDivE diversity estimator described in this chapter. Comments were provided on a draft of the paper by D.S and M.J.B. The author is otherwise responsible for vast majority of the work in this research chapter.

3.1 Introduction

The fossil record is our only empirical sample of past biodiversity, providing a critical resource for understanding macroevolutionary and macroecological processes in deep time (Benton 2015). Numerous abiotic and biological drivers have been proposed to explain apparent patterns of fossil diversity, but it has long been recognised that these patterns are heavily distorted by uneven sampling intensity through time from geological biases that affect the temporal distribution of fossils and formations (Raup 1975; Smith 2001; McGowan and Smith 2008; Benson et al., 2010; Close et al., 2020b), differing preservation potential across organisms and environments, and heterogeneity in collection practice, reporting and even geopolitics (Hunter and Donovan 2007; Shaw et al., 2021; Raja et al., 2022). These factors are often interlinked and are also geographically variable in their manifestation (Dunhill et al., 2014, 2017). Therefore, the known fossil record is not only an incomplete sample of the total fossil record (itself a biased fraction of past diversity as a whole), but that incompleteness is also inconsistent through time and across space (Benson et al., 2021).

Significant attention has been devoted to correcting diversity estimates for temporal variation in sampling intensity (Alroy 2010; Close et al., 2018), but it has also been demonstrated that variation in the palaeogeographic distribution of the fossil record through time imposes an equally severe distortion on patterns of diversity even after correction for uneven sampling intensity (Vilhenna and Smith 2013; Close et al., 2017, 2020a, 2020b). Furthermore, fossil diversity is itself geographically variable due to the spatially non-random distribution of factors influencing species richness, for example the locations of reefs and epeiric seaways, or climatically structured latitudinal diversity gradients (Close et al., 2020a; Hagen et al., 2021). Recent studies of global fossil diversity have calculated point-wise diversity estimates from temporally standardised, spatially even subsamples of fossil data (e.g., Close et al., 2020a, 2020b), allowing the mosaic of global diversity to be decomposed into its regional components while accounting for the distortion induced by spatial sampling bias (Benson et al., 2021). Focusing on diversity alone, however, is limiting as it is ultimately a dynamic product of origination and extinction rates (Silvestro et al., 2018). Standing diversity, as determined by these rates at any point in time, then interacts with a spatiotemporally variable sampling rate to produce the fossil record. A drop in apparent diversity may result from a drop in origination or sampling rate just as much as from an increase in extinction rate, while a

relatively flat diversity trajectory could mask cryptic phases of turnover resulting from concurrent pulses of origination and extinction. A few studies have used geographically or latitudinally restricted datasets to gain regional views of diversification rates through time (Powell et al., 2015; Di Martino et al., 2018; Dunhill et al., 2018; Song et al., 2020; Condamine et al., 2021; Allen et al., 2023) and methods have been developed for formal spatial standardisation of fossil samples (Close et al., 2020a, Antell 2022), but there are no methods to generate fossil datasets that are spatially uniform through time, hindering investigation of diversity dynamics at different spatial scales and between different geographic regions.

In this chapter I present a subsampling algorithm to produce spatially standardised fossil occurrence datasets which remain consistent in their geographic position and extent through time, along with a method of calculating sampling-corrected diversity in a Bayesian framework to complement the inference of sampling-corrected origination, extinction and preservation rates in the software packages PyRate and LiteRate (Silvestro et al., 2014a, 2014b, 2019; Gjesfled et al., 2020; Koch et al., 2021). I apply these methods to a composite dataset of marine fossil occurrences spanning the Late Permian to Early Jurassic, an interval characterised by a dramatic backdrop of interlinked palaeogeographic shifts (Tanner 2017), climatic fluctuations (Chen and Benton 2012) and three extinction events: the end-Permian mass extinction (EPME; Chen and Benton 2012); the Carnian Pluvial Episode (CPE; Dal Corso et al., 2020); and the end-Triassic mass extinction (ETME; Dunhill et al., 2017), alongside a series of other less well understood biotic upheavals (e.g. the Smithian-Spathian Event and the Middle Norian Climate Shift; Goudemand et al., 2019; Kent and Clemmenson 2021). Global trends are heavily biased by the regional distribution of the fossil record and regional diversity dynamics themselves are strongly heterogenous even during supposedly global biotic events, indicating that global trends are not simply an upscaling of regional processes. This regional variability reflects the unique biogeographic histories of each study region, demonstrating the importance of geographic context in the assembly and transformation of biodiversity through deep time and highlighting how our view of the history of global diversity remains biased by the uneven spatial distribution of the fossil record.

3.2 Methods

3.2.1 Spatial standardisation workflow

To produce spatially standardised fossil occurrence datasets which remain geographically consistent through time, I designed a subsampling algorithm which enforces a consistent spatial distribution of occurrence data between time bins, while maximising data retention and permitting flexible regionalisation (Fig. 3.1). This method takes some inspiration from the spatial standardisation procedure of Close et al. (2020b), which returns subsamples of occurrence data with threshold minimum spanning tree (MST) lengths within a given time bin. An average diversity estimate can be taken from this 'forest' of minimum spanning trees, selecting only those of a target tree length to ensure spatially standardised measurements. It does not produce a single dataset across time bins, however; rather a series of discontinuous, bin-specific datasets which cannot then simply be concatenated as the spatial extents of each bin-specific forest are not standardised (despite each individual MST being so), even when MSTs pertain to a specific geographic region (e.g., a continent or to a particular latitudinal band). This prevents estimation of rates, because such analyses require datasets that span multiple time bins and remain geographically consistent and spatially standardised through the time span of interest. The workflow presented here overcomes this shortcoming and consists of three main steps.

1. First, the user demarcates a spatially discrete geographic area (herein the spatial window) and a series of time bins into which fossil occurrence data is subdivided. Occurrence data falling outside the window in each time bin are dropped from the dataset, leaving a spatially restricted subsample (Fig. 3.1A). Spatial polygon demarcation is a compromise between the spatial availability of data to subsample and the region of interest to the user but allows creation of a dataset where regional nuances of biodiversity may be targeted. Careful choice of window extent can even aid subsequent steps by targeting regions that have a consistently sampled fossil record through time, even if the extent of that record fluctuates. To account for spatially non-random changes in the spatial distribution of occurrence data arising from the interlinked effects of continental drift, preservation potential and habitat distribution (Close et al., 2020a), the spatial polygon may slide to track the location of the available sampling data through time. This drift is performed with two conditions. First, the drift is unidirectional so that the sampling of data remains consistent relative to global geography,

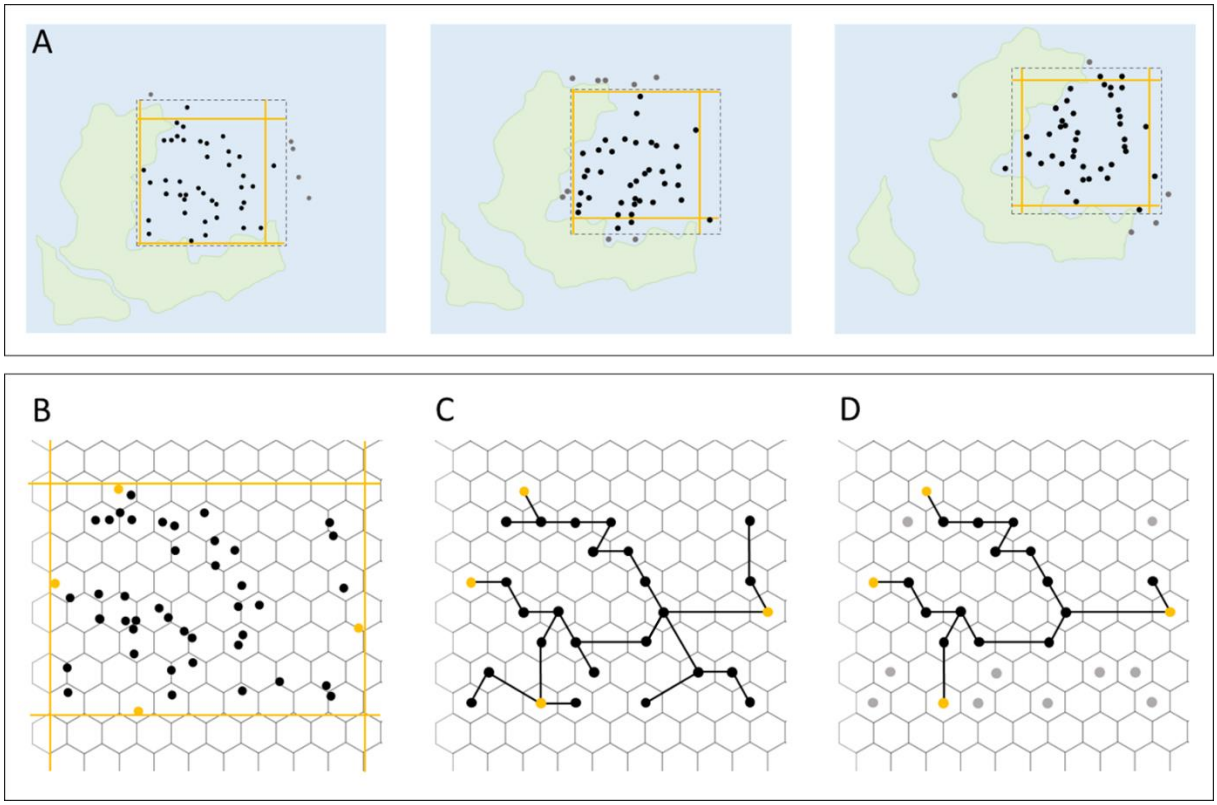


Fig. 3.1. Component steps of the spatial standardisation workflow. (A). A spatial window (dotted lines) is used to demarcate the spatial region of interest, which may shift in a regular fashion through time to track that region. Data captured in each window is clipped to a target longitude-latitude range (orange lines). **(B).** The data forming the longitude latitude extent is marked, then masked from further subsampling. **(C).** Data is binned using a hexagonal grid, the tally of occurrences in each grid cell taken, and a minimum spanning tree constructed from the grid cell centres. **(D).** The cells with the smallest amount of data are iteratively removed from the minimum spanning tree until a target tree length is reached.

rather than allowing the window to hop across the globe solely according to data availability and without biogeographic context. Second, spatial window translation is performed in projected coordinates so that its sampling area remains near constant between time bins, avoiding changes in window size that could induce sampling bias from the species-area effect, but still sampling the impacts of other spatial influences on diversification, for example increasing ocean basin size permitting increased regional diversity under the same effect.

2. Next, subsampling routines are applied to the data to standardise its spatial extent to a common threshold across all time bins using two metrics: the length of the minimum spanning tree (MST) required to connect the locations of the occurrences; and the longitude-latitude extent of the occurrences. MST length has been shown to measure spatial sampling robustly as

it captures not just the absolute extent of the data but also the intervening density of points, and so is highly correlated with multiple other geographic metrics (Close et al., 2017). MSTs with different aspect ratios may show similar total lengths but could sample over very different spatial extents, inducing a bias by uneven sampling across spatially organised diversity gradients (Close et al., 2017); standardising longitude-latitude extent accounts for this possibility. The standardisation methods can be applied individually or serially if both MST length and longitude-latitude range show substantial fluctuations through time. Data loss is inevitable during subsampling and may risk degrading the signals of origination, extinction and preservation. To address this issue, subsampling is designed to retain the greatest amount of data possible in two ways. During longitude-latitude standardisation, the range containing the greatest amount of data is preserved. During MST standardisation, occurrences are spatially binned using a hexagonal grid to reduce computational burden and to permit assessment of spatial density (Fig. 3.1B). The grid cells containing the occurrences that define the longitude-latitude extent of the data are first masked from the subsampling procedure so that this property of the dataset is unaffected, and then the occurrences within the grid cells at the tips of the MST are tabulated. Tip cells with the least data are iteratively removed (removal of non-tip cells may have little to no effect on the tree topology) until the target MST length is achieved (Fig. 3.1D), with tree length iteratively re-calculated to include the branch lengths added by the masked grid cells.

For both methods, the solution with the smallest difference to the target is selected and so both metrics may fluctuate around this target from bin to bin, with the degree of fluctuation depending upon the availability of data to exclude; larger regions that capture more data are more amenable to the procedure than smaller regions. Similarly, the serial application of both metrics reduces the pool of data available to the second method, although longitude-latitude standardisation is always applied first in the serial case so that the resultant extent will be retained during MST standardisation. Consequently, the choice of standardisation procedure and thresholds must be tailored to the availability and extent of data within the sampling region through time, along with the resulting degree of data loss. This places further emphasis on the careful construction of the spatial window in the first step. Threshold choice is also a compromise between data loss and consistency of standardisation across the dataset and so it may be necessary to choose targets that standardise spatial extent well for most of the temporal

range of a dataset, rather than imposing a threshold that spans the entire data range but causes unacceptable data loss in some bins.

3. Once the time-binned, geographically restricted data have been spatially standardised, the relationship between diversity and spatial extent is scrutinised (Figs. S7–S13). After standardisation, it is expected that residual fluctuations in spatial extent should induce little or no change in apparent diversity. Bias arising from temporal variation in sampling intensity may still be present, so diversity is calculated using coverage-based rarefaction (also referred to as shareholder quorum subsampling; Alroy 2010; Chao and Jost 2012; Chao et al., 2014), with a consistent coverage quorum from bin to bin (Figs. S14–S20). While coverage-based rarefaction has known biases, it remains the most accurate non-probabilistic means of estimating fossil diversity (Close et al., 2018). As such, it is the most appropriate method to rapidly assess the diversity of a region-level fossil dataset. The residual fluctuations in spatial extent may then be tested for correlations with spatially standardised, temporally corrected diversity. If a significant relationship is found, then the user must go back and alter the standardisation parameters, namely the spatial window geometry and drift, the longitude-latitude threshold, and the MST threshold. Otherwise, the dataset may be considered suitable for further analysis.

This spatial standardisation workflow was implemented in R with a custom algorithm, `spacetimestand()`, along with a helper function `spacetimewind()` to aid the initial construction of spatial window. `spacetimestand()` can then accept any fossil occurrence data with temporal constraints in millions of years before present and palaeocoordinates in decimal degrees. Spatial polygon construction and binning is handled using the *sp* library (Pebesma and Bivand 2005), MST manipulation using the *igraph* and *ape* R packages (Csardi and Nepusz 2006; Paradis and Schliep 2019), spatial metric calculation using the *sp*, *geosphere* and *GeoRange* R packages (Boyle 2017; Hijmans 2019), hexagonal gridding using the *icosa* R package (Kocsis 2021), and diversity calculation by coverage-based rarefaction using the `estimateD()` function from the *iNEXT* R package (Hsieh et al., 2020).

3.2.2 Data acquisition and cleaning

Fossil occurrence data for the Late Permian (260 Ma) to Early Jurassic (190 Ma) were downloaded from the PBDB on 28/04/21 with the default major overlap setting applied (an occurrence is treated as within the requested time span if 50% or more of its stratigraphic duration intersects with that time span), in order to minimise edge effects resulting from

incomplete sampling of taxon ranges within the study interval of interest (the Permo-Triassic to Triassic-Jurassic boundaries). Other filters in the PBDB API were not applied during data download to minimise the risk of data exclusion. Occurrences from terrestrial facies were excluded, along with plant, terrestrial-freshwater invertebrate, and terrestrial tetrapod occurrences (as these may still occur in marine deposits from transport), and 1712 occurrences from several minor and poorly represented soft-bodied invertebrate and planktonic phyla. Finally, non-genus level occurrences were removed, leaving 104,741 occurrences out of the original 168,124. Based on previous findings (Close et al., 2020b), siliceous occurrences were not removed from the dataset, despite their variable preservation potential compared to calcareous fossils. To increase the temporal precision of the dataset, occurrences with stratigraphic information present were revised to substage- or stage-level precision using a stratigraphic database compiled from the primary literature (available in published chapter electronic supplement). To increase the spatial and taxonomic coverage of the dataset, the PBDB data were supplemented by an independently compiled genus-level database of Late Permian to Late Triassic marine fossil occurrences (Song et al. 2018). Prior to merging, occurrences from the same minor phyla were excluded, along with a small number lacking modern coordinate data, leaving 47,661 occurrences out of an original 51,054. Absolute numerical first appearance and last appearance data (FADs and LADs) were then assigned to the occurrences from their first and last stratigraphic intervals, based on the ages given in A Geologic Timescale 2020 (Gradstein et al., 2020). Palaeocoordinates were calculated from the occurrence modern-day coordinates and midpoint ages using the plate rotation model of Getech PLC. Finally, occurrences with a temporal uncertainty greater than 10 million years and occurrences for which palaeocoordinate reconstruction was not possible were removed from the composite dataset, leaving 145,701 occurrences out of the original 152,402. At the time of writing, the *fossilbrush* R package (Flannery Sutherland et al., 2022b) was not complete and so unavailable for automated data cleaning and stratigraphic revision.

In the total dataset, the age uncertainty for occurrences was typically well below their parent stage duration, aside for the Wuchiapingian and Rhaetian where the mean and quartile ages are effectively the same as the stage length (Fig. S21). This highlights the chronostratigraphic quality of the composite dataset, particularly for the Norian stage (~ 18-million-year duration) which has traditionally been an extremely coarse and poorly resolved interval in Triassic-aged macroevolutionary analyses. Taxonomically, most occurrences are

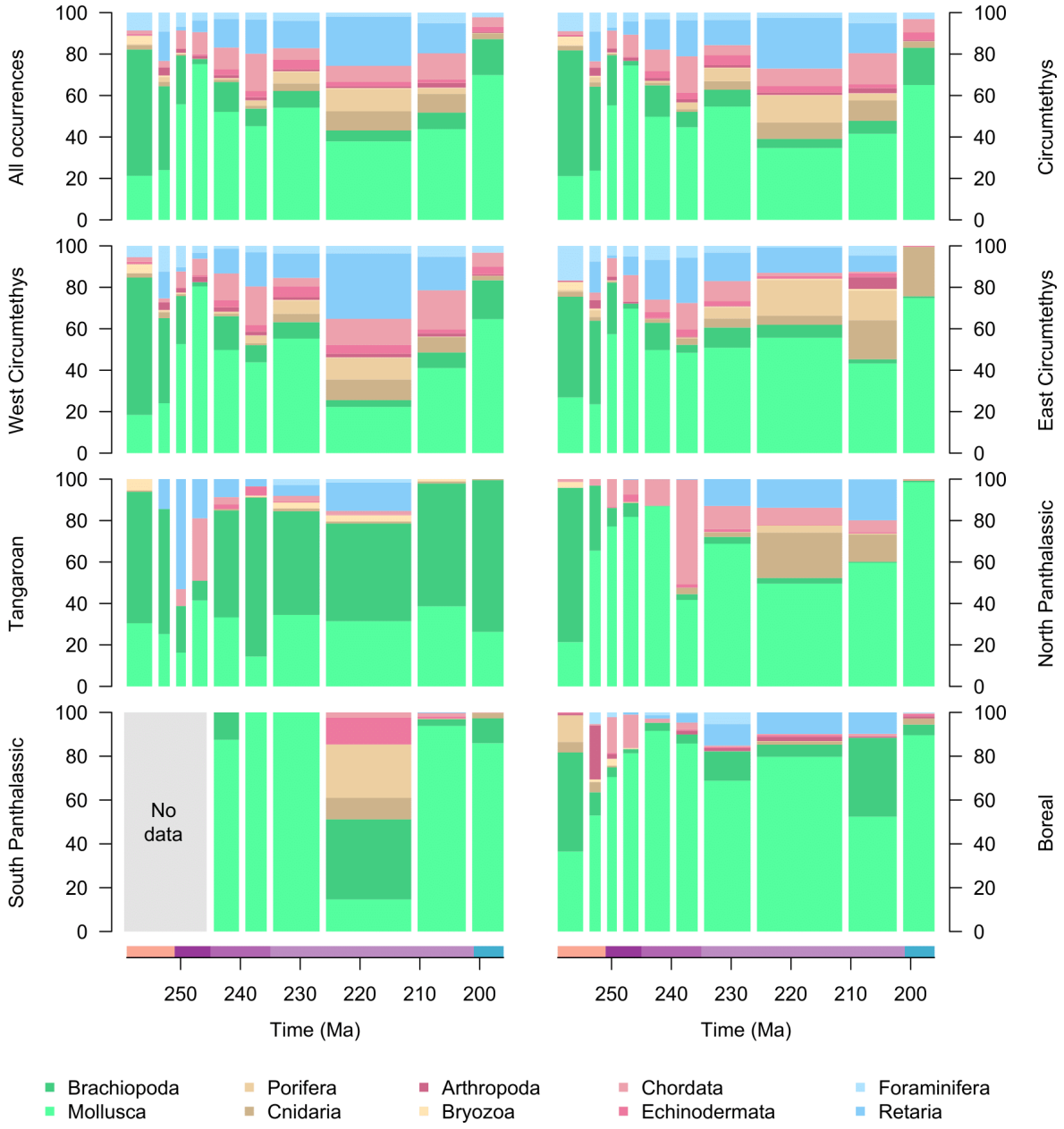


Fig. 3.2. Taxonomic compositions of Late Permian to Early Jurassic fossil occurrence data. Comparison of the total dataset to the West Circumtethys dataset indicates the dominance of the latter within the downstream analyses. The prevalence of cephalopods also highlights the potential for taphonomic bias reflecting the taxonomic affiliations of the fossil occurrences, indicating that regional and even global views of diversity dynamics through the study interval may largely reflect the tempo of just a few clades.

molluscs (Fig. 3.2), which is unsurprising given the abundance of ammonites, gastropods and bivalves in the PBDB, but introduces the caveat that downstream results will be driven primarily by these clades. Foraminiferan and radiolarian occurrences together comprise the next most abundant element of the composite dataset, demonstrating that good coverage of

both the macrofossil and microfossil records is still achieved, along with broad taxonomic coverage in the former despite the preponderance of molluscs.

3.2.3 Spatiotemporal standardisation

A largely stage-level binning scheme was used during spatial standardisation for several reasons. First, the amount of data in a stage-level bin will be greater than in a substage-level bin, providing a more stable view of occurrence distributions through time and increasing the availability of data for subsampling. Spatial variation at substage level might still affect the sampling of diversity, but the main goal of this study is to analyse origination and extinction rates where taxonomic ranges are key rather than pointwise taxonomic observations. Consequently, substage variation in taxon presences likely amounts to noise when examining taxonomic ranges, making stage-level bins preferable in order maximise signal.

During exploratory standardisation trials, a large crash in diversity and spatial sampling extent consistently emerged during the Hettangian (201.3–199.3 Mya; Figs. S7–13). No significant relationships with spatially standardised diversity were found when the Hettangian bin was excluded from correlation tests, indicating its disproportionate effect in otherwise well-standardised time series. Standardising the data to the level present in the Hettangian would have resulted in unacceptable data loss so this issue was instead accounted for by merging the Hettangian bin with the succeeding Sinemurian bin, when sampling returns to spatial extents consistent with older intervals. While this highlights a limitation of the method, as the Hettangian is less than 2 Ma in length it is reasonable to expect it would have a minor effect on taxonomic ranges in the long term, despite the magnitude of the sampling crash, and that any taxa surviving through the interval will be recorded in the much better sampled Sinemurian. This merging was performed for standardisation, but not for any downstream analyses to prevent the magnitude and timing of recovery from the ETME from being affected.

The occurrence data were plotted onto palaeogeographic maps to identify biogeographic regions that could feasibly be subsampled consistently through time (see 3.3.1). I identified five such regions which broadly correspond to major Permo-Jurassic seaboards and ocean basins: Circumtethys, Boreal, North Panthalassic and South Panthalassic, along with an unexpected set of marine occurrences from the Australian and New Zealand fossil record, here termed the Tangaroan (so named for the Maori god of the oceans). As Circumtethys is an extremely large region compared to the others, it was subdivided into eastern and western

sampling subdomains. While the extent of spatial regions reflects a compromise between biogeographic discretion and data availability and can theoretically be arbitrary, most of the regions share a degree of correspondence with bioregions for the Permo-Triassic predicted from abiotic drivers of marine provinciality (Kocsis et al., 2018a), suggesting that they are biologically realistic to a certain extent. The major exception to this is the east-west division of Tethys compared to the north-south divide recovered by Kocsis and colleagues (Kocsis et al., 2018a, 2018b, 2021), as this was a compromise between biogeographic realism and data availability through time. This highlights a limitation of the method where analytical outcomes may reflect sampling strategy over regional biogeographic history.

All regions extend for the full temporal range of the composite dataset, aside from the South Panthalassic, where adequate levels of sampling encompass the Late Triassic to Early Jurassic. Spatial windows were constructed for each region using `spacetimewind()`, then data were subsampled into each region under the described binning strategy using `spacetimestand()`. Four treatments were conducted for each spatially binned dataset: no standardisation, standardisation by MST length, standardisation of longitude-latitude extent and standardisation with both methods. For each treatment in each region, bin-wise diversity was calculated using coverage-based rarefaction at coverage levels of 40%, 50%, 60% and 70% (Figs. S7–S13). The relationships between diversity at each level of coverage with MST length, longitude range and latitude range were interrogated using one-tailed Pearson's product moment and Spearman's rho tests of correlation (Tables S3-S18), with Benjamini-Hochberg correction for multiple comparisons (Benjamini and Hochberg 1995). Spatial standardisation protocols for each region were then adjusted to eliminate significant correlations as needed (see electronic supplement for published chapter).

3.2.4 Rate data and preservation model

Origination, extinction, and preservation rates were jointly estimated in a Bayesian framework using PyRate (v3.0). As outlined in the first chapter, PyRate implements realistic preservation models that can vary through time and among taxa, yielding substantial increases in rate estimation accuracy over traditional methods (Silvestro et al., 2019; Warnock et al., 2020). PyRate has been criticised recently for only performing well when data availability is high and consistently sampled (Cerny et al., 2021). This criticism, however, was founded on simulated data with an underlying phylogenetic structure based on the parameters of a tree of

ornithischian dinosaurs, whose fossil record is known to be inconsistent (Tennant et al., 2018) and is at odds with the findings of simulations covering a broader range of turnover and preservation rates (Warnock et al., 2020). PyRate is demonstrably subject to the pitfall of spatial variability in the fossil record, with regional analyses of the crocodylomorph fossil record indicating declining diversity (Mannion et al., 2015), while global analysis with PyRate spuriously recovers increasing diversity driven by expansion of the geographic range of their fossil record (Mannion et al., 2015; Close et al., 2020a; Solorzano et al., 2020). The standardisation procedure used here avoids this issue and the marine fossil record is well-sampled compared to the scenarios where PyRate otherwise begins to perform poorly. Therefore, PyRate is a suitable method for inferring diversity dynamics in this instance.

I analysed datasets from the unstandardised, MST-standardised and MST + longitude-latitude-standardised treatments; as MST length is the most important control on spatial extent, the dataset with just longitude-latitude standardisation was expected to retain significant spatial bias. Ten age-randomised input datasets for each region and data treatment were generated in R with locality-age dependence (all occurrences from a locality are given the same randomised age), using collection number as a proxy for locality for PBDB-derived occurrences and geological section names for occurrences from the independent dataset. Locality-age dependence is both logically desirable as locality occurrences strictly represent a geographically localised and temporally discrete fauna (in idealised terms an assemblage from a single bedding plane) and has been shown to improve precision in age estimates in other Bayesian dating procedures using fossil data (King and Rucklin 2020).

The best fitting preservation model (homogenous, HPP; non-homogenous, NHPP; or time-variable homogenous Poisson process TPP) for each dataset was identified by maximum likelihood using the `-PPmodeltest` function of PyRate, with the best fitting model identified using the Akaike Information Criterion (Akaike 1973). In addition to testing between the HPP, NHPP and TPP preservation models, three TPP models of differing complexity were also considered: one with stage-level bins, one with stage-level bins plus subdivision of the Norian stage into three sub-bins (the informal divisions Lacian, Alaunian and Sevatian), and one with substage level bins and subdivision of the Norian stage into three sub-bins. For all datasets, the last binning scheme was found to be the best fitting, despite the greater number of model parameters (individual time bin preservation rates) that it introduces. As well as using the TPP

model of preservation through time (PyRate option `-qShift`) with substage-level bins and threefold subdivision of the Norian, the preservation rate was also allowed to vary according to a gamma distribution on taxon-wise preservation rates (here discretized into eight rate multipliers using PyRate options `-mG ncat 8`; Yang 1994; Silvestro et al., 2014). While there is currently no way to test between preservation models with and without the gamma parameter in PyRate, it is a recommended addition due to the known empirical variability of preservation rates among taxa, especially for taxonomically diverse datasets and because it includes a single additional parameter in the model (Silvestro et al., 2014b). In each analysis, the bin-wise preservation rates were assigned a gamma prior with fixed shape parameter set to 2, while the scale parameter was itself assigned a vague exponential hyperprior and estimated during the analysis (PyRate option `-pP 2 0`). This hierarchical approach provides a means of regularization while allowing the prior on the preservation to adapt to the dataset (Silvestro et al., 2019). Finally, rate shifts outside the covered range of the data were excluded in each analysis to avoid edge effects during parameter estimates (PyRate option `-edgeShift`).

3.2.5 Rate estimation

Regardless of the chosen preservation model, a PyRate analysis is parameter rich as the individual origination and extinction times for each taxon are jointly estimated along with the overall rates, in addition to the computational burden introduced by using rjMCMC to sample parameters across models with different numbers of rate shifts. This produces high computational burden, and models for larger sampling regions could not be estimated efficiently using PyRate's standard Metropolis Hastings algorithm for MCMC. PyRate can alternatively use an efficient Gibbs algorithm to sample from the posterior distribution of the parameters, producing preservation-corrected estimates of origination and extinction times that are virtually identical to those from the Metropolis Hastings algorithm, but with a coarse birth-death model that involves a dramatic loss of resolution in the resulting rate curves (Moharrek et al. 2022). A second program, LiteRate, was developed to permit origination and extinction rate estimation for taxonomically large datasets (Gjesfeld et al., 2018; Koch et al., 2021), gaining computational efficiency by implementing the same birth-death model used by PyRate with the rjMCMC and Metropolis Hastings algorithm, but without estimation of the complex preservation model. As ranges in a fossil dataset are expected to be truncated by variation in preservation rate through time, times of origination and extinction would be inaccurately estimated if LiteRate were run directly with a fossil dataset.

To overcome these methodological issues, I used a two-step procedure to permit efficient parameter estimation for taxonomically large fossil datasets. First, PyRate with the Gibbs algorithm was used to jointly estimate the parameters of the preservation model and the preservation rate-corrected estimates of origination and extinction times for each taxon. The origination and extinction time estimates were then supplied to LiteRate for estimation of rates and rate shifts from its birth-death model. In summary, PyRate is used to perform the computationally expensive task of estimating the complex preservation model parameters and taxon-specific origination and extinction times using the computationally efficient Gibbs algorithm, while LiteRate is used to estimate the high-resolution birth-death model, rates and rate shifts for the taxonomically large dataset.

PyRate analyses for each region were run across each set of ten age-randomised replicates using the Gibbs sampling algorithm (PyRate option `-se_gibbs`) with default rjMCMC for five million generations, aside for the Tangaroan (10 million) and South Panthalassic (20 million), with MCMC sampling generations set to produce 10,000 samples of the posterior. Output datasets were assessed using Tracer (v1.7.1; Rambaut et al., 2018) to determine suitable burn-in values by visually inspecting the MCMC trace, and to check for convergence by ensuring minimum effective sample sizes of > 200 on all post-burnin model parameters for each analysis. Mean origination and extinction times were derived using the `-ginput` function of PyRate with a 10% burn-in, before being supplied to LiteRate. LiteRate analyses for each region were run across the 10 sets of mean origination and extinction times for 200 million generations, aside for the South Panthalassic (250 million). To incorporate age uncertainty into each analysis, logs from each age-randomised replicate were combined respectively for PyRate and LiteRate using the `-combLog` function of PyRate, taking 100 random samples from each log post 10% burn-in, to give 1000 samples of the posterior across all age-randomised replicates. Rates were then plotted at 0.1 million-year intervals and statistical significance of rate shifts recovered by the rjMCMC assessed using Bayes factors (log BF > 2 = positive support, log BF > 6 = strong support; Kass and Raftery 1995) using the `-plotRJ` function of PyRate (Figs. S22–S42).

3.2.6 Probabilistic diversity estimation

Traditional methods of estimating diversity do not directly address uneven sampling arising from variation in preservation, collection and description rates, and their effectiveness is highly

dependent on the structure of the dataset (Close et al., 2018). I use an alternative method developed by co-author D.S, mcmcDivE, to infer corrected diversity trajectories based on the sampled occurrences and on the preservation rates through time and across lineages as inferred by PyRate. mcmcDivE implements a hierarchical Bayesian model to estimate corrected diversity across arbitrarily defined time bins. The method estimates two classes of parameters: the number of unobserved species for each time bin and a parameter quantifying the volatility of the diversity trajectory.

The sampled number of taxa (i.e., the number of fossil taxa, here indicated with x_t) in a time bin is assumed be a random subset of an unknown total taxon pool, indicated with D_t. The goal of mcmcDivE is to estimate the true diversity trajectory $\mathbf{D} = \{D_1, D_2, \ldots, D_t\}$, of which the vector of sampled diversity $\mathbf{x} = \{x_1, x_2, \ldots, x_t\}$ is a subset. The sampled diversity is modelled as a random sample from a binomial distribution (Starrfelt and Liow 2016) with sampling probability p_t:

(1)

$$x_t \sim \mathrm{Bin}(D_t, p_t)$$

The sampling probability is obtained from the preservation rate (q_t) estimated in the initial PyRate analysis. If the PyRate model assumes no variation across lineages, the sampling probability is taken from the Poisson process $p_t = 1 - \exp(-q_t \times \delta_t)$, where δ_t is the duration of the time bin. When using a gamma model in PyRate, however, the q_t parameter represents the mean rate across lineages at time *t* and the rate is heterogeneous across lineages based on a gamma distribution with shape and rate parameters equal to an estimated value α. To account for rate heterogeneity across lineages in mcmcDivE, an arbitrarily large vector of gamma-distributed rate multipliers $g_1, \ldots, g_R \sim \Gamma(\alpha, \alpha)$ is drawn and the mean probability of sampling in a time bin computed as:

(2)

$$p_t = \frac{1}{R}\sum_{i=R}^{R} 1 - \exp(-q_t \times g_i \times \delta_t)$$

While q_t quantifies the mean preservation rate in PyRate (i.e. averaged among taxa in a time bin *t*), the mean sampling probability p_t will be lower than $1 - \exp(-q_t \times \delta_t)$ (i.e., the probability expected under a constant preservation rate equal to q_t) especially for high levels of rate heterogeneity, due to the asymmetry of the gamma distribution and the non-linear relationship between rates and probabilities. Corrected diversity is then sampled from its posterior through MCMC. The likelihood of the sampled number of taxa is computed as the probability mass function of a binomial distribution with D_i as the 'number of trials' and p_i as the 'success probability'. To account for the expected temporal autocorrelation of a diversity trajectory (Silvestro et al., 2021), a Brownian process is used as a prior on the log-transformed diversity trajectory through time. Under this model, the prior probability of D_t is:

(3)

$$P(\log(D_t)) \sim \mathcal{N}(\log(D_{t-1}), \sqrt{\sigma^2 \times \delta_t})$$

where σ^2 is the variance of the Brownian process. For the first time bin in the series, $D_{t=0}$, a vague prior $\mathcal{U}(0, \infty)$ is used. Because the variance of the process is itself unknown and may vary among clades as a function of their diversification history, it is assigned an exponential hyper-prior $\mathrm{Exp}(1)$ and estimated using MCMC. Thus, the full posterior of the mcmcDivE model is:

(4)

$$\underbrace{P(D, \sigma^2 | x, q, \alpha)}_{\text{posterior}} \propto \underbrace{P(x|D, q, \alpha)}_{\text{likelihood}} \times \underbrace{P(D|\sigma^2)}_{\text{prior}} \times \underbrace{P(\sigma^2)}_{\text{hyperprior}}$$

where $\mathbf{D} = \{D_0, D_1, \ldots, D_t\}$ and $\mathbf{q} = \{q_0, q_1, \ldots, q_t\}$ are vectors of estimated diversity, sampled diversity, and preservation rates for each of *T* time bins. Parameters *D* and σ^2 are estimated using MCMC to obtain samples from their joint posterior distribution. To incorporate uncertainties in q and α, these parameters are randomly resampled during the MCMC from their posterior distributions obtained from PyRate analyses of the fossil occurrence data. While in mcmcDivE a posterior sample of q_t and α precomputed in PyRate is used for computational tractability of the problem, a joint estimation of all PyRate and mcmcDivE parameters is in principle possible, particularly for smaller datasets.

Table 3.1. Settings of birth-death-preservation simulations of the mcmcDivE model

	Speciation		Extinction		Preservation	
ID	Rate	Shifts	Rate	Shifts	Rate	Lineage heterogeneity
A	0.4, 0.01	25	0.05, 0.01	25	0.25, 2.5	-
B	0.4, 0.01	25	0.05, 0.01	25	0.25, 2.5; 0	-
C	0.4, 0.1, 0.01	20, 10	0.05, 0.3, 0.01	15, 10	0.25, 2.5	-
D	0.4, 0.1, 0.01	20, 10	0.05, 0.3, 0.01	15, 10	0.25, 2.5	0.5
E	0.2, 0.4	10	0.01, 0.9, 0.1	10, 7	0.25, 2.5	-
F	0.2, 0.4	10	0.01, 0.9, 0.1	10, 7	0.25, 2.5	0.5

100 datasets were generated and run from 35 arbitrary time units in the past. Preservation rate shifts occurred at 23, 15, 8, 5.3, 2.6 in each setting

3.2.6 Simulated and empirical diversity analyses

The performance of the mcmcDivE method was assessed using 600 simulated datasets obtained under different birth-death processes and preservation scenarios. The settings of the six simulations (A–F) are summarized in Table 3.1, with 100 datasets simulated from each setting. Since the birth-death process is stochastic and can generate a wide range of outcomes, only simulations with 100 to 500 species were accepted, although the resulting number of sampled species decreased after simulating the preservation process. From each birth-death simulation, fossil occurrences were sampled based on a heterogeneous preservation process. Each simulation included six different preservation rates with rate shifts set to 23, 15, 8, 5.3 and 2.6 time units before the present. To ensure that most rates were small (i.e., reflecting poor sampling), preservation rates were randomly sampled as:

(5)

$$q \sim \exp(\mathcal{U}(\log(0.25), \log(2.5)))$$

Two of the five scenarios (D, F) included strong rate heterogeneity across lineages in addition to the rate variation through time by assuming that preservation rates followed a gamma distribution with shape and rate parameters set to 0.5. This indicates that if the mean

preservation rate in a time bin was 1, the preservation rate varied across lineages between <0.001 and 5 (95% interval). In one scenario (B), preservation rate was set to 0 (i.e., a complete gap in preservation) in addition to the temporal rate changes used in the other scenarios. Specifically, the preservation rate was set to 0 in two intervals between 15 and 8 Ma and between 5.3 and 2.6 Ma.

The occurrence data was analysed using PyRate to estimate preservation rates through time and infer the amount of rate heterogeneity across lineages, running 10 million MCMC generations using the TPP preservation rate model with gamma-distributed heterogeneity. mcmcDivE was then ran for 200,000 MCMC iterations assuming bins of 1 Ma duration to estimate corrected diversity trajectories while resampling the posterior distributions of the preservation parameters inferred by PyRate. To summarize the performance of mcmcDivE mean absolute percentage error was computed as the absolute difference between true and estimated diversity averaged across all time bins and divided by the mean true diversity through time, then a one-tailed t-test used to determine whether the mean absolute percentage error for the mcmcDivE estimate was significantly smaller than those for the other diversity estimation methods in each set of 100 simulations. The coefficient of determination (R^2) between estimated and true diversity was also computed to assess how closely the estimated trends matched the true diversity trajectories. Performance of the mcmcDivE estimates was compared with a curve of raw sampled diversity (i.e., number of sampled species per 1 Ma time-bin), a range-through diversity trajectory based on first and last appearances of sampled species, and sampling-corrected trajectories estimated using coverage-based rarefaction obtained by `estimateD()` and the squares extrapolator (Alroy 2018).

The simulation results show that mcmcDivE provides accurate results under most settings and significantly better estimates (significantly smaller mean absolute percentage error; $p < 0.0001$ for all six sets of simulations) of the diversity through time compared with raw diversity curves, range-through diversity trajectories or sampling-corrected estimates from coverage-based rarefaction, or extrapolation by squares (Figs. 3.3, 3.4). The mean absolute percentage error averaged 0.13 (95% CI: 0.04–0.29) in simulations without across lineage rate heterogeneity, with a high correlation with the true diversity trajectory: $R^2 = 0.93$ (95% CI: 0.72–0.99). The diversity estimates remained accurate even in the presence of time intervals with zero preservation (Fig. 3.3B).

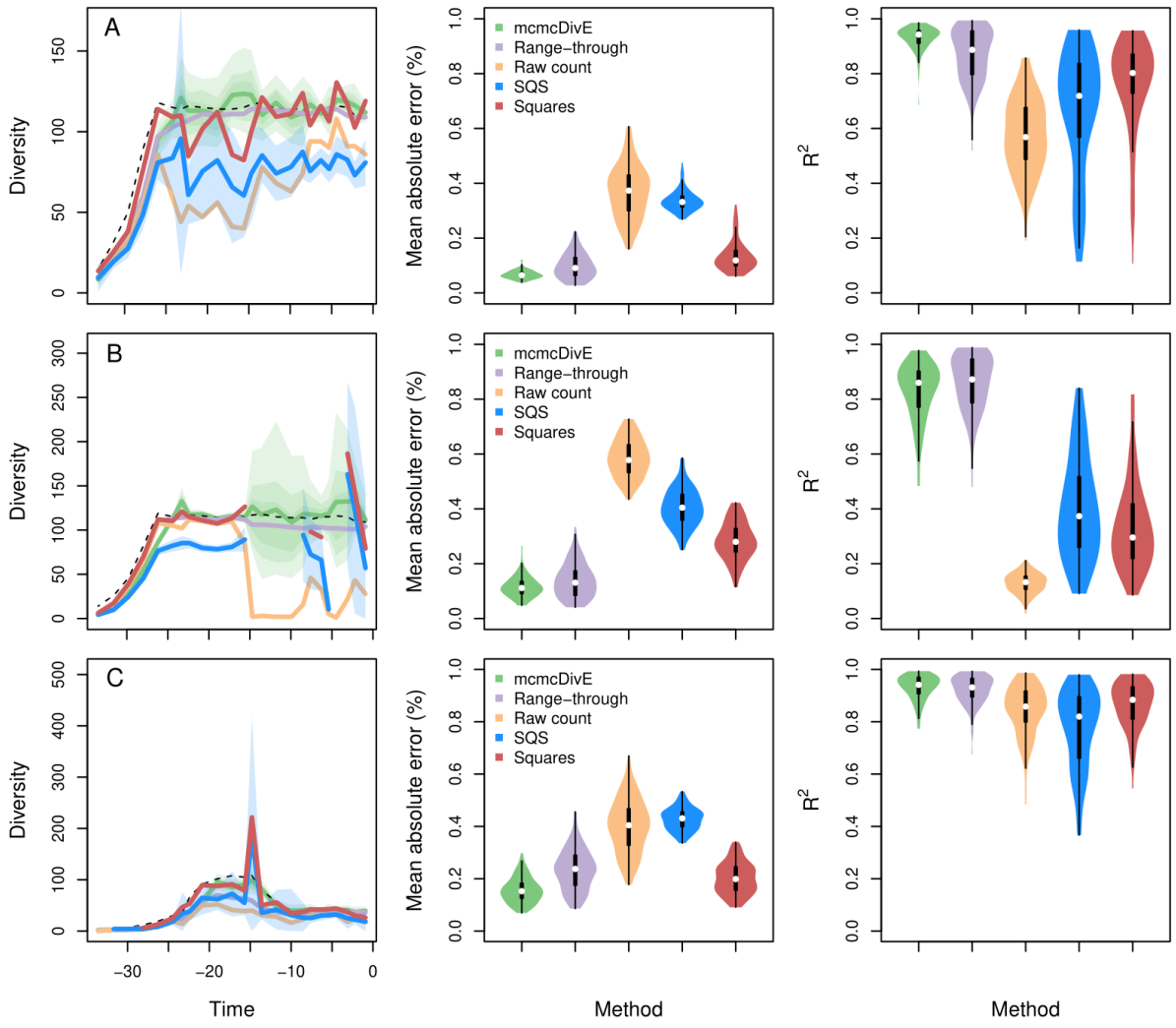


Fig. 3.3. Validation of the mcmcDivE method (simulations A–C). Diversity trajectories from mcmcDivE compared to other popular diversity estimators and violin plots of their mean squared errors. Refer to Table 3.1 for the simulation settings pertaining to parts A–C.

Simulations with rate heterogeneity across lineages (Fig. 3.4D, 3.4F) yielded higher mean absolute percentage errors (0.43, 95% CI: 0.24–0.55) while maintaining a strong correlation with the true diversity trajectory $R^2 = 0.95$ (95% CI: 0.85–0.99). This indicates that, while the absolute estimates of diversity are on average less accurate in the presence of strong rate heterogeneity across lineages (in addition to strong rate variation through time), the relative changes in diversity through time are still accurately estimated. The increased relative error in these simulations is mostly linked with an underestimation of diversity throughout, which has been observed in other probabilistic methods to infer diversity in the presence of rate heterogeneity across lineages (Close et al. 2018). This, however, does not hamper the robust

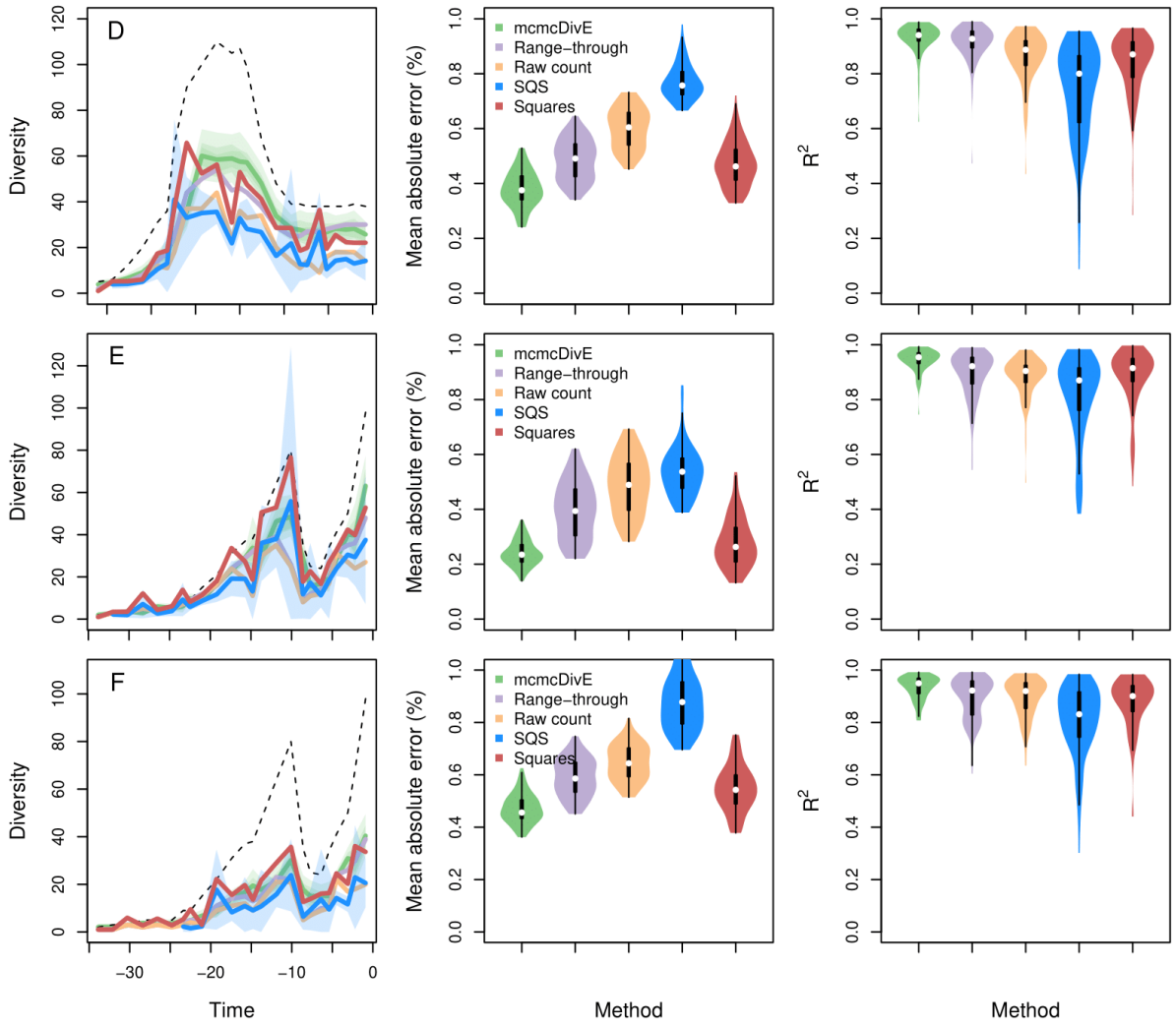


Fig. 3.4. Validation of the mcmcDivE method (simulations D–F). Diversity trajectories from mcmcDivE compared to other popular estimators and violin plots of their mean squared errors. Refer to Table 3.1 for the simulation settings pertaining to parts D–F.

estimation of relative diversity trends using this method (Figs. 3.4D, 3.4F). After validating the accuracy of the model, regional analyses of Triassic marine diversity were run for 1000,000 MCMC iterations at 1 Ma intervals, then diversity estimates summarised by calculating the median of the posterior samples and the 95% credible intervals (Figs. S22-S42).

3.2.7 Turnover estimation

Counts of unique taxa within a sample (geographic area or time bin) are a measure of diversity while the degree of taxonomic differentiation between two samples constitutes a measure of turnover. Taxonomic turnover through time, measured by comparison of the taxon pools in

successive pairs of time bins, avoids the pitfall of cryptic turnover hidden within diversity or diversification rate curves as high extinction and origination rates will result in marked taxonomic differentiation. The modified Forbes index (Forbes*; Alroy 2015b) with relative abundance correction (RAC; Brocklehurst et al., 2018) robustly accounts for both incomplete sampling in each sample and differing abundance distributions between a pair of samples, both of which can bias the apparent degree of similarity (Brocklehurst et al., 2018). RAC is a potentially computationally expensive method so the RAC-corrected Forbes* metric (converted to dissimilarity as 1 – Forbes*) was calculated using an efficient, parallelised C++ function with an Rcpp wrapper (Eddelbuettel and Francois 2011) in R. Occurrences in each region were first binned at stage level, then with twofold subdivision of the Anisian, Ladinian and Carnian and threefold subdivision of the Norian, using the occurrence midpoint ages. RAC-Forbes* dissimilarity was then calculated for each region between successive pairs of time bins with 100 null trials, and 100 sampling standardisation trials at a sampling quorum of 0.5 for each null trial and empirical estimate.

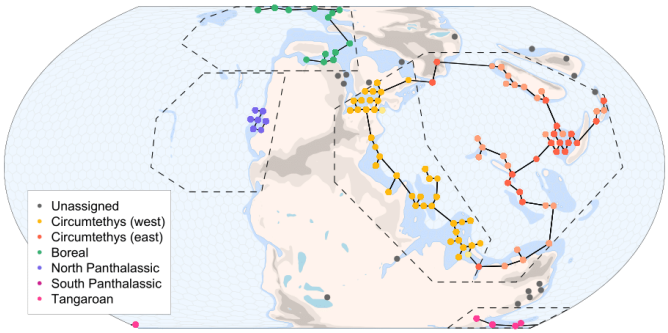
3.3 Results

3.3.1 Spatial standardisation

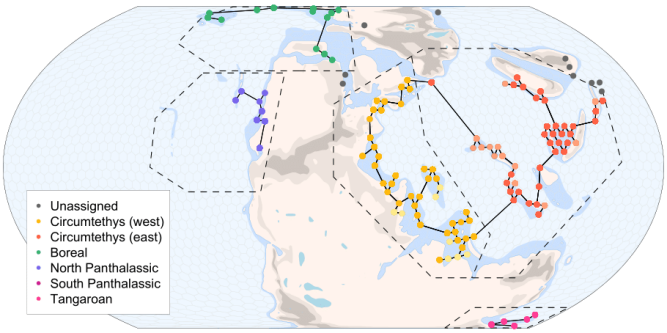
Regional samples of fossil occurrence data for the West Circumtethys, East Circumtethys, North Panthalassic, South Panthalassic, Boreal, and Tangaroan regions were captured using sliding spatial windows, spatially binned using a hexagonal grid (Fig. 3.5) then spatially standardised by longitude-latitude range and minimum spanning tree (MST) length. The resulting samples of fossil occurrence data are geographically consistent through time and free from spatial sampling biases which can substantially distort trends in apparent diversity. When coupled with diversity estimation methods which correct for heterogenous sampling between time bins, this workflow permitted estimation of diversity dynamics unaffected by spatiotemporal sampling bias, allowing regional diversity dynamics to be interrogated.

The spatial standardisation workflow successfully reduced variance in MST length and longitude-latitude range whilst enforcing a consistent geographic distribution of data through time in each region, although the degree of reduction is dependent on the dataset and target extent, with a noticeable increase in standardisation efficacy with increasing region size (Table

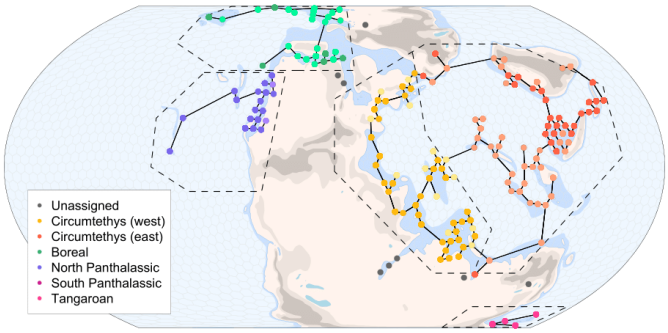
Wuchiapingian (259.5 - 254.2 Ma)



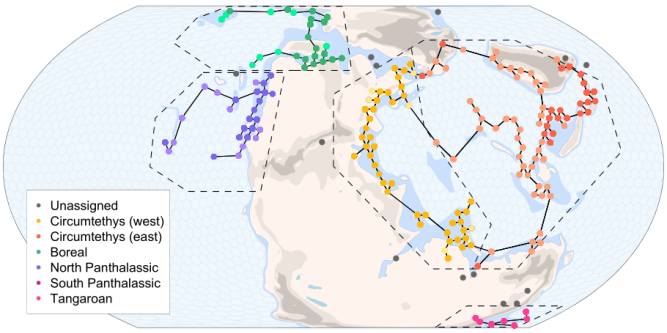
Changhsingian (254.2 - 251.9 Ma)



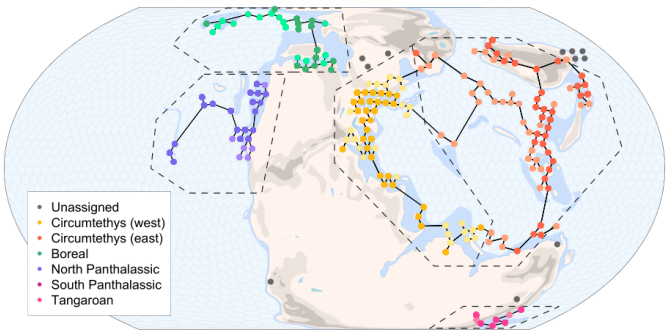
Induan (251.9 - 249.9 Ma)



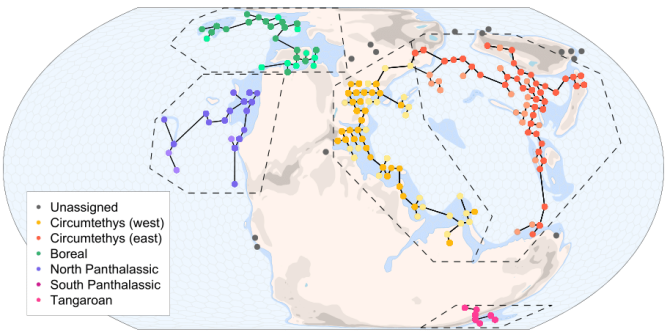
Olenekian (249.9 - 246.7 Ma)



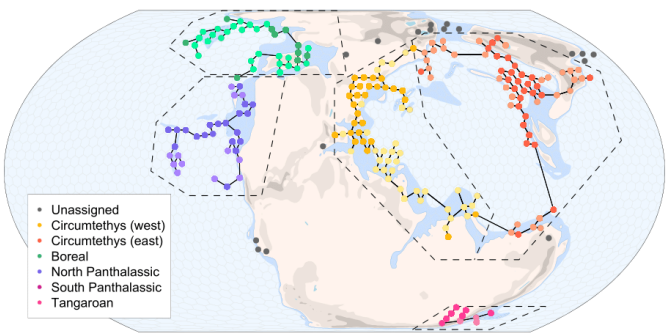
Anisian (246.7 - 241.5 Ma)



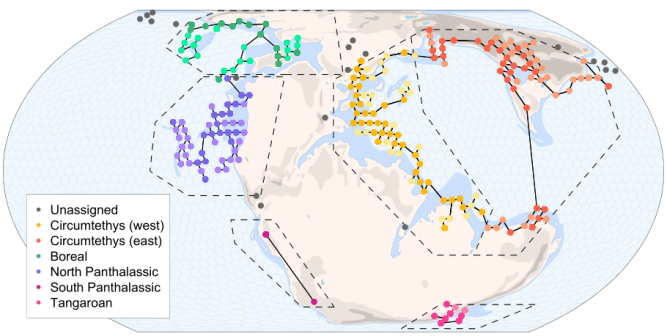
Ladinian (241.5 - 237 Ma)



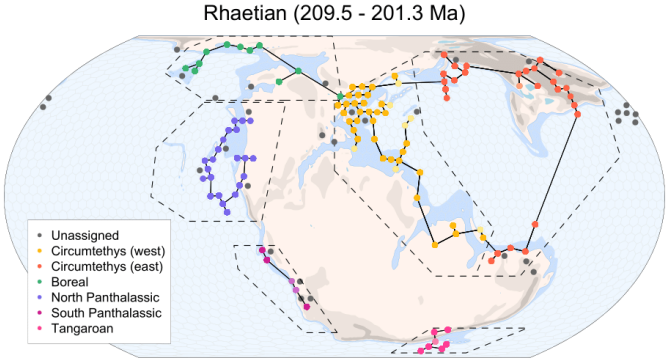
Carnian (237 - 227.3 Ma)



Norian (227.3 - 209.5 Ma)



Rhaetian (209.5 - 201.3 Ma)



Early Jurassic (201.3 - 192.9 Ma)

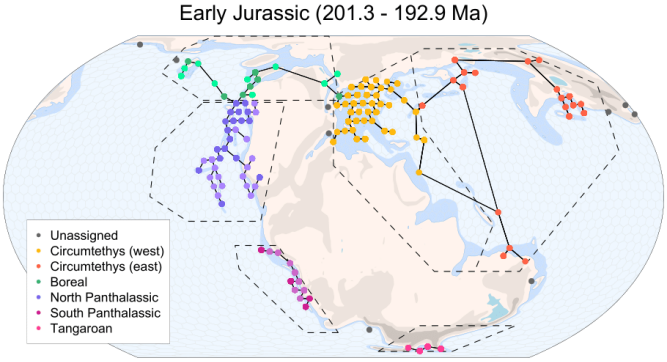


Fig. 3.5. Spatial standardisation for the Late Permian to Early Jurassic sampling regions. Spatially standardised datasets from each region were used to estimate diversity dynamics corrected for heterogenous sampling through time and across space. Dotted black lines demarcate the sliding spatial windows defining each sampling region, solid black lines their MSTs, coloured points the occupied grid cells comprising the MST, and points with the darker hue those retained after MST standardisation.

Circumtethyan regions were all less than 1.4% after spatial standardisation as a target length suitable for all bins could be chosen. By comparison, the standard deviation around the target rose for the smaller North Panthalassic region as the target threshold was chosen to improve data retention for the vast majority of the bins (Early Triassic–Early Jurassic) but was still significantly higher than the unstandardised spatial extent of the data in the Late Permian, resulting in greater variance when the full time span is considered (Table 3.2; Fig. S11). Similarly, the South Panthalassic region is well standardised from the Norian to the Early Jurassic, allowing the signal of the TJME to be scrutinised, despite the vastly reduced spatial extent at the start of the dataset (Table 3.2; Fig. S13). The constraints imposed by standardisation for both spatial metrics are also apparent in the Tangaroan region, where MST standardisation was reasonably effective throughout the Triassic but declined in quality when longitude-latitude standardisation was first applied (Table 3.2; Fig. S12).

Prior to standardisation, the relationship between region-level (RL) diversity by coverage-based rarefaction and spatial extent was significant for multiple regions across the measured quorum levels (Fig. 3.6). After standardisation, significant correlations with RL diversity were broadly eliminated but some are present for comparisons which were previously insignificant. Regardless, at least one dataset was produced for each region where RL diversity at each quorum level showed no significant correlations with spatial extent, except for the North Panthalassic region. Here, significant Spearman correlations were still present at some quorum levels (Table S8), but the lack of a consistent correlation across quorum levels, along with their weak statistical significance, suggests that the apparent relationships are not robust. Further, all correlations were rendered insignificant when the two Late Permian bins were excluded from the analyses (Fig. 3.6, Table S7), indicating that the remainder of the dataset is otherwise well standardised (Table 3.2).

Table 3.2. Target extent and efficacy for MST and longitude-latitude standardisation

	Target extent	Unstandardised SD (%)	Standardised SD (%)
MST length (km)			
West Circumtethys	23000	16.7	1.2
East Circumtethys	23000	33.7	1.3
Boreal	10000	35.6	9
North Panthalassic[1]	12000	29.6	3.4
North Panthalassic	12000	49.9	28.6
Tangaroan	3000	29.5	22.6
South Panthalassic[2]	2500	42.7	12.2
South Panthalassic	2500	88.7	70.1
Longitude range (°)			
West Circumtethys	60	15.8	5.3
East Circumtethys	95	5	3.6
Boreal	30	15.2	9.1
North Panthalassic[1]	30	34.9	34.8
North Panthalassic	30	53.2	25.6
Tangaroan	30	35.1	14.6
South Panthalassic[2]	90	9.4	9.4
South Panthalassic	90	42.4	42.1
Latitude range (°)			
West Circumtethys	75	16.4	14.5
East Circumtethys	90	12.7	11.6
Boreal	35	14.4	16.7
North Panthalassic[1]	30	16.3	15.4
North Panthalassic	30	31	23.4
Tangaroan	10	28.7	20.1
South Panthalassic[2]	35	15.2	15.2
South Panthalassic	35	46.8	45.6

SD = standard deviation relative to target. [1]Early Triassic to Early Jurassic only. [2]Rhaetian-Early Jurassic only

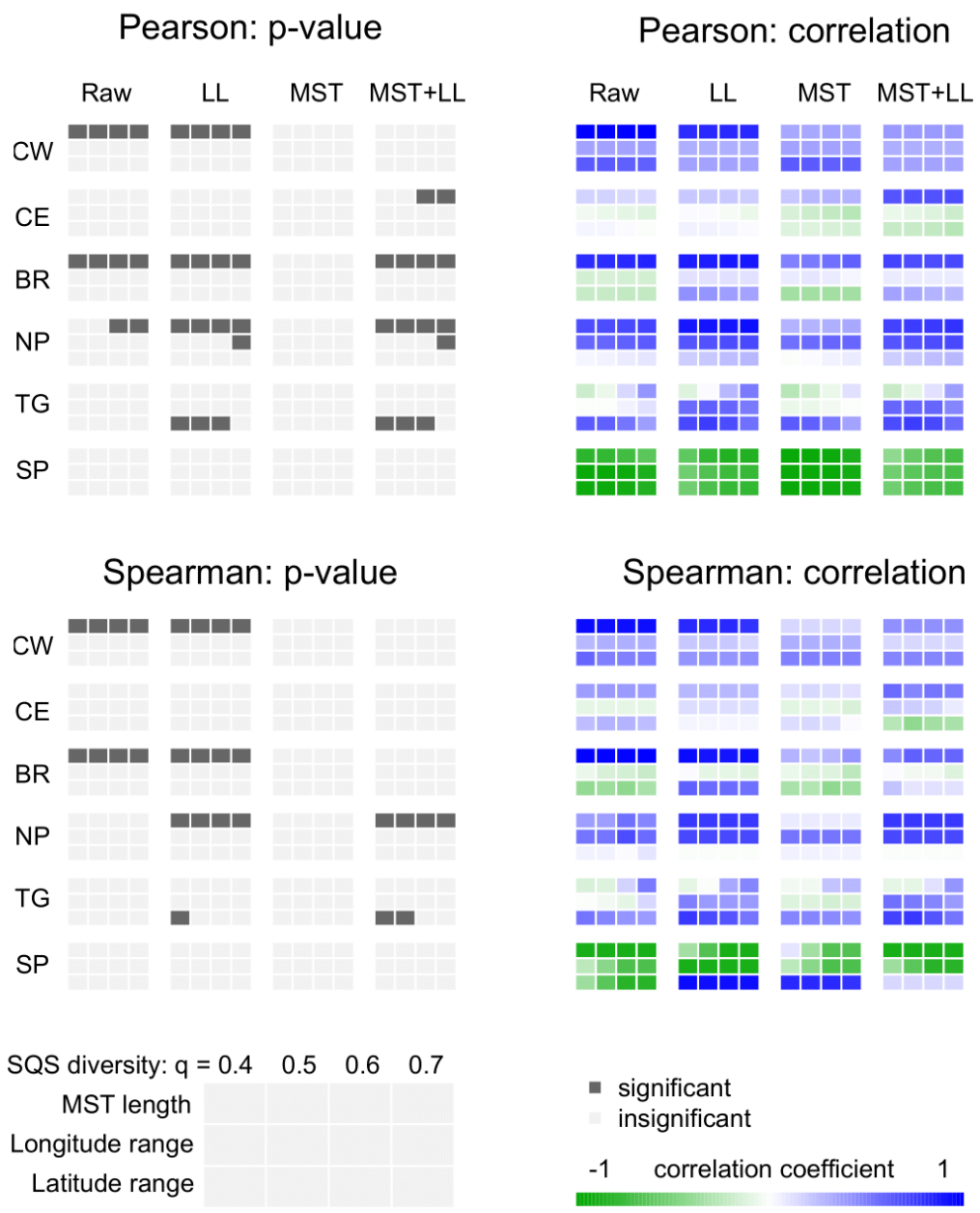


Fig. 3.6. Correlation tests between sampling-corrected diversity and spatial extent. One-tailed Pearson and Spearman correlations under each spatial standardisation treatment, for each sampling region, with false discovery correction. Without any spatial standardisation (raw), spatial sampling extent frequently shows significant relationships with diversity from shareholder quorum subsampling (SQS) across several quorum (q) levels. Standardisation by minimum spanning tree length (MST) is the most effective means of mitigating this bias, compared to standardisation of longitude-latitude extent (LL). CW = West Circumtethys, CE = East Circumtethys, BR = Boreal, NP = North Panthalassic, TG = Tangaroan, SP = South Panthalassic. For the North Panthalassic region, the Permian was omitted from the time series correlations as an unavoidable increase in spatial sampling extent which otherwise resulted in significant Spearman correlations in longitude and latitude extent at some quorum levels. Exact p-values and correlation coefficients are available in Tables S3–S18.

3.3.2 Probabilistic origination, extinction and diversity

Origination and extinction rates (Fig. 3.7) and probabilistic diversity trajectories (Fig. 3.8) clearly differ between sampling regions even during well documented global events. The signal of the end-Permian mass extinction is clear cut in the Circumtethyan and North Panthalassic regions, but the onset of elevated extinction rates occurred a couple of million years earlier in the latter (Fig. 3.7F) and may reflect the age uncertainty of the fossil occurrences. An increase in extinction rate also occurred in the Tangaroan region, although support for an increase in extinction rate here is barely positive rather than strong (Fig. S38). In the Boreal region, however, there is high uncertainty in the magnitude of the extinction rate increase at the Permo-Triassic boundary and the median origination rate remained consistently higher than the median extinction rate (Fig 3.7B), producing a subdued extinction signal with very little change in diversity. Further paroxysms in extinction rates took place in the North Panthalassic, Boreal and Tangaroan regions throughout the Early Triassic, again with positive rather than strong support in the latter. There was a clear spike in extinction rate in West Circumtethys at the CPE (Fig. 3.7C), along with more subdued increases in the Boreal and East Circumtethys regions (Fig. 3.7B, 3.7D). Elsewhere, extinction rates showed little change through the CPE, while elevated extinction rates instead occurred at the end of the Carnian in the Circumtethyan, Boreal and North Panthalassic regions (Fig. 3.7B-F). Distinct extinction signals were recovered in all regions at the end of the Triassic, but this was somewhat reduced in West Circumtethys due to a concurrent spike in origination rate (Fig. 3.7C), while in the Tangaroan region the rate shift significance is again merely positive rather than strongly supported.

Shifts in origination and extinction rate occurred frequently throughout the duration of each region, with strong support for their statistical significance. Away from major extinction events where there were clear shifts in the median rate, most of these shifts mark minor fluctuations in the background extinction rate or periods where sharp rate changes were inferred but with high uncertainty on their magnitude and timing. Probabilistic diversity also displayed marked short-term fluctuations (Fig. 3.8), punctuated by sharp peaks and crashes marking major periods of biotic turnover where concurrent disparities between extinction and origination rates (i.e., a sharp change in net diversification rate) may be noted (Figs. 3.7, 3.8).

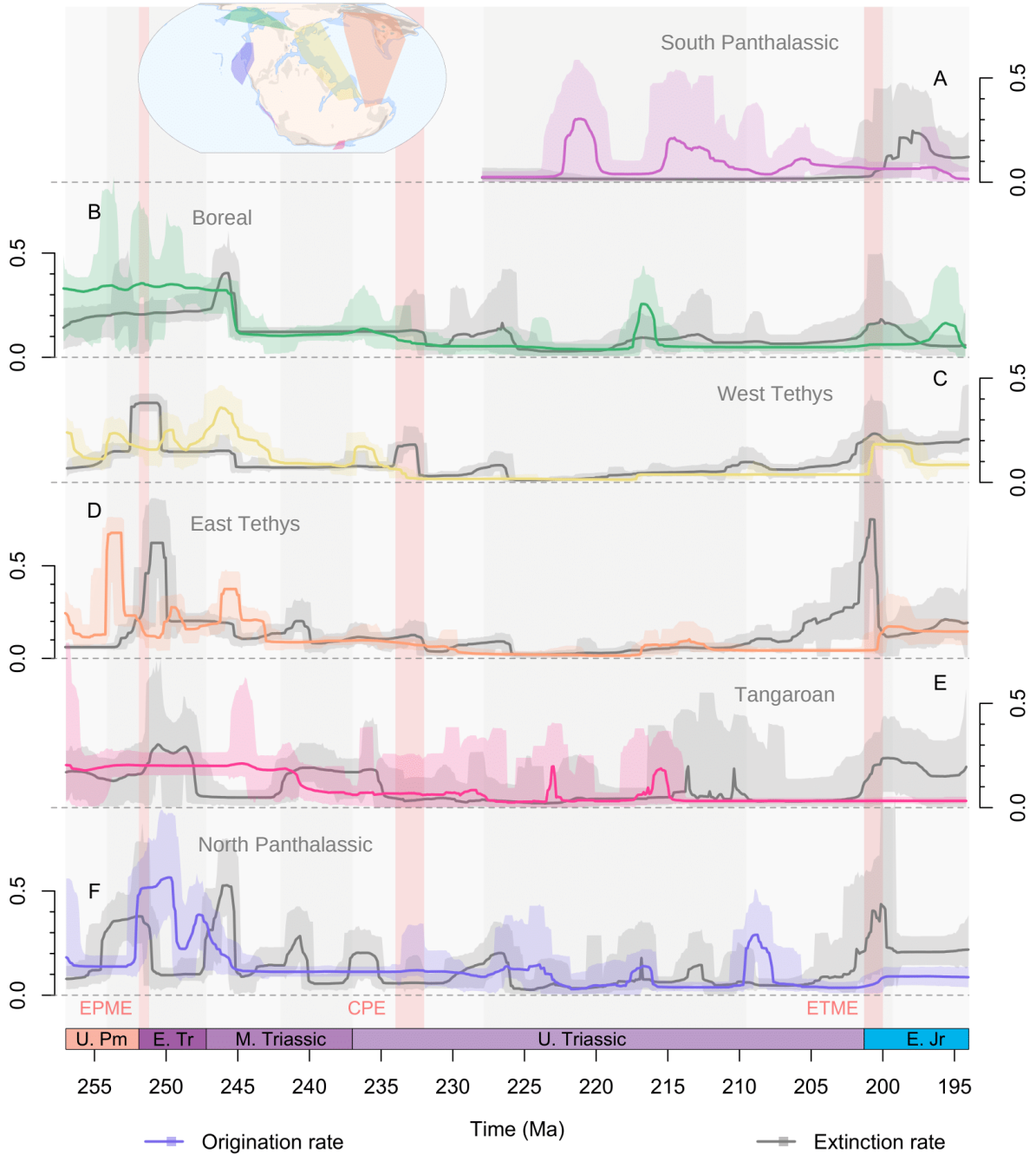


Fig. 3.7. Late Permian to Early Jurassic probabilistic origination and extinction rates. Extinction and origination rates differ markedly between sampling regions through the study interval, even during globally pervasive biological upheavals like the end-Permian mass extinction (PTME), Carnian Pluvial Episode (CPE), or end-Triassic mass extinction (TJME). Deconvolving the diversification rates into the component origination and extinction processes is necessary to capture pulses of cryptic turnover where net diversification and so the change in diversity is only small. Displayed results are derived from PyRate applied to datasets standardised by minimum spanning tree length, meaning that the results are unbiased by heterogenous sampling through time and across space. **(A).** South Panthalassic. **(B).** Boreal. **(C).** West Circumtethys. **(D).** East Circumtethys. **(E).** Tangaroan. **(F).** North Panthalassic.

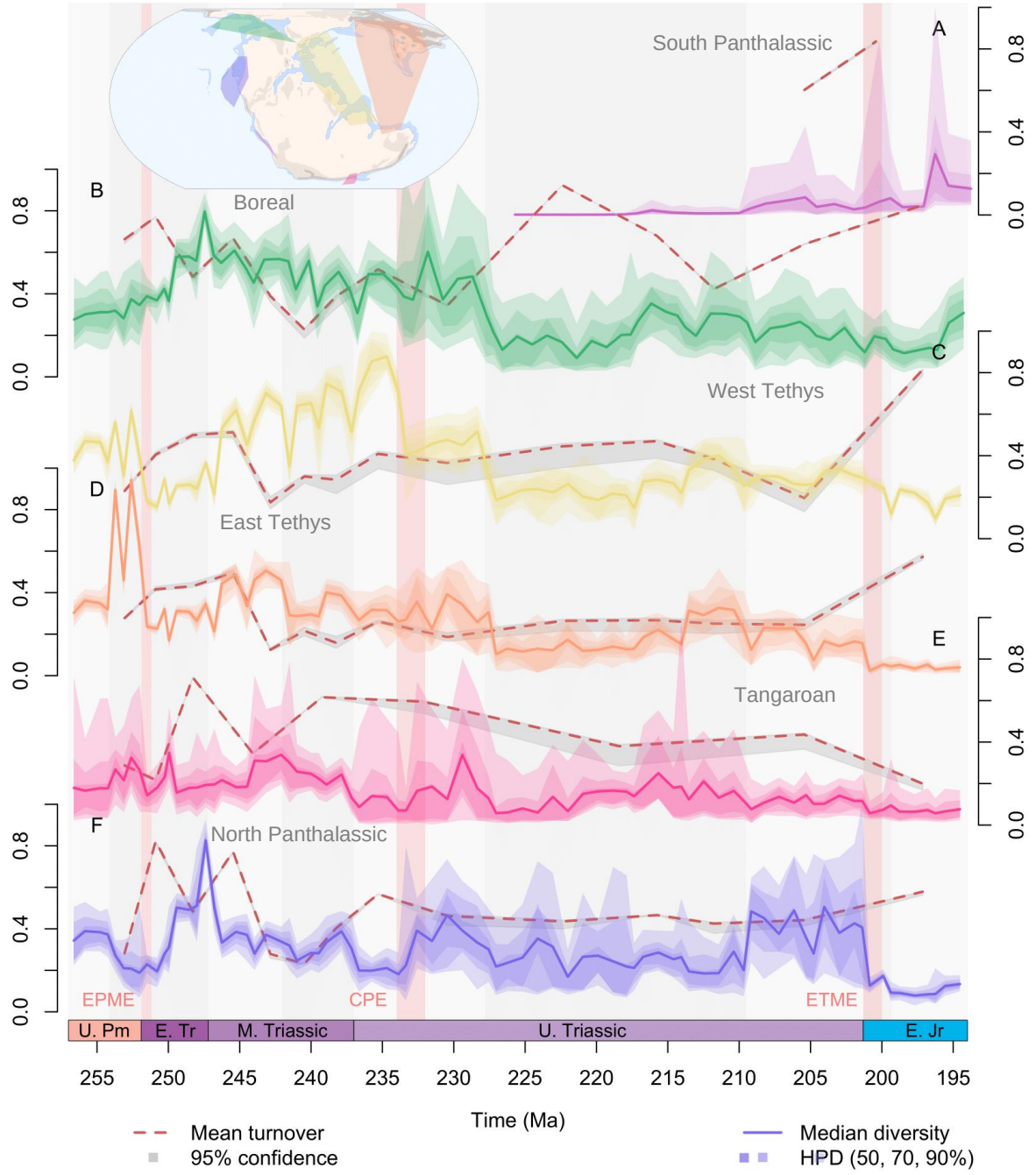


Fig. 3.8. Regional probabilistic diversity curves and median turnover. Diversity trajectories with 50, 75 and 95% highest posterior densities, and magnitude of turnover with 95% confidence intervals, differ substantially between different parts of the globe and across key biotic upheavals highlighted with red bars: the end-Permian mass extinction (PTME), Carnian Pluvial Episode (CPE), and end-Triassic mass extinction (TJME). As in previous figures, displayed results are derived from datasets standardised by minimum spanning tree length. All diversity values are rescaled so that the highest value in each 95% confidence interval is equal to 1. Each turnover value and confidence interval record the relative abundance-corrected Forbes* dissimilarity relative to the preceding bin. **(A).** South Panthalassic. **(B).** Boreal. **(C).** West Circumtethys. **(D).** East Circumtethys. **(E).** Tangaroan. **(F).** North Panthalassic.

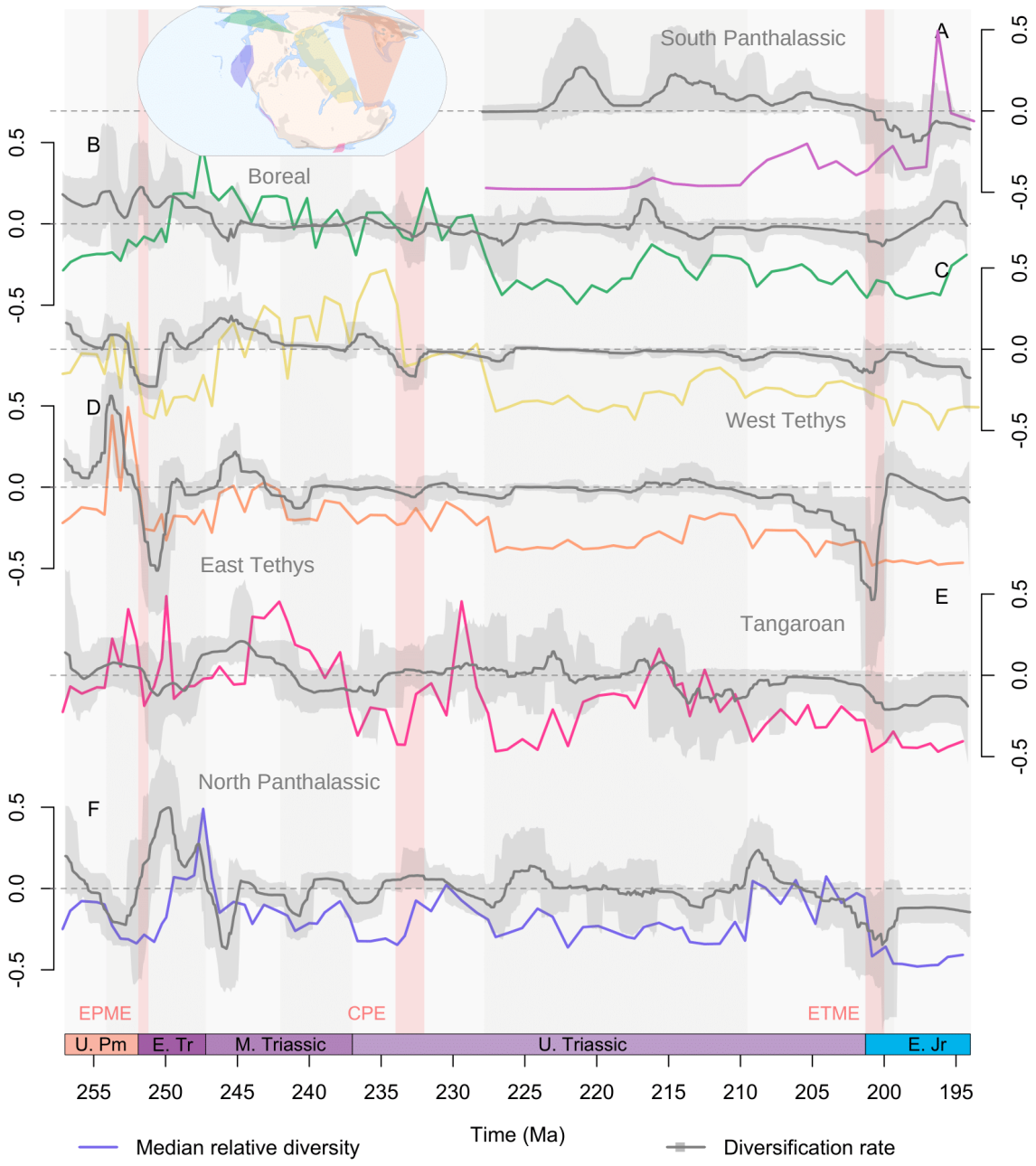


Fig. 3.9. Regional diversification rates and diversity through the Late Permian to Early Jurassic. Regional probabilistic median diversification rates with their 95% highest posterior densities, and median probabilistic diversity. Diversification rates and the diversity curves resulting from these rates differ markedly between sampling regions, demonstrating that the accumulation of biodiversity is heterogenous around the globe through the study interval, even during globally pervasive biological upheavals like the end-Permian mass extinction (EPME), Carnian Pluvial Episode (CPE), or end-Triassic mass extinction (ETME). Y-axes are not provided for diversity as these values are normalised by their highest values (see Fig. 3.8 for axes and confidence intervals). **(A).** South Panthalassic. **(B).** Boreal. **(C).** West Circumtethys. **(D).** East Circumtethys. **(E).** Tangaroan. **(F).** North Panthalassic.

3.3.3 Turnover

As with region-level diversity, within-region turnover shows regional differences in both the magnitude and pattern of dissimilarity through time (Fig. 3.8). The most pronounced shifts in turnover occurred in the Early Triassic in the aftermath of the PTME in all regions, but even here there were distinct regional differences. Turnover spikes occurred across the Permo-Triassic boundary in all regions, aside for the Tangaroan where the spike manifested in the Olenekian rather than in the Induan. West and East Circumtethys showed comparable trends through the Late Permian to Carnian, and in both cases turnover throughout the Middle to Late Triassic was lower compared to the Early Triassic. From the Norian onwards, however, West Circumtethys showed steadily greater dissimilarity through time into the Sinemurian, while dissimilarity steadily declined in East Circumtethys before spiking across the Triassic-Jurassic boundary. In the Boreal and North Panthalassic regions, more prominent changes in turnover occurred throughout the Late Triassic, and both regions showed generally greater turnover than in the Circumtethyan regions. Dissimilarity through time was also more pronounced in the Tangaroan and South Panthalassic regions, which may reflect the impact of reduced sampling in both regions leading to greater incompleteness.

3.4 Discussion

3.4.1 Limitations and impact of spatial standardisation

While the spatial standardisation workflow was successful in minimising spatial bias, its utility is potentially restricted to large geographic samples and may not scale to smaller regions or clades, as increasing spatial or taxonomic granularity would increase the patchiness of sampling through time. Instead, local stratigraphic sections will continue to provide the data required to analyse diversity dynamics at local scales with high temporal resolution. As with the choice of sliding window geometry, the choice of target standardisation extent is dependent on multiple factors, including the availability of data in the initial subsample and potential trade-offs between the length contributed to the MST by its component grid cells versus the amount of data contained in those grid cells. Consequently, there may be scope to develop a Pareto-optimal solution to the subsampling workflow using multi-criterion MSTs (i.e., minimum spanning trees that are constructed to satisfy multiple dataset properties in a trade-

off) to optimise spatial extents and maximise data retention simultaneously. Demarcating spatial regions using sliding windows is subjective, but it might be possible to identify spatial regions more objectively using network approaches that detect biogeographic continuity through time (Kocsis et al., 2018a, 2018b; Rojas et al., 2021). Alternatively, the spatial regions could be designed arbitrarily to assess the robustness of emergent patterns to sampling region choice, or designed systematically to assess spatially structured patterns, for example using even latitudinal bands.

Prior to standardisation, significant correlations between spatial extent and diversity corroborate previous findings that variation in the former distorts the latter (Close et al., 2020a, 2020b). Not all correlations were significant, however, suggesting that at a regional scale the otherwise strong relationship noted at the global level (Close et al., 2020b) begins to decouple. Nonetheless, spatial variation in a fossil dataset will still affect measured diversity, even if the net changes in diversity and spatial extent are uncorrelated and so it remains important to reduce this spatial variation to isolate true origination and extinction rates. Significant correlations between diversity and spatial extent after standardisation are unexpected, but these cases are infrequent and may be spurious given that spatial variation is heavily curtailed, substantially reducing its impact on diversity through time.

The limited degree of qualitative change between rate curves compared to diversity curves for each data standardisation treatment (Figs. S43–S49) shows that taxonomic ranges in the fossil record are more robust to spatial sampling bias than standing diversity. The quantitative differences between standardisation treatments, however, demonstrate that spatial sampling bias still affects origination and extinction rates. As spatial sampling extent grows so too does the likelihood that some of the increasing number of sampled fossil occurrences will be of new taxa, inflating observed fossil diversity under the species-area effect (Close et al., 2020b). Greater spatial extent also conceivably increases the likelihood that some of the sampled occurrences will represent the FADs and LADs for their lineages. Thus, as spatial extent fluctuates through time, fluctuation in the capture of FADs and LADs will distort origination and extinction rates even if the FADs and LADs themselves are relatively unbiased. Preservation rate, however, may still vary independently of the spatial extent of a sample (for example, the influence of geographically localised, but taxonomically well-sampled lagerstätten; Muscente et al., 2017) and so will continue to distort FADs and LADs, along with their potential for discovery as the spatial extent of the fossil record fluctuates. Further, while

a taxon has an absolute FAD and LAD, these values may vary regionally through heterogeneity in the time required for a taxon to disperse from its point of origin into a new region and in the timing of its extirpation (Allen et al., 2023). Thus, the fossil record may fail to capture absolute taxon FADs and LADs through fluctuation in spatial sampling extent and instead only preserve regional originations and extirpations, biasing individual taxon ranges.

Not only does spatial sampling bias affect rate estimates, but spatial variation in sampling intensity also biases the composition of the 'global' fossil record. The differences between diversity from the total Circumtethyan dataset to those from its eastern and western subdomains demonstrates variation in diversity dynamics at different spatial scales (Fig. 3.7, 3.8). Data from West Circumtethys comprise the largest portion of the composite dataset and the region shows a similar taxonomic composition to that of the total dataset (Fig. 3.2). This is not unexpected given the historical intensity of sampling in Europe (Raja et al., 2022) but suggests that data sampling from West Circumtethys exerts a disproportionate influence on global diversity trends at least for this study interval. The regional heterogeneity recovered here further demonstrates that the quasi 'global' signal from West Circumtethys is not representative of diversity dynamics elsewhere. Consequently, major biotic events described from the supposedly global fossil record must be scrutinised to determine the degree to which they manifest at a regional level, or whether they are primarily West Tethyan phenomena.

3.4.2 Regional heterogeneity during major biotic events

Regional diversity dynamics all support the Permo-Triassic mass extinction as a global event (Figs. 3.7, 3.8), but extinction intensity shows a degree of latitudinal structuring between regions. The greatest deficits in origination rates and diversity crashes occurred in the Circumtethyan and North Panthalassic regions (Fig. 3.7C, 3.7D, 3.5F, 3.8C, 3.8D, 3.8F), which strongly sample the equator and tropics, in contrast to more modest deficits and declines in the high latitude Boreal and Tangaroan regions. This is consistent with the geological evidence for extreme ocean warming at low latitudes (Sun et al., 2012), along with the flattening of the latitudinal diversity gradient across the equator and tropics in the earliest Triassic (Song et al., 2018) Field data indicates that diversity loss in the Boreal region during the EPME was catastrophic, however, which is at odds with the lack of change in the curves from mcmcDivE and may reflect a limitation of applying probabilistic approaches that infer time-averaged processes from more strongly episodic data. Recovery from the PTME was also regionally

heterogenous, suggesting that varying abiotic conditions and taxonomic compositions across different regions influenced ecosystem stabilisation in the aftermath of the extinction. Extinction rates remained high throughout the earliest Triassic, but soon dipped below a relatively constant origination rate in the high latitude Tangaroan and Boreal regions (Fig. 3.7B, 3.7E). The credible intervals on origination rates in the latter, however, indicate that spikes of origination may have taken place in the earliest Triassic, indicating that the PTME and its aftermath may have manifested as pulses of turnover rather than a steady increase in diversity. Steady recovery was instead seen in the Circumtethyan regions, with modest spikes in median origination rate in the wake of the extinction pulse (Fig. 3.7C, 3.7D). In the North Panthalassic region, however, massive spikes in origination exceeding extinction took place in the immediate aftermath of the PTME. Although this pattern may be influenced by the change in the spatial extent of the data, the confidence interval on extinction rate still clearly picks out the PTME, while the peaks in origination rate fall fully within the well-standardised portion of the dataset (Fig. 3.7F). This confirms rapid and strong recovery from the event in this region and is well supported by the existence of widespread and exceptionally diverse marine assemblages just three million years after the PTME in the North Panthalassic region (Brayard et al., 2017; Smith et al., 2021; Dai et al., 2023). These differences may indicate different ecological dynamics underpinning the recovery at different latitudes, with re-entry of surviving or opportunistic lineages into newly vacated niches at low latitudes versus chaotic patterns of turnover at high latitudes, driven by the invasion of survivors in ecologically stressed refugia (Song et al., 2018).

The timing and placement of pulses of origination and extinction throughout the Middle Triassic are variable and do not correspond to any proposed global events. This heterogeneity continued through the Carnian and may reflect the role of regionally unique macroecological influences on diversity, along with the regionally variable quality of the fossil record. Sedimentological evidence for regionally synchronous environmental upheaval during the CPE is globally pervasive (Simms et al., 2015; Jiang et al., 2019; Dal Corso et al., 2020) and four distinct pulses of volcanism and carbon isotope excursion, linked to the eruption of the Wrangellia Large Igneous Province, can be identified with confidence during the CPE in both East (Lu et al., 2021) and West Circumtethys (Mazaheri-Johari et al., 2021). Only West Circumtethys, however, shows the signal of biotic crisis during the CPE, with strongly negative diversification rates and a sharp crash in diversity (Fig. 3.9C). A diversity crash is also well

supported in the Tangaroan region, but the diversification rates show high uncertainty (Fig. 3.9E), while negative diversification occurred in the Boreal and East Circumtethys regions but without a substantial crash in diversity (Fig. 3.9B, 3.9D). Conversely, diversity increased sharply in the North Panthalassic region with an accompanying pulse of strong diversification (Fig. 3.9F). Intriguingly, there is more consistent evidence in each region for a diversity crash at the Carnian-Norian boundary in all regions, bar the South Panthalassic which does not extend to this interval. While there is some geological evidence in East Circumtethys for genuine environmental fluctuations at the end of the Carnian (Jin et al., 2019), it may instead be the case that the temporal resolution of many of the occurrences in each region is driving this signal. Even though most of the data is constrained to substage level, for stages divided into an early and a late substage (as is the case for the Carnian), FADs of early substage occurrences and LADs of late substage occurrences will still coincide with stage level divisions and so may continue to drive apparent changes in rates and diversity across these boundaries. This suggests that the Permo-Jurassic data in the PBDB may be approaching its analytical limit, even when coupled with model-based estimation methods that can account for temporal uncertainty. There was no strong change in turnover in any region across the CPE or the Carnian-Norian boundary. While there is still dissimilarity ranging from 0.2–0.5, there were no sharp increases in turnover that would otherwise be expected because of a sudden crash in diversity. Consequently, the ecological signature of turnover throughout the Carnian appears subdued compared to that across the Permo-Triassic boundary.

Compared to the EPME, the signal of the ETME is more complex. The onset of negative diversification rates at the Triassic-Jurassic boundary was abrupt in all regions, aside from East Circumtethys where they became steadily more negative throughout the Rhaetian (Fig. 3.8D) and with only weak support in the Tangaroan. Given the proposed mechanism of FAD-LAD sampling, the sharp contraction in spatial extent during standardisation is expected to mute origination and extinction rates during the Hettangian, suggesting that the strongly negative diversification signal is genuine. Diversity loss around the TJ boundary was only substantial in the North Panthalassic region (Fig. 3.F) but reduced in the others, further indicating that it is a poor proxy for diversification dynamics. In the Boreal and West Circumtethys regions, turnover increased only modestly across the boundary (Fig. 3.7B, 3.7C), following on from steadily increasing dissimilarity throughout the Late Triassic, suggesting that the ecological impact of the event merely represented the zenith of long-term turnover starting well before

the extinction boundary. In the Tangaroan region, however, turnover declined across the event (Fig. 3.7E), showing that the change in faunal composition of the region across the extinction boundary was not as marked compared to earlier change taking place throughout the Late Triassic.

High-resolution records of the ETME from stratigraphic sections confirm that the event was complex, with multiple pulses of extinction separated by a few hundred thousand years (Wignall et al., 2020), and mercury anomalies indicating that hostile environmental conditions resulting from continued eruptive phases of the Central Atlantic Magmatic Province (CAMP) extended into the Hettangian by a similar degree (Thibodeau et al., 2016; Percival et al., 2017; Beith et al., 2021), matched by the persistent negative diversification rates in each region throughout the Hettangian. This is therefore unusual given the more muted changes in diversity across the event. Analysis of the Phanerozoic fossil record as a series of eco-evolutionary units based on taxon co-occurrences through time has shown that the TJME had a significant impact at the ordinal level, with prominent ecological restructuring particularly among reef communities, but little impact at the family or generic levels (Muscente et al., 2018). Thus, while strong ecological and environmental change certainly took place at the Triassic-Jurassic boundary in concert with CAMP volcanism (Davies et al., 2017), this may have been predicated on relatively small generic changes suggestive of the loss of keystone species. Similarly, previous work using PBDB data and traditional bin-based approaches found little change in functional richness at the TJ boundary and a reduced impact of the ETME in the Tethys and Boreal oceans compared to the Panthalassic despite high diversity loss across all three regions (Dunhill et al., 2018), supporting those aspects of these results and further suggesting that the ecological and taxonomic severities of the ETME were somewhat decoupled.

3.4.3 Reef demise and regional extinction

There is strong correspondence between global diversity in deep time and the history of reef ecosystems (Sheehan 1985; Muscente et al., 2018), with reefs acting as cradles of biodiversity and evolutionary innovation throughout the Phanerozoic (Kiessling et al., 2010) but displaying high sensitivity to strong environmental disturbances such as those during mass extinctions (Kiessling and Simpson 2010). I tentatively identify two key instances of this relationship from these analyses. The strongest evidence for biotic upheaval during the CPE comes from West Circumtethys (Fig. 3.8C), driven by the decline of carbonate platforms and hyper-diverse reef

assemblages in the European geological record (Jin et al., 2020). On this basis, it has been proposed previously that not only was the CPE a primarily West Tethyan phenomenon, but also that the apparent scale of the diversity crash was exacerbated by the loss of these assemblages and environments (Hallam 1981; Sepkoski and Raup 1981). The evidence for environmental perturbation during the CPE is globally distributed (Dal Corso et al., 2020), however, and there is evidence for diversity decline in other regions to some extent. In a global diversity curve, the loss of ecologically diverse West Tethyan reef systems may be viewed as a statistical artefact, but decomposition of the global signal into regional subsets transforms this artefact into an empirical aspect of Tethyan biogeographic history. As a modern analogue, the Great Barrier Reef is individually one of the most diverse habitats on the planet and its decline is viewed as a genuine and catastrophic aspect of the current global diversity crisis (Plaisance et al., 2011), rather than as a regional anomaly.

A similar pattern was present in East Circumtethys during the Late Permian where the development of ecologically diverse reef systems across a regionally extensive carbonate platform (Payne et al., 2006) coincided with a large pulse of origination and a corresponding diversity zenith, followed by catastrophic extinction and diversity loss at the EPME (Fig. 3.7D, 3.8D). These regional analyses highlight the spatial component of the correspondence between reef systems and Phanerozoic marine diversity, with the regional loss of reef systems contributing substantially to global marine diversity crises. Thus, while large evolutionary events may have global signatures in the fossil record, they may also display regional epicentres due to the interactions of spatially non-random controls on diversity with diversity drivers operating at global scales. Across the ETME, reefs were widely distributed, and so their relationship with global diversity approached a global trend, rather than displaying any distinct regionalisation, with previous studies confirming the severity of the event for reefs globally (Kiessling et al., 2007; Dunhill et al., 2018; Palfy et al., 2021).

3.5 Conclusions

My approach to examining the fossil record provides a powerful way to decompose global diversity trends into their regional components, but its scope remains reliant on the availability of high-quality occurrence data. As such, these methods will be well suited to examining major

biotic events in other transects of geological history, for example the poorly resolved Late Devonian mass extinction. Full resolution of some events, however, may be hindered by the current quality of fossil occurrence data. Continued analytical gain will come from refinement of occurrence ages, either through the literature-based approach applied here or through stratigraphic modelling approaches like CONOP.SAGA (Fan et al., 2020). Similarly, a regional view of Triassic diversity dynamics will be aided by improved spatial coverage of the fossil record, although this remains contingent on the availability of fossiliferous sedimentary rocks around the globe. Otherwise, a nuanced understanding of the differences between diversification signals at the section level will continue to provide a fine-controlled means of decomposing global biotic history into its regional components.

Chapter Four

Intrinsic versus Extrinsic Drivers of the Origins of Scleractinian Coral Reefs

Scleractinian coral reefs originated in the Triassic during a remarkable period of ecological recovery following the greatest mass extinction in geological history. Their emergence has been attributed to climate change and ecological succession from older, dominant reef builders, but these hypotheses remain unexplored using deep time climate reconstructions or modern macroevolutionary methods for quantifying the rates, timing, and drivers of diversification. Here, I use ecological niche modelling and Bayesian approaches to quantifying origination and extinction dynamics from fossil occurrence data to test hypothesised abiotic and biotic drivers of Triassic reef succession and the rise of scleractinian corals. I recover patterns of coral and sponge diversification that are broadly concordant with trends in the abundances of the reefs they constructed, although the drivers of their origination and extinction rates remain difficult to confidently identify. Similarly, trends in habitable area estimated from ecological niche modelling are affected by the choice on plate tectonic reconstruction, necessitating consideration of the choice of the palaeogeographic boundary conditions within which the fossil record is interpreted. While these results highlight the uncertainties and challenges faced in analyses of the coral fossil record during this interval, they indicate that abiotic rather than biotic dynamics were the likely drivers of long term Triassic reef succession, but that there is still an outstanding question of what drove shorter, more abrupt changes in the ecological incumbencies of different reef builders.

Author contributions: This chapter is based on a collaboration with A. Farnsworth, D. Lunt and J. Landwehrs. A. F. and D. L. ran the HadCM3L climate models. J. L. ran the CLIMBER-3α climate models. The author conducted all other analyses. Comments were provided on a draft of this chapter by M. J. Benton. The author is otherwise responsible for the vast majority of the work in this research chapter.

4.1 Introduction

Reefs have been cradles of marine biodiversity for the last five hundred million years of evolutionary history (Kiessling et al., 2010; Close et al., 2020b). Modern-type coral reefs built by scleractinian corals originated during wider recovery of the marine biosphere following the end-Permian mass extinction (EPME) 251.9 million years ago (Kershaw et al., 2011; Martindale et al., 2019). Biostromes built by algal and microbial frameworks proliferated in the aftermath of the catastrophe in the Early Triassic before sponges took over as the first true reef builders in the Middle Triassic – the Wetterstein phase of Triassic reef evolution (Fig. 4.1). Scleractinian corals then underwent a major phase of diversification in the Late Triassic to emerge as dominant reef builders for the first time in Earth history, the Dachstein phase of Triassic reef evolution (Flugel 2002; Martindale et al., 2019) and a critical step towards the marine biodiversity 'explosion' during the Mesozoic as ecosystems transitioned from Palaeozoic- to Modern-type faunas (Wu and Benton 2022).

While the transition from Wetterstein to Dachstein reefs is a well-established event, the scleractinian diversity dynamics underpinning this shift are obscured by the complex nature of their fossil and molecular records. Simpson et al. (2011) considered these records to broadly support one another but their root age for Scleractinia was substantially underestimated based on their choice of a conservative choice of root age prior and sampling of exclusively photosymbiotic, zooxanthellate (z) lineages all of which originated from the Jurassic onwards; later studies have demonstrated that inclusion of non-photosymbiotic, azooxanthellate (az) lineages yields root ages in the Middle to Late Ordovician (Stolarski et al., 2011; Seiblitz et al., 2020). The gap in their phylogeny between root age and the oldest scleractinian divergences consequently spanned the Early Triassic to Late Jurassic, hindering phylogenetic inference of diversification rates (itself problematic, see 1.2.1) through this clearly dramatic interval in their evolutionary history. In parallel, their fossil-based analyses used traditional methods for inferring diversity dynamics come with known limitations (see 1.2.2), while their poor stage-level resolution masked the nuances of their rise to prominence through the coarsely subdivided Late Triassic. Later work, reanalysing the data from Simpson et al. (2011) subsequently identified substantial conflict between the fossil and phylogenetic records of scleractinian diversification which could not be resolved by accounting for differences in origination mode, likely due to poor sampling within extant phylogenies and challenging fossil taxonomy (Silvestro et al., 2018).

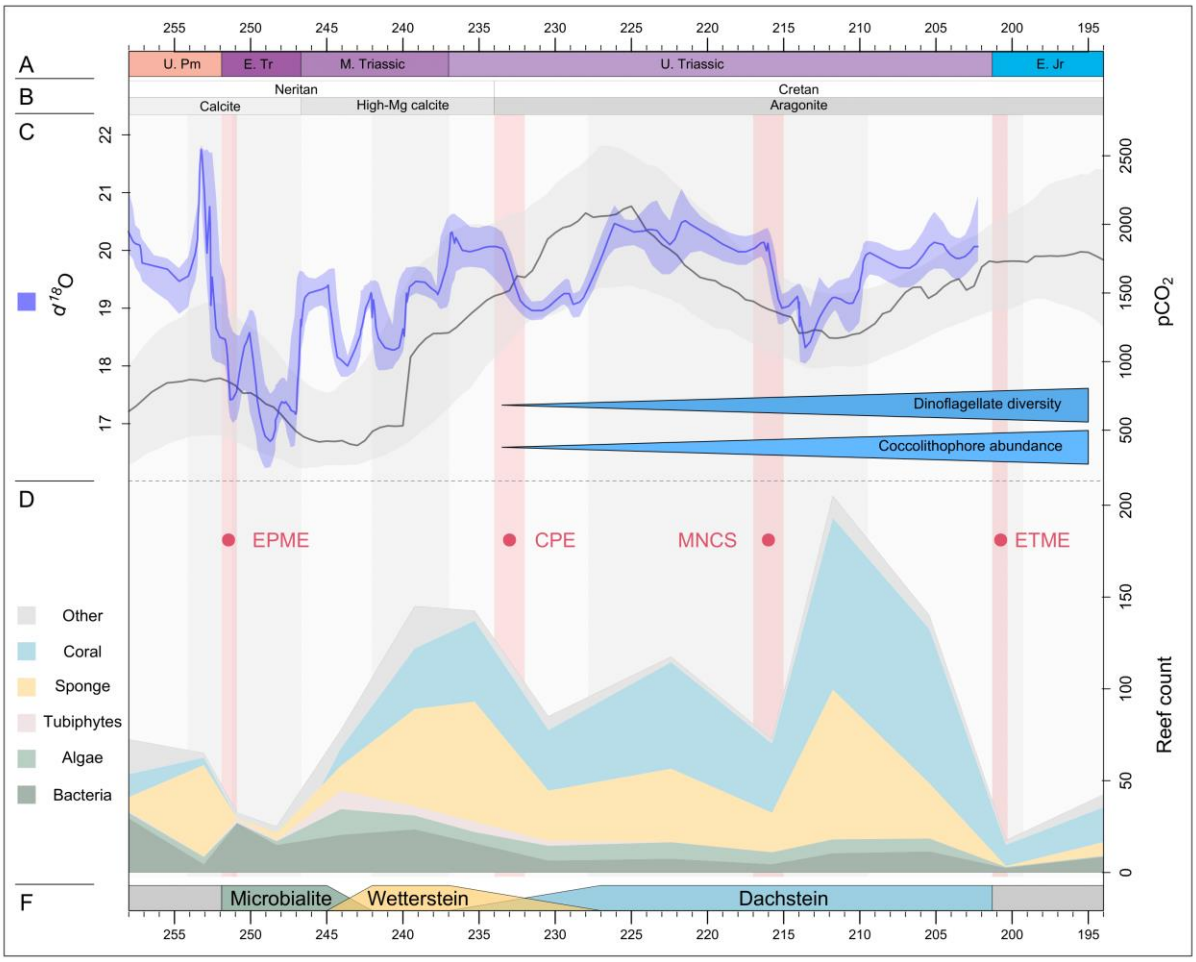


Fig. 4.1. Reefs and the Earth-system through the Late Permian to Early Jurassic. (A). The Permian-Triassic timescale (time units in Mya). **(B).** Ocean carbonate chemistry phases. **(C).** Oxygen isotopes (as a proxy for temperature) and atmospheric CO_2, along with trends in pelagic plankton diversity. **(D).** Tallies of reefs through the Late Permian to Early Jurassic divided by builder type, tabulated following Kiessling (2010), where the primary and secondary reef building taxa are upweighted and downweighed respectively. **(F).** Major reef builder phases. Several climatic upheavals are highlighted in red: the end-Permian mass extinction (EPME), the Carnian Pluvial Episode (CPE), the Middle Norian Climate Shift (MNCS) and the end-Triassic mass extinction (ETME).

Our unclear picture of scleractinian diversification dynamics hinders our understanding of the factors that may have influenced their initial rise to ecological prominence. Several factors could explain the Wetterstein–Dachstein transition during the Late Triassic, including falling CO_2 and cooling climate promoting the acquisition of photosymbionts by scleractinians (Kiessling et al., 2009), but these have never received formal testing. Similarly, little attention has been given to the role of biotic competition in this succession despite their potential importance in coral reef establishment and evolution through the Phanerozoic (Kiessling et al., 2010; Spiridonov and Lovejoy 2022). Reef recovery following the EPME also presents an

intriguing inversion of a trend observed in corals within reef ecosystems stressed by anthropogenic climate change and habitat disturbance, where they are outcompeted by environmentally hardier associations of algae and sponges (Chadwick et al., 2010; Bell et al., 2013; Chaves-Fonnegra et al., 2017; Bell et al., 2018). Despite this observation, however, the potential role of diversity dependence as a driver of the Wetterstein-Dachstein transition has never been considered. Besides the lack of formal hypothesis testing for the rise of scleractinian reefs, severe spatial sampling biases in the geological record (Close et al., 2020b; Jones and Eichenseer 2022) may limit detection of the 'true' signal of diversification from fossil record alone, necessitating re-evaluation using methods which can circumvent these limitations. While the Triassic fossil record fortuitously shows the most complete spatial coverage of any geological interval besides the near present (Jones et al., 2021), alterations to its chronostratigraphic timescale necessitate revision and improved resolution of the rate and timing of reef builder diversification, particularly through the Carnian Pluvial Episode (CPE) and the often undivided, seemingly monolithic 18 million years of the succeeding Norian stage (Martindale et al., 2019; Dal Corso et al., 2020; Sun et al., 2020).

In this chapter I investigate shallow marine reef builder dynamics from the Late Permian to Early Jurassic using analytical approaches that address the limitations and problems outlined above. I analyse taxonomically revised coral and sponge fossil data from the Paleobiology Database and Paleoreefs Database, with substage precision through the critical Carnian and Norian stages, and differentiation of zooxanthellate and azooxanthellate coral lineages. I infer sampling-corrected diversification rates and diversity using PyRate (Silvestro et al., 2014, 2019) and mcmcDivE (Flannery-Sutherland et al., 2022) which overcome the shortcomings of traditional analytical methods, then investigate whether hypothesised abiotic and biotic drivers for the Wetterstein to Dachstein transition can explain their diversification dynamics. For biotic drivers, I test the hypothesis that diversity dependence between the analysed clades controlled their diversification rates, versus a range of abiotic drivers, namely atmospheric CO_2 concentration. O_2 availability, nutrient flux from weathering, sea level, and continental fragmentation. As the geological record is spatially biased, I also use ecological niche modelling as an independent, spatially explicit analytical approach to investigate the hypothesised influence of global temperature on the rise of scleractinian reefs by leveraging advances in deep time climate modelling to characterise how the Earth-system may have promoted or constrained reef builder proliferation.

4.2 Materials and methods

4.2.1 Modern and fossil coral, sponge, and reef occurrence data

Present day reef occurrence data, and scleractinian and sponge occurrence data were downloaded from ReefBase (http://www.reefbase.org/main.aspx) and the Ocean Biodiversity Information System (www.obis.org) respectively on 23/10/22 and processed in R (v.4.0.4). Azooxanthellate versus zooxanthellate states were assigned to the scleractinian occurrences using the Coral Trait Database (Madin et al., 2016), then occurrences without photosymbiotic states (non-species level, or without matches in the Coral Traits Database) excluded. Occurrences were inspected for suspect geographic coordinates using the CoordinateCleaner R package (Zizka et al., 2019), removing those flagged as potentially erroneous, along with any non-specimen occurrences, or occurrences detected by machine learning and data mining. All non-reef coral communities were additionally excluded from the ReefBase dataset, along with all occurrences occurring on the land or at depths below 200 metres from both datasets, as these occurrences do not inform on reef builder occurrences within the photic zone.

Late Permian (Wuchiapingian) to Early Jurassic (Sinemurian) fossil reef occurrences and fossil coral and sponge occurrences were download from the PaleoReefs Database (PARED; Kiessling and Krause 2022) and the Paleobiology Database (PBDB) respectively on 02/01/22 and processed in R (v.4.0.4). Chronostratigraphic ages in both datasets were updated according to A Geologic Timescale 2020 (Gradstein et al., 2020) using the `chrono_scale()` function of the *fossilbrush* R package (Flannery-Sutherland et al., 2022b) and the constraints for Carnian and Norian-aged occurrences in PARED revised to substage resolution where possible using secondary literature data (Sun et al., 2020). Some PARED reefs have linked PBDB collection numbers, so PARED and PBDB entries representing the same collection were assigned the same stratigraphic age range to maintain consistency between the two datasets. Next, any PBDB occurrences with stratigraphic ranges exceeding the duration of the longest geological interval in study window (the Norian, 18 Ma), or with stratigraphic ranges greater than 10 Ma that also spanned multiple international geological stages, were discarded. PBDB occurrences at the genus level were examined for potential taxonomic errors using the `check_taxonomy()` function of *fossilbrush*.

PARED reefs record the higher taxonomies of their primary and secondary reef builders. These fields can differ taxonomically, making a PARED entry comparable to a PBDB collection containing two occurrences classifiable at higher taxonomic levels using the reef

builder identities. For example, a PARED reef may have sponges and corals recorded as its primary and secondary reef builders respectively, equivalent to a PBDB collection containing the occurrences Anthozoa *indet.* (class-level) and Porifera *indet.* (phylum-level). I created a corresponding higher taxonomy field for PBDB entries that aligned with the PARED primary reef builder taxonomy scheme, then a duplicate dummy field to serve as the 'secondary' taxonomic assignment. The PARED and PBDB datasets were then combined, treating the PARED reef types as equivalent to the PBDB environments. Finally, any entries from terrestrial environments or those with missing chronostratigraphic ages or present-day coordinates were excluded.

4.2.2 Palaeogeographic models and palaeocoordinates

Palaeocoordinates were calculated under two palaeogeographic models in R (v.4.0.4): PALEOMAP (Scotese and Wright 2018) and Getech (Getech Plc.). While the PALEOMAP and Getech plate rotation models and implemented rotation methods are fundamentally different (see below), the discrepancies in palaeolatitude and palaeolongitude for fossil occurrences are generally small (Fig. 4.2), indicating that the broad spatial trends in the coral, sponge and reef records through time will be somewhat comparable. PALEOMAP palaeocoordinates were derived directly from a time-continuous plate rotation model via the GPlates API (118) using the `palaeorotate()` function of the *palaeoverse* R package (Jones et al., 2023) to calculate point rotations from the occurrence midpoint ages.

Getech palaeocoordinates were obtained from stage-level rotations of a relatively coarse 2.5° x 3.75° spatial grid as the time-continuous plate rotation model was unavailable for use. Consequently, palaeogeographic uncertainty arising from stratigraphic uncertainty could not be assessed. To improve spatial resolution, however, stage-level point rotations were approximated by first taking the offset between an occurrence coordinate and its parent grid cell centre in the present, then calculating the rotation experienced by that parent grid cell as the change in angle between the parent and its immediate northward neighbour in the present to the desired reconstruction stage (i.e., from $\Delta^{o} = 0$ in the present, to some Δ^{o} in the past; for cells along the top of the grid, the southward neighbour was used instead). The final reconstructed coordinate is then taken as the offset from the rotated grid cell centre, along the angle of rotation. Comparison of the results from this procedure to rotations based purely on

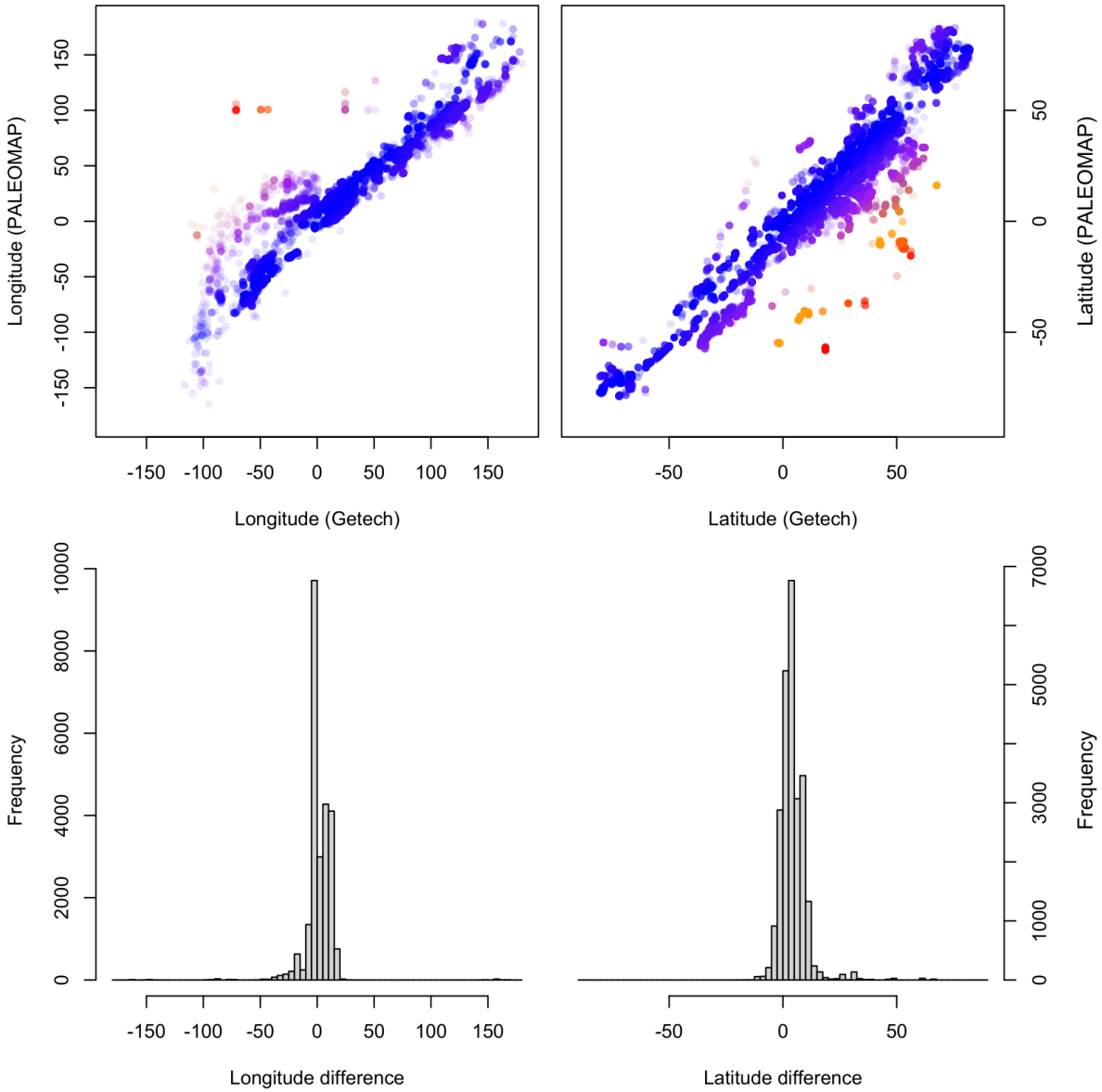


Fig. 4.2. Differences in fossil occurrence latitudes and longitudes between palaeogeographies. Discrepancies in palaeolatitude and palaeolongitude between the Getech and PALEOMAP plate rotation models are generally small, indicating that the broad spatial trends in the coral, sponge and reef records through time are somewhat comparable.

the nearest grid cell centre showed that both methods returned similar results, with discrepancies in location falling within the expected range of cell centre offsets (Fig. S50).

4.2.3 Climate simulation data

To quantify the ecological niches of corals and sponges, along with the distributions of their suitable environments through geological time, I used simulation data from two climate models: a fully coupled atmosphere-ocean general circulation model (GCM) HadCM3L

(Valdes et al., 2017), and an Earth-system model of intermediate complexity (EMIC) CLIMBER-3α (Montoya et al., 2005). Both models have shown skill in reconstructing broad-scale climatic patterns (Montoya et al., 2005; Chiarenza et al., 2020), making them suitable for the global-scale niche modelling approach attempted here.

The HadCM3L GCM has a spatial resolution of 2.5° x 3.75° latitude by longitude with 19 vertical levels in the atmosphere and 20 vertical levels in the ocean. Sea-ice is calculated on a zero-layer model on top of the ocean surface. Vegetation is predicted as a fraction for each grid box using the MOSES 2.1 land surface scheme (Cox et al., 1999) and the dynamic vegetation model TRIFFID (Top-Down Representation of Interactive Foliage and Flora Including Dynamics; in equilibrium mode HadCM3BL-M2.1aE, in fully dynamic mode HadCM3BL-M2.1aD). Stage-specific solar luminosity was calculated following Gough (1981). Boundary conditions for ice sheet configuration, topography and bathymetry were downscaled from 0.5 x 0.5° latitude by longitude to model resolution. Each simulation was run for 7,500 model years, having been initialised from a pre-industrial state, to allow surface and deep ocean levels to reach equilibrium and achieve an equilibrium state with close to zero net energy imbalance at the top of the atmosphere. This is fundamental as ocean circulation can take many thousands of model years to establish (Valdes et al., 2021). Climate means are calculated from the last 100 years of each simulation. Compared to Valdes et al. (2017), the model underlying the simulations presented here contains several updates, mostly designed to improve the representation of polar amplification in past warm climates, using similar methods to Sagoo et al. (2013).

The CLIMBER-3α EMIC has a spatial resolution of 7.5° x 22.5° latitude by longitude in the atmosphere and 3.75° x 3.75° latitude by longitude in the ocean, with 11 vertical levels in the ocean and a single statistical-dynamic atmospheric layer using vertically averaged prognostic equations for temperature and humidity. Sea ice is calculated on a two-layer model with one layer for ice and one for snow. Vegetation is prescribed using proxy-based estimates of climatic zonation and the land surface scheme BATS (Biosphere-Atmosphere Transfer Scheme) (Dickinson et al., 1986). Stage-specific solar luminosity was calculated from its present-day value following Bahcall et al. (2001). Boundary conditions of topography and bathymetry were downscaled from 0.5 x 0.5° latitude by longitude to model resolution (ice sheets were not included). Each experiment was run for 5000 model years, having been initialised from a pre-industrial ocean temperature and salinity profile, and climate means calculated from the last 500 years of each simulation.

The climate model simulations (Table 4.1). cover a range of palaeogeographic boundary conditions and atmospheric CO_2 values taken from the curve of Foster et al. (2017). One set of climate data come from HadCM3BL-M2.1aE simulations with palaeogeographic boundary conditions taken from stage-level Getech digital elevation models (DEMs) reconstructed under highstand sea level (Lunt et al., 2016). Potential uncertainty associated with both the modelling method and underlying DEM is then captured by two other sets of simulations using HadCM3BL-M2.1aD and CLIMBER-3α with palaeogeographic boundary conditions taken from PALEOMAP DEMs (Landwehrs et al., 2021a, 2021b; Scotese and Wright 2018). These DEMs are supplied at five-million-year intervals but are mostly considered to represent the average palaeogeography for the entire stratigraphic stage within which the reconstruction age falls (Scotese 2020, pers. comm.), e.g., the 245 Ma DEM is representative of the Anisian stage (247.2 – 244.6 Ma). The exceptions are for stages longer than 10 million years, in the case of this study the Carnian and the Norian, where the DEM intervals correspond to early and late, and early, middle and late subdivisions respectively.

HadCM3BL-M2.1aE simulations with Getech palaeogeographic boundary conditions and HadCM3BL-M2.1aD simulations with PALEOMAP palaeogeographic boundary conditions were run using CO_2 values taken from the proxy estimates of Foster et al. (2017), based on the stage midpoint ages for Getech-based simulations and the DEM reconstruction ages for PALEOMAP-based simulations. Additional HadCM3BL-M2.1aE simulations were then run for the early Carnian, early Norian and late Norian, using the Getech stage-level boundary conditions, but substage-appropriate CO_2 values. CLIMBER-3α simulations with PALEOMAP boundary conditions were run at with several CO_2 values for each DEM (Landwehrs et al., 2021a, 2021b), so the closest combinations of palaeogeographic boundary conditions and atmospheric conditions was selected according to the proxy estimates of Foster et al. (2017). A single HadCM3BL-M2.1aE simulation with present day geography and pre-industrial atmospheric CO_2 was also used. All HadCM3L simulations are archived in the Providing Unified Model Access (PUMA) repository of the National Centre for Atmospheric Science Computer Modelling Service (NCAS-CMS). All CLIMBER-3α simulations are available from the GFZ-Potsdam archive (Landwehrs et al., 2021b).

Table 4.1. Interval-wise climate simulation atmospheric and palaeogeographic effects

	HadCM3L		CLIMBER-3α
Interval	Getech[1]	PALEOMAP[2]	PALEOMAP
Wuchiapingian	810.9 ppm, 256.8 Ma	810.9 ppm, 255 Ma	800 ppm, 255 Ma
Changhsingian	856.3 ppm, 253.0 Ma	810.9 ppm, 255 Ma	800 ppm, 255 Ma
Induan	878.7 ppm, 250.8 Ma	878.7 ppm, 250 Ma	700 ppm, 250 Ma
Olenekian	796.7 ppm, 248.3 Ma	878.7 ppm, 250 Ma	500 ppm, 250 Ma
Anisian	418.5 ppm, 244.1 Ma	418.5 ppm, 245 Ma	400 ppm, 245 Ma
Ladinian	1033.5 ppm, 239.2 Ma	1033.5 ppm, 240 Ma	800 ppm, 240 Ma
Early Carnian	1120.0 ppm, 233.5 Ma	1491.9 ppm, 235 Ma	1400 ppm, 235 Ma
Late Carnian	1613.5 ppm, 233.5 Ma	1613.5 ppm, 230 Ma	1900 ppm, 230 Ma
Early Norian	2058.0 ppm, 217.9 Ma	2059.0 ppm, 225 Ma	2000 ppm, 225 Ma
Middle Norian	1481.0 ppm, 217.9 Ma	1481.0 ppm, 215 Ma	1600 ppm, 215 Ma
Late Norian	1223.3 ppm, 217.9 Ma	1222.3 ppm, 210 Ma	1300 ppm, 210 Ma
Rhaetian	1503.1 ppm, 205.4 Ma	1503.1 ppm, 205 Ma	1500 ppm, 205 Ma
Hettangian	1728.9 ppm, 200.4 Ma	1728.9 ppm, 200 Ma	1700 ppm, 200 Ma
Sinemurian	1783.6 ppm, 196.2 Ma	1783.6 ppm, 195 Ma	1700 ppm, 195 Ma
Present	-	280 ppm, 0 Ma	-

[1] PUMA codes: tfgsh, tfgsi, tfgsj, tfgsk, tfgsl, tfgsm, tfgxb, tfgsn, tfjia, tfgso, tfjib, tfgsp, tfgsq, tfgsr
[2] PUMA codes: teyDz1, teyDz1, teyDy1, teyDy1, teyDx1, teyDw1, teyDv1, teyDu1, teyDt1, teyDs1, teyDr1, teyDq1, teyDp1, teyDo1, teyDn1, teyfa

As the goal of this study is to examine shallow marine habitable area through geological time, each set of simulations was upscaled to increase the precision at which the photic zone could be distinguished bathymetrically, and to reduce the mismatch between occurrence locations and the extent of the land-sea mask. This upscaling is reasonable given the original downscaling of the DEMs for climate model reconstruction, as the sub-grid topography and bathymetry of those DEMs will still be preserved in original, lower resolution climate model outputs (Chiarenza et al., 2019). Upscaling was performed in R (v.4.0.4) from model resolution to 1° x 1° using nearest-neighbour interpolation to avoid edge effects in oceanic grid cells adjacent to the model boundary cells (the land surface). This resolution was chosen as a compromise between the resolution of the climate model versus the resolution of the original DEMs, following upscaling procedures used by previous analyses (Chiarenza et al., 2019; Jones et al., 2021). In addition, the nearest suitable land or sea cell was used fill any remaining

discrepancies between the climate model land-sea mask and the land-sea mask calculated from the downscaled DEMs at 1° x 1° resolution.

4.2.4 Coral, sponge and reef habitat suitability modelling

Heterogeneous sampling through time means that fossil data for a clade may not capture the true extent of its fundamental abiotic niche (Varela et al., 2011). Rather than calibrating niche models using incomplete and spatially sparse fossil records, models were instead calibrated using the present-day HadCM3L climate simulation outlined above and modern coral, sponge, and reef occurrences, following the approach of Jones et al. (2021). While this comes with the caveat that past and modern abiotic niches may not be equivalent, it has the major advantage of providing a greater quantity of data for model calibration that is much more likely to capture the full extent of the modern abiotic niche, along with greater comparability of model hindcast estimates of habitability between geological intervals (i.e., the niche space remains constant, enabling the effects of changing geographic space to clearly emerge). I chose the Maximum Entropy method (MaxEnt; Phillips et al., 2006, 2017) for ENM calibration as this method has repeatedly been shown to outperform other methods when only data on taxon presences are available, as is the case here (Valavi et al., 2021). Briefly, MaxEnt is a machine-learning algorithm that maximises the contrast between the environmental conditions observed at presence locations, versus the 'background' environmental conditions observed across a spatially broader set of locations, with the resultant model providing a likelihood estimate for the density of presence records (interpreted as the degree of habitat suitability) at a given location conditioned on its environmental data (Phillips et al., 2017). This model can then be projected across the full spatial extent of a set of environmental data to predict the geographic distribution of suitable habitats. As the probabilistic prediction is continuous, this then necessitates the choice of a threshold at which to binarize the projection into habitable versus non-habitable area to formally assess model performance (i.e., its specificity and sensitivity, or the prevalence of true negative and true positive predictions in relation to the original presence data).

I extracted environmental variables from the modern climate simulation with ecological relevance to corals, sponges and reefs in R (v.4.0.4): summer, winter, and annual irradiance; coldest quarter, warmest quarter, coldest month and warmest month sea-surface temperature; mean annual salinity, upwelling and sea-surface temperature; and annual standard deviation of sea-surface temperature. While sponges will be broadly unaffected by solar irradiance and

sensitivity, the other variables will still capture aspects of their fundamental abiotic niche through thermophysiological constraint. The choice of variables is therefore appropriate for all four groups, enabling the construction of comparable niche models and fair comparison of their predicted habitable areas through geological time. These variables were investigated for multicollinearity with the finding that seasonal and monthly sea-surface temperatures were highly correlated (Fig. 4.3), so the monthly means were selected for subsequent modelling. While it is beneficial to select variables which show only a limited degree of covariance (typically a Pearson correlation coefficient < 0.7), this strategy should not be followed blindly as two correlated variables may have determinant effects on an ecological niche, for example minimum and maximum temperature constraints, and MaxEnt can actually take advantage of covariance to optimise model performance (De Marco and Nobrega 2018; Feng et al., 2019; Sillero and Barbosa 2020). The final selected variables, namely minimum and maximum solar irradiance, and the coldest and hottest month temperatures, were consequently chosen to maximise model simplicity and performance (see model calibration and evaluation below), despite displaying a stronger degree of covariance. Other effects like nutrient or oxygen availability would have also ontrolled their distribution over ecological timescales, but thermophysiology is prioritised here to test the hypothesis that cooling drove reef succession in the geological long term

The modern occurrence data were divided into scleractinian reef, sponge, z-coral and az-coral subsets. Each subset was then spatially rarefied to the resolution of the environmental data (1° × 1°) to prevent overweighting of heavily sampled grid cells during model training. 714, 2871, 722 and 1143 unique localities were left for scleractinian reefs, sponges, z-corals and az-corals respectively. The present-day climate simulation data was first masked to exclude grid cells >200 metres depth to match the bathymetric distribution of the occurrence data. 2000 background cells were then randomly sampled from across the unmasked portion of the grid for each occurrence subset, with sampling weights proportional to cell area to correct for latitudinal bias in sampling. Sampling weights for cells containing occurrence data were set to zero to ensure that they were not included within the background data. Ecological niche models were constructed for each set of occurrences using MaxEnt with the environmental data outlined above using the `enmtools.maxent()` function of the *ENMTools* R package. Following previous recommendations (Elith et al., 2010; Kearney et al., 2010), only linear and quadratic features were permitted to approximate realistic response curves (i.e., smooth, unidirectional change in the probability of presences with a corresponding change in each

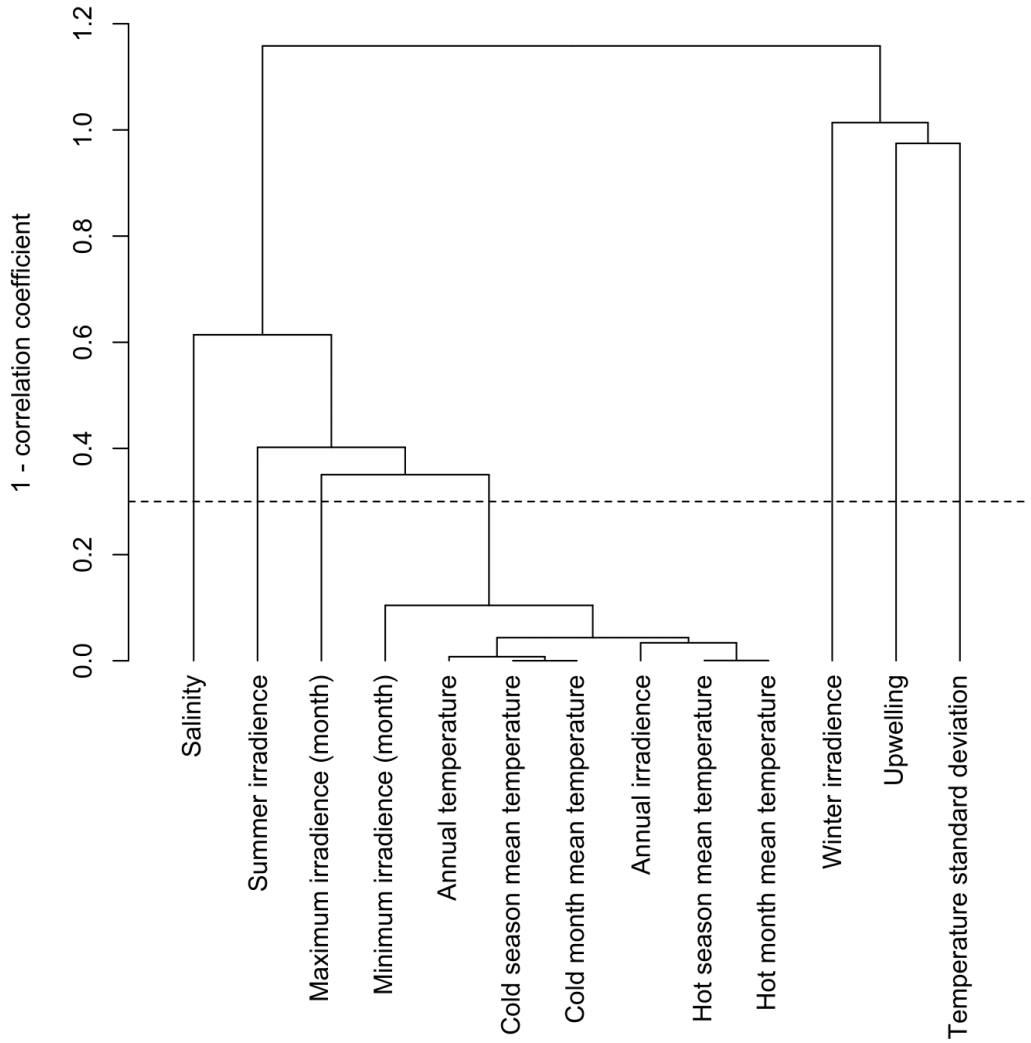


Fig. 4.3. Dendrogram of the covariance between the investigated environmental variables. Final variables selected for ecological niche modelling were minimum and maximum solar irradiance, and the coldest and hottest month temperatures.

environmental variable). Each subset was randomly divided into five folds, withholding 20% of the points each time to enable cross validation. During this procedure, occurrence and background points were divided separately then recombined to ensure that the proportions of both types of points were equal between folds.

Calibrated models for each fold were projected to the full spatial extent of the modern environmental data (Fig. S51) and their predictive capacities assessed using the Area Under the Curve (AUC) statistic and the continuous Boyce index, using the `evaluate()` and `ecospat.boyce()` functions of the *dismo* and *ecospat* R packages respectively (Hijmans et al., 2022; Broennimann et al., 2023). These metrics indicated favourable performance between folds, so final ecological niche models were then calibrated using all available

Table 4.2. Evaluation of ecological niche model performance

Model	AUC	Boyce	mSSS	Z95P
coral reefs	0.95	0.95	0.30	0.43
z-corals	0.93	0.86	0.28	0.43
az-corals	0.90	0.94	0.18	0.20
sponges	0.82	0.93	0.44	0.24

occurrences for each group to maximise the information informing model calibration, with statistical evaluation again confirming good model performance (Table 4.2). Finally, two binarization thresholds were calculated from the continuous, probabilistic model projections to differentiate habitable from non-habitable grid cells: the threshold that includes 95% of the occurrences (L95P) and the threshold that maximises the sum of sensitivity and specificity (mSSS). L95P is related to the least presence threshold but assumes that the lowest 5% of occurrences represent environments which are tolerable, but not favourable to the modelled group. Excluding these occurrences using the 95% threshold therefore helps to reduce potential model noise.

The final set of calibrated ecological niche models (scleractinian reefs, z-corals, az-corals, sponges) was hindcasted to the three sets of Permo-Jurassic climate simulations (Figs. S52–S93). The model hindcasts were binarized at their associated L95P and mSSS thresholds, then grid cells <300 metres depth masked. This deeper threshold was chosen compared to the bathymetric range of the training occurrences (> 200 metres depth) to include the magnitude of short-term sea level fluctuations through geological time and uncertainty in palaeogeographic bathymetry estimates. The predictive performance of the hindcasts was then assessed against the empirical fossil record of corals and coral reefs using the method of Jones et al. (2021). For each hindcast, the same number of cells as spatially rarefied fossil occurrences were randomly sampled from across the grid, repeating this procedure 1000 times. The percentage of habitable points between the random samples versus empirical occurrences was then compared using one-sided Wilcoxon rank-sum tests with the null hypothesis that there was no significant difference in predictive accuracy between the two.

4.2.5 Diversity dynamics

Genus and species level diversification dynamics of Late Permian to Early Jurassic shallow-water corals and sponges, plus subsets for azooxanthellate scleractinians and zooxanthellate scleractinians, were estimated from their fossil occurrence data using PyRate and mcmcDivE (Silvestro et al., 2014, 2019; Flannery-Sutherland et al., 2022a). The subsets of genus and species level sponge and coral occurrences were extracted from the combined dataset, excluding entries with uncertain generic assignments (? And " " modifiers in the identified name) along with a small number of taxonomically dubious entries. While this study focuses on diversity dynamics amongst known reef-building clades in shallow water, photic communities, the marine crisis at the Capitanian-Wuchiapingian boundary resulted in a paucity of reef environments in the Late Permian, despite the survival of coral and sponge taxa more broadly. Exploratory analyses showed that restriction of data to solely shallow-water reef taxa caused the signal of the EPME to be largely lost for both groups (Fig. 4.4) as true reefs were rare in the Late Permian following the end-Capitanian reef crisis (Wang et al., 2017; Huang et al., 2019), resulting in patchy sampling of coral diversity when restricted to reef facies only. To account for this, all Permian-aged occurrences were retained to accurately quantify its impact, but all non-reef occurrences were excluded from the Triassic onwards. This ensured that the corresponding Triassic diversity dynamics pertained strictly to reef taxa, with exploratory analyses demonstrating that Triassic coral and sponge diversity curves derived from reef and non-reef occurrences are largely the same (Fig. 4.4). Azooxanthellate vs zooxanthellate states were then assigned to the Triassic corals at genus level using the database of Kiessling and Kocsis (2015), which is based primarily on morphological proxies for photosymbiosis.

Fifty replicate datasets were generated for each analysis (sponges, corals, z-corals, az-corals) in R (v4.0.4) where the age of each occurrence was randomly fixed within its stratigraphic range, whilst ensuring that occurrences from same locality received the same randomised age. The maximum likelihoods of the three preservation model options were estimated and compared by Akaike's Information Criterion using the `-PPmodeltest` function of PyRate (v3.0). The TPP model with rate shift boundaries largely corresponding to geological stage boundaries was identified as best fitting in all cases (PyRate option `-qShift`). Each PyRate analysis was run with the default Metropolis Hastings sampling algorithm for 50 million generations, sampling every 5000, using rjMCMC to propose temporal

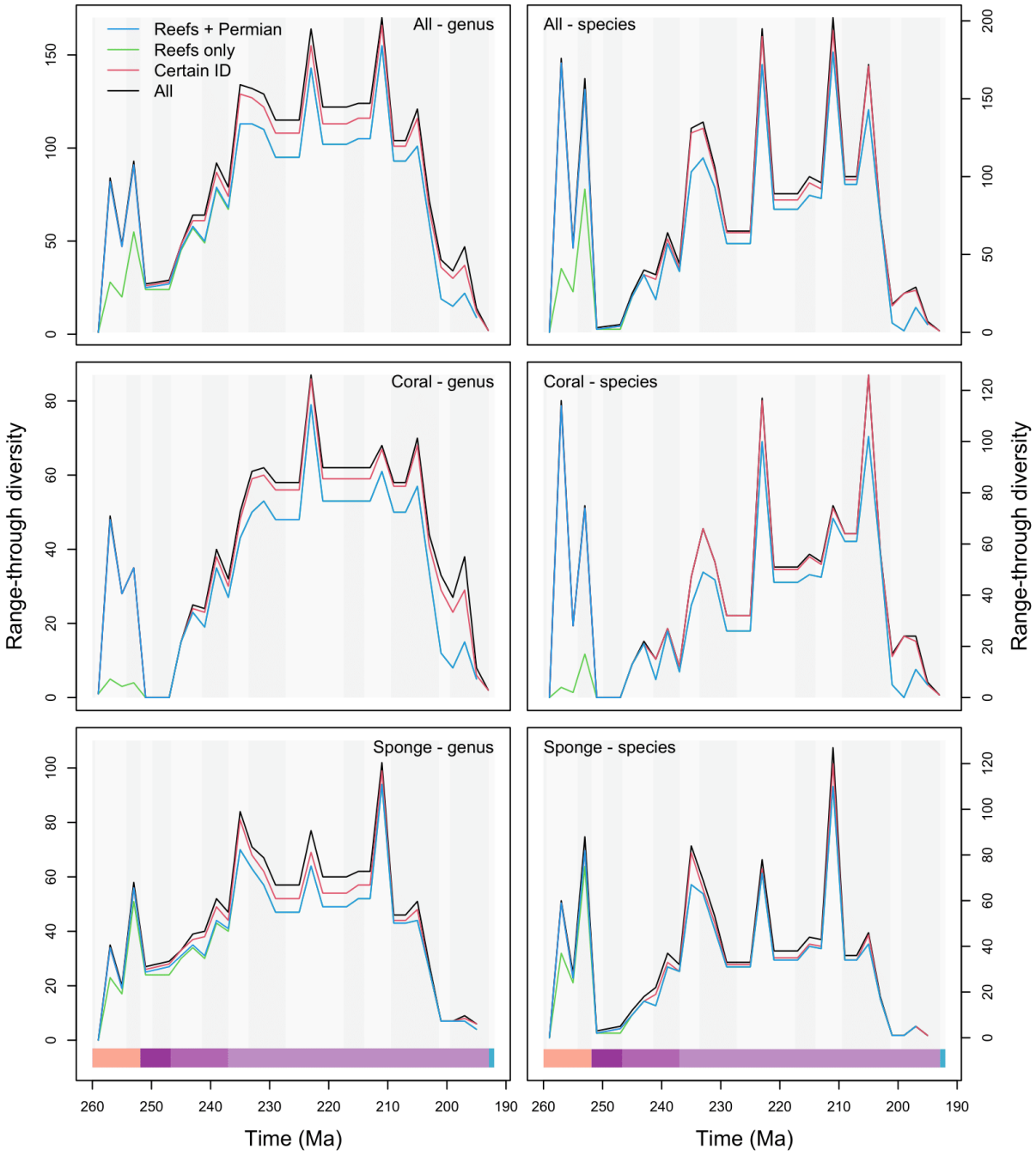


Fig. 4.4. Range-through diversity trajectories from coral and sponge PBDB occurrences. Diversity trajectories are largely unaffected by subsampling of occurrences, aside from during the Late Permian where reef sites are rare following the reef crisis at the end of the Capitanian.

shifts in origination and extinction rates. Origination and extinction times for each lineage were obtained from the individual MCMC logs using the `-ginput` function of PyRate with a 10% burn-in to exclude the non-converged portion of the chain. Individual log files from each age-randomised replicate were then combined with the same burn-in, whilst randomly sampling 1000 post-burn-in generations to reduce the size of the combined log. Origination and

extinction rates through time were then obtained using the `-plotRJ` function from the combined PyRate log. Diversity-through-time analyses were run using mcmcDivE for one million generations sampling every 1000 at the genus level only, using the estimated preservation rate from the combined PyRate log and the 50 age-randomised occurrence datasets. At each stage of the workflow, log files were analysed using Tracer (Rambaut et al., 2018) to ensure that convergence had been achieved (ESS > 200 for all model parameters) and to identify suitable burn-in values for log file combination.

4.2.6 Selection of diversification drivers

Potential explanatory variables for a clade's diversification dynamics may be biotic factors like self-limiting intraclade competition or dependence on the diversity of a separate clade of competitor, predator, or prey taxa, or abiotic environmental factors like ocean temperature and sea level. Biotic effects on diversification dynamics were parameterised using the diversity trajectories of the four focal groups, namely: the effect of each group's own diversity trajectory on their diversification dynamics, plus; the effects of sponges on all corals; all corals, z-corals, and az-corals on sponges; az-corals and sponges on z-corals; and z-corals and sponges on az-corals. mcmcDivE accounts for unsampled diversity, but its estimates contain a high degree of uncertainty and are not infallible, given the potential for highly volatile diversity trajectories (see results). Instead, range-through diversity curves were produced by averaging across the 50 sets of lineage origination and extinction times for each analysis previously obtained using its `-ginput` function.

Several geochemical and palaeogeographic variables were considered as abiotic effects: of the former, CO_2 and O_2 availability and nutrient flux from weathering; and of the latter, sea level and continental fragmentation. Estimates of atmospheric CO_2 were taken from the compilation of Foster et al. (2017). Modification of atmospheric CO_2 has been repeatedly inferred to impact marine diversity through both regulation of long-term climate, and as a physiological control on photosymbiotic corals (Kiessling 2009). Atmospheric O_2 will partially determine oxygen concentration in the oceans, a key metabolic control on marine communities and nutrient availability through its redox impact and has previously been inferred to affect patterns of reef diversity (Weidlich et al. 2003). Many proxy- and model-based estimates of Phanerozoic atmospheric O_2 are available and recently received critical review by Mills et al. (2023), whose consensus curve is used here. Other biogeochemical processes will affect the oxygenation state of the oceans, so a sulphur isotope curve (Present et al. 2022) was used as a

proxy for ocean redox conditions. Phosphate is a key limiting nutrient in marine ecosystems and so may affect the potential for such ecosystems to support diversity. This condition was approximated using the terrestrial phosphate runoff estimates of Sharoni and Halevy (2023), calculating the weighted mean from their simulation results.

Continental fragmentation may affect reef builder diversity not just through controlling the extent of shallow marine shelf area, but also by controlling the extent of upwelling zones favouring reef development. Zaffos et al. (2017) provide a continental fragmentation index that has since been used by multiple palaeodiversity analyses. This curve, however, utilises the plate tectonic model of Seton et al. (2012). The continental fragmentation index was therefore recalculated in R (v4.0.4) for the PALEOMAP plate tectonic model. PALEOMAP plate tectonic rotations were obtained at million-year intervals through the Phanerozoic using the GPlates API, discarding any plates which were not present in all time slices and cleaning the remaining plate polygons using functions from the *sf* R package (Pebesma 2018; Pebesma and Bivand 2023). For each time slice, the perimeters of all plates were summed, overlapping plates merged, then the perimeters of the merged plates summed. The circumference of a circle with area equal to the total area of the unmerged plates was subtracted from both perimeters, then the adjusted perimeter of the merged plates divided by the adjusted perimeter of the unmerged plates. This procedure yields an index where 1 indicates total continental fragmentation (no plate overlap) and 0 a state where all continental blocks are contiguously arranged such that their ocean-facing perimeter is minimised (i.e., a perfectly circular supercontinent; Zaffos et al., 2017). The index could not be calculated for the Getech palaeogeography as the plate tectonic models underlying the DEMs were not available for study.

Finally, sea level directly affects shelf area through continental flooding, influencing the extent of shallow marine habitats and so their diversity through the species-area effect. Various estimates for Phanerozoic sea level are available, but these typically fail to account for one or more of regional tectonic effects, geoid variation, or glacial rebound affecting isotopic or stratigraphic signals of eustacy. Instead, I utilise the sea level curve of Van der Meer et al. (2022). This curve utilises strontium isotopes as a reliable proxy for the rate of seafloor production through the Phanerozoic, adjusting for strontium flux resulting from terrestrial weathering. In turn, the rate of sea floor production, combined with empirically constrained assumptions regarding the relationship between seafloor age and depth, the average distribution of seafloor crust age, and the global hypsometry of the continental margins, enables the changing volume of the ocean basins to be estimated throughout the Phanerozoic. Changing

ocean basin volume, after adjusting for variable loss of water during glaciated intervals, necessitates displacement of water onto the continental shelf and so yields a measure of long-term sea-level change. This approach accounts for all major influences on sea level and provides an estimate of global sea level that is free from the effects of regional signals of eustacy (Van der Meer et al., 2022).

4.2.7 Multivariate birth-death modelling

To assess the drivers of coral and sponge diversification, I used a multivariate birth death (MBD) model (Lehtonen et al., 2016), where the rates of a birth-death process are modelled as linear or exponential functions of time-varying predictors, using user-supplied explanatory variables and times of origination and extinction for a set of lineages. The drivers of diversity may themselves differ through time, particularly during mass extinction intervals (Foster et al., 2023). As my study interval includes two such events, analyses were restricted to the interval covering the Triassic succession of reef builders (246 to 202 Ma) to examine their long-term diversification dynamics under 'normal' conditions, although my selected drivers may also have conceivably varied during other biotic upheavals within this interval like the CPE. As the raw variables were obtained at different timesteps, each was interpolated at 0.1 Ma intervals from 246 to 202 Ma, then normalised to the range [0,1] to remove any effects of scale, following previous recommendations (Lehtonen et al., 2017). The sets of origination and

Table 4.3. AIC and Bayes Factor evaluation of exponential versus linear MBD models

Model	AIC	dAIC	Marginal likelihood	Bayes Factor
z-coral - exponential	1144.61	16.10	-556.48	10.10
z-coral - linear	1160.71		-561.51	
az-coral - exponential	346.09	10.49	-168.28	6.26
az-coral - linear	356.58		-171.41	
coral - exponential	1497.77	14.56	-732.39	13.17
coral - linear	1512.33		-738.98	
sponge - exponential	1813.09	36.01	-883.21	29.26
sponge - linear	1849.10		-897.84	

Marginal likelihood estimated using log-likelihood (lnL) harmonic mean. BF = 2 · (lnL(exp) – lnL(lin))

extinction times obtained previously with `-ginput` were serially tested for linear then exponential correlations with the biotic and abiotic variables outlined above using PyRateMBD (Lehtonen et al., 2016). Each analysis was run for 10 million generations, sampling every 1000. Log files were analysed in Tracer to confirm convergence and identify a suitable burn-in value, then combined across all analyses with a 10% burn-in, and correlation parameters summarised with the `-plot` function of PyRateMBD. The relative fits of the linear versus exponential models for each group were compared using Bayes factors, calculated as the difference in the harmonic mean of their log-likelihoods (Newton and Raftery 1994; Lehtonen et al., 2017). As the harmonic mean estimator has been criticised due to the possibility of infinite variance (Neal 1994; McEwen et al., 2022), model fit was additionally distinguished using AIC for MCMC samples (Raftery et al., 2007). Both tests were implemented in R (v4.0.4) and in all cases the exponential model was identified as better-fitting (dAIC > 2, BF > 6; Table 4.3).

4.3 Results

4.3.1 Habitable area through geological time

Evaluation on training data withheld during model calibration demonstrated that each niche model performed well in capturing the present-day distributions of their taxa (Table 4.2). In all time bins with fossil occurrence data available for niche model evaluation using the Wilcoxon signed rank test, predictive performance across all hindcasts was significantly better than random (Tables S19–S26). Trends in habitable area for each of the examined groups did not change substantially depending on whether the L95P or mSSS threshold for binarization was chosen (Fig. 4.5); results from the latter are presented for the remainder of this chapter. Trends in habitable area are also remarkably similar across all four groups between niche models constructed with climate simulation data from HadCM3L versus CLIMBER-3α (Fig. 4.6), demonstrating a degree of insensitivity to variation in climate model choice. Conversely, there are clear differences in trends in habitable area between clades and between niche models constructed from climate simulations with different palaeogeographic boundary conditions (Fig. 4.6). Sponges consistently show the greatest extent of habitable area through time in each set of climate model hindcasts, frequently occupying almost the total extent of shallow marine shelf area. Broadly speaking, az-corals then show the next greatest extent of habitable area

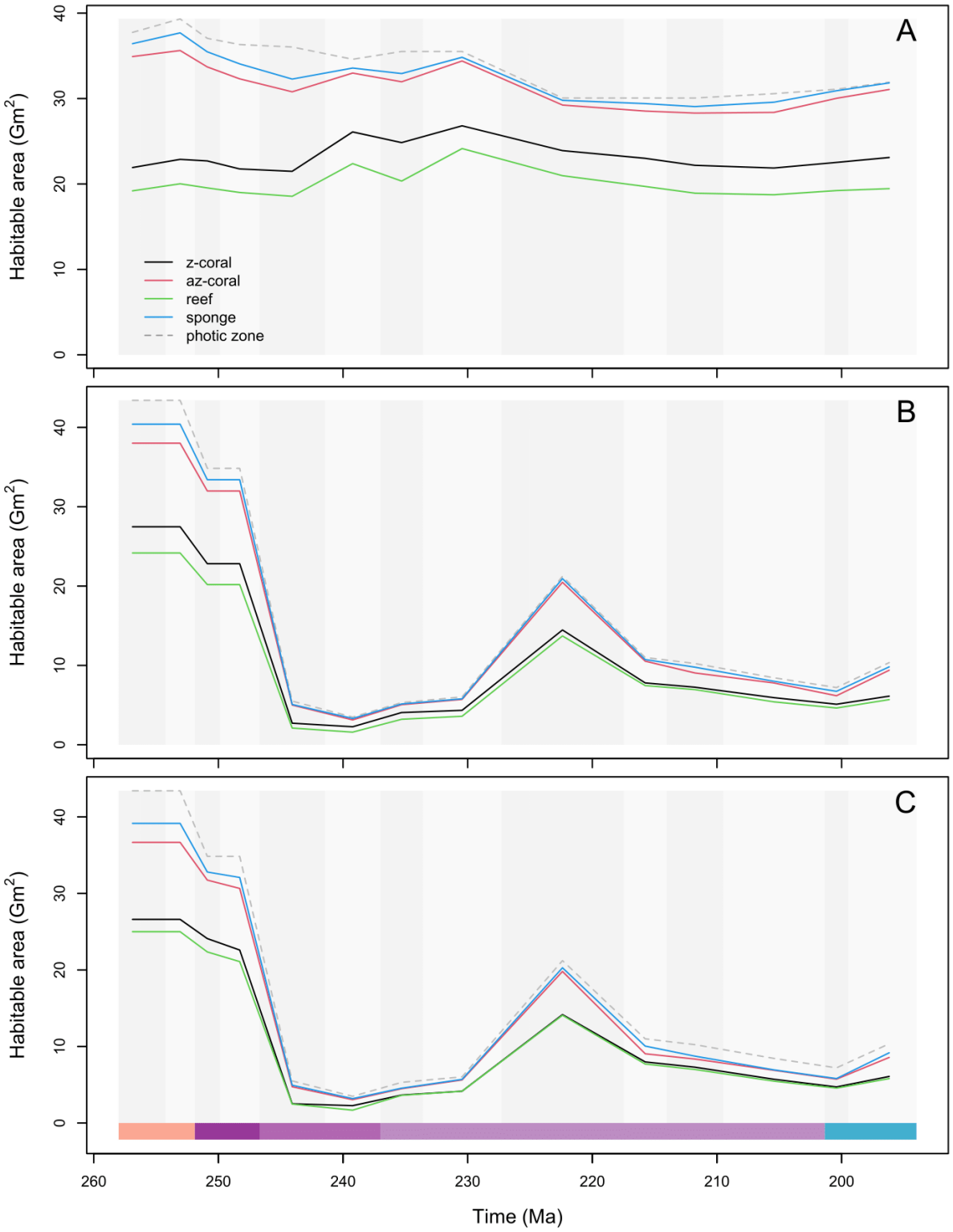


Fig. 4.5. Trends in habitable area between clades, climate models and palaeogeographies. Habitable area was measured after binarization using mSSS. For each map, the dotted grey line represents the total area of the photic zone, measured at the -300m isobath. **(A).** HadCM3L + Getech. **(B).** HadCM3L + PALEOMAP. **(C).** CLIMBER-3α + PALEOMAP.

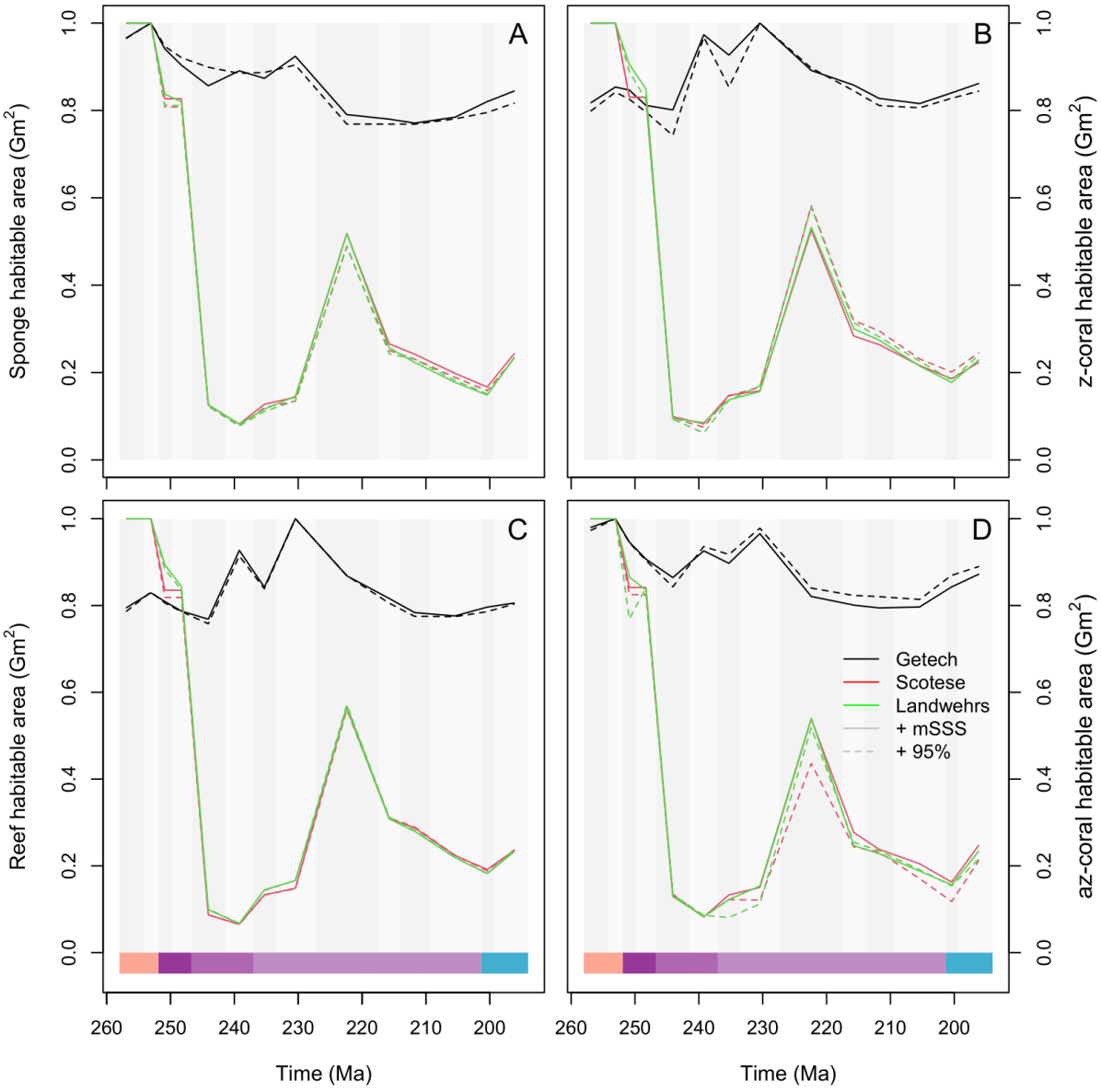


Fig. 4.6. The effect of binarization threshold on habitable area through time. Each time series was normalised with respect to its highest value so that the relative trends between habitable area measured from hindcasts binarized using mSSS and L95P could be more easily compared. **(A).** Sponge habitable area. **(B).** z-coral habitable area. **(C).** Reef habitable area. **(D).** az-coral habitable area.

through time, although this varies somewhat between hindcast sets and time bins, while z-corals and coral reefs generally show the lowest extents.

In the Getech hindcasts, sponge and az-coral habitable area changes only marginally though the Permian, before entering a phase of progressive decline through the Early Triassic and Anisian; z-corals and coral reefs show very little change by comparison (Fig. 4.6A). All four groups then show comparable patterns through the remainder of the study interval.

Habitable area increases in the Ladinian (although this increase is greater for z-corals and reefs), decreases slightly in the early Carnian, and reaches a zenith in the late Carnian. These stepwise increases in habitable area through the Ladinian and Carnian take place in the face of only minor changes in shallow marine shelf area. Habitable area then declines smoothly through the remainder of the Late Triassic before recovering slightly through the Hettangian and Sinemurian, more closely tracking changes in shallow marine shelf area. Conversely, in both sets of PALEOMAP hindcasts (Fig 4.6B, 4.6C), trends in habitable area for all four groups are strongly concordant. Habitable area declines from the Late Permian to Ladinian, increases only marginally through the Carnian, then spikes sharply to an early Norian zenith, before stepwise decline through the remainder of the Late Triassic, and slight recovery during the Early Jurassic. While the geological record highlights the reef crisis taking place at the ETME, the niche models for the Rhaetian and the Hettangian reflect long term climate states, rather than the upheaval during the mass extinction itself where a crash in habitable area is expected.

Not only do trends in absolute extent of habitable area differ between palaeogeographic reconstructions, distribution of habitable area varies also. In the PALEOMAP hindcasts, there is a shift in peak area distribution from the southern to northern tropics through the study interval across all clades, (Fig. 4.7). Sponges and az-corals maintained a broader latitudinal distribution to either side of this peak from the Wuchiapingian to the Early Triassic compared to z-corals and reefs. Habitable area then sharply narrows to mid-high latitude peaks between 30-50ºN in all clades, before a shift in all peaks to the equator during the Carnian. A broad latitudinal distribution is re-established during the early Norian, followed by a second round of narrowing through the early and late Norian. Wider latitudinal distributions are then re-established from the Rhaetian onwards, again with sponges and az-corals showing broader ranges compared to z-corals and coral reefs. Conversely, the Getech hindcasts do not support such a progressive northward shift in habitable area (Fig. 4.8). Habitable area for all four groups is more widely distributed through the study interval and maintains similar bounds within each group, with sponges and az-corals showing broader distributions overall. z-corals and reefs show clear peaks in their distributions through the Late Permian to Early Triassic, centred in the southern tropics. Rather than gradually migrating northwards, however, these peaks then evenly distribute across their group latitudinal ranges in the Anisian before re-establishing themselves in the northern tropics from the Ladinian onwards, accompanied by broadening latitudinal extent through the Ladinian to early Norian. By contrast, sponges and az-corals display distinctly bimodal distributions from the Late Permian to the Early Triassic, with peaks

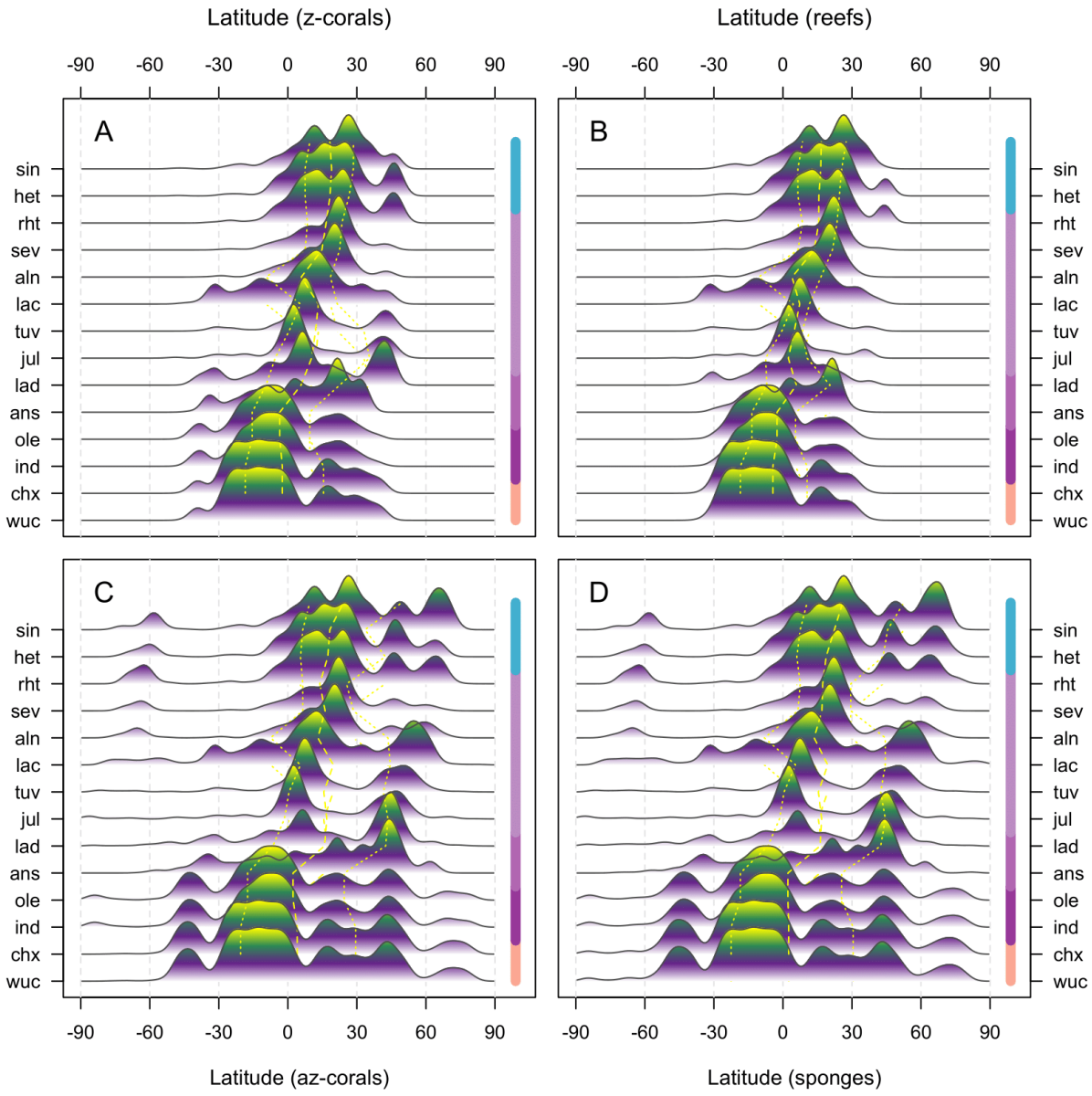


Fig. 4.7. Latitudinal distribution of habitable area using PALEOMAP palaeogeographies. Habitable area was measured from mSSS-binarised hindcasts using the HadCM3L + PALEOMAP climate simulations, then summarised as a kernel density estimate across latitude. y-axis labels refer to the stages of the Late Permian (Wuchiapingian – wuc) to Early Jurassic (Sinemurian – sin). **(A).** z-corals. **(B).** Coral reefs. **(C).** az-corals. **(D).** Sponges.

centred on or even beyond the limits of the tropics. These bimodal peaks then evenly distribute during the Anisian, then reform as singular peaks in the northern tropics through the Ladinian to early Norian, matching the trend displayed by z-corals and reefs. Bimodality in latitudinal extent is re-established for the remainder of the Triassic, however, before shifting to encompass a broader distribution of northerly latitudes during the Hettangian and Sinemurian.

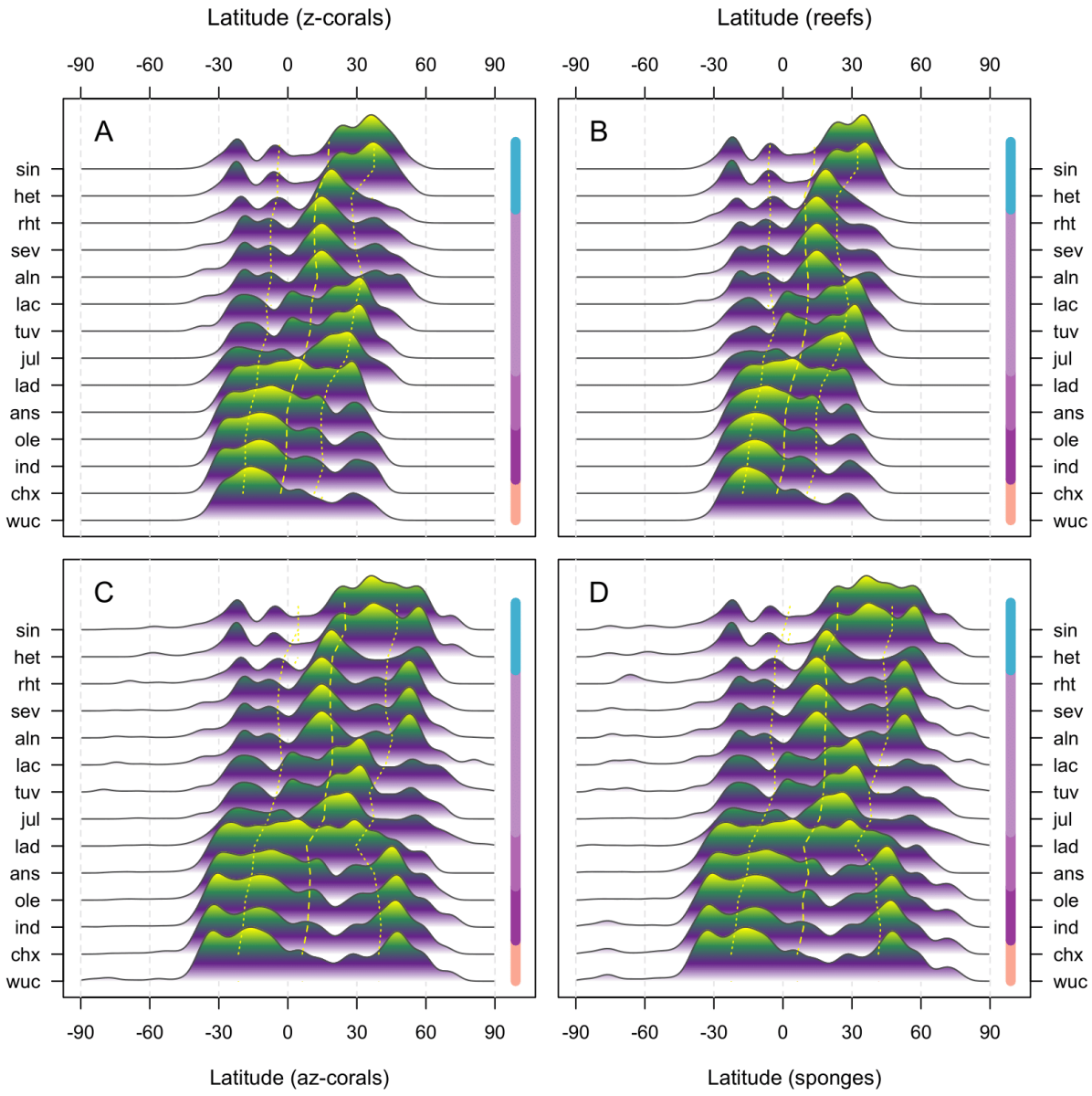


Fig. 4.8. Latitudinal distribution of habitable area using Getech palaeogeographies. Habitable area was measured from mSSS-binarised hindcasts using the HadCM3L + Getech climate simulations, then summarised as a kernel density estimate across latitude. y-axis labels refer to the stages of the Late Permian (Wuchiapingian – wuc) to Early Jurassic (Sinemurian – sin). **(A).** z-corals. **(B).** Coral reefs. **(C).** az-corals. **(D).** Sponges.

4.3.2 Reef builder diversification dynamics

Genus level sponge origination rate spiked briefly in the Changhsingian (~254 Ma) before crashing to well below their extinction rate at the EPME (Fig. 4.9A). Genus level extinction rate remained largely constant, while the effects of the mass extinction manifest much more clearly at the species level (Fig. 4.10A). Net positive diversification was restored in the Induan

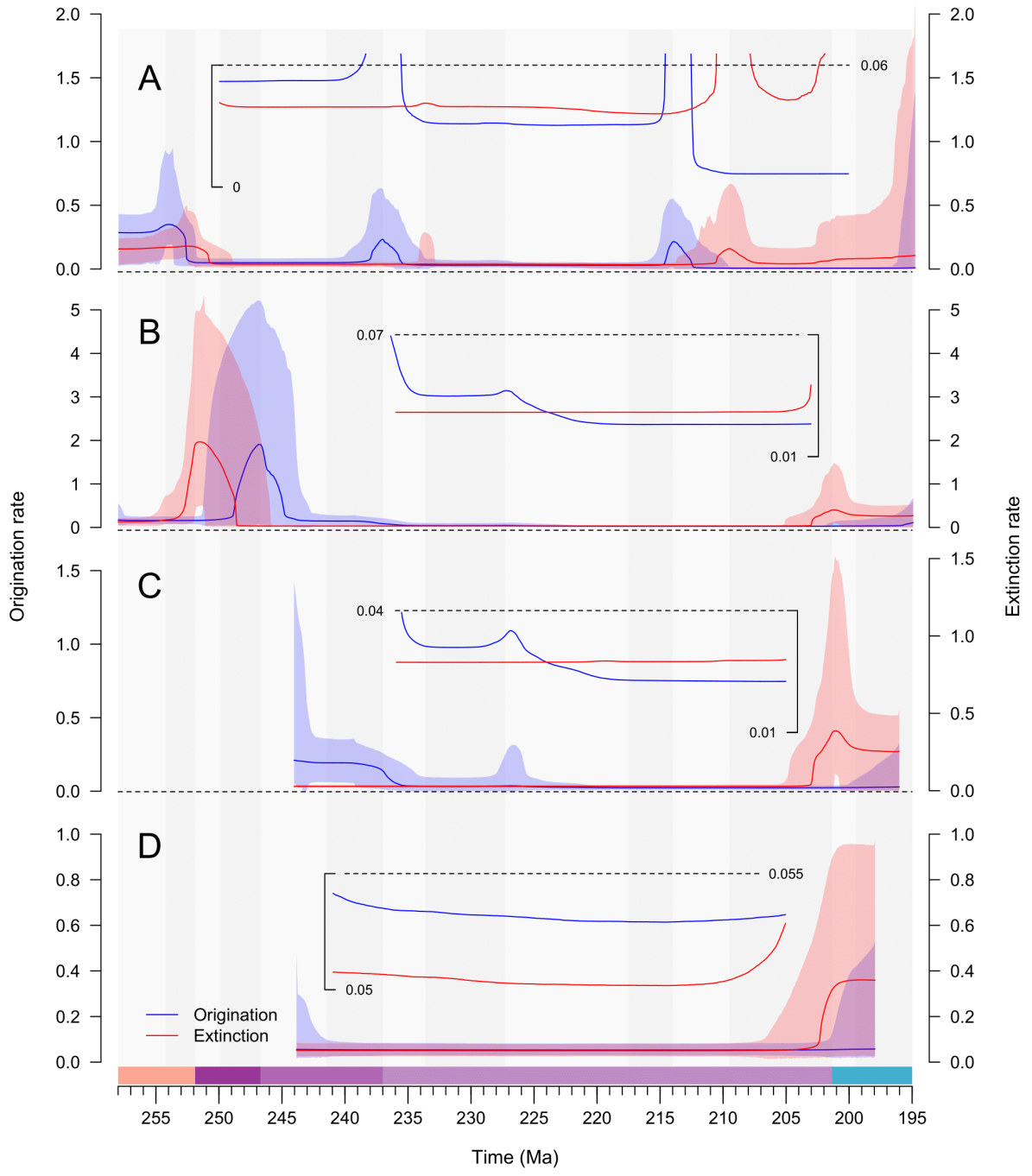
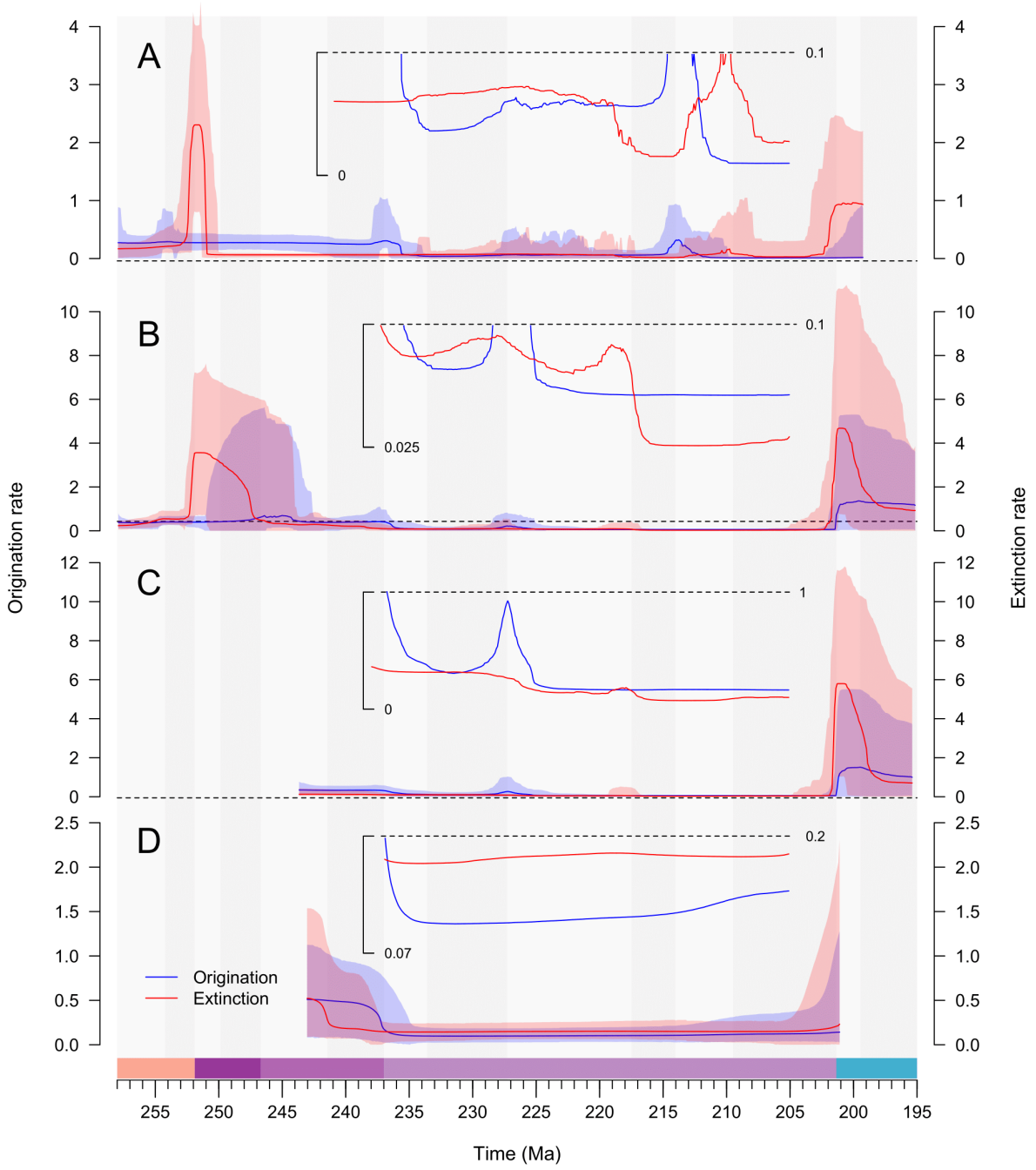


Fig. 4.10. Species-level origination and extinction rates for corals and sponges. The inset graphs show their median rates in intervals where background rates appear flat compared to rate changes during mass extinctions at the EPME and ETME. **A).** Sponges. **(B).** Corals. **(C).** z-corals. **(D).** az-corals.

(~250 Ma), with genus and species level extinction and origination rates remaining near constant through the end of the Middle Triassic (~237 Ma). Sponges experienced a second pulse of diversification at the beginning of the Late Triassic, followed by a decline in their origination rate to marginally lower than their extinction rate at the genus level through the



early Carnian to middle Norian (235 – 214 Ma). Their marginally negative net diversification rate during this interval is reflected at the species level (Fig. 4.10A) but with a more dynamic origination rate in the face of a stable extinction rate. Both extinction and origination rates spike in the late Norian (~213 Ma), producing an initially net positive diversification rate, but extinction remained high after origination subsided to its lowest rate through the remainder of the Triassic. Genus level extinction rate increased only marginally at the genus level across the ETME, although the confidence interval on the rate allows for a potentially catastrophic increase in rate (Fig. 4.9A) which is more congruent with the empirical geological observation

of rapid sponge demise across the mass extinction (Kiessling et al., 2007). Again, the effects of the mass extinction manifest more clearly in the species level extinction rate.

Coral diversification was steady during the Late Permian (259 – 252 Ma) before sharp spikes in genus and species level extinction rates at the EPME. Extinction rate gradually declined through the Early Triassic (251 – 248 Ma) while origination rate steadily increased to a peak at the beginning of the Middle Triassic (~247 Ma). Genus level extinction rate was then near constant for almost the remainder of the Triassic, while origination rate remained high through the Middle Triassic, declined slightly through the Carnian, and spiked again at the beginning of the Norian (~227 Ma). Origination then stabilised for the remainder of the Triassic, producing constant, weakly negative diversification from the early Norian (~225 Ma). At the species level, extinction rate was more dynamic and while origination rate shows the same pattern as at the genus level, positive diversification took place from the Middle Norian onwards (~216 Ma). Extinction rate then spiked well above origination rate across the ETME, although species level origination rate also displays a clear increase during this interval. When corals are subdivided into photosymbiotic zooxanthellate and non-photosymbiotic azooxanthellate groups, the former clade almost perfectly match the total coral trends described above, with the notable exception of a weakly positive speciation rate through the entirety of the Norian. Meanwhile azooxanthellate corals showed a near constant weakly positive diversification at the genus level until the middle Rhaetian (205 Ma), followed by a sharp increase in extinction rate around the onset of the ETME. This trend is inverted at the species level with a weakly negative diversification rate through the Triassic, produced by an initial crash in origination rate at the beginning of the Carnian followed by a gradual increase through the Norian and Rhaetian. Set against the stark extinction rates recovered during times of environmental upheaval, and the elevated origination rates marking the appearance of new clades in the succeeding recovery phases, diversification dynamics in intervening periods of apparent quiescence appear muted. Nonetheless, the seemingly minor fluctuations in background extinction and origination rates over protracted timescales are responsible for the marked changes in standing diversity, measured from the overlap of lineage durations through time (Fig. 4.11). This does not account for unsampled lineages, however, unlike the preservation-corrected diversity curves (mcmcDivE) for corals and sponges. The diversity trajectory for sponges recovered by this method suggests much more rapid diversification through the Ladinian followed by a series of stepwise diversity declines at the beginning, middle and end of the Carnian. Reduced diversity through the remainder of the Triassic was

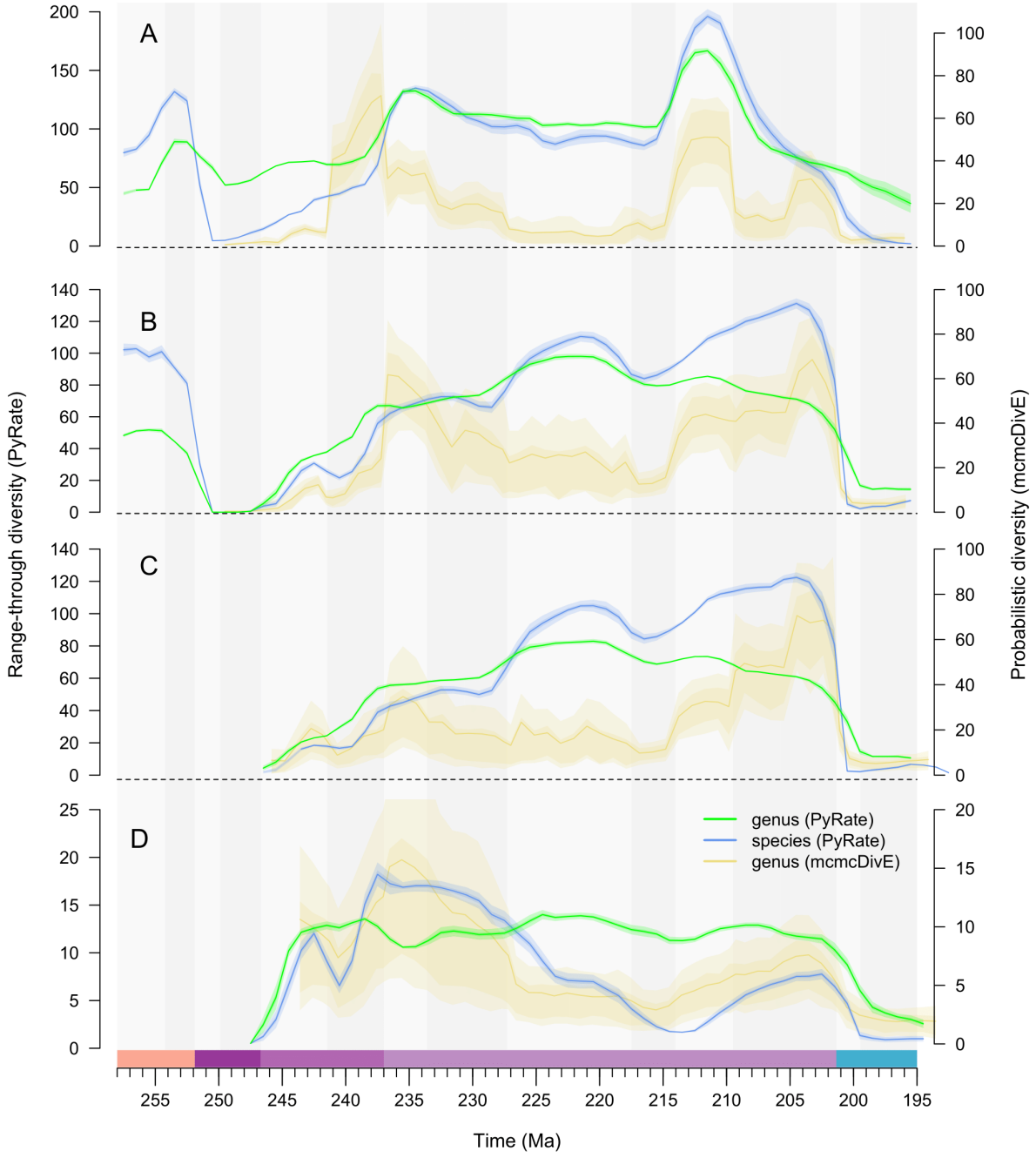


Fig. 4.11. Coral and sponge diversity estimates. Mean genus and species diversity and their 95% confidence intervals, calculated across 50 sets of lineage origination and extinction times for each group. A. Probabilistic diversity estimates from mcmcDivE for corals and sponges. Median probabilistic genus diversity and their 95% highest posterior densities. Range-through mean genus diversity from PyRate is shown as a grey line in each plot. As in previous analyses, the results for corals and z-corals are nearly identical so the former are omitted here. **A).** Sponges. **(B).** Corals. **(C).** z-corals. **(D).** az-corals.

punctuated by sudden diversity spikes in the late Norian and late Rhaetian, prior to a final crash across the ETME and persistently low diversity in the Early Jurassic. Conversely

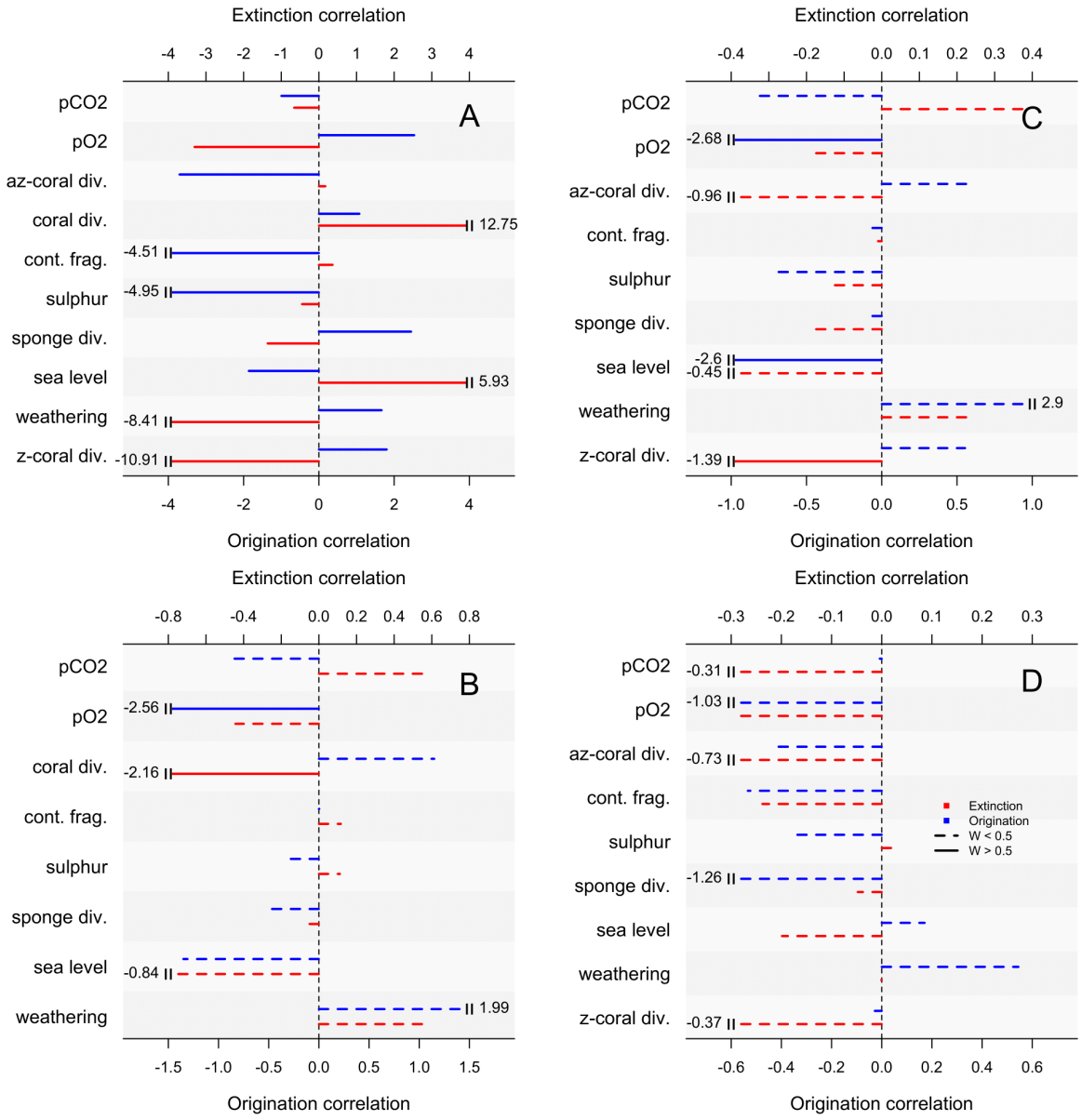


Fig. 4.12. Multivariate birth-death modelling of sponge and coral diversification drivers. All results shown are from exponential correlations with origination and extinction rates from 246 to 202 Ma. Solid lines indicate statistical significance (W > 0.5). **(A).** Sponges. **(B).** Corals. **(C).** z-corals. **(D).** az-corals.

zooxanthellate coral diversity fluctuated through the Middle Triassic to a modest peak around 235 Ma, followed by gradual decline through the remainder of the Carnian. Their diversity fluctuated again through the early Norian, increased stepwise from the late Norian to late Rhaetian peak, then terminated abruptly in a crash at the ETME. Azooxanthellate coral diversity was substantially lower through the study interval by comparison and only matched the trends of zooxanthellate corals up to the end of the Carnian. Their diversity dipped again at the beginning of the Norian, followed by a muted rise through to the late Rhaetian, and a decline across the ETME.

Exponential MBD models were identified as better-fitting by both Bayes factors and AIC for all groups (Table 4.3). When corals are examined as a whole, their origination rates show a negative correlation with atmospheric oxygen concentration and their extinction rates show a negative correlation with their own diversity (Fig. 4.13). When subdivided into azooxanthellate and zooxanthellate subgroups, the former displays no significant correlations while the latter show the same correlations with atmospheric oxygen and self-diversity, in addition to a negative correlation between sea level and origination rate. Contrastingly, all drivers show significant correlations for either sponge origination or extinction rates. Sponge origination rate shows negative correlations with atmospheric CO_2, az-coral diversity, continental fragmentation, sulphur isotope composition and sea level, and positive correlations with atmospheric oxygen concentration, weathering, the diversities of corals and z-corals, and their own diversity. Sponge extinction rate shows positive correlations with coral diversity, sea level and continental fragmentation, and negative correlations for all other drivers.

4.4 Discussion

4.4.1 Biotic controls on the emergence of scleractinians

Following the cryptic origin and divergences of their higher-level clades through the Palaeozoic (Stolarski et al., 2011; Seiblitz et al., 2020), there is then a substantial gap in the fossil record to the sudden appearance of ecologically and phylogenetically diverse scleractinians in the Anisian, aside from a couple of disputed Palaeozoic scleractiniomorph corals (Scrutton and Clarkson 1991; Scrutton 1993, 1996; Scrutton et al., 1997; Ezaki 1997, 2000). Scleractinians were ancestrally solitary, non-mineralising, azooxanthellate and deep marine with multiple lineages convergently acquiring photosymbionts and the associated z-coral trait complex of coloniality and enhanced biomineralisation and energy generation (Allemand et al., 2010; Davy et al., 2012; Campoy et al., 2020), resulting in an approximately even balance of zooxanthellate and azooxanthellate taxa across their extant phylogeny (Kiessling and Kocsis 2015). z-corals, however, displayed elevated origination rates compared to az-corals in the Anisian (Fig. 4.9C–D, 4.10C–D), suggesting that their ecological traits offered major competitive advantages during their initial entry into shallow marine environments. The benefits of photosymbionts to emerging scleractinians is congruent with the warm, oligotrophic conditions presented by West Tethys where they first appeared (Fig. 4.13; Frankowiak et al., 2021), which map onto the ecological preferences of z-corals in contrast to

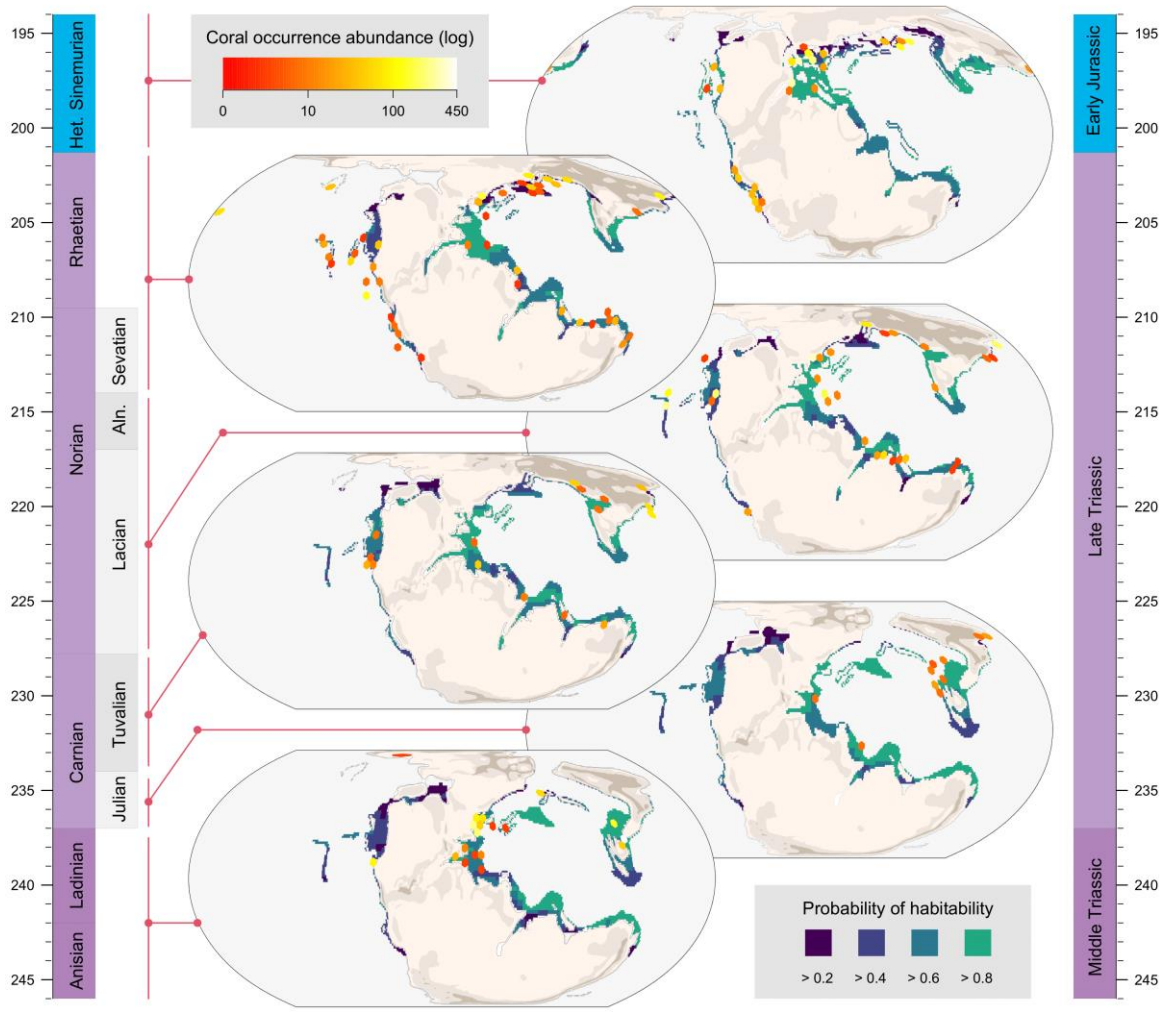


Fig. 4.13. Geographic distribution of corals from the Middle Triassic to Early Jurassic. Counts of coral abundances were calculated after spatially binning their occurrences using a hexagonal grid. These counts are superimposed on the distribution of habitable area through time under Getech palaeogeographic boundary conditions.

the cool, nutrient-rich deepwater conditions preferred by az-corals (Kiessling and Kocsis 2015). These conditions may have favoured pre-adapted lineages of scleractinians or promoted the simultaneous acquisition of photosymbionts by separate, ancestrally azooxanthellate lineages. Either scenario is compatible with the elevated origination rates recovered here for z-corals, but the sheer lack of scleractinian fossils in well-sampled, Palaeozoic-aged facies from within the photic zone favour the latter, suggesting that shallow marine, photosymbiotic lineages were genuinely absent in these environments until the Triassic and that their emergence was instead an adaptive radiation into the shallows.

The biotic benefits of photosymbionts appear to have been crucial to the ecological success of scleractinians in the Triassic generally, but the importance of biotic factors in the rise of Dachstein reefs is mixed. Despite the competitive relationships displayed by corals and sponges in modern-day reef systems under times of ecological stress, no diversity dependent effects from sponges on coral diversification rates were detected, nor when these rates were divided into their z and az components (Fig. 4.12B–D). As such, the gap imposed by the Wetterstein reef phase between the appearance of scleractinians in the Anisian and the first Dachstein reefs in the Carnian was likely not the result of ecological suppression by sponges. Coral extinction rates declined as they became more diverse (Fig. 4.12B) which could suggest that their ecological incumbency during the Dachstein reef phase promoted resilience to extinction, but equally this could be an artefact of the longer taxonomic durations displayed by Late Triassic corals due to the longer stage durations within this epoch. Conversely, sponge origination and extinction rates are positively correlated with coral diversity as whole, indicating that sponge turnover intensified as corals increased in ecological prevalence (Fig. 4.12A). Their origination and extinction rates increased and decreased respectively with rising z-coral diversity, however, showing that both clades diversified in tandem rather than antagonistically. Together these relationships do not support any strong competitive effects between reef builder clades during the Wetterstein-Dachstein transition.

4.4.2 Abiotic controls on the Wetterstein-Dachstein transition

Fluctuations in coral and sponge diversification rates during the turnover between Wetterstein and Dachstein reef phases were small compared to the marked shifts during mass extinctions (Fig. 4.9, 4.10). Nonetheless, multivariate birth-death modelling identified significant correlations with several abiotic drivers. The reef-building z-corals showed a declining origination rate in response to rising sea level (Fig. 4.12C), which is congruent with the empirical observation that the Dachstein reef phase took place during an interval of long-term low sea level. This result is ecologically unexpected, given that rising sea level should promote diversification through increasing the extent of favourable shallow marine habitats by flooding low lying parts of the land surface, but is explicable given that eustatic sea level change alone may not fully capture empirical trends in shelf area, and that other palaeogeographic effects may have controlled the timing of the Late Triassic reef bloom (see 4.4.3). z-coral origination rate showed a negative correlation with atmospheric oxygen concentration (Fig. 4.12C), but the biological mechanism for declining origination rate in response to elevated oxygen concentration resulting from dissolution from the atmosphere into the ocean is unclear given

that high oxygen availability should be favourable to reef communities. As no correlation was detected with sulphur isotope ratios as an alternative proxy for oceanic redox conditions, this suggests that the correlation displayed with oxygen is instead spurious, while the lack of a correlation between coral diversification rates and atmospheric CO_2 concentration fails to support the hypothesis of Kiessling (2009) that declining atmospheric CO_2 during the Carnian and Norian promoted scleractinian diversification and reef construction.

Sponges appear to show more dynamic responses to abiotic drivers than corals. Rising sea levels were associated with increased extinction rate and decreased origination rate (Fig. 4.12A), a complementary set of effects that indicate sea level exerted strong control over sponge diversification, although this is again in conflict with the expected scenario of rising sea level resulting in increasing taxonomic diversity. Similarly, as with corals, their responses to redox conditions and oxygen availability are conflicting and lack a clear biological mechanism. Conversely, the effects of enhanced continental weathering increase their origination rate and decrease their extinction rate (Fig. 4.12A) can be readily interpreted as elevated nutrient input from surface runoff promoting their diversification. Intriguingly, the opposite set of effects is seen for continental fragmentation, despite fragmentation increasing the length of the coastline for runoff. While these factors may be expected to positively covary, weathering and runoff are also affected by long term climate change, as well as the distribution of continental landmasses in relation to climate zones (see 4.3.3), potentially explaining the differential effects of continental fragmentation and runoff on sponge diversity dynamics.

4.4.3 The role of palaeogeography in the rise of coral reefs

The Wilson cycle has shaped shallow marine habitat extent through the Phanerozoic by reduction of coastline length and shelf area during supercontinent phases, with increasing continental fragmentation and shelf area from the Middle Triassic to the present acting as a major control on the spatial distribution of scleractinian coral reefs (Zaffos et al., 2017; Jones et al., 2022). The ecological niche modelling results presented here, however, show that uncertainties in palaeogeographic boundary conditions (Fig. 4.5) present major challenges for investigating the relationships between the Earth-system and the biosphere. Short term eustatic variation and its uncertainties affect the reconstructed extent of shallow marine shelf area even during supercontinent phases (Kocsis and Scotese 2021; Van de Meer et al., 2022). Further, the PALEOMAP and Getech palaeogeographies are reconstructed not just under different eustatic conditions but also at different timesteps, resulting in cases where boundary conditions

for one stage were re-used for alternative stages or substages during climate modelling. In the case of the former, this re-use encompasses the latest Permian to Early Triassic, where relatively minor changes took place, but in the case of the latter, the use of stage-level reconstructions for substages of the Carnian and the Norian could mask the impacts of more prominent changes in palaeogeography through this 30-million-year duration when scleractinians and their reefs rose to prominence.

Despite their broader timestep, the Getech reconstructions are built from a wealth of proprietary and public data and display realistic topography and bathymetry, including high mountains and mid-ocean ridges, and so may be more reliable than the PALEOMAP reconstructions (A. Farnsworth, pers. comm.). The PALEOMAP DEMs additionally suffer from two potentially related shortcomings. Firstly, they display a vertical resolution of 40 metres, meaning that sudden eustatic changes between DEMs that exceed this threshold may produce extreme fluctuations in continental or marine shelf area. Secondly, the Triassic is subject to some of the greatest degrees of mismatch between the locations of marine fossil occurrences and the extent of PALEOMAP coastlines (Kocsis and Scotese 2021). This effect may be responsible for the extreme crashes in habitable area during the Middle Triassic and through the Norian (Fig. 4.6B, 4.6C), reducing confidence in niche model estimates predicted on these palaeogeographic reconstructions. For these reasons, Triassic reef evolution will be evaluated here under Getech boundary conditions.

There is a degree of similarity between trends in habitable area across reefs, sponges, z-corals and az-corals, and the absolute extent of shallow marine shelf area (Fig. 4.5A), indicating that fundamental palaeogeographic limits on habitat space played a role in shaping the distributions of these groups, but the differences between shelf area and habitable area are also informative. Under highstand conditions shallow shelf area was lowest in the Norian, but habitable area across all groups was lowest in the Anisian (Fig. 4.5A). At first glance this is at odds with the Anisian emergence of Wetterstein-type reefs and corals but can be rationalised as a part of the general recovery of marine ecosystems in the Middle Triassic following the ecological instability of the Early Triassic. In essence, the low extent of habitable area in the Anisian did not suppress sponge and coral diversification as these clades were emerging from an interval of even more stringent environmental constraints that are clearly evidenced by the empirical geological record, but not captured by ecological niche modelling or palaeogeographic reconstruction alone. Sponge habitable area was highest in the Changhsingian but reached a second peak in the late Carnian A (Fig. 4.5A). This peak is also

present for z-corals and reefs with the crucial difference that this peak represented the greatest extent of their habitable area during the study interval. Consequently, the climatic conditions of the late Carnian in combination with the spatial distribution of shallow marine shelf area during this interval appear to have formed a window of maximum opportunity for reef construction that facilitated the Wetterstein-Dachstein transition.

4.4.4 Methodological limitations

Inference of habitable area using Earth-system simulations and ecological niche modelling are subject to caveats which may limit the scope of the results presented here. As well as the issues surrounding the combined effects of palaeogeographic and eustatic uncertainty, the spatiotemporal resolution of DEM sets will impact the sensitivity of analyses which rely on Earth-system models, although this problem is being gradually addressed by the availability of increasingly finely resolved reconstructions (e.g., the landscape evolution models of Salles et al (2023). Climate simulations are run at substantially lower resolutions than their underlying DEMs due to computational limitations and upscaling simulation outputs to original geographic resolution is only a partial solution to this problem. Atmospheric CO_2 concentrations used for deep time climate modelling come with broad confidence intervals (Foster et al., 2017). One solution to this problem is to conduct sensitivity tests for CO_2 concentration whilst using fixed boundary conditions. This approach was performed explicitly in the CLIMBER-3α simulations of Landwehrs et al. (2021) and while niche models were only hindcast to a single threshold for each interval to ensure comparability across all three sets of climate simulations, the Induan and Olenekian simulations share the same underlying boundary conditions (the 250 Ma PALEOMAP DEM) at different CO_2 thresholds. Multiple CO_2 thresholds are also available in the Getech HadCM3L simulations, namely the Carnian and Norian substages which re-use stage-level boundary conditions with interval-appropriate CO_2 concentrations. Latitudinal distributions are virtually unchanged, but absolute extent of habitable area is affected in all groups through these intervals, indicating that that downstream niche model estimates are sensitive to the choice of CO_2 threshold in the underlying climate simulation (Fig. 4.8). A further aspect of changing CO_2 between intervals is that each climate simulation is run until near or total model equilibrium is achieved, yielding a view of stable, long-term climate but not of sudden climate change across events like the EPME, CPE or ETME, or their effects on habitable area.

Besides the underlying limitations outlined above, ecological niche models in turn may not account for other controls on true species distributions beyond their thermophysiological niche (Soberón and Peterson 2005), for example pH, benthic oxygenation, or aragonite saturation state. Aragonite saturation was of secondary importance in predicting coral niche space in previous studies (Couce et al., 2012; Jones et al., 2019, 2022), but these focused on the last glacial maximum to the present, in contrast to the high CO_2 conditions of the Permo-Triassic biosphere. Oxygen limitation and ocean acidification, particularly across the EPME and ETME, where coral taxon extinctions exceed 95% (Stanley and Shepherd 2018), are expected to be strong controls on coral distribution – effects which the niche models used here do not capture. These caveats are important to consider but can be practically difficult to diagnose or overcome, particularly in palaeontological applications of ecological niche modelling where environmental variables and empirical fossil records are often limited in their availability and spatiotemporal resolution (Varela et al., 2011). Calibration of a niche model on extant taxa followed by hindcasting to past geological intervals also makes the potentially flawed assumption that the ecological niche of the examined clade does not change through time. Nonetheless, hindcast predictive performance was still significantly better than random in all sponge and coral niche models (Tables S19–S26), suggesting that their present-day niches remain somewhat informative of their deep time distributions.

Inference of diversification rates from fossil data poses challenges generally (see 1.2.3), but the results for reef builder diversification dynamics present additional issues. PyRate and mcmcDivE yield contrasting views of coral and sponge diversity through the Triassic, but while the former does not account for unsampled lineages this does not necessarily mean that it is outperformed by the latter. Genera, as broader taxonomic units, are more robust to both sampling bias and variable opinions between systematists (Alroy 1996) but may fail to reflect changes in biodiversity at other taxonomic levels. Genus and species level diversity curves clearly record the termination of all Palaeozoic corals at the EPME, while the curves for sponges demonstrate that some sponge genera survived across the extinction event, maintaining a level of cryptic diversity in the Early Triassic despite the total extinction of all species level taxa (Fig. 4.4, 4.11). As PyRate uses the range-through-time (RTT) diversity concept, it provides a minimum bound on diversity with a fully congruent set of origination and extinction rates, but poor sampling can result in discrepancies between different taxonomic levels, leading to systematic underestimation of species level diversity. Sponge genus durations necessitate a minimum level of cryptic species diversity (at least one species per genus) through

the Early and Middle Triassic, yet the lack of species-level sponge occurrences in this interval means that species level diversity does not meet this requirement. This could arise due to a genuine absence of fossils (geological bias), failure of observation (anthropogenic bias), or a paucity of fossils identifiable at the species level (taphonomic bias), leading to a low preservation rate in PyRate in all three cases.

Low preservation rates may in turn complicate diversity estimation using mcmcDivE due to the discrepancy between sampled-in-bin (SIB) diversity versus RTT diversity (see 1.2.3). Preservation rate from PyRate is used to estimate true diversity from pointwise observations of lineages across an arbitrarily defined number of intervals, but if fossil occurrences do not record the presences of all lineages spanning those intervals (decoupling of SIB and RTT), then mcmcDivE will systematically underestimate diversity. This helps explain the discrepancies seen in coral and sponge diversity through the late Carnian to early Norian, where fossil occurrences are more sporadically sampled (Fig. 4.11). While the simulations in the previous chapter indicated mcmcDivE still provides accurate estimates of relative changes in diversity compared to SQS, it may be the case that the coral and sponge preservation rates vary too substantially through time and between taxa for the method to provide reliable results in this instance. Nonetheless, this does not mean that PyRate's lower bound RTT diversity curve or accompanying origination and extinction rates are any better. The sample of taxa in a PyRate analysis is assumed to comprise a complete phylogeny with an unknown, underlying topology. Estimated birth and death rates will be largely unaffected when incomplete sampling is random, but any preservation rate will violate this randomness as taxa are increasingly less likely to be sampled as their total stratigraphic durations decrease. PyRate was still able to accurately infer birth and death rates in simulated datasets, violating the assumption of randomness, but this was only explicitly considered under an NHPP preservation model (Silvestro et al., 2014), rather than the more complex TPP preservation model. Consequently, it is less clear whether the birth-death estimates remain accurate when the proportion of unsampled taxa is expected to non-randomly vary through time due to bin-wise variation in preservation rate.

4.5 Conclusions

The results presented in this chapter demonstrate that the probabilistic diversification dynamics for Triassic corals and sponges are broadly congruent with the trends in the reefs that they constructed, suggesting that the fossil record does contain a genuine, useful signal of their

emergence during the Triassic. Their diversification dynamics also provide evidence that the influence of the Earth-system on their long-term taxonomic diversities can help explain the transition from Wetterstein-type to Dachstein-type reefs. Nonetheless, there is an outstanding disconnect between gradually changing, long term drivers examined here versus the relatively abrupt switch from sponge-dominated to scleractinian-dominated reefs. For example, the late Carnian peak in reef and z-coral habitable area was still preceded by steady increases through the Ladinian and early Carnian, and scleractinians in the Middle Triassic included taxa bearing the ecological traits key to their latter success (Kolodziej et al., 2018; Stanley and Helmle 2010). This begs the question of why scleractinians did not show at least show a more gradual rise to ecological prominence during this interval, particularly given the lack of a signal for ecological suppression of corals by sponges during the Wetterstein reef phase.

While the reason for the delay cannot be evaluated clearly from the results presented here, a range of geological and geochemical evidence provides a scenario for future work. The tectonic and bathymetric setting of the carbonate platforms supporting the sponge- and microbe-cemented Wetterstein-type reefs were favourable to calcification generally (Flugel 2002; Martindale et al., 2019), but fossil evidence and data from calcification experiments on modern corals indicate that scleractinians are only dominant reef builders when the Mg/Ca ratio of seawater sits in the aragonite domain (Stanley and Hardie 1998, 1999; Ries 2009). Middle Triassic seawater chemistry may have instead favoured the production of high-Mg calcite by hypercalcifying sponges, rather than scleractinian aragonite (Ries 2009). Green algae and microbes with ‘green’ plastids containing chlorophyll *a*, were ecologically dominant in Middle Triassic marine systems (Saito et al., 2016) within the prevailing, Neritan style carbonate factory where production was situated solely in shallow shelf and platform areas (Zeebe and Westbroek 2003; Ridgwell and Zeebe 2005), but redox shifts in the Earth-system are hypothesised to have driven their displacement in the Late Triassic by dinoflagellates and coccolithophores with ‘red’ plastids containing chlorophyll *c*, reflecting their differential macronutrient and transition metallomic requirements (Quigg et al., 2003; Falkowski et al., 2004; Tosti et al., 2014). The dinoflagellate radiation included likely ancestors of the photosymbiont clades found in extant scleractinians (Payne and Van de Schootbrugge 2007; Stanley and Schootbrugge 2018; Mangerud et al., 2019), while the coccolithophores shifted the seawater carbonate saturation state into the aragonite domain for the first time in the Mesozoic and switched the global carbonate factory from the archaic Neritan type to the modern-day Cretan type where production is evenly split between benthic corals and pelagic

nannoplankton (Zeebe and Westbroek 2003; Ridgwell and Zeebe 2005; Preto et al., 2013; Dal Corso et al., 2020). Increasingly well resolved stratigraphic data suggest that these fundamental transitions in seawater chemistry and phytoplankton diversity coincided with the collapse of microbe- and sponge-dominated Wetterstein-type carbonate factories due to enhanced terrestrial runoff and eustatic fluctuation during the CPE, with the rise of coral-dominated Dachstein reefs taking place in its aftermath (Flügel and Stanley 1984, Simms and Ruffell 1989, Flügel 2002; Gattolin et al., 2015; Dal Corso et al., 2015, 2020). Together these lines of evidence permit construction of the hypothesis that the ocean revolution across the CPE mediated the switch from 'green' Cretan oceans favouring high Mg calcitic sponges to 'red' Neritan oceans favouring aragonitic corals and photosymbiosis. If confirmed, this hypothesis would help reconcile the effects of long-term abiotic drivers for the rise of scleractinian reefs with the abrupt nature and timing of the Wetterstein-Dachstein transition.

Finally, this chapter demonstrates that significant challenges remain in reconstructing the spatiotemporal patterns and drivers responsible for the rise of scleractinian reefs. The quality of their fossil record may fundamentally limit the ability of modern macroevolutionary approaches to infer diversification dynamics, while attempts to circumvent the spatial inadequacies of their record are hampered by palaeogeographic uncertainty. Future work may benefit from focusing on traditional statistical modelling methods to identify predictors of changing reef-type abundances (Raja et al., 2023), while improvements in the PALEOMAP DEMs to record high stand sea level will permit proper evaluation of eustatically-driven trends in habitable area (Kocsis and Scotese 2021). This will yield a clearer view of how climate and palaeogeography shaped Triassic scleractinian evolution.

Chapter Five

Characterising Spatial and Climatic Trends of the Archosauromorph Radiation

The diversification of archosauromorph reptiles from the Late Permian to Late Triassic is a classic example of an adaptive radiation. It has received significant attention from phylogenetic, morphological and ecological perspectives, but its spatial signature has been little studied. In this chapter, I reconstruct the geographic origins of early archosauromorphs using a taxonomically updated, time-calibrated super tree. I then couple landscape connectivity analysis and high-resolution reconstructions of deep time climate and topography in a novel workflow (TARDIS; Traverses And Routes of Dispersal In Spacetime) to infer the phylogeographic structure of their radiation across the ecologically anachronous stage presented by Pangaea. I use rates and routes of dispersal of archosauromorphs, in conjunction with ancestral state estimates of their climate tolerances, to examine biogeographic influences on their early evolutionary success. I demonstrate that the deepest archosauromorph divergences took place in Northern Pangaea (present day Europe) but later cladogenetic events were geographically disparate and frequently involved trans-equatorial dispersals. This indicates that extreme climates across the supercontinent were not such robust biogeographic barriers as has been suggested and demonstrates the role that landscape connectivity analysis can play in reconstructing the spatial dynamics of evolutionary radiations beyond the bounds of the fossil record.

Author contributions: This chapter is based on work conducted in collaboration with A. Farnsworth and A. Elsler and supervised by M.J. Benton. A.F. provided the HadCM3L climate models. A.E. provided the supertree of early tetrapods and database of their chronostratigraphic first and last appearances. The author conducted all other analyses. Comments were provided on a draft of this chapter by M.J.B. The author is otherwise responsible for the vast majority of the work in this research chapter.

5.1 Introduction

Crown archosauromorphs (Archosauria) have accrued immense taxonomic and morphological diversity across the terrestrial, freshwater, and marine realms since their emergence in the Triassic (Jetz et al., 2012; Brusatte et al., 2015; Benson 2018; Jagielska and Brusatte 2021; Stubbs et al., 2021). Their success, however, was founded on the earlier diversification of stem archosauromorphs. From a cryptic origin in the Middle to Late Permian (Evans and King 1993; Ezcurra et al., 2014; Ezcurra et al., 2020), stem archosauromorphs rapidly entered a host of ecological niches left newly vacant the end-Permian mass extinction (EPME), followed by sustained increases in disparity and diversity through the Middle Triassic as pioneering lineages explored the bounds of ecomorphospace under stabilised environmental conditions (Brusatte et al., 2011; Roopnarine and Angielczyk 2011; Foth et al., 2016, 2021; Ezcurra and Butler 2018; Kammerer et al., 2019; Ezcurra et al., 2021). Improved sampling and chronostratigraphic constraint of key assemblages (see Pradelli et al., 2021) have shown that archosaurs attained modest diversity during the Middle Triassic. It was during the Carnian Pluvial Episode (CPE) in the Late Triassic, however, that they rose to prominence as more archaic stem archosauromorph and tetrapod lineages declined (Bernardi et al., 2018; Dal Corso et al., 2020; Foffa et al., 2022), setting the stage for divergent ecomorphological innovation between bird-line (Avemetatarsalia) and crocodile line (Pseudosuchia) archosaurs through the remainder of the Triassic (Brusatte et al., 2008; Apaldetti et al., 2018; Stubbs et al., 2021).

This remarkable macroevolutionary origin took place through multiple climate upheavals across the ecologically extreme biogeographic stage provided by the supercontinent Pangaea, factors that are considered to have shaped the success of archosauromorphs (Benton et al., 2018; Bernardi et al., 2018; Simões et al., 2022). Climate simulations and the lithological record indicate the presence of hostile desert environments spanning the tropics (Parrish et al., 1982; Parrish 1993; Woods 2005; Sellwood and Valdes 2006; Tanner 2017), which have been linked to marked provinciality across the equator and bimodal distributions of tetrapod diversity centred on more clement seasonal conditions at higher latitudes (Sidor et al., 2005, 2013; Ezcurra 2010; Sahney et al., 2010; Whiteside et al., 2011; Brocklehurst et al., 2017; Brocklehurst et al., 2018; Allen et al., 2020; Romano et al., 2020; Dunne et al., 2021). The provincialism displayed by the Pangaean tetrapod fossil record implies the existence of non-geographic barriers to dispersal, most notably the ‘tropical dead zone’ extending across the equator in the wake of the EPME (Sun et al., 2012; Bernardi et al., 2018). Contrary to this

picture of a climatically segregated supercontinent, the oldest archosauromorph body and trace fossils spanned Europe, Tanzania, and South America, prior to a near global distribution in the Early Triassic (Bernardi et al., 2015, 2022). Stem archosauromorphs clearly dispersed widely, but the timing and pathways of their biogeographic expansion are unknown, making it challenging to determine whether they dispersed under relaxed climatic constraints, or were more climatically robust and capable of dispersal than is currently assumed.

The climatic tolerances of crown archosauromorphs have been given greater consideration, with the consensus that the constraints imposed by climate on thermophysiology and dispersal potential were key controls on the biogeographic structure and differential evolutionary success of various pseudosuchian and avemetatarsalian lineages following the CPE (Lovelace et al., 2020; Dunne et al., 2021, 2023; Chiarenza et al., 2022; Griffin et al., 2022; Hartman et al., 2022; Olsen et al., 2022; Simões et al., 2022). These tolerances, however, have only ever been quantified for the Late Triassic as a whole, masking our understanding of how complex, varied climates through this interval controlled the distribution and diversity of different clades of crown archosauromorphs, while virtually nothing is known about prior radiation of stem archosauromorphs. There is also a circularity in current approaches to quantifying climatic tolerances as fossil occurrences are interpreted as marking the precise geographic region and habitat a clade inhabited, excluding other possibilities where they have not been found (Liu et al., 2018; Dunne et al., 2021). Both interpretations are vulnerable to the effects of spatially biased sampling, and previous workers have called for the use of ancestral state estimation based on phylogenetic relationships to help overcome the limitations imposed by incomplete sampling (Brusatte et al., 2011; Dunne et al., 2021).

In this chapter I reconstruct the geometry of the early archosauromorph radiation across Pangaea using a novel landscape connectivity framework that couples ancestral state estimates of the climatic breadth and geographic roots of the archosauromorph radiation with high resolution reconstructions of topography and climate from Earth-system models. The minimal geometry required to satisfy the implications of their phylogeographic history provides a means of inferring their missing climatic tolerances beyond the bounds of the empirical fossil record. I use these results to characterise the cryptic Permian radiation of basal archosauromorphs and to re-evaluate the climatic controls on their success in the Triassic. Specifically, I test the null hypothesis that reconstructed climatic tolerances of early archosauromorphs do not differ substantially from measurements taken from the empirical distribution of the fossil record, and

the prevailing assumption that extreme Pangaean climates, particularly the Early Triassic tropical dead zone, were hard barriers to dispersal. Finally, I reevaluate the current view that bird-line archosaurs displayed greater climatic tolerances than crocodile-line archosaurs to determine whether this observation is an artefact of previous data treatment strategies and the failure to account for reconstructed ancestors.

5.2 Materials and methods

5.2.1 Time-scaled phylogeny and fossil occurrences

A phylogeny of archosauromorph reptiles was subsampled from an unpublished, informal supertree of early tetrapods within the Early Tetrapod Database (Benton et al., 2013; Benton 2015b; Elsler, 2019). Chosen taxa included all Permian and Triassic archosauromorphs, plus a small number of Jurassic taxa subtending Triassic-aged nodes. The tree of Ezcurra et al. (2017) forms the primary scaffold for Archosauromorpha with a variety of formal phylogenies informing subclade-specific topologies, alongside modifications to update taxon-level relationships and addition of new species surveyed from the literature up to September 2018 (see Elsler 2019). The tree was trimmed to genus level from species level then new genera surveyed from the literature up to May 2022 were added by the author for this work.

While the interrelationships of the early archosauromorphs are broadly well-resolved (Nesbitt 2011; Ezcurra 2016; Ezcurra et al., 2017), recent fossil discoveries and re-analysis of enigmatic taxa with modern imaging techniques have spurred revisions to the placements of basal avemetatarsalians, with substantial implications for how archosaurs rose to prominence (Ezcurra et al., 2020; Kammerer et al., 2020; Muller and Garcia 2020; 2023; Foffa et al., 2022). Consequently, four final phylogenies were considered: the original tree with the traditional placements of lagerpetids and silesaurids; a tree with the modified placement of lagerpetids; a tree with the modified placement of silesaurids; and a final tree with both modifications. The resolved topology of lagerpetids in Elsler (2019) was favoured over the clade-wide polytomy presented by Ezcurra et al. (2020). The OTUs in Muller and Garcia (2020) were added to the topology of silesaurids in Elsler (2019), one of which additionally features in Ezcurra et al. (2020). The topology of Elsler (2019) was then used for the trees with the traditional placement of silesaurids, and the topology of Muller and Garcia used for the alternative placement of lagerpetids. Finally, a few tips now considered to be *nomina dubia* or of non-archosauromorph

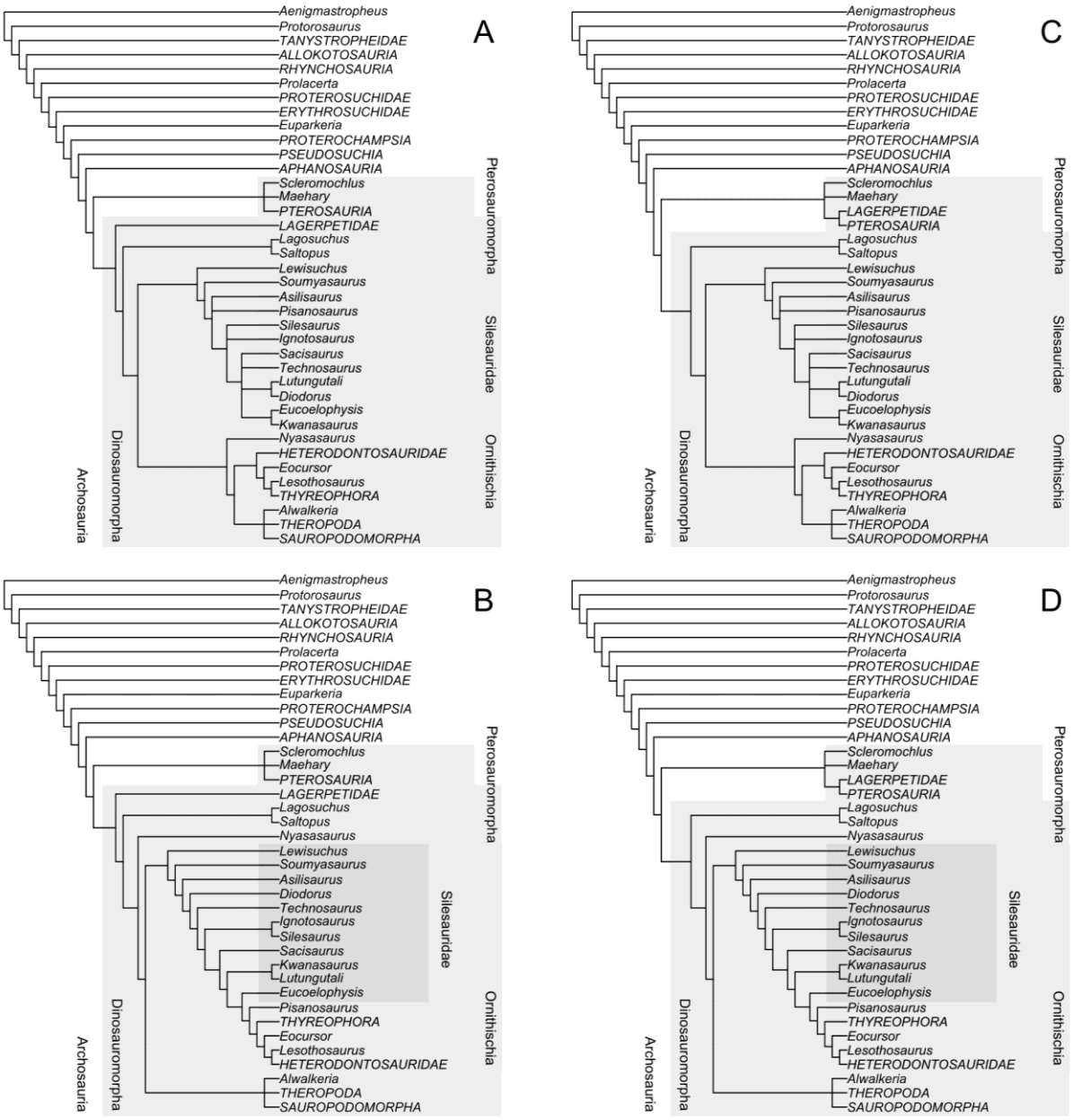


Fig. 5.1 Comparison of archosauromorph phylogenetic topologies. The different phylogenetic topologies considered for inference of biogeographic origins, differing in their placement of classically dinosauromorph clades. **(A).** Traditional archosauromorph phylogeny based on Ezcurra et al (2017). **(B).** Alternative placement of lagerpetids in Muller and Garcia (2020). **(C).** Alternative placement of silesaurids in Ezcurra et al., (2020). **(D).** Both alternative placements within the same tree.

affinity were removed from the trees. While the interrelationships of theropod, ornithischian and sauropodomorph dinosaurs are currently under debate (e.g., Cerny and Simonoff 2023), I do not investigate multiple topologies for these groups and instead retain the traditional topology of Seeley (1887).

Fossil occurrence data for all genera in the final set of trees were downloaded from the Paleobiology Database (PBDB) on 01/08/22 or manually added in the case of more recently named taxa, and updates made to a small number of synonymies on the basis on Elsler (2019). Occurrence chronostratigraphy was revised using the database of Elsler (2019), then updated to GTS2020 standard using the `chrono_scale()` function of the *fossilbrush* R package (Flannery-Sutherland et al., 2020b). Finally, palaeocoordinates were calculated under the Getech plate rotation model using an interpolation method to yield higher precision (see 4.2.2).

Each phylogeny was time-scaled using the fossilised birth-death (FBD) model with the clockless tip dating method of Zhang et al. (2016). Blank 'dummy' nexus files for this procedure were generated in R using the `createMrBayesTipDatingNexus()` function of the *paleotree* R package (Bapst 2021). Default function settings were retained as these settings reflect best methodological practise (Matzke and Wright 2016; Bapst 2021). The root age prior was set as an exponential with an offset of 10 Ma from the maximum age of the oldest empirical observation in the tree, chosen based on recent estimates of a Middle-Late Permian age for the root of Archosauromorpha (Ezurra et al., 2020). Tip priors were set as uniform distributions bounded by the maximum and minimum ages of their FAD occurrences, aside from the youngest taxon in the tree (*Allkaruen*, Pterosauria, Toarcian of Argentina) whose tip prior was held constant as its FAD maximum age to permit *post hoc* conversion of relative branch lengths to calendar time. Tip dating was implemented in MrBayes (v3.3.7; Ronquist and Huelsenbeck 2003; Altekar et al., 2004) using Markov Chain Monte Carlo (MCMC) to estimate the parameters of the FBD model from their posterior distributions. Four replicate analyses were run for each tree for 250 million generations, sampling every 1000, with each replicate analysis comprising four Metropolis-coupled MCMC chains to enable more efficient exploration of parameter space. Model parameters were summarised across all chains, with convergence identified in all analyses by potential scale reduction factors approaching one (Gelman and Rubin 1992). Lending further support to the robustness of these analyses, mean estimated divergence times and their associated highest posterior densities (HPDs) across all four archosauromorph phylogenies were highly similar, aside from a few nodes associated with the alternative positions of lagerpetids and silesaurids (see 5.3.1).

5.2.2 Geographic origin ancestral state estimation

Geographic origins are often inferred using discrete spatial areas as optimisable character states (typically major continental blocks or bioregions). Pangaea does not present the typical

geographic barriers that enable ready designation of bioregions and while such regions could be defined based on assemblage similarity to avoid subjectivity (e.g., Sidor et al., 2013; Button et al., 2017), such approaches may return spurious results by inserting 'barriers' that are false and represent gaps in sampling (Brocklehurst et al., 2018; Dunne et al., 2018). Further, as the origin of a clade is ascribed to a potentially vast geographic extent, its environmental content cannot be precisely inferred, nor do such analyses consider how variation in geographic and environmental constraints across bioregions may have affected dispersal potential along their borders (although see van Dam and Matzke 2016). Instead, geographic points of origin were estimated as spatially continuous coordinates using the *geo* model of O'Donovan et al. (2018), which considers range evolution through phylogeny as a random walk across the surface of the globe (i.e., between Cartesian coordinates on a sphere). The parameters of the *geo* model (internal node and tip points of origin and the rates of dispersal) are estimated by MCMC from a time-scaled phylogeny and the longitude-latitude coordinates of its tip observations, supplied as either single observations or a set of observations from which the estimated tip state is sampled according to its probability. In its original implementation, the *geo* model considers all regions of the sphere to be equally accessible and so cannot account for how past geographic configurations may have fundamentally limited clade distributions, although estimated locations may remain plausible subject to the quality of the phylogeny under consideration (O'Donovan et al., 2018). An updated version of the model used here, however, can additionally be supplied with sets of geographic masks, where each mask enables inaccessible regions to be designated within discrete windows through time (e.g., geological stages; A. Meade pers. comm.). An ongoing limitation of the model is that it can only consider a single tree topology, precluding the inclusion of topological uncertainty into an analysis, although polytomies are permitted (of which several are present in the trees analysed here).

Geographic origin ancestral states were estimated for each time scaled archosauromorph phylogeny with an experimental implementation of the *geo* model in BayesTraits (v.4.0; Pagel et al., 2004; A. Meade pers. comm.). Geographic masks were generated from Early Permian (Kungurian) to Middle Jurassic (Bajocian) stage level palaeogeographic digital elevation models (DEMs) from Getech PLC. DEMs were masked into land and sea (accessible versus inaccessible) regions at their native resolution (0.5° × 0.5°) and upscaled by a factor of two (0.25° × 0.25°) in R (v.4.0.4) to improve the performance of continuous coordinate proposal in the *geo* model analysis, then BayesTraits-compatible mask files were generated using a bespoke C++ utility (A. Meade, pers. comm.). Getech palaeocoordinates for fossil occurrences

were checked to ensure that they fell within the bounds of the accessible portion of their stage-specific palaeogeographic masks and adjusted to the nearest land cell where needed. While the *geo* model can accept a sample of coordinate tip states, a matching set of single tips must also be provided to initiate the analysis: these were supplied directly in the case of singleton taxa, as great circle midpoints in the case of pairs of tip observations, or the centroid in the case of three or more tip observations. Two models were fitted for each tree with six replicate analyses per model: one where the rates of movement varied uniformly across the tree according to a Brownian motion prior (BM); the other where variable rates (VR) across the tree were proposed during the analysis using reverse jump MCMC (Venditti et al., 2011). Each analysis was run for one billion iterations, discarding the first 50% as burn-in, then sampling every 50,000. During each analysis, stepping-stone sampling was used to calculate the marginal likelihood of the fitted model, with 1000 equally spaced stones through the post-burnin portion of the chain and 100,000 iterations per stone.

Post-processing of BayesTraits log files was performed in R (v.4.0.4). MCMC traces were plotted to visually assess stationarity (Fig. 5.2), then formally tested using the `effectiveSize()` function of the *coda* R package (Plummer et al., 2006) to determine effective sample sizes (ESS). Convergence between replicate analyses was tested using the `gelman.diag()` function of *coda* (Gelman and Rubin 1992) to calculate potential scale reduction factors (PSRF). Finally, the marginal likelihoods of the BM versus VR models were averaged across replicate analyses, then Bayes factors calculated to identify the better performing model (Kass and Raftery 1995). The VR models were identified as better fitting in all cases (logBF = 8), but likelihood traces and PSRFs indicated inadequate convergence between chains (PSRF > 1.1), despite each individual chain reaching stationarity (ESS > 200). 90% of the estimated ancestral states nonetheless had PSRFs approaching one, with 95.5% of terminal nodes and 85.4% of internal nodes displaying unimodal posterior distributions of proposed origin points. Consequently, only a handful of nodes appear to be responsible for the signal of non-convergence in the VR models and these are frequently cases where singular geographic points of origin could not be readily estimated. As such, the VR analyses are taken here as largely reliable, particularly given their strong Bayes factor support (BF > 6, Kass and Raftery 1995). Replicate log files from each VR analysis were concatenated, sets of ancestral origination points spatially binned using a 1° × 1° grid, and the resulting frequencies smoothed with a Gaussian kernel. Node estimates were summarised by taking the cell-centre coordinate

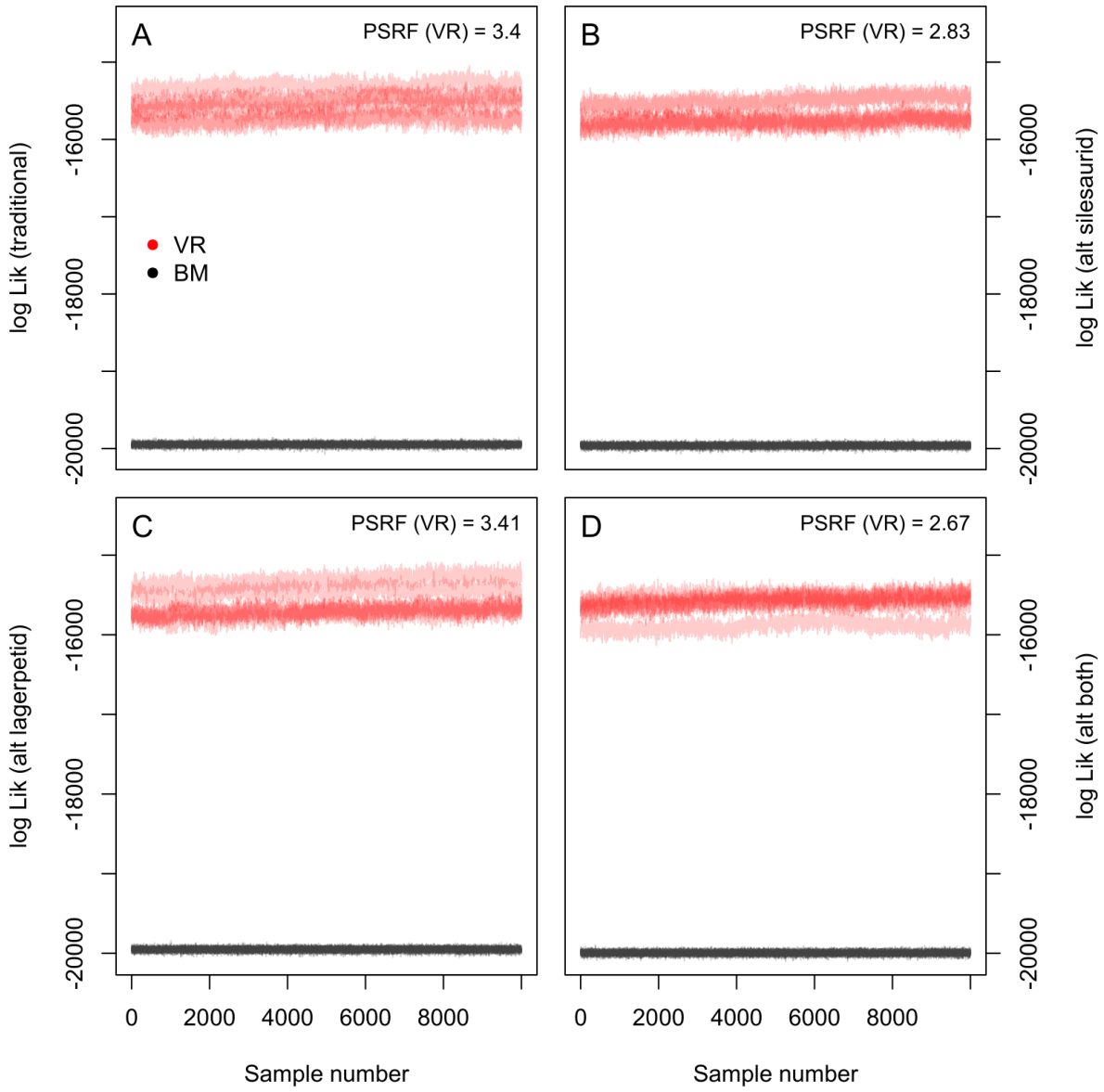


Fig. 5.2. Likelihood traces from the BayesTraits *geo* model. In all cases, the variable rates (VR) model is substantially better fitting than the Brownian motion (BM) model. Each VR run displays stationarity, although between-chain convergence is low with potential scale reduction factors (PSRFs) deviating from ≈ 1. **(A).** Traditional archosauromorph topology. **(B).** Alternative topology with silesaurids as basal ornithischians. **(C).** Alternative topology with lagerpetids as pterosauromorphs. **(D).** Alternative topology with the alternative positions of silesaurids and lagerpetids.

of highest density cell in the grid (i.e., the highest probabilistic estimate for the node geographic origin), and of the ranges and 95% highest density intervals for the estimated longitudes and latitudes. As distance is not constant across latitude, the ancestral longitude-latitude points were projected to the Mercator coordinate system, then lengths of the major and minor axes of their 95% confidence ellipse calculated to provide true distance confidence intervals.

5.2.3 Climate tolerance ancestral state estimation

The investigated climatic conditions for early archosauromorphs were selected from the BIOCLIM scheme, a series of **bio**logically relevant **clim**atic variables developed for species distribution modelling in terrestrial habitats (Nix 1986; Elith et al., 2006; Hijmans and Graham 2006). These variables are calculated from monthly data on maximum and minimum temperature and precipitation data and include common metrics like mean annual temperature and annual precipitation, along with descriptors of their ranges (i.e., minima and maxima) and seasonal variations. The full set of BIOCLIM variables was calculated from the monthly temperature and precipitation data from each climate simulation using the `bioclim()` function of the *dismo* R package (Hijmans 2022). Sets of BIOCLIM variables were then investigated for multicollinearity using Pearson's product moment tests of correlation, dendrograms constructed from the resulting set of pairwise correlations, and clusters of highly correlated variables identified at a cut-off off at a dendrogram node height of 0.3 (corresponding to 70% positive correlation). Redundant variables within each cluster were discarded to reduce multicollinearity, leaving the maximum and minimum, annual mean and mean diurnal range of temperature, isothermality, annual precipitation, precipitation seasonality, and precipitation in the wettest and driest months as the final set of climatic conditions for initial analysis.

The selected climatic conditions for fossil archosauromorphs were extracted from stage-level deep time climate simulations from the HadCM3L general circulation model with Getech palaeogeographic DEMs as boundary conditions (see 4.3.2), using the geographic origin ancestral state estimates derived for each taxon with the BayesTraits *geo* model. Climatic conditions were spatially sampled from stage-specific climate simulations with respect to tip age, in proportion to the density of longitude-latitude estimates for each node, then sets of conditions summarised using their mean and 95% confidence intervals. Maximum likelihood ancestral state estimates and 95% confidence intervals were then calculated for the internal nodes using the `fastAnc()` function of the *phytools* R package (Revell 2012) and the extracted climatic conditions for the fossil tips. To validate these estimates, climatic conditions for the internal nodes were extracted using the same procedure as above, then the 95% confidence intervals of the geographic origin and ancestral state estimates of climate tolerances compared to determine whether they overlapped.

The climate space occupied by ancestral archosauromorphs, and their fossil representatives was ordinated using principal components analysis (PCA). The selected BIOCLIM variables capture key aspects of their overall climate space, but some still display a modest degree of correlation. Consequently, the PCA loadings were examined to remove redundant variables, then the PCAs repeated using mean annual temperature, temperature seasonality, mean diurnal temperature range, annual precipitation and precipitation seasonality. Climatic disparity through geological time was summarised using three metrics to capture the extent and dispersal of archosauromorph tips and nodes in climate space (the sum of ranges, the sum of variances and the mean distance from the centroid), chosen for their relative robustness to sample size effects and outliers in ordination spaces (Ciampagliao et al., 2001). The climate tolerances in each time bin were randomly subsampled 100 times to their bin-specific data quantity to provide bootstrap confidence intervals for each summary statistic. Separate summaries were produced using the estimated ancestral states only.

5.2.4 Spatiotemporal landscape graphs

The geographic distance between the points of origin for a pair of ancestor-descendant nodes divided by their intervening branch length gives a measure of ancestral dispersal rate (O'Donovan et al., 2018). The locations of those points were derived whilst accounting for palaeogeographic context as above, but this does not extend to their interconnecting paths. Previously, geodesic (or great circle) paths between a pair of coordinates have been used for distance, direction, and rate calculation (e.g., O'Donovan et al. 2018; Gardner et al., 2019; Avaria-Llautureo et al., 2021). This accounts for necessity of travel across a surface of a sphere but not the additional distance incurred by changes in elevation along the path, although this will be a second order effect relative to great circle distance itself on continental to global scales. More importantly, blind use of geodesics makes the implicit assumptions that all intervening areas of the sphere are accessible and that ease of movement across its surface is isotropic, while such paths may instead cut across impermeable terrains like oceans or high mountains, or through regions with other environmental conditions that are otherwise unfavourable to dispersal. The solution to the problems posed by geodesic paths is a concept derived from landscape connectivity analysis, the least cost path (LCP), which marks the route of lowest resistance between two points through an anisotropic landscape.

Landscape connectivity analysis originated as a tool for optimising urban planning and transit systems (Warntz 1957, 1965; McHarg 1967; Turner and Miles 1971; Tomlin 1990) but

has since been applied to ecological problems involving the movement of organisms through spatially heterogenous, anisotropic landscapes (Taylor et al., 1993; Ricketts 2001). LCPs which minimise expended effort and/or risk model how organisms are expected to optimise their movement through their habitats, whether for acquisition of resources like nutrition or reproductive partners, or for hazard avoidance (Gallagher et al., 2017; Balbi et al., 2019). Consequently, geographic 'nearness' and accessibility in ecological systems are often better characterised using LCPs rather than Euclidean paths, with the observations that LCPs can accurately predict empirical travel routes across organisms in a variety of habitats and genetic distances between populations (Michels et al., 2001; Wang et al., 2009; Green et al., 2020). The utility of LCPs has also been recognised in archaeological and palaeoanthropological contexts to identify likely routes of dispersal and trade (Ericson and Goldstein 1980; Van Leusen 1993; Verhagen et al., 2019), including potential hominid dispersal pathways during the early Pleistocene (Carotenuto et al., 2016; Beyin et al., 2019; Husson et al., 2022).

Landscape connectivity problems (i.e., quantifying the cost of moving through geographic space) are modelled using graphs (Urban and Keitt 2001; Etherington 2016), where vertices represent physical locations, edges the available connections between those locations, and edge weights the cost of transit for those connections (termed their resistance). The weights associated with graph edges may reflect the true geographic distances between locations but can alternatively quantify other, potentially non-Euclidean aspects of node connectivity, like constrained travel between nodes (e.g., between urban locations along a road network) or disparities in their conditions (e.g., temperature gradient between the foot and summit of a mountain). A landscape, represented as a raster grid, is converted to a lattice graph with grid cells as vertices and edges connecting orthogonally adjacent or orthogonally and diagonally adjacent cells (4-degree Rook's case vs 8-degree Queen's case). In an idealised, isotropic landscape, all edges will be symmetrically traversable and equally weighted; travel is possible in both directions along an edge and incurs the same cost regardless of edge or direction. For anisotropic landscapes, edge weights vary according to the changing cost of traversal between grid cells and may also be asymmetric for a given edge; the cost of transit from *a* to *b* is not equivalent to *b* to *a*, with the scenario that travel may be possible in one direction only. The LCP between a pair of vertices in a graph comprises the set of intervening edges with the lowest cumulative resistance and can be identified using Djikstra's algorithm (Djikstra 1959), yielding the optimal pathway through the original landscape raster.

Despite their utility in ecological analysis and applications to problems in the recent geological past, LCPs have not been applied in wider palaeobiological contexts to examine the biogeographic histories of fossil taxa (although landscape connectivity has been considered in the context of discrete area phylogeography; van Dam and Matzke 2016). Part of the problem is that LCPs are calculated on fixed landscapes, namely that of the present day, or if they are set in a palaeogeographic reconstruction, modified to reflect changes in sea level or ice cover, but only at a single time. Palaeobiological dispersal, however, may take place across much greater time spans with substantial intervening changes in landscape architecture. To estimate phylogeographic paths and rates of dispersal for archosauromorphs, I build on existing developments in landscape connectivity analysis with a workflow termed TARDIS (Traverses And Routes of Dispersal In Spacetime), although the tools developed here can be developed for analysis of pathways between any given pairs of spatiotemporally resolved points (longitude, latitude, age) regardless of their derivation. The core concept behind TARDIS is extension of the two-dimensional lattice graph used in any landscape connectivity analysis (Fig. 5.3A) to a three-dimensional lattice graph, where each horizontal layer of vertices and edges represents a changing landscape across discrete time frames, with unidirectional edges linking spatially homologous points from older to younger graph layers (Fig. 5.3D). Pathways within this lattice can therefore traverse geographic space or move with the passage of time, while reverse jumps in time are logically impossible. From here, any tool applicable to directed graphs can be applied to analyse spatial properties of the lattice and of the pathways between vertices. Graph construction and analysis is implemented in several custom R functions, with the *raster*, *igraph*, *geosphere*, *nngeo*, and *cppRouting* R packages as dependencies (Csardi and Nepusz 2006; Hijmans et al., 2019; Gombin et al., 2020; Dorman 2022; Larmet 2022; Hijmans 2023).

5.2.5 Functions for TARDIS landscape graphs

Landscape rasters for TARDIS can be supplied at any spatial resolution and extent, although these properties must remain uniform between time frames. Raster resolution determines the number of vertices in each graph layer with a choice of 4-degree or 8-degree adjacencies between vertices. In the case of global rasters, edges spanning the antimeridian are included. Raster resolution, number of time frames and choice of cell adjacency are thus the primary factors affecting final graph size, which may impose potential memory constraints on TARDIS for analyses with high spatiotemporal resolutions. Symmetric edge weights are assigned as the

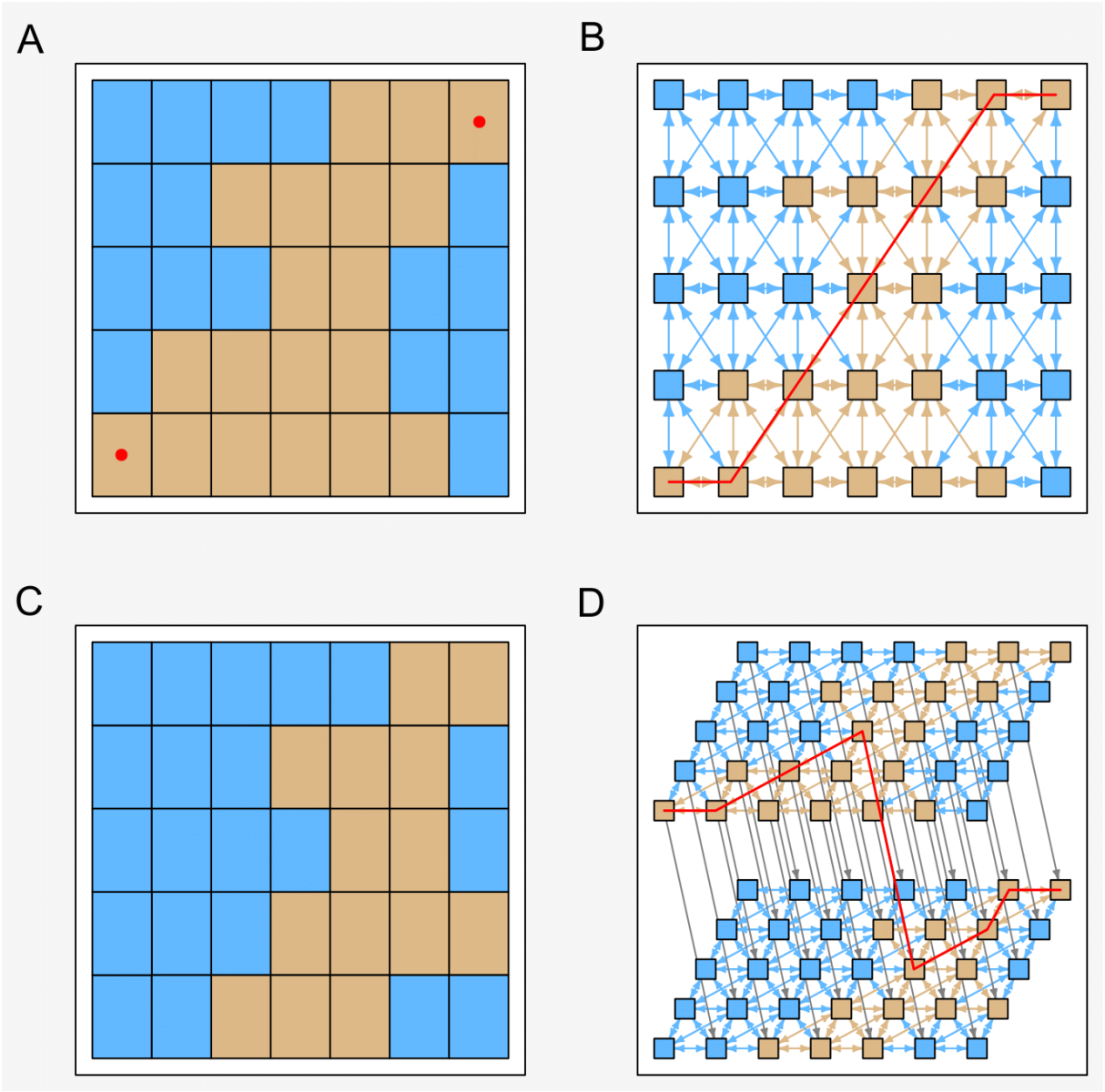


Fig. 5.3. Illustration of a TARDIS spatiotemporal lattice graph. (A). A hypothetical raster landscape comprising a single contiguous landmass (tan) surrounded by ocean (blue) with a pair of source and target points (red). **(B).** Lattice graph representation of the raster grid in (A) where grid cells are vertices with Queen's case connections to their neighbours, and the least cost path between the source and target points. **(C).** A second raster landscape which may be considered as the evolution of (A) after millions of years of topographic remodelling and continental drift. **(D).** The three-dimensional lattice graph used by TARDIS, connecting the same pair of source and target points, but through time as well as across space.

distances between grid cell centres. If all values in a landscape raster are identical then topographic changes do not apply and solely the great circle distances between cells are considered. Alternatively, the landscape raster can be a DEM recording topography and bathymetry, in which case great circle distances are adjusted using Pythagoras' theorem to

account for the additional distances incurred along the path by the change in elevation between adjacent grid cells. Conversion of landscapes to graph layers is implemented with the function `geog_link()`.

A set of mask rasters can be supplied to `geog_link()` during graph construction to designate impermeable cells in the landscape, resulting in removal of any edges to and from the corresponding vertices in the graph. As the use of masks may produce islands in a landscape which cannot then be accessed from other parts of the lattice graph, the closest grid cells between islands are identified and connected with additional edges weighted as the intervening great circle distance. In this way, islands are eliminated, but the shortest, most direct transitions are enforced under the assumption that they are most likely. While masks incur these potential complications, they are useful not only to block movement through untraversable portions of the landscape (e.g., oceans in the case of cursorial organisms), but also to eliminate unnecessary connections and so avoid memory constraints, besides the option of using 4-degree instead of 8-degree adjacencies. The resultant graph, with potential mask-induced non-adjacent connections contains the minimal set of geographic pathways required to traverse between any unmasked locations in the input landscape. Mask rasters are checked using the utility function `mask_check()`, which is called automatically by `geog_link()`, but can also be used separately to critique the unmasked links between islands prior to landscape construction.

Minimal geographic routes are useful as they solve the issue posed by naively using great circle connections between points without considering the connectivity of the intervening landscape. Other factors may affect traversal, however, so a separate set of weights can be calculated for the graphs produced by `geog_link()` using a set of rasters of the same resolution as the original landscape raster that describe alternative properties of the landscape. The data from these rasters, in addition to the horizontal and vertical distances along graph edges can then be used to construct custom weighting schemes for the graph which represent the cost of transit through the landscape for the biological system of interest. Weighting schemes are highly flexible, allowing for any number of calculations using the geographic distance and landscape data and any combination of the source and destination cells defining each graph edge, and permitting asymmetric traversal costs to be represented (e.g., a reduced cost when moving from less hospitable to more hospitable cells and vice versa). Edges connecting islands of cells resulting from the use of masks are assigned an arbitrarily high weight exceeding the highest empirical weight in the graph to ensure that a longer, but more

likely pathway will be preferentially recovered. While capturing complex landscape connectivity is desirable, weighting strategy strongly affects LCP calculation and may ultimately be subjective, necessitating careful choice informed by biological realism (Koen et al., 2012), although the original distance weights are stored separately so that the empirical length of an LCP can be determined. Graph weighting is performed by the function `geog_weight()` and the resulting resistance surfaces can be extracted as a raster from the resultant 2D lattice for visualisation using the utility function `resistance_surface()`.

Once a set of temporally discrete landscapes has been converted to 2D lattice graphs, they can be sequentially linked to produce the final 3D lattice graph considered by TARDIS (Fig. 5.3D). In the simplest case, vertical, unidirectional edges link spatially homologous cells between successive time frames. Masked cells in any time layer do not receive any inbound or outbound edges. Vertical edges spanning all time frames all receive uniform edge weights so that costs of traversal relate entirely to the changes within landscapes. Vertical edges are appropriate for relatively short time spans where geographic locations remain constant within the reference frame of the landscape extent. For global analyses through longer periods of time, however, geographic locations will shift position due to continental drift (Fig. 5.3C). To account for this effect, the changes in cell positions from first to last time frame can be supplied during construction of the 3D lattice (Fig. 5.3D). Not all cells need to have connections defined, nor will this be possible if masked cells are present in any of the layers, but at least one connection is required between layers to ensure that the passage of time is continuous through the graph. Landscape graph layers are linked through time using the function `time_link()`.

5.2.6 Pathway analysis in spatiotemporal graphs

LCP analysis within the 3D lattice is performed by matching a pair of spatiotemporal coordinates (longitude, latitude, age) to their nearest spacetime vertices, obtaining the intervening vertices comprising their LCP using Djikstra's algorithm, then summing the elevation-adjusted great circle distances between successive pairs of vertex coordinates. This functionality is provided with the function `tardis()`, but several considerations must be noted here. The LCP, subject to chosen graph weights, is optimal but there may be many alternative paths with costs approaching that of the LCP, while organisms themselves are unaware of the geometry of the theoretical LCP, do not disperse with a predetermined destination in mind, and do not necessarily behave as optimal agents during transit (Mueller

and Fagan 2008; Pinto and Keitt 2009). Instead, their movement may be better described as a random walk directed by landscape connectivity. As such, LCPs provide useful lower bounds on dispersal distances and geometries but may be closely redundant with multiple other pathways in the same graph and the actual distances covered during the dispersal between a pair of ancestor-descendant nodes is likely much greater.

The full extent of any detours cannot be inferred without well dated fossil evidence and may be impossible to determine. Instead, the difficulty of a journey can be captured by the commute time for a pair of vertices (the average number of steps required for a random walker to make a round trip between two) which can be calculated from the resistance properties of the graph (Chandra et al., 1989). Mathematical derivations of commute times are currently only available for undirected, strongly connected graphs (Boley et al., 2011, Padgham 2019), but the average time taken to traverse between two vertices can instead directly simulated using a Markovian random walk through the graph (Husson et al., 2022). Similarly, detours of a given magnitude can be calculated around an LCP, providing a mean of assessing path redundancy. I therefore provide the function `detours()` to calculate an isochrone around an LCP encompassing all nodes that fall within a percentage additional cost relative to that of the LCP itself (the least cost wide path) to provide a sensitivity measure of plausible routes around the optimum. I also provide the function `random_walk()` to conduct Markovian random walks between two spatiotemporal locations. As traversing the unidirectional edges between graph layers may cause a random walker to irreversibly overshoot the layer containing the target vertex, random walks are conducted piecewise between the vertices bounding the portions of the overall LCP in each time frame. The numbers of steps prior to hitting the target vertex with respect to time frame are averaged across multiple trials, then the overall distance summed across time frames. Conceivably, the length of random walk could imply unrealistically rapid rates of movement, but these lengths should be interpreted as an approximation of the resistance between a pair of vertices rather than as true distances of traversal.

5.2.7 Routes and rates of phylogeographic dispersal

To infer early archosauromorph phylogeographic routes of dispersal, I converted the stage-level Getech palaeogeographic DEMs for the intervals covering the temporal extent of the four time-scaled trees (Kungurian–Bajocian) into landscape connectivity graphs using `geog_link()`. The 2D lattice graphs were then weighted to account for the effects of

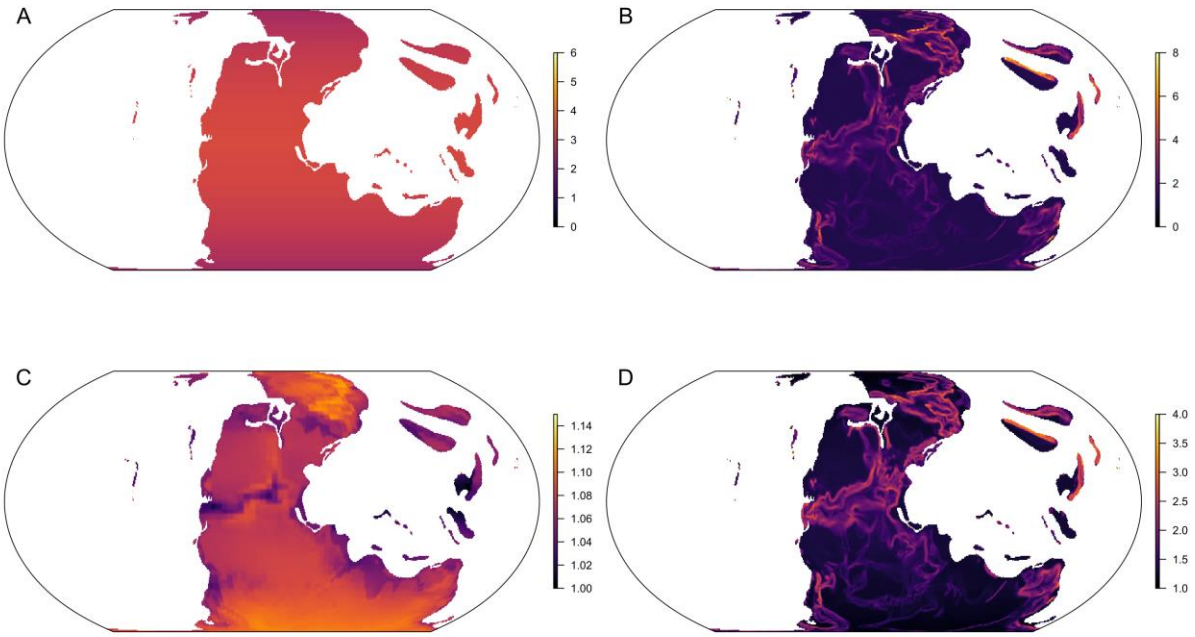


Fig. 5.4. Resistance surfaces for TARDIS from different weighting schemes. Resistances are displayed for the Roadian stage of the Middle Permian. In each case, resistance units are relative and arbitrary. **(A).** Resistance based on great circle distance between grid cells. **(B).** Resistance based on difference in elevation between grid cells. **(C).** Resistance based on Euclidean distance from the centroid of archosauromorph climate space. **(D).** Resistance derived from multiplying together the components in A–C.

changing topography and climate. As various early archosauromorph clades displayed markedly different climate tolerances (Dunne et al., 2021, 2023), it is not straightforward to construct a weighting scheme for TARDIS that would precisely parameterise all costs of dispersal with respect to their phylogeny. Instead, a compromise approach was taken where the minimal set of BIOCLIM variables outlined above (see 5.2.3) was extracted from the HadCM3L general circulation models for all vertices in the TARDIS graph, ordinated using PCA, and the Euclidean distance from the position of each vertex to the centroid of archosauromorph climate space calculated. Euclidean distances in climate space were then multiplied by the great circle distances and changes in elevation between grid cell centres to give a compound weighting metric describing the topographic and climatic difficulty of traversal relative to the archosauromorph climatic average. Each component of the compound metric was scaled to unit variance prior to multiplication to avoid their vastly different magnitudes and ranges from biasing the final distances in favour of any single component. This strategy was compared to weighting schemes from each individual component, using `resistance_surface()` for visualisation (Fig. 5.4).

As outlined in the previous chapter, incomplete spatial sampling of fossil clades may lead to underestimation of their climatic niches. To investigate this possibility, I downloaded modern occurrence data for the extant phylogenetic bracket of Triassic archosauromorphs, squamates and crocodilians, from the Global Biodiversity Information Facility on 09/06/23 using functions from the *rgbif* R package (Chamberlain et al., 2023). Birds were not included due to their derived thermophysiologies, while squamates were used due to the limited diversity and distribution of crocodylians. Occurrences were truncated to human-observed or specimen-based records only, then their climate states for the minimal set of BIOCLIM variables extracted from a present-day HadCM3L general circulation model simulation (see 4.2.3). The sets of climate data for modern reptiles, the modern Earth, Permo-Triassic archosauromorphs, and the Permo-Triassic Earth were then ordinated in a single PCA analysis. Plotting the first two principal components of this space demonstrates that the distributions of Permo-Triassic and modern climates are quite similar (Fig. S94). In some ways this is unsurprising as while atmospheric CO_2 concentrations and palaeogeographies differ between these two timeframes, resulting in radically different spatial distributions of climate, there is only so much absolute climatic variation that the Earth-system can generate. Within this space of all possible climates, modern reptiles are more widely sampled compared to Permo-Triassic archosauromorphs and the centroids of the two groups occupy moderately different positions, with modern reptiles extending more strongly into warmer and wetter tropical environments. This highlights a limitation of the centroid-based approach taken here, but the distribution of Permo-Triassic archosauromorphs nonetheless covers a broad extent of modern-day reptile climate space, suggesting that their fossil record does reflect the broad thermophysiological constraints to which all reptiles are subject.

Weighted graph layers were linked together into a 3D lattice graph using `time_link()`, with interlayer vertex connections calculated as the changing positions of spatially homologous cells using the interpolation method for the Getech plate rotation model (see 4.2.2). LCP paths and their empirical distances were calculated with Djikstra's algorithm for the highest density geographic origination points for each pair of ancestor descendent nodes in each time-scaled archosauromorph phylogeny using `tardis()`. LCP distances were then divided by their corresponding time-scaled branch lengths to measure the rate of traversal through archosauromorph phylogeny. To visualise the potential extent of unsampled climate space traversed during their dispersal across Pangaea, the minimal set of BIOCLIM climatic

conditions were sampled along each phylogeographic pathway, then projected into the PCA space derived from tip-state and ancestrally estimated climatic tolerances (see 5.2.3).

Finally, to detect potential signals of event-driven adaptation and dispersal through the Late Permian to Early Jurassic, I used phylogenetic ride regression to search for shifts in evolutionary rates through archosauromorph phylogeny using functions from the RRPhylo R package (Castiglione et al., 2018). This method was chosen as it allows for evolutionary rates to be modelled from pre-supplied tip and ancestral states rather than estimating the latter during model fitting, as is the case here.

5.3 Results

5.3.1 Alternative positions of stem dinosaurs

Progressive removal of the traditional stem dinosaur clades had virtually no effect on the mean divergence times and uncertainties for Archosauria, Avemetatarsalia, Ornithodira and Dinosauromorpha (Fig. 5.5A), increased divergence times for Dinosauriformes, Saurischia, Sauropodomorpha and Theropoda by just one million years (Fig. 5.5B), and increased the estimates for Pterosauria, Pterosauromorpha and Ornithischia by four, seven and twenty-six million years, respectively (Fig. 5.5C, 5.5D). Inclusion of lagerpetids within Pterosauromorpha reduced their own divergence time and associated uncertainty by roughly three million years, and the divergence time uncertainty for Pterosauromopha by almost ten million years. Revision of silesaurids decreased the divergence time uncertainty for Ornithischia by around seventeen million years and the divergence times of its deepest split, Heterodontosauridae – Genasauria, by five million years, accompanied by discrepant decreases and increases to the uncertainties of these groups by ten and eight million years, respectively.

Alternative placements of traditional dinosaur outgroup clades affected estimated geographic origins for three parts of the tree (Fig. 5.6). Archosauria displays a bimodal origin split between northern Pangaea and western Gondwana, with revision of silesaurids and lagerpetids pushing support in favour of the former and the latter regions, but a fully unimodal origin in northern Pangaea with revision of both groups. This same series of changes consequently affects the bimodal distributions displayed by Pseudosuchia and its earlier diverging subclades (see 5.3.2). With a traditional stem dinosaur topology, Pterosauromorpha

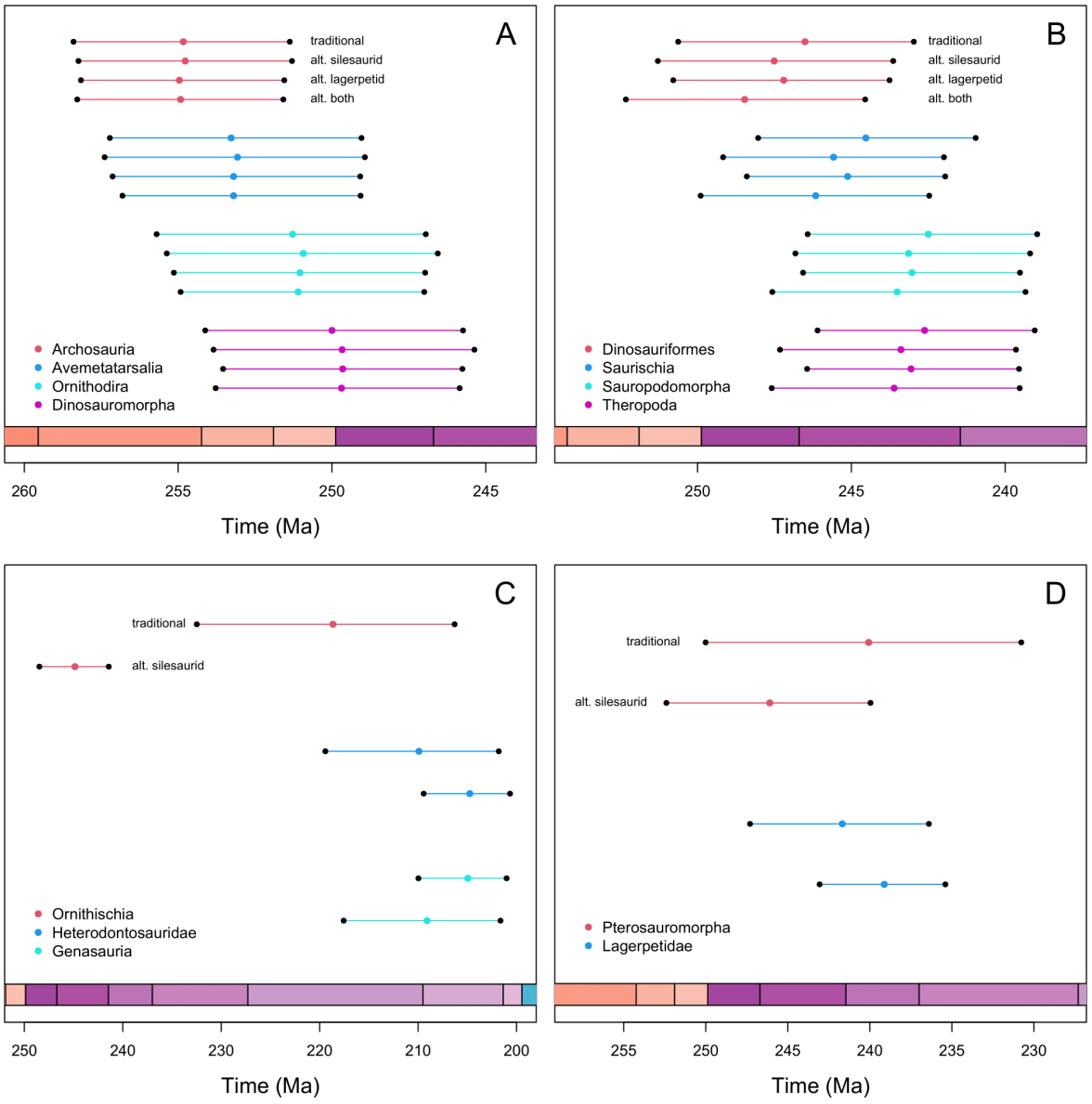


Fig. 5.5. Effects of alternative topologies on bird-line archosaur divergence times. Mean divergence times and their 95% highest posterior density intervals. **(A).** Impact on the deepest archosaur divergences within the bird line. **(B).** Impact on dinosaur-line divergences. **(C).** Impact on ornithischian divergences. **(D).** Impact on pterosaur-line divergences.

displays a bimodal origin split between northern Pangaea and western Gondwana but favouring the latter. When silesaurids are included within Ornithischia, a northern Pangaean origin is favoured, while the alternative position of lagerpetids induces a fully unimodal origin in western Gondwana. Finally, the alternative position of silesaurids collapses a highly uncertain, multimodal estimate for Genasauria to a unimodal distribution spanning the tropics.

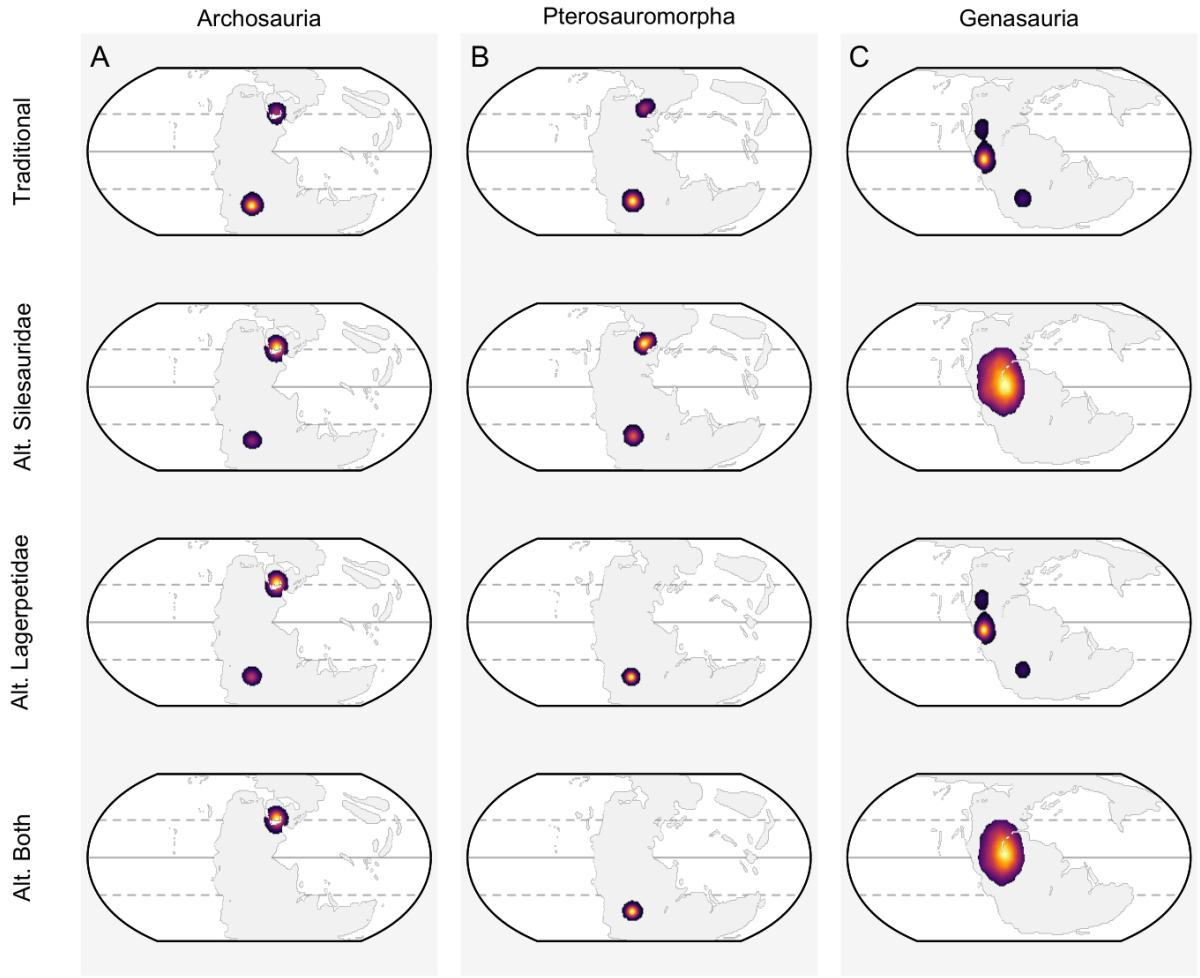


Fig. 5.6. Effects of alternative topologies on bird-line archosaur geographic origins. Kernel density estimates of geographic origination points from the BayesTraits *geo* model. **(A).** Impact on the origin of Archosauria. **(B).** Impact on Pterosauromorpha. **(C).** Impact on Genasauria.

While a consensus has yet to be reached on the alternative positions of silesaurids and lagerpetids, those positions substantially reduce the spatiotemporal uncertainty in origination estimates for several key avemetatarsalian clades and divergence time uncertainties between all four trees more generally are minimal. Spatial uncertainty within each tree is also generally low, with most nodes displaying unimodal points of origin and confidence intervals of 2000 km or less (Fig. S95). Consequently, results from the tree with the alternative positions of both clades are presented for the remainder of this chapter (Fig. 5.7, 5.8, 5.9).

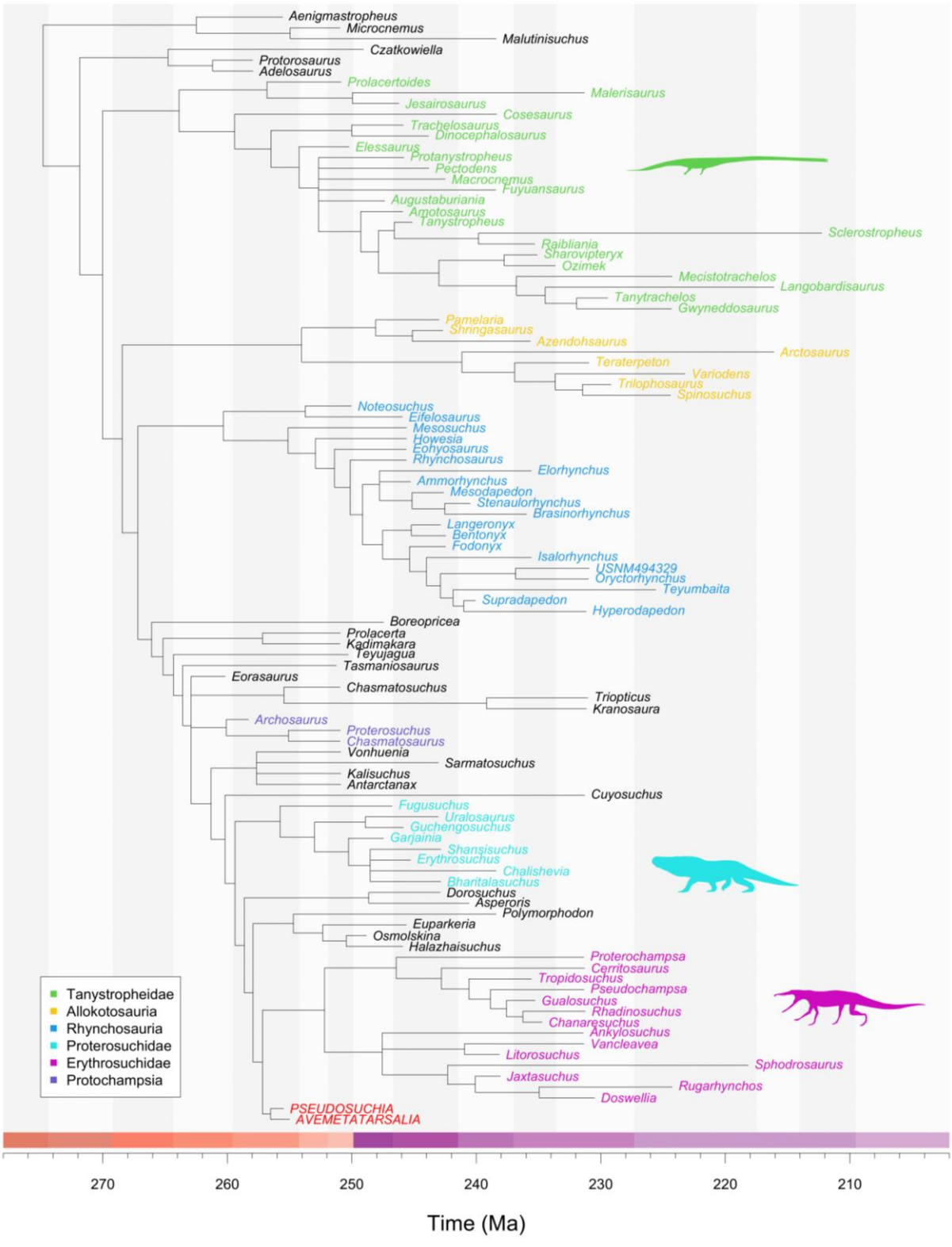


Fig. 5.7. Basal and higher archosauromorph phylogeny. Part of a time-scaled Permian–Triassic archosauromorph phylogeny from MrBayes. The constituent clades of crown-group Archosauromorpha are given in red. Please refer to Fig. 5.7 and Fig. 5.8 for their respectively topologies. Grey bars demarcate Permian and Triassic stages, with subdivisions of the Carnian and Norian.

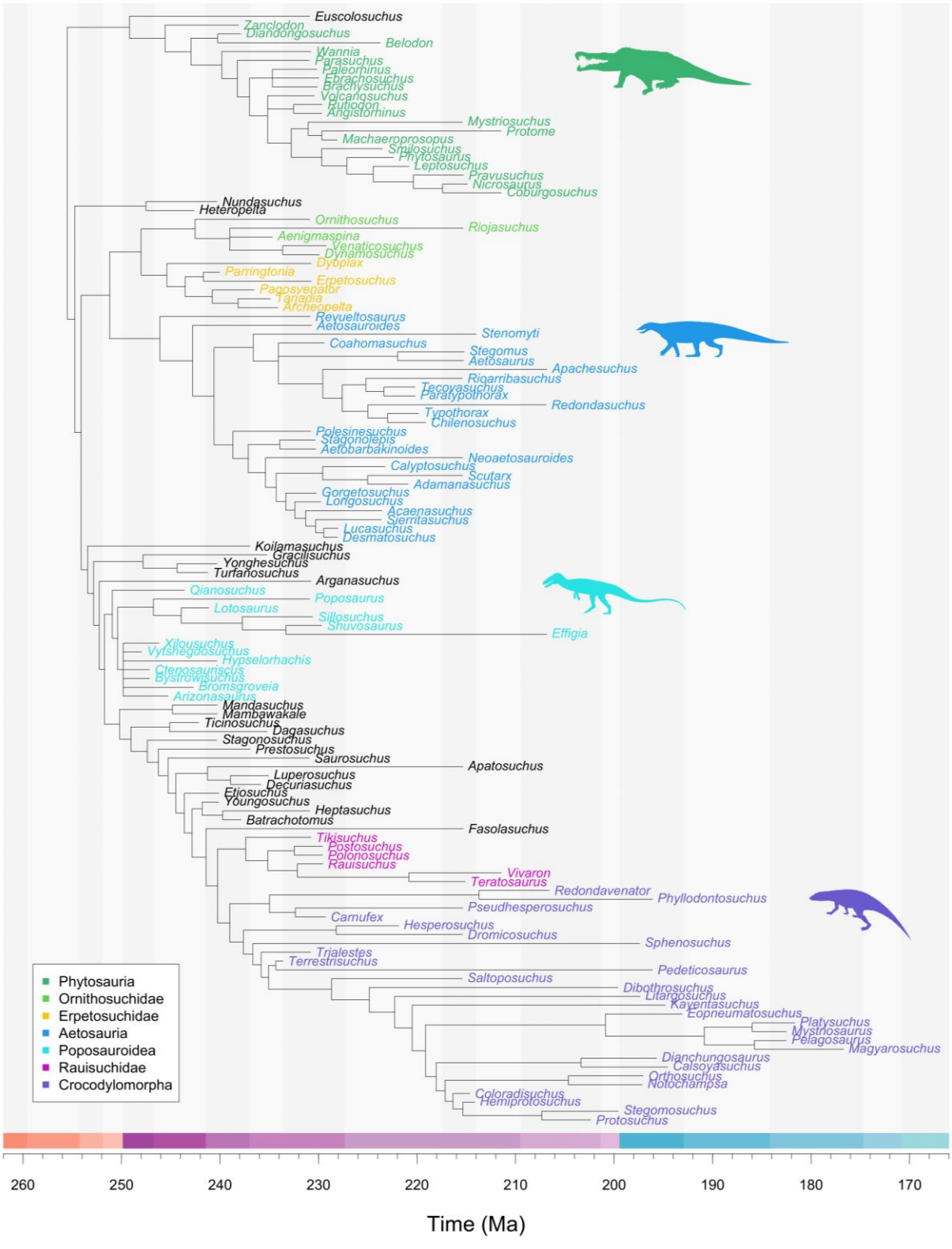


Fig. 5.8. Early pseudosuchian phylogeny. Part of a time-scaled Permian–Triassic archosauromorph phylogeny from MrBayes. The suchian crown-group is incompletely represented, containing only Triassic lineages plus a few Jurassic lineages selected to improve estimates of ancestral spatiotemporal origins. Grey bars demarcate Permian and Triassic stages, with subdivisions of the Carnian and Norian.

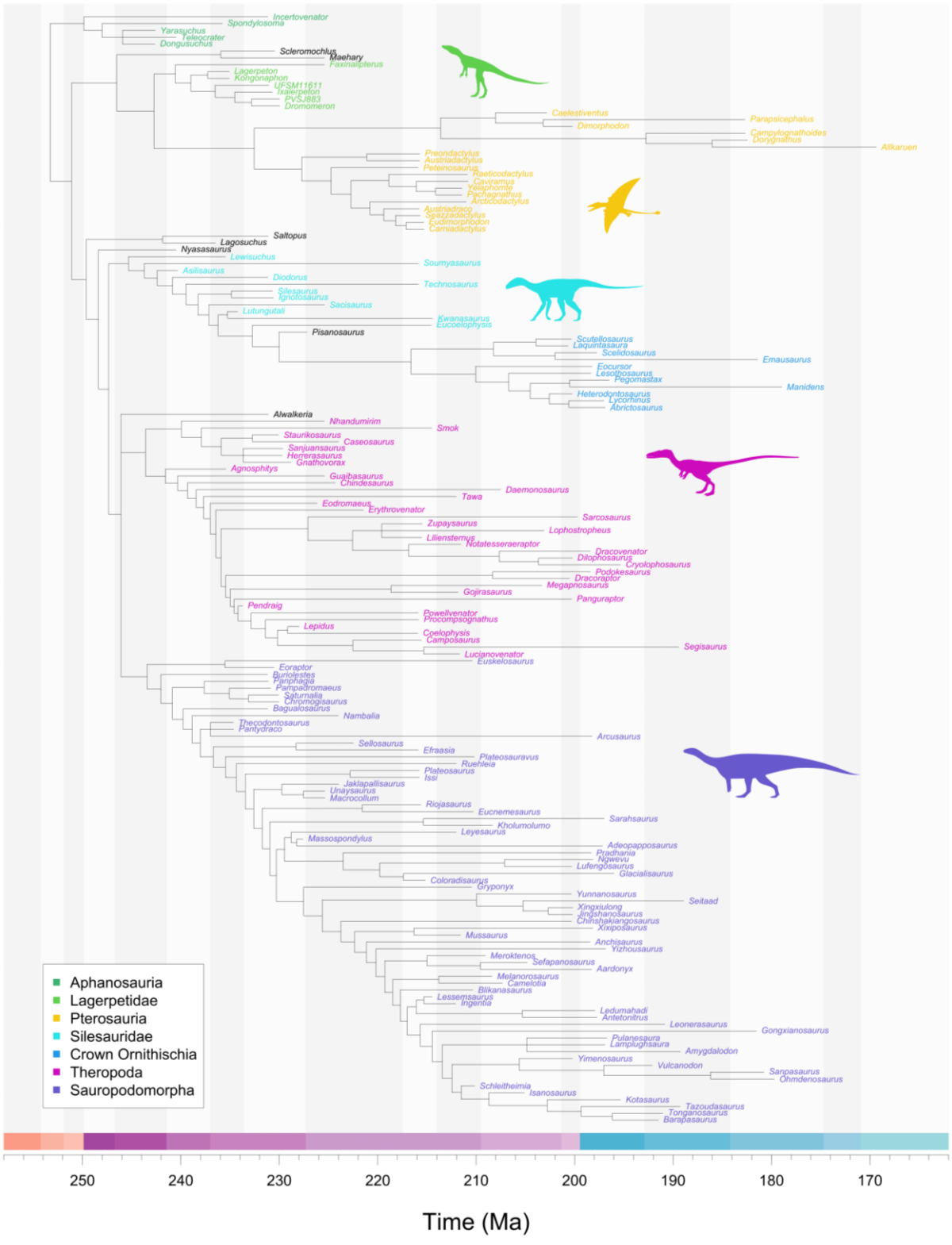


Fig. 5.9. Early avemetatarsalian phylogeny. Part of a time-scaled Permian–Triassic archosauromorph phylogeny from MrBayes. Pterosaurs and dinosaurs are incompletely represented, containing only Triassic lineages plus a few Jurassic lineages selected to improve estimates of ancestral spatiotemporal origins. Grey bars demarcate Permian and Triassic stages, with subdivisions of the Carnian and Norian.

5.3.2 Archosauromorph origins in space and time

The divergence times of higher archosauromorph clades frequently predate their oldest representative fossils by tens of millions of years, resulting in the steady acquisition of cryptic diversity from the latest Early Permian to middle Late Permian within the European portion of northern Pangaea between 30–40º latitude (Fig. 5.10). Archosauromorphs emerged in the latest Kungurian, followed by the divergence of crocopodans in the late Roadian, archosauriforms in the latest Wordian and archosaurs in the latest Wuchiapingian (Fig. 5.10). Smaller, basal archosauromorph and archosauriform groups conform to this pattern, including proterosaurids, tanystropheids, proterosuchids and euparkeriids, whose younger fossil records show significantly broader geographic distributions spanning northern Pangaea and southern Gondwana (Fig. 5.11). Erythrosuchids had a similarly broad distribution, but their point of

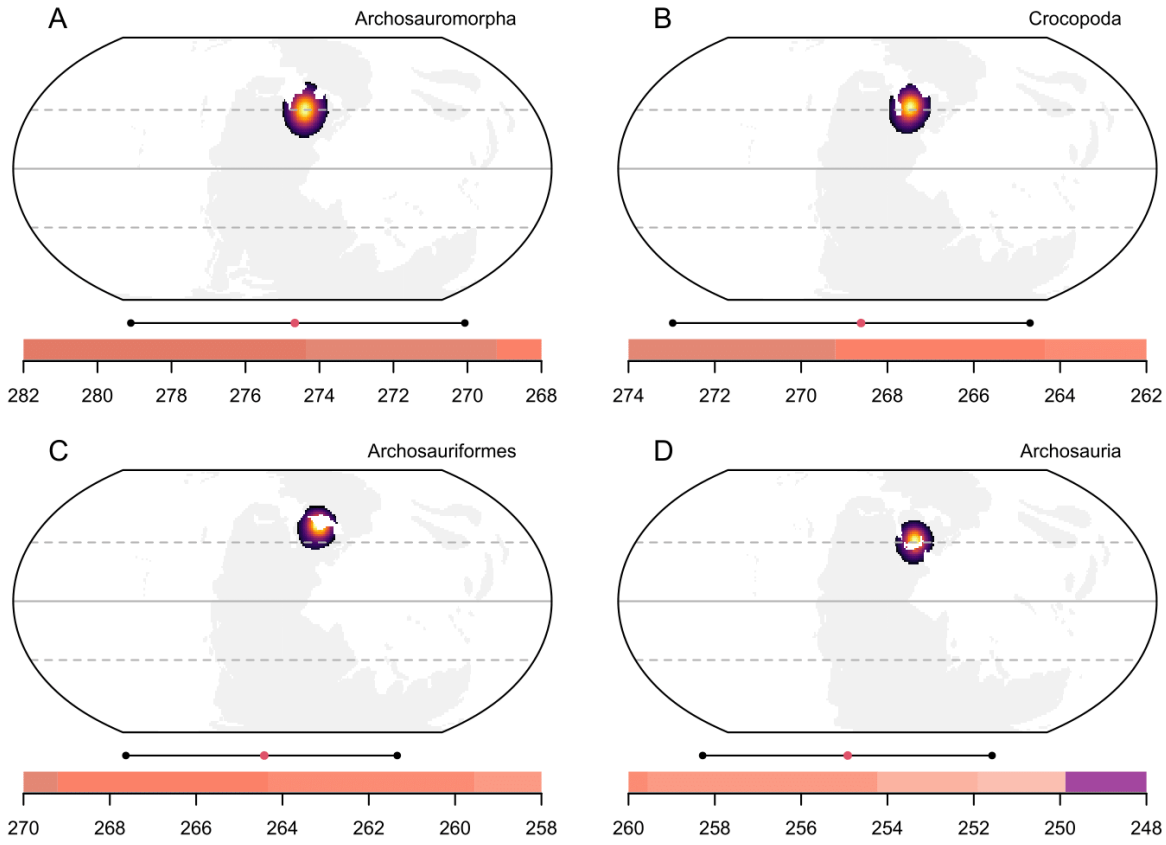


Fig. 5.10. Spatiotemporal origins of higher level archosauromorph clades. Mean divergence times and their highest posterior densities, along with kernel density estimates of their spatial points of origin at their mean divergence time. Coloured ribbons mark the stages of the Permian and Triassic. Map lines demarcate the equator and tropics. Notably, all divergence times are gradual, but remain constrained to northern Pangaea. **(A).** Archosauromorpha. **(B).** Crocopoda. **(C).** Archosauriformes. **(D).** Archosauria.

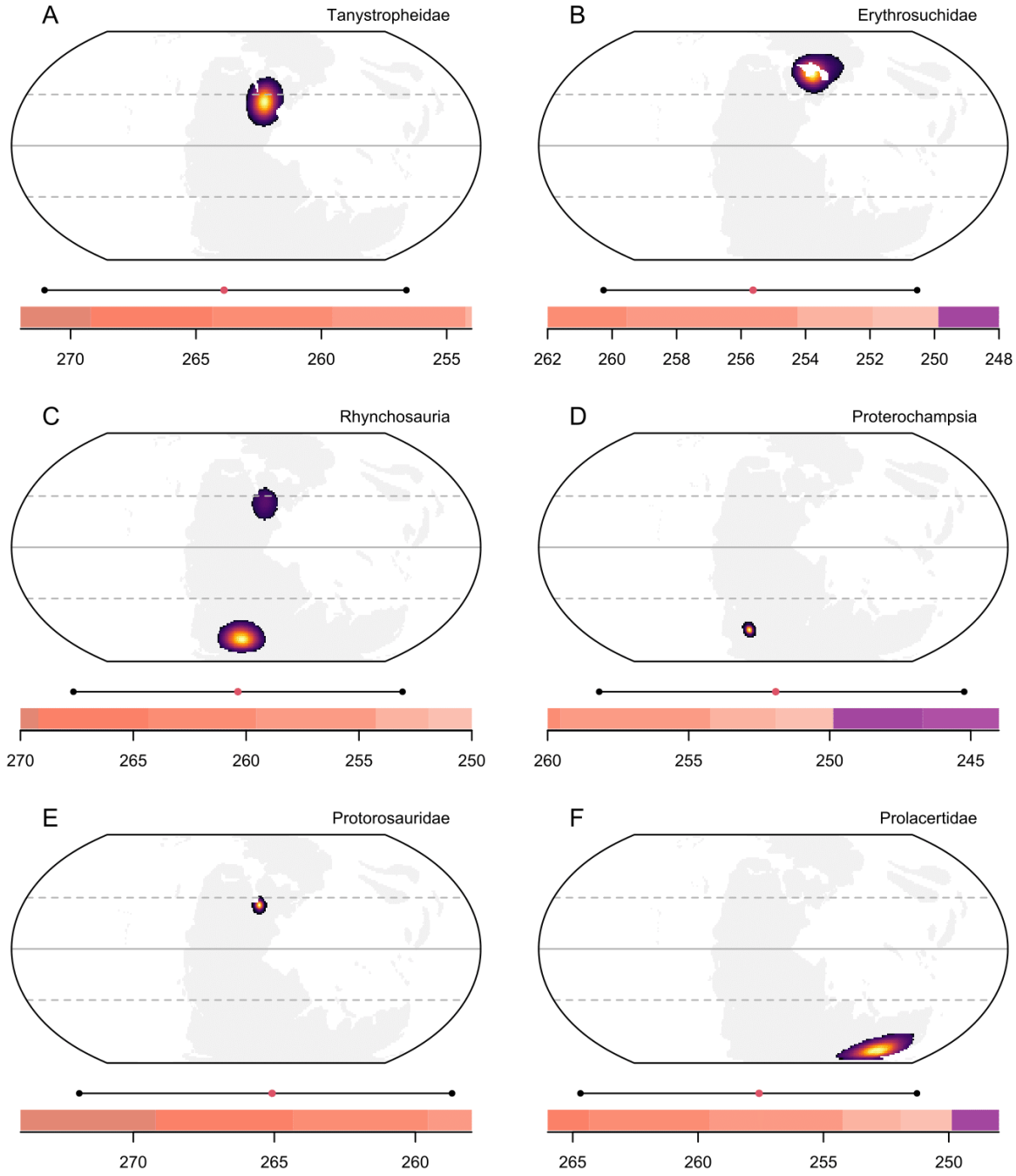


Fig. 5.11. Spatiotemporal origins of early-diverging archosauromorph clades. Mean divergence times and their highest posterior densities, along with kernel density estimates of their spatial points of origin at their mean divergence time. Coloured ribbons mark the stages of the Permian and Triassic. Map lines demarcate the equator and tropics. All divergences show substantial temporal uncertainty but demonstrate the high phylogeographic diversity of early archosauromorphs. **(A).** Tanystropheidae. **(B).** Erythrosuchidae. **(C).** Rhynchosauria. **(D).** Proterochampsia. **(E).** Protorosauridae. **(F).** Prolacertidae

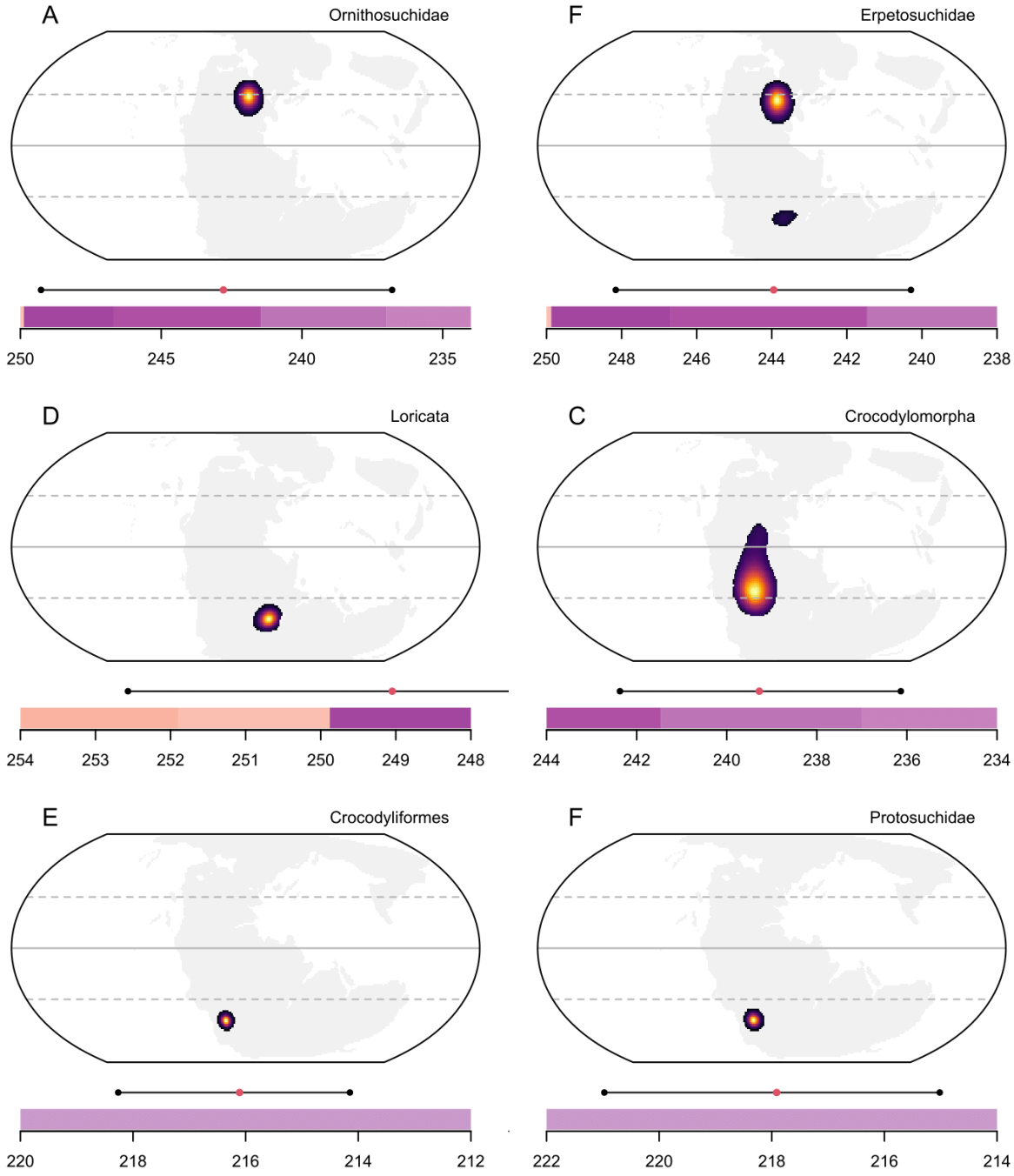


Fig. 5.12. Spatiotemporal origins of early pseudosuchian clades. Mean divergence times and their highest posterior densities, along with kernel density estimates of their spatial points of origin at their mean divergence time. Coloured ribbons mark the stages of the Permian and Triassic. Map lines demarcate the equator and tropics. **(A).** Pseudosuchia. **(B).** Suchia. **(C).** Poposauroidea. **(D).** Phytosauria. **(E).** Aetosauria. **(F).** Rauisuchidae.

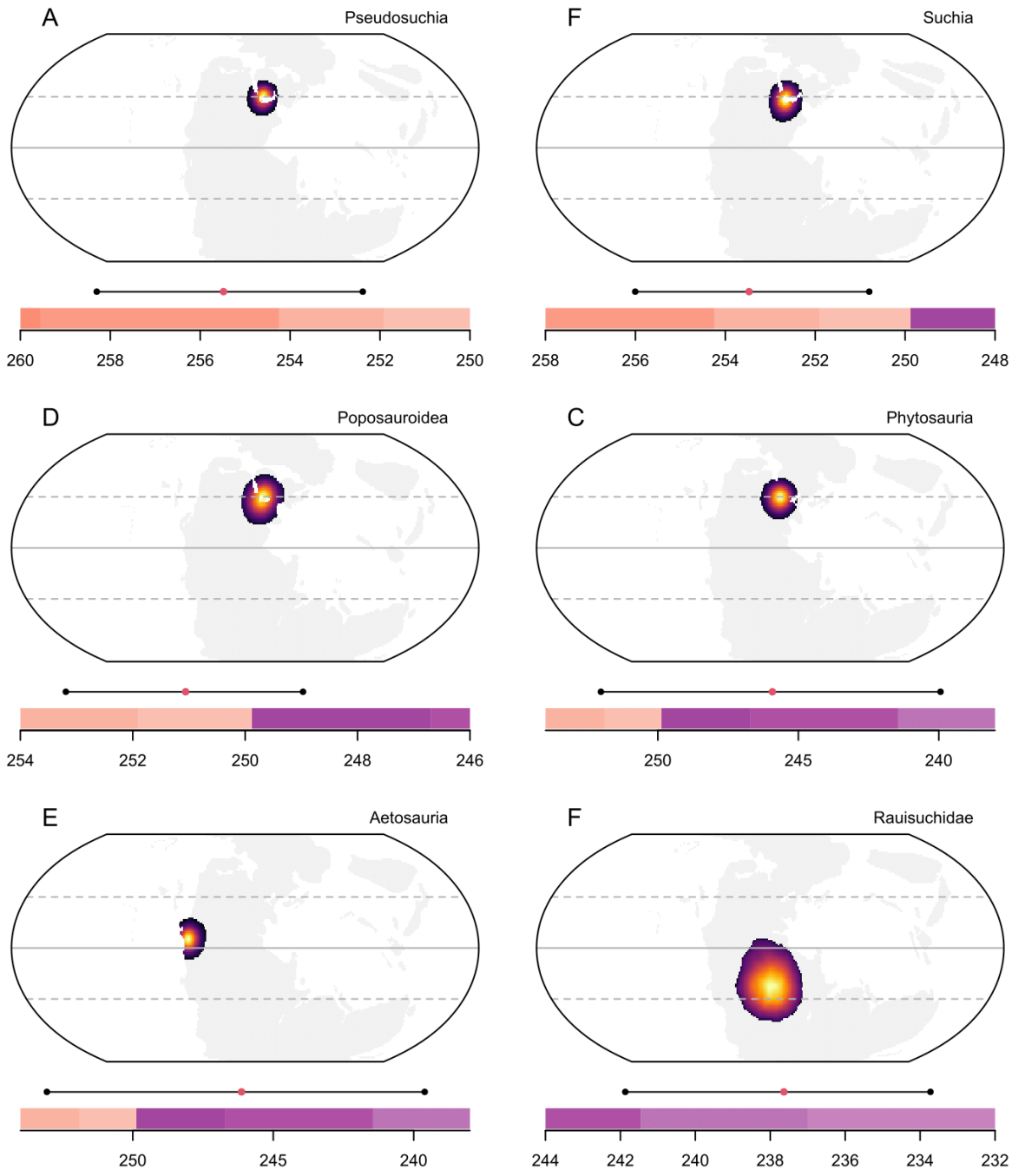


Fig. 5.13. Spatiotemporal origins of later pseudosuchian clades. Mean divergence times and their highest posterior densities, along with kernel density estimates of their spatial points of origin at their mean divergence time. Coloured ribbons mark the stages of the Permian and Triassic. Map lines demarcate the equator and tropics. **(A).** Ornithosuchidae. **(B).** Erpetosuchidae. **(C).** Loricata. **(D).** Crocodylomorpha. **(E).** Crocodyliformes. **(F).** Protosuchidae.

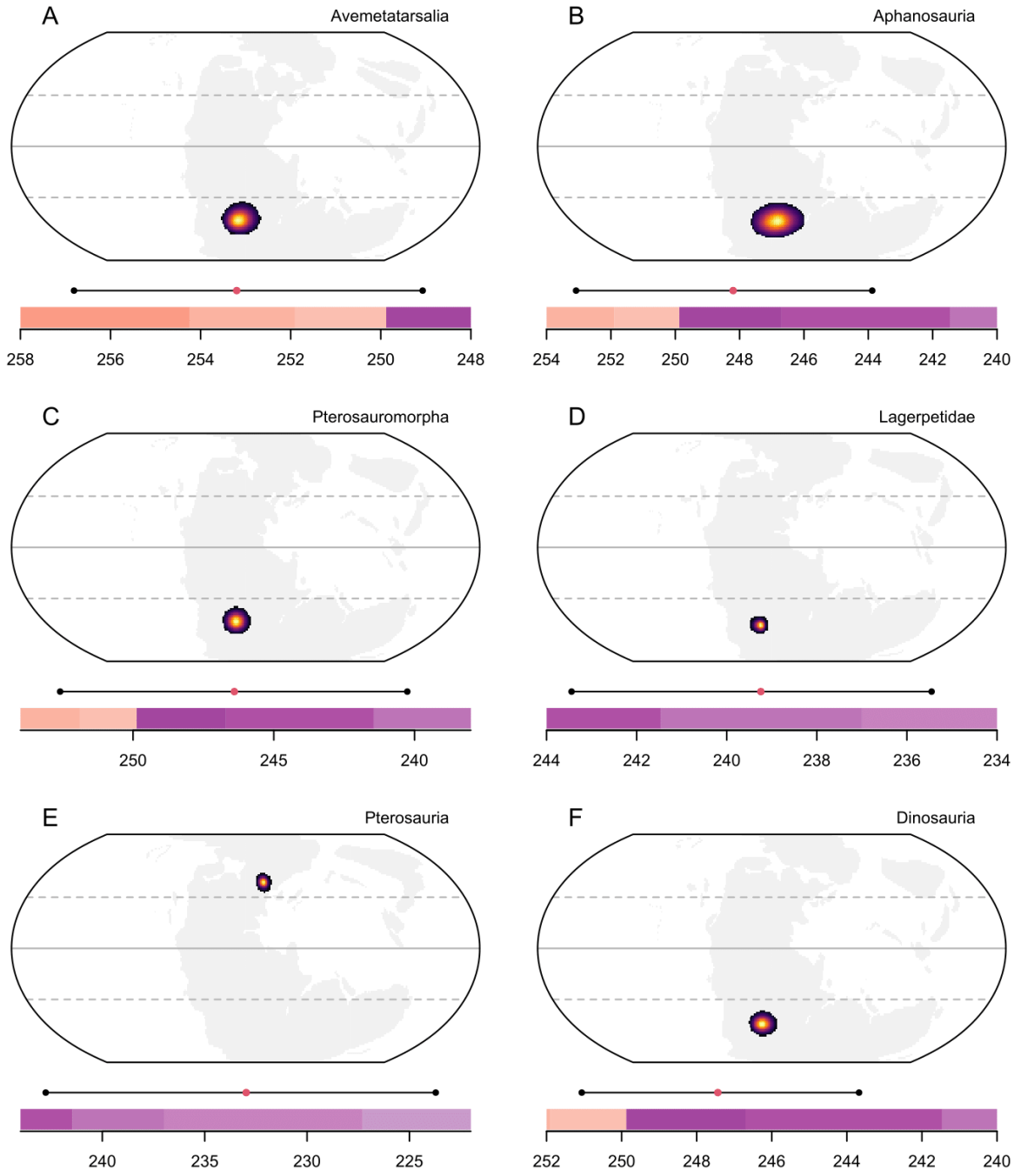


Fig. 5.14. Spatiotemporal origins of early avemetatarsalian clades. Mean divergence times and their highest posterior densities, along with kernel density estimates of their spatial points of origin at their mean divergence time. Coloured ribbons mark the stages of the Permian and Triassic. Map lines demarcate the equator and tropics. **(A).** Avemetatarsalia. **(B).** Aphanosauria. **(C).** Pterosauromorpha. **(D).** Lagerpetidae. **(E).** Pterosauria. **(F).** Dinosauria.

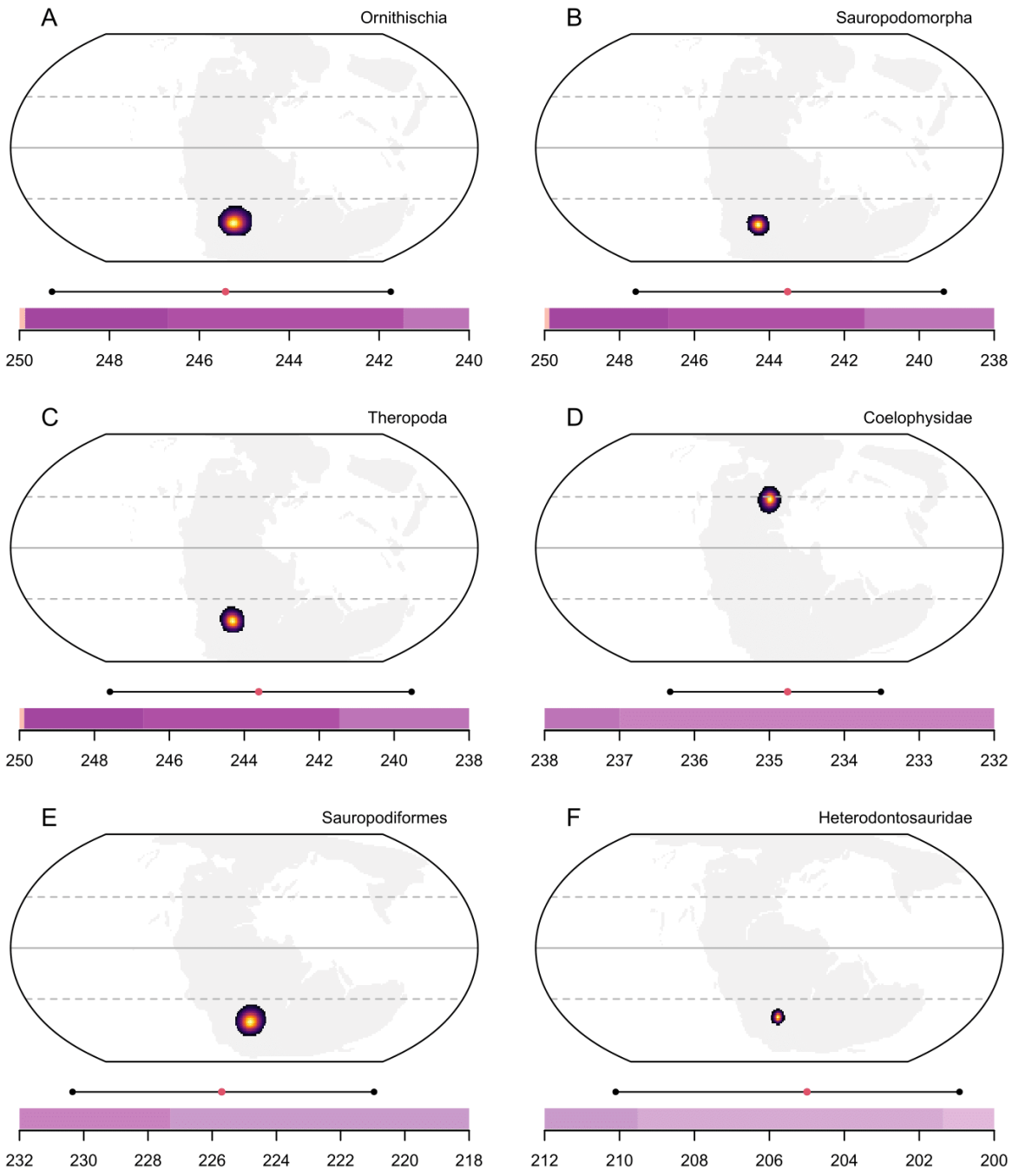


Fig. 5.15. Spatiotemporal origins of later avemetatarsalian clades. Mean divergence times and their highest posterior densities, along with kernel density estimates of their spatial points of origin at their mean divergence time. Coloured ribbons mark the stages of the Permian and Triassic. Map lines demarcate the equator and tropics. **(A).** Ornithischia. **(B).** Sauropodomorpha. **(C).** Theropoda. **(D).** Coelophysidae. **(E).** Sauropodiformes. **(F).** Heterodontosauridae.

origin was still relatively close to the European centre of early archosauromorph divergences (Fig. 5.11). There were nonetheless some exceptions to this general trend of Permian-aged divergences in northern Pangaea. Prolacertids appeared in far eastern Gondwana in the early Wuchiapingian (Fig. 5.11) despite their sister and ancestral nodes originating in northern Pangaea, and proterochampsians appeared in the Late Permian to Middle Triassic with a tightly constrained point of origin in western Gondwana. Even though their crocopodan ancestors also originated in northern Pangaea in the late Wordian, rhynchosaurs show a bimodal estimate split between northern Pangaea and southwestern Gondwana in the late Capitanian, although the latter region is heavily favoured.

Subject to alternative placements of lagerpetids and silesaurids, archosaur origins are well constrained to the northern Pangaean cradle of the early archosauromorph radiation. Pseudosuchians also originated in the late Wuchiapingian in northern Pangaea, but their subsequent divergences towards the crocodylomorph crown group are more geographically varied (Fig. 5.12). Suchians originated in the same location just a few million years later and early diverging pseudosuchian clades conform to this ancestral pattern, including phytosaurs, poposauroids and ornithosuchids. Conversely, the ancestors of other basal pseudosuchian clades dispersed away from Northern Pangaea through the Changhsingian to Ladinian, with aetosaurs originating in North American Pangaea and rauisuchids in central Pangaea, while erpetosuchids display a greater degree of uncertainty in their geographic origin (Fig. 5.13). There was then a notable southward shift in the main theatre of crownward pseudosuchian divergences, encompassing loricatans in the early Olenekian, crocodylomorphs in the Ladinian, and crocodyliforms in the middle Norian (Fig. 5.13).

In contrast to pseudosuchians, avemetatarsalians emerged in western Gondwana in the Changhsingian, with their higher-level divergences taking place in relatively rapid succession from the Early to Middle Triassic within this same ancestral hub (Fig. 5.14), including aphanosaurs, ornithodirans, dinosaurs, and lagerpetids, while pterosaurs originated nearly five million years later and over ten thousand kilometres away in Northern Pangaea during the Carnian. Higher dinosaurian divergences (Ornithischia, Theropoda, Sauropodomorpha) also took place in western Gondwana (Fig. 5.15), with successive sauropodomorph clades (guaibasaurids, massospondylids, sauropodiformes, anchisaurians) continuing to emerge in here through the late Ladinian to late Norian. Conversely, the stage of theropod diversification rapidly shifted in the early Carnian from western Gondwana, where the basal herrerasaurids

originated, to tropical northern Pangaea with the emergence of coelophysoids and neotheropods (Fig. 5.15). Finally, divergences through the basal silesaurid grade of ornithischians all took place in mid latitude western Gondwana from the middle Anisian to middle Ladinian. There is still a substantial gap in time to the emergence of heterodontosaurids and genasaurians in the Rhaetian, with the former displaying a precise point of origination in the same region as basal ornithischians, compared to the broader tropical origin displayed by the latter (Fig. 5.6, 5.15).

5.3.3 Climatic disparity through early archosauromorph phylogeny

Patterns of climatic disparity based on geographic origins versus ancestral state estimates are broadly concordant across multiple metrics (Fig. 5.16). Disparity increased steadily from the Kungurian to the Wuchiapingian as ancestral archosauromorphs ventured out into new regions of climate space (Fig. 5.17), before a small decline during the Changhsingian, a sharp increase in the Induan and a correspondingly sharp fall in the Olenekian (Fig. 5.16A, 5.16B, 5.16C). Climatic disparity from range-based metrics was highest during the Middle Triassic (Fig. 5.16A), matched by the most extensive exploration of climate space by early archosauromorphs (Fig. 5.17B), while the sum of variances and mean distance from the centroid record a slight decline indicative of shift in average taxon positions towards the centre of climate space (Fig. 5.16B, 5.16C). All metrics indicate a decrease in disparity in the early Carnian followed by a sharp spike in the late Carnian, coupled with a distinct shift in climate space occupation away from wetter, more thermally seasonal conditions towards drier conditions with greater precipitation seasonality (Fig. 5.17C). Disparity then showed a general decline through the Norian and a final spike in the Rhaetian, with the bounds of occupied climate space remaining relatively stable.

Disparity metrics show varying degrees of sensitivity to sample size and the presence of outliers (Ciampagliao et al., 2001), while sampled tips tend to show a greater degree of dispersion in these spaces compared to their reconstructed ancestors (Brusatte et al., 2011). The numbers of tips and nodes in the early archosauromorph phylogeny considered here are roughly equal so there is consequently a degree of difference in the resulting disparity patterns when the potential impact of sporadic tip sampling is reduced by examining ancestral climate space occupation, as well as between the node-only patterns from geographic origin versus ancestral state estimates. Range-based metrics still support a Kungurian-Wuchiapingian rise, a Changhsingian decline, and the Late Triassic trends presented above, but fluctuation through

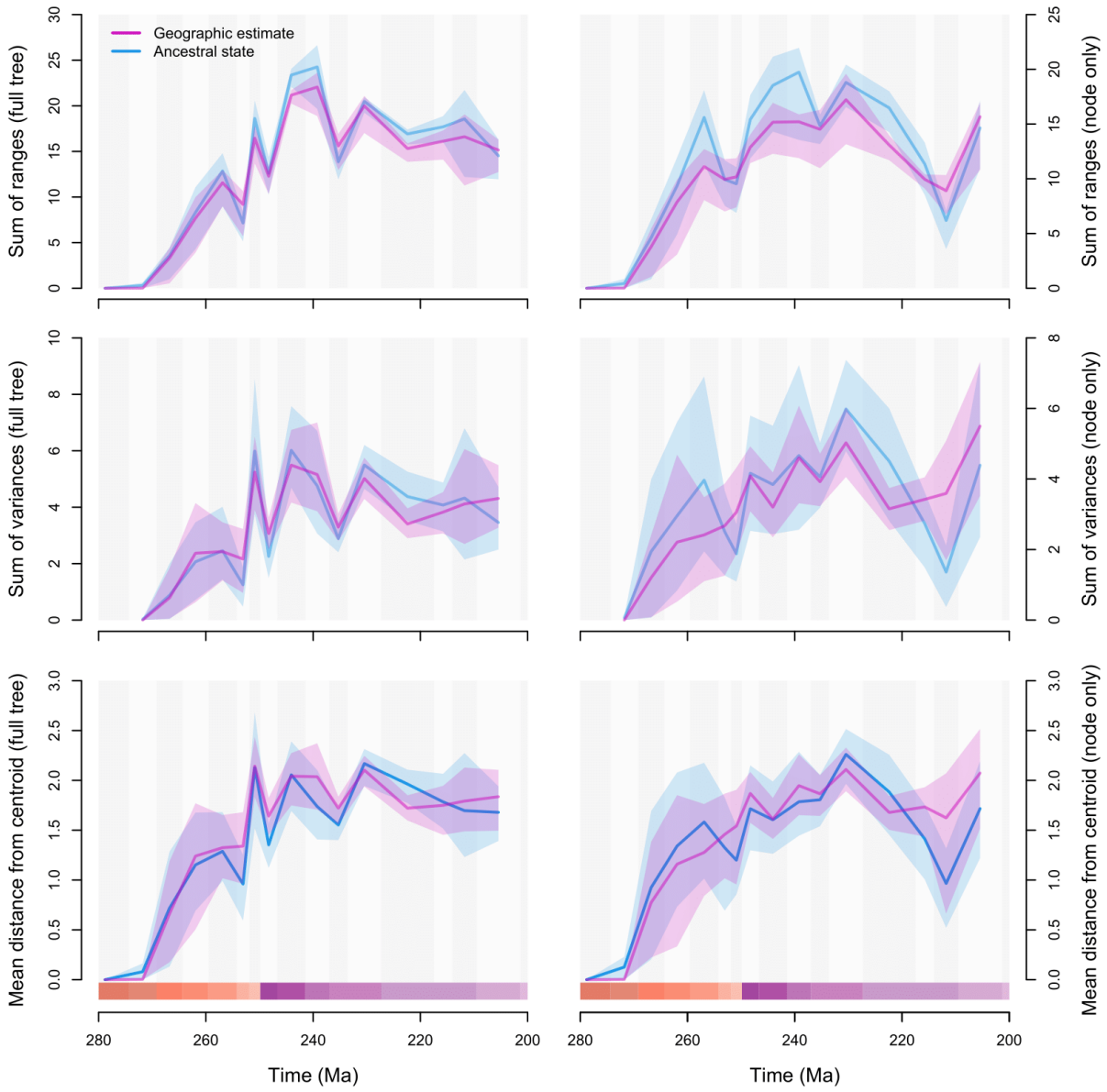


Fig. 5.16. Climatic disparity of Late Permian to Late Triassic archosauromorphs. Climatic disparity through time measured from an ordination space of node and tip climatic tolerances. The mean is displayed for each metric, along with 95% bootstrapped confidence intervals. Grey bands mark the stages of the Permian and Triassic, along with Carnian and Norian subdivisions. **(A).** Sum of ranges from node and tip states. **(B).** Sum of variances from node and tip states. **(C).** Mean distance from the centroid from node and tip states. **(D).** Sum of ranges from nodes only. **(E).** Sum of variances from nodes only. **(F).** Mean distance from the centroid from nodes only.

the Early Triassic is replaced by a second phase of steady increase that continued into the Middle Triassic. The sum of variances and mean distance from the centroid broadly recover these revised trends, although the Middle Triassic rise is more modest. When the pattern from geographic origin estimates is examined across all metrics, there is no decline during the

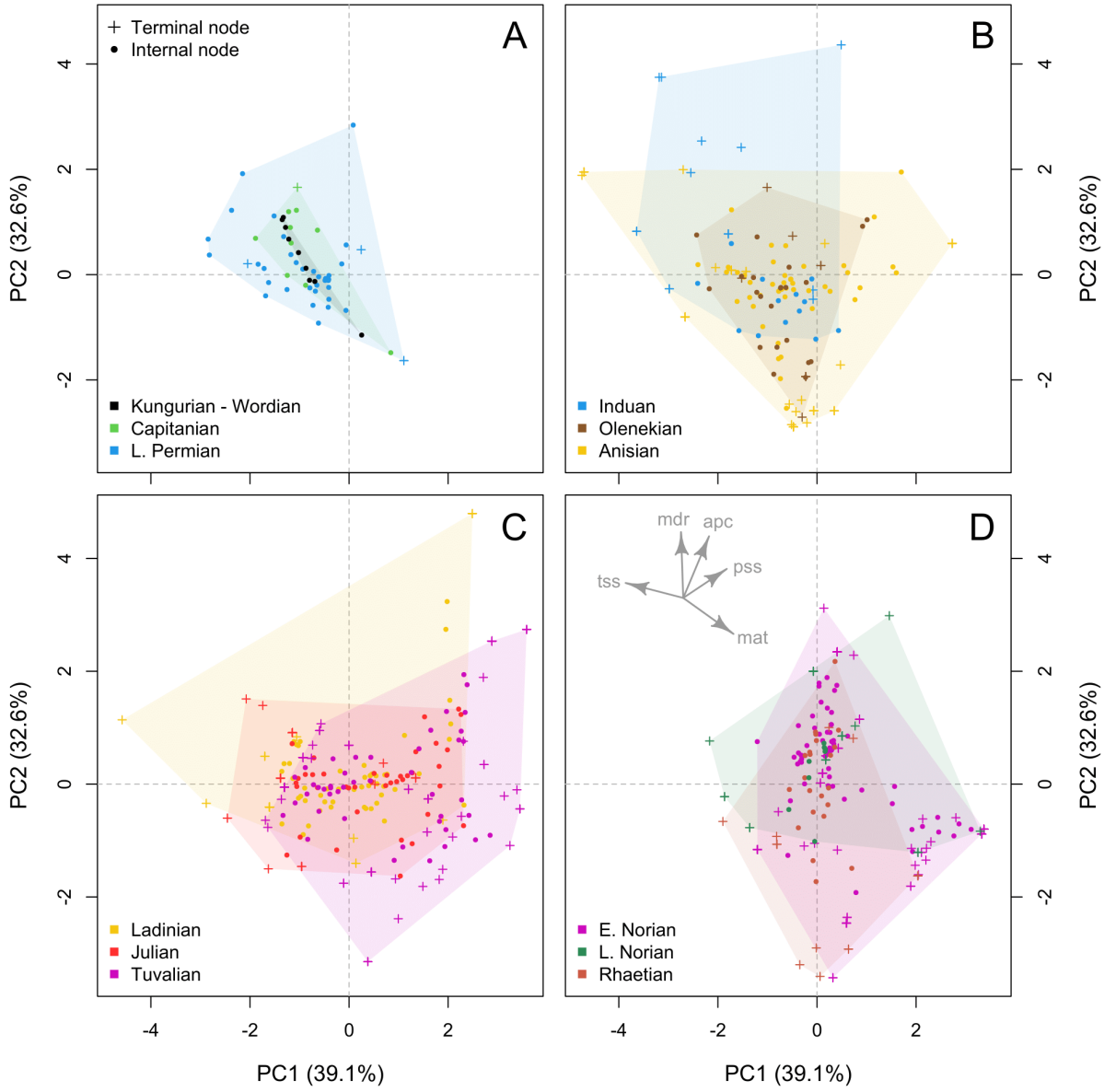


Fig. 5.17. Climate space occupancy of Late Permian to Late Triassic archosauromorphs. Ordination spaces using the first two principal components of a PCA analyses of archosauromorph tip state and ancestral state estimates for five selected climatic variables: temperature seasonality (tss), mean diurnal temperature range (mdr), annual precipitation (apc), precipitation seasonality (pss) and mean annual temperature (mat). **(A).** Permian. **(B).** Early Triassic to Anisian. **(C).** Ladinian to late Carnian. **(D).** Remainder of the Late Triassic.

Changhsingian, but instead a stepwise pattern of increasing climate disparity from the Kungurian to late Carnian, followed by the previously observed trends through the remainder of the Triassic.

5.3.4 Taxonomic patterns of climate space occupancy

While there is a general trend of increasing climatic disparity from the Kungurian to the Carnian, punctuated by a marked shift in overall climate space occupation from the early to the late Carnian, and followed by Norian decline and Rhaetian resurgence, these patterns are not taxonomically pervasive. The earliest diverging archosauriforms (i.e., stem archosaurs) drove the steady expansion of occupied climate space through the Permian (Fig. 5.17D), prior to a dramatic shift into cooler, wetter and more thermally seasonal environments driven by the sudden emergence of a diverse range of fossil tips in the Induan (Fig. 5.17E). A greater variety of early-diverging archosauromorph clades contributed to a second burst of widening climate space occupation during the Middle Triassic. Basal archosauriformes continued their expansion into cooler and wetter environments, while tanystropheids and rhynchosaurs radiated from their initially conservative region of ancestral archosauromorph climate space into warmer, drier environments during the Anisian (Fig. 5.17A, 5.17C). Rhynchosaur climate space occupation then declined through the Ladinian and early Carnian prior to their total extinction in the early Norian. Tanystropheids and stem archosaurs maintained broader climatic disparities, but their patterns of occupation fluctuated unpredictably through the Ladinian and early Carnian. Tanystropheids played only a minor role in the shift towards drier, more pluvially seasonal conditions seen at broader taxonomic scales, corresponding with their diminished taxonomic and climatic diversity through the Norian. Rhynchosaurs and basal archosauriformes were more successful in shifting their climate space occupation during the late Carnian but were all eliminated within ten million years.

As the taxonomic and climatic diversity of basal archosauromorph clades waned, the archosaur crown groups steadily expanded (Fig. 5.18, 5.19). Pseudosuchians and avemetatarsalians were initially conservative in their climate space occupancy, falling within the bounds of basal archosauromorph space more generally. Pseudosuchians then undertook wider dispersals into a range of different climatic conditions through the Middle Triassic compared to their sister clade, and notably were the primary drivers of the tree-wide shift towards drier, more pluvially seasonal conditions (Fig. 5.18). There is still taxonomic nuance to this shift, however, with phytosaurs displaying an early trend in the Ladinian to late Carnian towards warmer climates with higher rainfall (Fig. 5.18A), aetosaurs radiating from the early to late Carnian into warmer environments across a range of aridities (Fig. 5.18B), and the

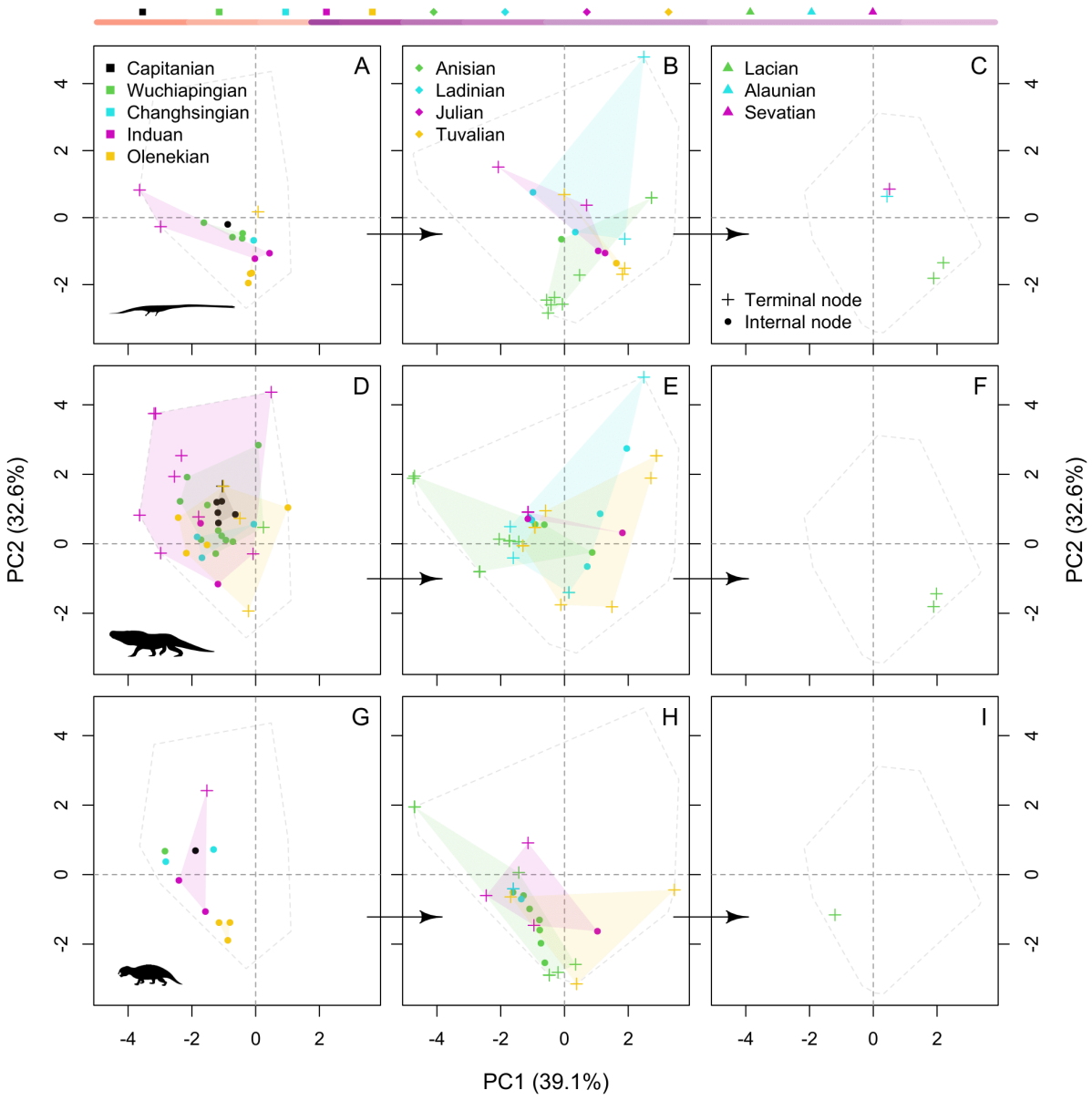


Fig. 5.18. Climate space occupancy of early archosauromorphs. Ordination spaces subsampled from the first two principal components of a PCA analyses of archosauromorph tip state and ancestral state estimates for five selected climatic variables (see Fig. 5.16 for variables and loadings). **(A-C).** Tanystropheids. **(D-F).** Basal archosauriformes. **(G-I).** Rhynchosaurs.

crocodylomorph crown group into the warmest and driest regions of climate space (Fig. 5.18C). Phytosaur climate space then diminished gradually through the Norian. Crocodylomorphs show no clear pattern by contrast, with wider occupancies in the middle Norian and Rhaetian, while aetosaurs enjoyed a renewed burst of climatic disparity during the middle Norian prior to their final decline through the remainder of the Triassic.

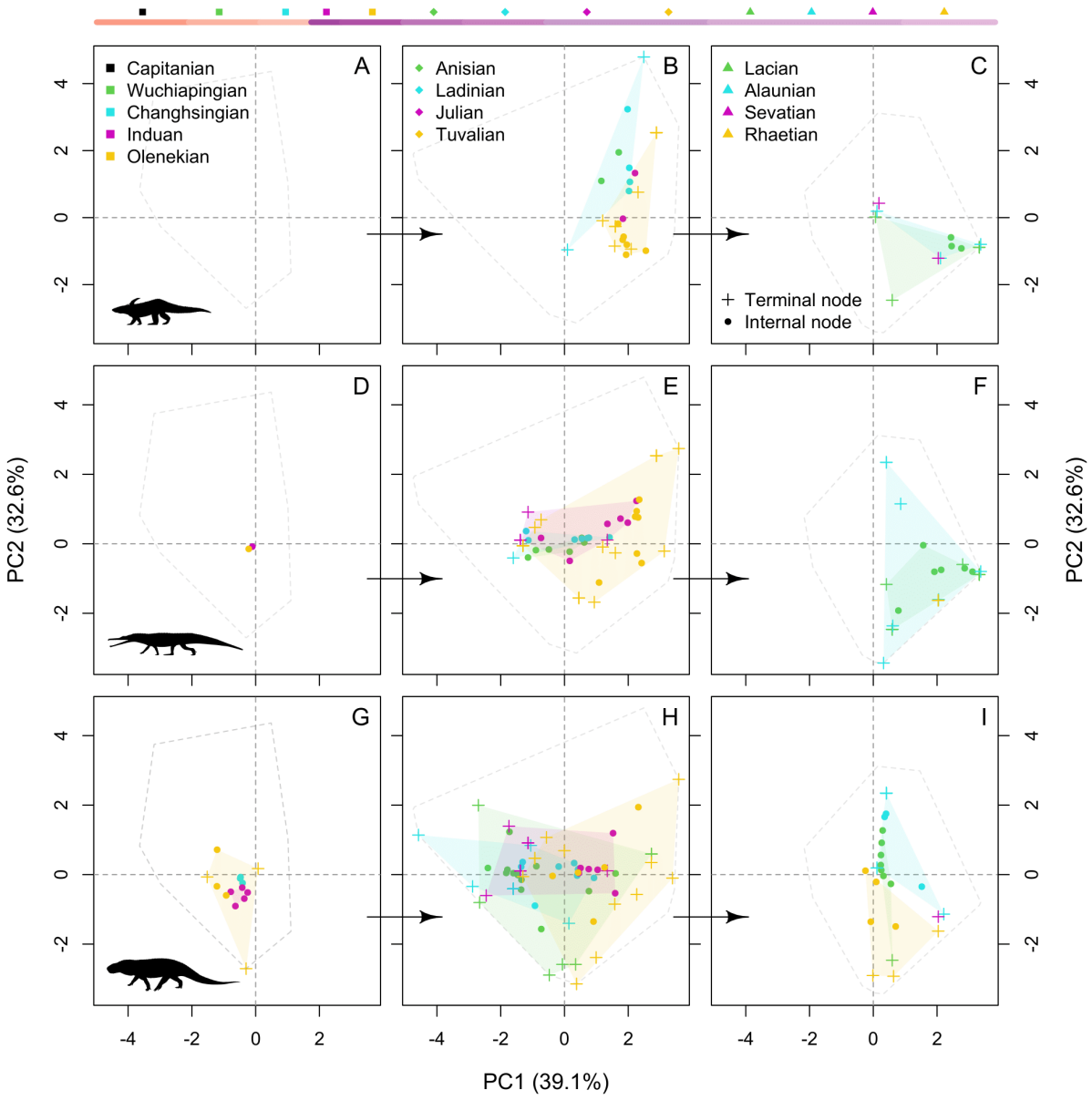


Fig. 5.19. Climate space occupancy of pseudosuchian clades. Ordination spaces subsampled from the first two principal components of a PCA analyses of archosauromorph tip state and ancestral state estimates for five selected climatic variables (see Fig. 5.16 for variables and loadings). **(A-C).** Phytosaurs. **(D-F).** Aetosaurs. **(G-H).** Crocodylomorphs.

Within the avemetatarsalian line, pterosauromorphs contributed only marginally to the broad shift in climate space occupation and otherwise show no clear patterns through the Triassic (Fig. 5.19A). Sauropodomorphs and theropods showed opposing trends through the Carnian with the former diversifying under wetter and cooler conditions, while the latter expanded into environments with overall greater, but still strongly seasonal precipitation. Ornithischians, when silesaurids are considered as a basal grade, initially occupied drier

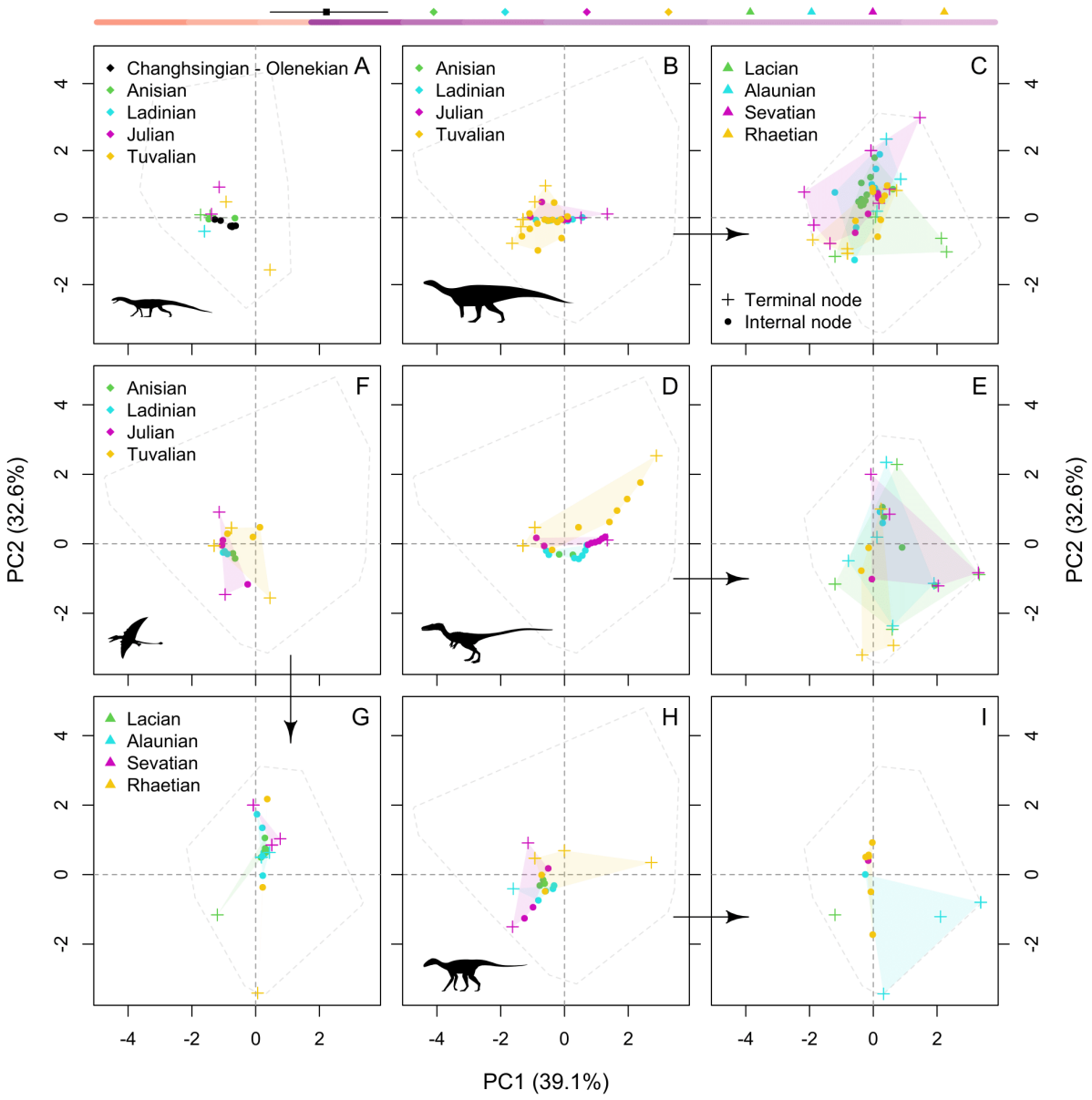


Fig. 5.20. Climate space occupancy of avemetatarsalian clades. Ordination spaces subsampled from the first two principal components of a PCA analyses of archosauromorph tip state and ancestral state estimates for five selected climatic variables (see Fig. 5.16 for variables and loadings). **(A).** Basal avemetatarsalians. **(B-C).** Sauropodomorphs. **(D-E).** Theropods. **(F-G).** Pterosauromorphs. **(H-I).** Ornithischians.

environments before shifting into slightly warmer and wetter conditions in the late Carnian. Major shifts in avemetatarsalian climate space instead took place in the Norian, with sauropods generally remaining restricted to cooler and wetter environments, while theropods followed in the wake of crocodylomorphs into drier conditions, but across a wider range of precipitation intensities. Intriguingly, this expansion was matched in the early Norian by ornithischians,

although the otherwise scant early fossil record of their crown group likely corresponds with their seemingly diminished climate space occupation for the remainder of the Triassic.

5.3.5 Phylogeographic routes and evolutionary rates

The geometries of early archosauromorph phylogeographic least-cost pathways highlight the tight spatial distribution of their oldest divergences and dispersals (Kungurian to Wordian) within their northern Pangaean cradle (Fig. 5.20). Long distance cryptic dispersals first took place in the Capitanian and continued through the remainder of the Permian in concert with expanding climate space occupation driven by early archosauromorph divergences (Fig. 5.16). These pathways generally took conservative routes that hugged the western Tethys coastline, but some archosauromorphs notably began crossing the northern branch of the Central Pangaean Mountains into North America in the Capitanian, a dispersal predating the oldest known archosauromorph taxon on the continent, the ctenosauriscid *Arizonasaurus babitti*, by nearly 15 million years and their oldest footprint records by almost the same degree (Bernardi et al., 2015).

Despite hostile conditions across terrestrial ecosystems in the aftermath of the EPME, transcontinental traversals continued through the Early Triassic, establishing a second centre of archosauromorph radiation in southern Gondwana situated south of the arid belt spanning the austral horse latitudes (Fig. 5.20, 5.21). During the Anisian, exchanges continued between the northern and southern hemispheres through the corridor situated between the Central Pangaean Mountains and the western Tethys coast, alongside the first transoceanic traversals into East Asia. The Ladinian and Carnian, however, was the most intense interval of phylogeographic exchange between the European, North American and southern Gondwanan centres of the archosauromorph radiation (Fig. 5.21). Traversals were less frequent during the Norian and Rhaetian and largely utilised the coastal corridors flanking the eastern and western margins of the Central Pangaean mountains, although some routes did cut through the more arid continental interiors during this time. Traversals also increased in frequency through the Central Pangaean Mountains themselves during the Late Triassic, following increasingly accessible intermontane biogeographic highways resulting from the incipient rifting of Pangaea (Fig. 5.21).

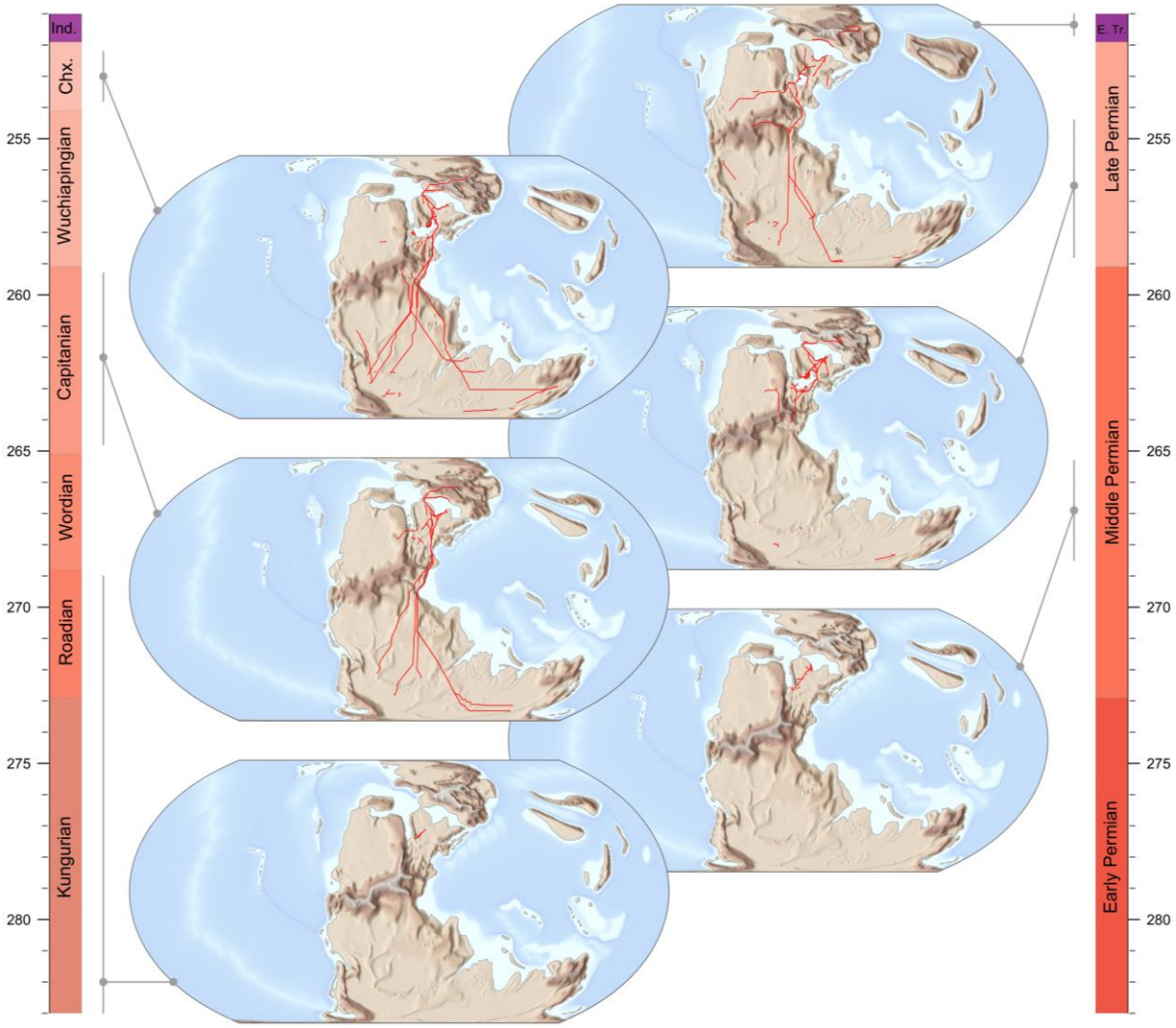


Fig. 5.21. Archosauromorph phylogeographic pathways from the Kungurian to Induan. Phylogeographic pathways derived from TARDIS for the earliest phases of the archosauromorph radiation across Pangaea. Pathways are superimposed onto false colour height maps of continental topography.

Phylogeographic rates of dispersal vary by several orders of magnitude with the majority in the range of 500 to 2000 kilometres per million years, but no statistically significant shifts in dispersal rate through their phylogeny were detected using phylogenetic ridge regression. Shifts in the evolutionary rate of climate tolerance adaptation were similarly sparse, with significantly slower evolutionary rates detected at the root of avemetatarsalians for all tested climatic parameters (mean annual temperature, temperature seasonality, mean diurnal temperature range, annual precipitation and precipitation seasonality), and significantly faster rates for annual precipitation in aetosaurs. While niche adaptation and dispersal display very

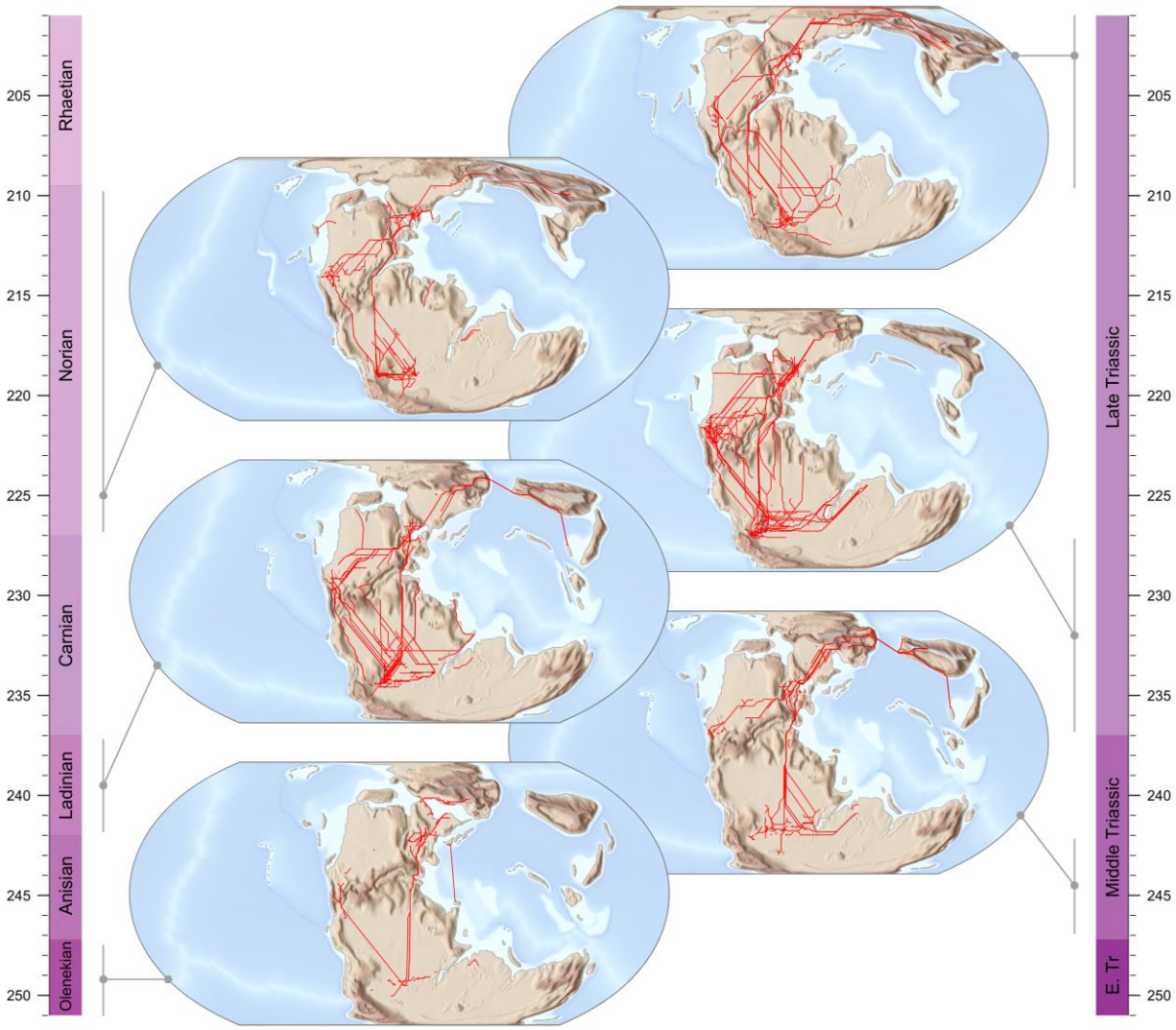


Fig. 5.22. Archosauromorph phylogeographic pathways from the Olenekian to Rhaetian. Phylogeographic pathways derived from TARDIS for the later phases of the archosauromorph radiation across Pangaea. Pathways are superimposed onto false colour height maps of continental topography.

few rate shifts, however, this does not mean that archosauromorphs did not encounter and overcome varied environmental conditions during their diaspora. Projections of phylogeographic pathways highlight wider occupancy of climate space by archosauromorphs than their empirical and ancestral states alone would demonstrate (Fig. 5.22, 5.23). During their constrained phase in the Kungurian to Wordian, explorations of climate space were not extensive (Fig. 5.22), until increased dispersal in the Capitanian resulted in traverses through environments with high temperature seasonality but low annual precipitation. Through the Late Permian to Early Triassic, archosauromorphs then traversed through climatic conditions nearly

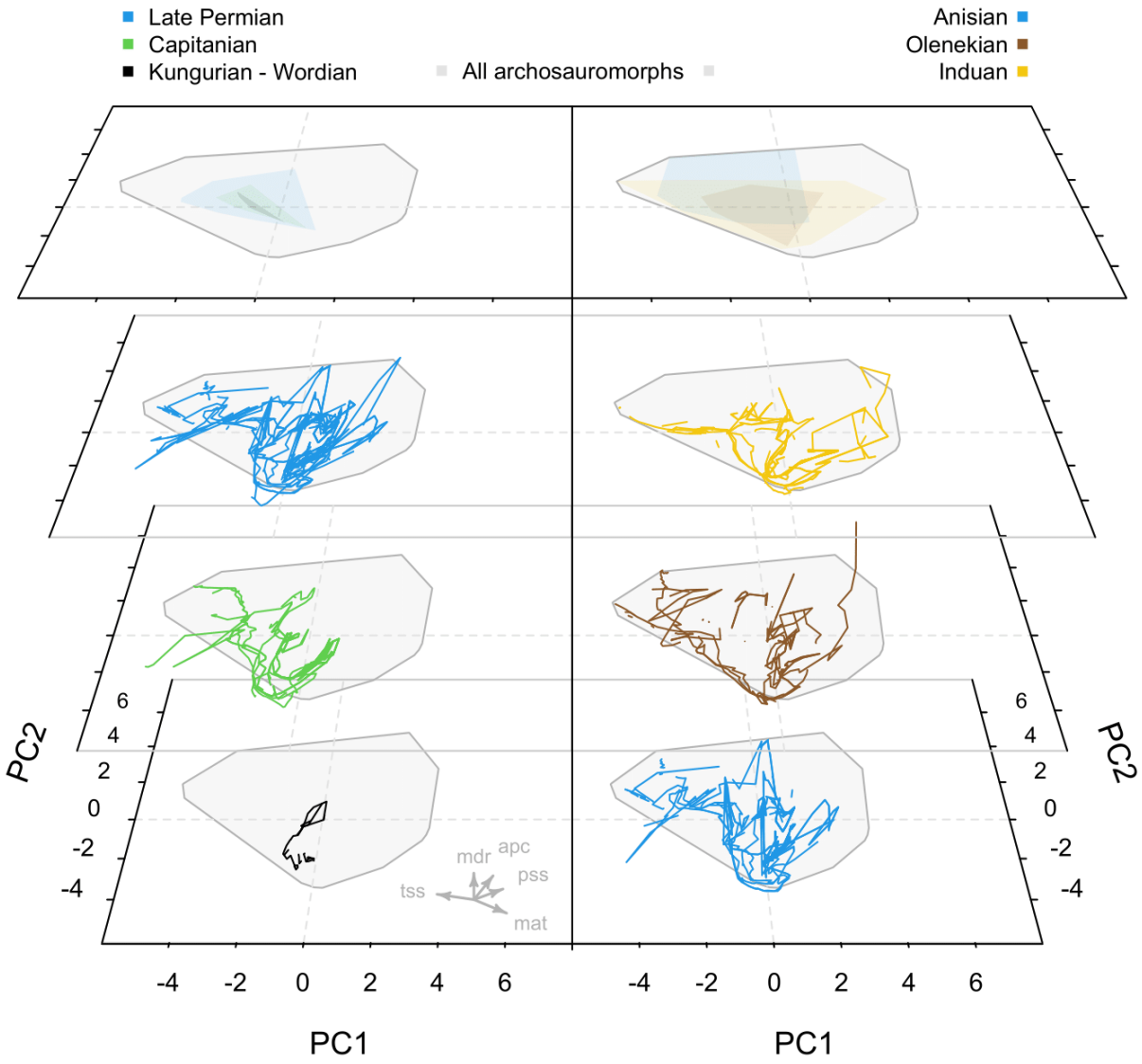


Fig. 5.23. Kungurian to Anisian climate space traversed by early archosauromorphs. Ordinations spaces corresponding to panels A and B in Fig. 5.16, displaying the projections of TARDIS phylogeographic paths through five selected climatic variables encountered along their lengths: temperature seasonality (tss), mean diurnal temperature range (mdr), annual precipitation (apc), precipitation seasonality (pss) and mean annual temperature (mat).

as varied as their total climate space occupancy from their origin to the end of the Triassic. Wide-ranging climatic tolerances persisted through the Middle Triassic and Carnian, while the Norian and Rhaetian were marked mostly by dispersals through environments with high mean annual temperatures and low temperature variation across a range of precipitation seasonalities.

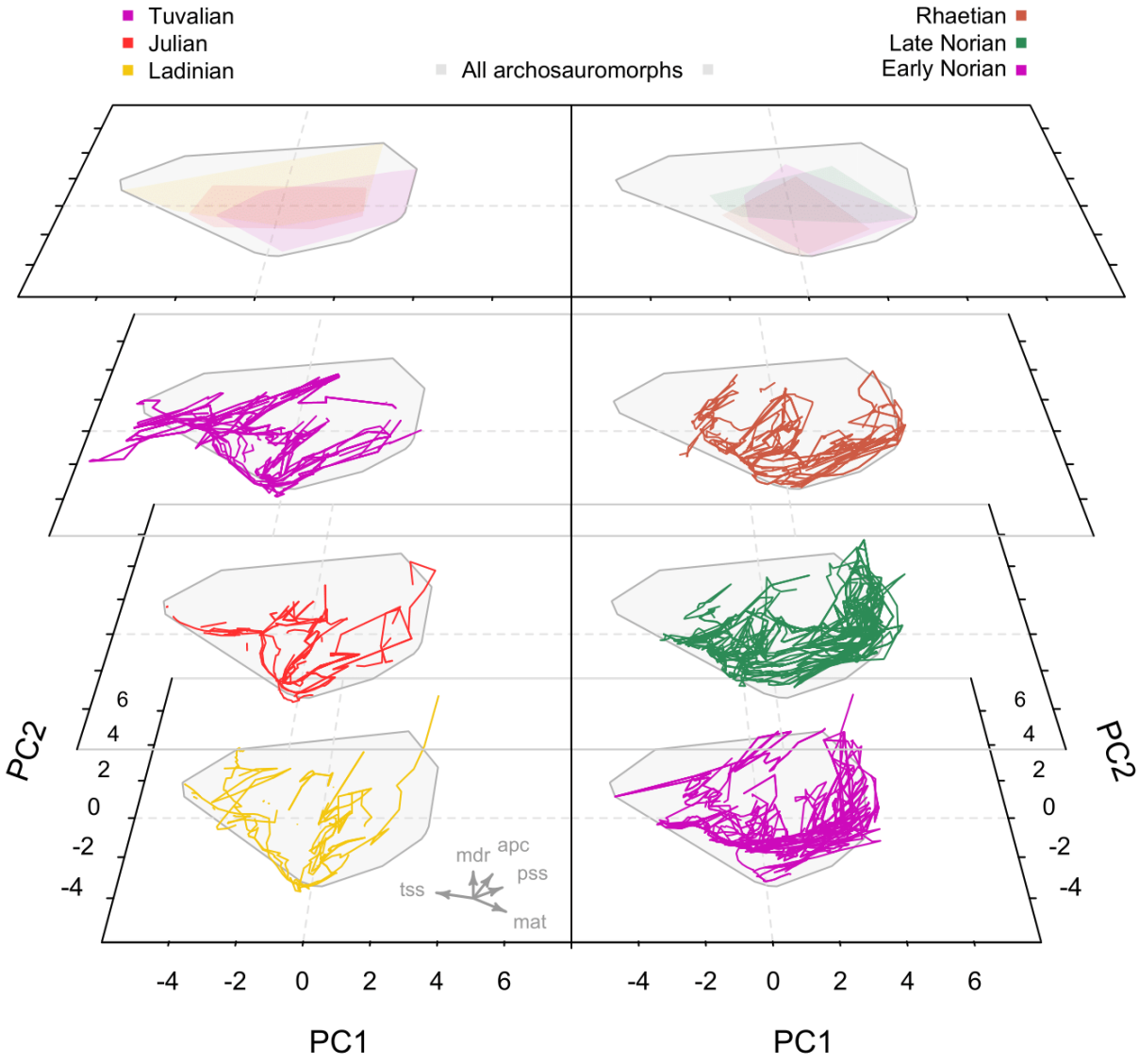


Fig. 5.24. Ladinian to Rhaetian climate space traversed by early archosauromorphs. Ordinations spaces corresponding to panels C and D in Fig. 5.16, displaying the projections of TARDIS phylogeographic paths through five selected climatic variables encountered along their lengths: temperature seasonality (tss), mean diurnal temperature range (mdr), annual precipitation (apc), precipitation seasonality (pss) and mean annual temperature (mat).

5.4 Discussion

5.4.1 Hidden climatic disparity of early archosauromorphs

The phylogeographic least cost pathways returned by TARDIS can be treated as hypotheses for past routes of dispersal, permitting biogeographic histories to be inferred even where there

are temporal or spatial gaps in the fossil record. Projecting these pathways back into the climate space occupied by early archosauromorph taxa provides further insight into the greater diversity of environmental conditions that they must have tolerated to have successfully radiated across Pangaea, complementing the insights which can be gained from estimating the tolerances of reconstructed ancestors (Brusatte et al., 2011; Dunne et al., 2021). Consequently, they may also be of utility in predicting the locations of future fossil sites that document key phases of the early archosauromorph radiation, subject to the availability of terrestrial sedimentary rocks along their lengths.

There are only five known Permian-aged archosauromorph taxa spanning the late Capitanian to Changhsingian (Bernardi et al., 2015). Their time-scaled phylogeny, however, indicates a major phase of higher-level cryptic diversification from the Kungurian to the Changhsingian (Fig. 5.7–5.9), congruent with the findings of previous authors (Ezcurra et al., 2018; Simões et al., 2022), that can only be investigated using phylogenetic comparative methods. When climatic disparity is quantified for this phase of their radiation using ancestral state estimation, it shows that early archosauromorphs were more climatically tolerant than their fossil record indicates at face value (Fig. 5.17A), highlighting a shortcoming of approaches which quantify climatic ranges for archosauromorph clades from fossil tips alone (e.g., Chiarenza et al., 2020; Dunne et al., 2021, 2023). This hidden disparity, however, only increased substantially from the Capitanian onwards even when the additional climatic tolerances required by their phylogeographic history are considered (Fig. 5.16, 5.23) and the earliest phases of their radiation were instead constrained to European Pangaea from the Kungurian to Wordian (Fig. 5.21). This apparent constraint could alternatively be ascribed to incomplete sampling, but this hypothesis can be set aside given the broader rejection of a Middle Permian gap in the tetrapod fossil record (Benton 2012), particularly by well-sampled lower latitude assemblages which do not contain any early archosauromorphs. Instead, the signal of constraint recovered here is supported by increasing dissimilarity detected by previous authors between equatorial and temperate tetrapod faunas by climate-driven vicariance across latitudes at the end of the Kungurian (Sahney and Benton 2008; Brocklehurst et al., 2017; Brocklehurst 2018).

While the earliest phase of the archosauromorph radiation appears to have been constrained, this does not hold true for the remainder of the Permian. Their phylogeographic history indicates that extensive, trans-equatorial dispersals commenced in the Capitanian with

the west coast of Tethys acting as a major biogeographic highway out of northern Pangaea (Fig. 5.21), supporting the suppositions of previous authors (Cisneros et al., 2012; Olroyd and Sidor 2017; Liu et al., 2021). Besides forming the geographically shortest route between southern Pangaea and the inferred origination point of archosauromorphs, the western Tethys coast may have additionally been favourable due to its modest topography compared to the Central Pangaean Mountains further east, and monsoonal moisture countering continental aridity revealed by alternation between red bed and playa lake facies across multiple sedimentary basins situated along the palaeocoastline (Angiolini et al., 2003; Schneider et al., 2006).

Ancestral state estimates indicate that archosauromorph climatic disparity increased during this interval, but conditions measured along TARDIS pathways show that they were not simply dispersing exclusively through environmentally clement corridors that matched these ancestral tolerances, but also across landscapes with more widely varying climatic conditions (Fig. 5.23). If climatic tolerances at the origin of the clade were already broad, then basal archosauromorphs would have been more capable of dispersal out of European Pangaea from the Kungurian onwards, a scenario rejected by the initial trend of geographic and climatic constraint recovered here. Instead, their surge in climatic tolerance during the Capitanian must have been adaptive. This in turn fits with inference of generally high evolutionary rates in basal archosauromorphs, linked to fluctuating climatic conditions breaking an incumbent trend of stabilising selection amongst reptiles as they tracked their environmental optima (Simões et al., 2022). Consequently, these results show that strong provincialism was not always a dominant aspect of Pangaean tetrapod biogeography and that leaky biogeographic barriers permitted interchange of basal archosauromorphs between different parts of Pangaea through the Middle to Late Permian.

5.4.2 Archosauromorph biogeography through climate upheaval

The geochemical record indicates that extreme temperatures persisted across low latitude Pangaea in the aftermath of the EPME that sat at the upper limits or even exceeded the thermophysiological tolerances of plants and animals, forming the so-called tropical dead zone (Sun et al., 2012; Bernardi et al., 2015; Li et al., 2022). The fossil record supports the exceptional rarity of tetrapods in these low latitudes, with abundant, yet unfossiliferous sedimentary rocks spanning the tropics through this interval suggesting that this is not simply an artefact of incomplete sampling (Sun et al., 2012; Bernardi et al., 2015; Li et al., 2022).

While ancestral state estimates are the only indicators of archosauromorph climate space occupation or routinely outstrip the bounds of tip-occupied space in the Permian (Fig. 5.16A), fossil tips show greater climatic disparity compared to ancestral state estimates in their substantially better sampled Early Triassic record (Fig. 5.16B). Consequently, their empirical geographic distribution may more faithfully reflect their preferred climatic ranges in the Early Triassic compared to the Permian. The geometries and climatic conditions measured along TARDIS pathways, however, indicate that early archosauromorph tolerances still exceeded the bounds of tip and ancestral state estimates (Fig. 5.23), with lengthy trans-equatorial traverses along the western Tethys coastal highway during the Induan and Olenekian (Fig. 5.21, 5.22) showing that even the ecologically hostile tropical dead zone was not an absolute barrier to archosauromorph dispersal.

Long pathways were still present under ameliorated climate conditions in the Anisian, but short pathways lacking significant geographic barriers between ancestral nodes became more prevalent (Fig. 5.22), supporting previous suppositions that regional sympatry was a key process underlying Middle Triassic tetrapod recovery and diversification (Ezcurra 2010). Following this interval of relative calm, intense climate change during the CPE drove a long-term shift in archosauromorph climatic space occupancy towards warmer conditions with reduced seasonality (Fig. 5.17C). Both stem and crown archosauromorphs participated in this shift, but the former experienced subsequent decline in the early Norian (Fig. 5.18). Notably, this trend is consistent within ecological guilds across phylogenetically disparate portions of early archosauromorph phylogeny. For example, the herbivorous rhynchosaurs showed diminishing climate space occupation from the late Carnian to early Norian (Fig. 5.17H, 5.17I) while aetosaurs radiated into warm, humid environments (Fig. 5.18B, 5.18C). Aetosaur and rhynchosaur climate spaces still overlapped moderately in the late Carnian, however, so it is unlikely that differential climatic tolerances were directly responsible for the extinction of the former. Instead, floral turnover following climate change appears to have driven their divergent ecological fates due to functional differences in their feeding morphologies (Singh et al., 2021).

Besides aetosaurs, a variety of other crown archosauromorphs radiated at the CPE, but previous workers have generally viewed the pseudosuchian lineages as displaying a greater degree of climatic and geographic constraint compared to avemetatarsalians during the Late Triassic (Mannion et al., 2015; Dunne et al., 2021). The inclusion of reconstructed ancestors does not have effect on the occupied bounds of climate space, but when climate space

occupancy is instead examined at stage to substage levels, pseudosuchians emerge as mostly responsible for the shift seen in archosauromorph climate space towards hotter, but more thermally stable environments and initially occupied a greater breadth of environmental conditions compared to avemetatarsalians in the wake of the CPE (Figs. 5.16C, 5.18). While avemetatarsalians may have cumulatively occupied a broader range of environmental conditions by the end of the Rhaetian (Dunne et al. 2021), pseudosuchians appear to have displayed greater ecological success during the late Carnian and early Norian, matching their exceptionally early burst in taxonomic diversity and widely distributed fossil record (Brusatte et al., 2015; Lacerda et al., 2015; Mannion et al., 2015; Stocker et al., 2017). In essence, just because pseudosuchians occupied a more conservative extent of climate space overall, that did not mean that they were strongly constrained when this niche was realised in empirical geographic space; hot-adapted reptiles may be expected to do well when they occupy a supercontinent whose surface area predominantly sits within the hot low latitudes. Following their South American origin (Fig. 5.14C) the modest climate space occupancy of early dinosaurs supports current proposals that their initial dispersal across austral Pangaea was constrained by latitudinally structured climatic belts (Griffin et al., 2022). Conversely, the sudden geographic jump to the Carnian of northern Pangaea with the origin of pterosaurs (Fig. 5.14E) shows that climatic constraints were variable across early avemetatarsalians, potentially due to their differences in locomotory modes permitting easier traversal of harsher environments by pterosaurs (Britt et al., 2018, Foffa et al., 2022).

5.4.3 Study limitations

While the results presented here illuminate the biogeographic origins and cryptic climatic breadth of the initial archosauromorph radiation, the methods used are subject to several caveats. Sporadic sampling and poor stratigraphic age constraints for many early archosauromorph fossils increase divergence time uncertainties, with node age ranges of 8–10 Ma for the earliest phases of their radiation and a near 25-million-year gap separating their origin from a widely sampled archosauromorph record in the Olenekian onwards. Consequently, it is challenging to conclusively link climatic events to their divergence and dispersal during this cryptic interval.

Higher-level archosauromorph interrelationships are generally well established, but the topologies and placements of individual clades vary slightly between different phylogenetic hypotheses (e.g., Ezcurra et al., 2018; Elsler 2019; Simões et al., 2022). These revisions may

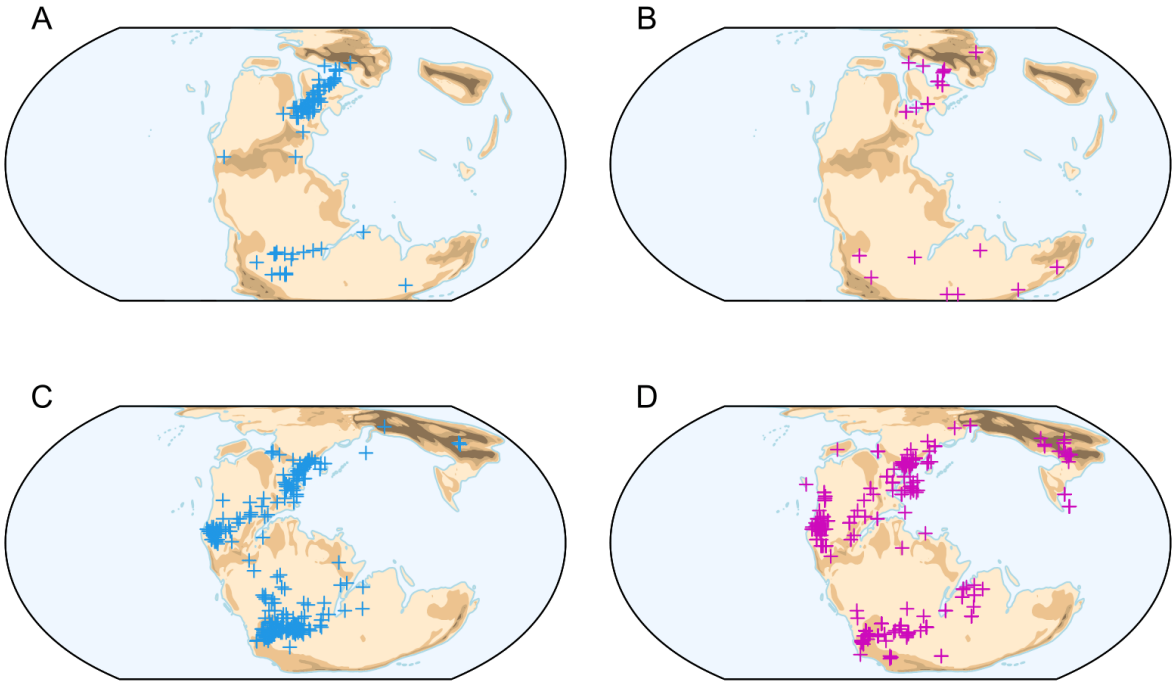


Fig. 5.25. Correspondence between tip locations and ancestral geographic origins. Throughout the Permian and Triassic, estimated geographic origins (blue) are generally located in the same broad areas as the observed locations of fossil tips (purple). **(A).** Late Permian to Early Triassic node locations on a compromise Induan base map. **(B).** Late Permian to Earl\y Triassic tip locations on a compromise Induan base map. **(C).** Middle to Late Triassic node locations on a compromise Norian base map. **(B).** Middle to Late Triassic tip locations on a compromise Norian base map.

conceivably alter estimated divergence times and locations, but more concerning is the observation that molecular phylogenies more faithfully fit the biogeographic data for their clades than morphological phylogenies (Oyston et al., 2022), leading to the possibility that current archosauromorph topologies may fundamentally misrepresent their phylogeographic history. The *geo* model is nonetheless limited to a single topology, meaning that uncertainties in phylogenetic relationships or divergence times cannot be easily incorporated into the final estimates of geographic origins. Similarly, there is no way to parameterise variable accessibility besides simple geographic masking to help constrain ancestral origination points, although it may still be more appropriate to impose as few assumptions as possible to reduce the risk of overparameterized ancestral state estimates.

There is then a conspicuous correspondence between major areas of fossil sampling and inferred divergence locations throughout early archosauromorph phylogeny (Fig. 5.25). It is

potentially unsurprising that many of the oldest archosauromorph fossils come from Europe as an historical aspect of sampling effort (Raja et al., 2022) that, in conjunction with the rapidity of early divergences along the path towards crown archosaurs, may be responsible for northern Pangaean signature of their geographic origins and deepest phylogenetic splits. In turn, once geographically proximate sedimentary sequences spanning southern Gondwana (the Uspallata Supergroup of Argentina, Santa Maria Supergroup of Brazil, and the Karoo Supergroup and Manda Beds of southern Africa), and the extensively sampled Chinle Formation and Dockum Group in the southern United States may be responsible for biasing archosauromorph divergences towards the southern Gondwanan and North American centres of origination in the Early and Middle Triassic respectively. As clade divergence times frequently predate the ages of their oldest fossils, it stands to reason that clade divergence locations may follow a similar spatial pattern dictated by the geographic distributions of their oldest fossils irrespective of their true biogeographic histories.

Finally, TARDIS bears several caveats. Careful choice of weighting scheme is paramount to ensure that the assumptions made about the constraints on dispersal are reasonable, but the climatic centroid-based approach taken will ultimately be biased against taxa which naturally sit away from this centroid in the first place. Similarly, the analysis conducted here has also made the inherent assumption of equal locomotor capabilities of all taxa examined, despite including tanystropheids with a contentious mixture of proposed terrestrial and aquatic life, and volant pterosauromorphs. Volancy would not have freed pterosauromorphs from the ecological and metabolic constraints imposed on habitability by climate, however and so they may have conformed to the broad spatial patterns of dispersal displayed by cursorial archosauromorphs, although the presence of the Triassic pterosaur *Caelestiventus* in hyper arid desert environments does highlight their early ecological diversity (Britt et al., 2018). Further, some tanystropheids have been found in clearly marine sediments while others inhabited continental lakes far removed from the oceans (Casey et al., 2007; Fraser et al., 2013), demonstrating that ancestral tanystropheids must have been capable of terrestrial dispersal. Consequently, the pathways recovered by TARDIS appear broadly reasonable at global scales, despite locomotor variation and uncertainty.

5.5 Conclusions

The spatial signature of the archosauromorph radiation is limited by sporadic sampling and variable age control of its fossils. Advancements will come with the discovery of new fossils or localities, or substantial improvements to radiometric and stratigraphic constraints on key formations, but there is no guarantee of this given the finite volume of accessible Triassic terrestrial sedimentary rocks. Consequently, our view of the archosauromorph radiation remains biased towards the narratives presented by the European, North American and southern Gondwanan sampling centres. These records nonetheless permit phylogeographic inference of their wider patterns of dispersal across Pangaea. TARDIS represents a novel approach to investigating these patterns by enabling tools rooted in landscape connectivity analysis to be applied over macroevolutionary timescales, making full use of the remarkable advances in reconstructing the stage over which these grand biogeographic events took place provided by Earth-system modelling. In combination in phylogeographic analysis, TARDIS is particularly powerful as it provides spatially explicit estimates of dispersal pathways that can extend beyond the bounds of the fossil record, and in turn a way of probing the unsampled environmental conditions that clades must have overcome during their evolutionary radiations. In the context of archosauromorphs, insights from TARDIS are not incompatible with previous assertions regarding the broader biogeographic structuring of tetrapod assemblages across Pangaea but reveal that the barriers to dispersal that generally separated the northern and southern hemispheres of the supercontinent through the Permian and Triassic were not absolute, and that archosauromorphs were more climatically tolerant than their fossil record or ancestral state estimates alone would suggest.

While TARDIS raises caveats, these also present potential avenues for improving the method in future work. The underlying landscape and climate data could be used to derive additional environmental factors that affected dispersal, for example the geometry of palaeoriver networks based on catchment geometries delineated from DEMs using flow accumulation algorithms (see Zhou et al., 2019). Such features could conceivably have permitted traversal though otherwise arid portions of Pangaea analogous to the way the river Nile creates a corridor of habitability through the Eastern Desert of Egypt. Similarly, climate data on shorter term climatic change (e.g., annual seasonality, decadal to millennial orbital variability) or unique, rapid climate events like the Carnian Pluvial Episode could be incorporated as additional graph layers in TARDIS to capture the sensitivity of dispersal to

climatic departures from time-averaged conditions. The least cost pathways derived using TARDIS fundamentally assume a process of dispersal between two points and so does not account for other biogeographic processes affecting populations like vicariance. There is clear scope, however, to use community detection algorithms to identify variable connectivity between spatially discrete neighbourhoods of grid cells to investigate such processes in a singular graphical framework routed in landscape ecology. Similarly, it will be valuable to compare how these results compare to inferences from traditional phylogeographic models like those implemented in BioGeoBEARS, or to mechanistic simulations of origination and dispersal of early archosauromorphs using gen3sis (Hagen et al., 2021), incorporating key thermophysiological constraints to control these processes. Ground truthing results from such simulations with the inferences from TARDIS and the remnant empirical records provided by fossils will a crucial step towards bridging the gap in our understanding of archosauromorph biogeography at the beginning of their evolutionary success story.

Chapter Six

General Conclusion

In this final chapter I summarise the advancements which this thesis has made and reflect on their context within wider palaeobiological research. I consider whether diversity can be reliably estimated in an incomplete fossil record given the challenges posed from its temporal, spatial and taxonomic heterogeneities, and how the techniques I have developed for data cleaning and subsampling help address these issues. I then discuss whether we can ever reliably infer the spatiotemporal controls on diversity and biogeography from an incomplete fossil record, even by leveraging advances from Earth-system modelling, tools from geographic information systems, and the phylogenetic requirements of relict taxon distributions. I conclude that advances in both areas will be made by coupling Earth-system simulations with mechanistic models of diversification, dispersal, and fossil sampling, giving extensive scope for future research endeavours that can circumvent the spatial limits of the fossil record.

Author contributions: This chapter was written and developed by J. T. Flannery-Sutherland. Comments were provided on a draft by M. J. Benton. The author contributed all the work in this conclusion.

6.1 Introduction

The fossil record is our only empirical sample of past biotic history. Fossil occurrence databases have provided unprecedented access to the wealth of fossil data acquired over decades of palaeobiological enquiry and are indispensable sources of information regarding the spatiotemporal distribution of life on Earth, but their growth has been accompanied by the accumulation of errors which can mar their utility and increasing awareness of the inherent limitations of their data. Even when high quality fossil occurrence data is available, it remains extremely difficult to confidently identify patterns of true biodiversity or to infer its drivers. Chief among the myriad issues which distort patterns of fossil diversity are pervasive spatial sampling biases. In this thesis I have explored ways in which these issues may be addressed using novel computational methods developed herein, tools from geographic information systems (GIS), and simulations from Earth-system models (ESMs) which extend beyond the spatial limits of the geological record. I have applied these approaches to several case studies of the transformation of the biosphere through the Late Permian to Early Jurassic. In some cases, these analyses have successfully uncovered the spatial nuances of the recovery and radiation of life through the Triassic; in others they have revealed methodological limitations and caveats that provide insight into where palaeobiology must go next.

6.1.1 Can patterns of diversity be estimated in an incomplete fossil record?

Any palaeobiological analysis is predicated on high quality data. In chapter two I provided cleaning methods which can scale to the task of identifying and correcting suspect entries in the largest fossil occurrence datasets. These will ease the burden of manually checking tens of thousands of individual entries prior to analysis and provide a basis for future development of more sophisticated statistical cleaning solutions. Cursory examination of the effects of automated error detection and resolution across the entirety of the Phanerozoic fossil record demonstrates that corrections had a quantitative impact on diversification patterns, but the qualitative results remained the same. More concerning is the finding that well-sampled clades which contribute the greatest amount of signal to large, polyphyletic occurrence datasets many also be the most error-prone, although the ratio of reliable signal versus confounding error may still be sufficient to reliably recover their qualitative diversity patterns.

Despite historic concerns over the quality of sampling in the total fossil record, palaeobiologists have recovered broadly stable patterns of global palaeodiversity from ever

larger compilations of fossil ranges and occurrences (Sepkoski et al., 1981; Sepkoski 1993), deployment of ever more complex methods of inferring sampling-corrected diversity trends (Alroy et al. 2001, 2008), and most recently from mechanistic approaches for estimating past diversity without reference to empirical fossil data (Cermeño et al., 2022). Such stability suggests that the total fossil record does provide at least a partial signal of the global trend in Phanerozoic diversity from which limited conclusions can be made, for example the sudden rise in diversity from the Jurassic onwards exceeding previous peaks in Palaeozoic by a notable margin. The findings of this thesis, however, demonstrate that other macroevolutionary trends and phenomena are much more challenging to infer in the face of incomplete sampling. Mass extinctions, evolutionary radiations and the records of individual clades through geological time must be examined with spatially sensitive analytical methods in order to draw meaningful conclusions about their underlying patterns and processes at broader scales. Further, the outstanding issues posed by palaeogeographic uncertainty affect even spatially explicit simulations of biodiversity, necessitating extensive sensitivity testing of simulated diversification dynamics across different sets of palaeogeographic boundary conditions to demonstrate that global trend are truly robust.

Global patterns are important as they can reveal fundamental changes to the capacities of ecosystem to host biodiversity (Sepkoski et al., 1981; Novack-Gottshall 2016; Zaffos et al., 2017; Benton et al., 2022) but even if such patterns are derived from the total fossil record, they are still fundamentally incomplete reflections of true macroevolutionary history that are biased towards historically well-sampled regions of the globe. In other words, the total fossil record can tell us something useful about the emergent global trend, but those trends are nonetheless specific to those well-sampled regions while the contributions of other parts of the globe remain unknown. Instead, palaeobiologists are rapidly recognising the value in spatially decomposing these patterns to reveal the regional complexity of global Phanerozoic diversity trends and their relationships to spatiotemporal variation in the tectonic arena that literally and mechanistically supports the biosphere (Close et al., 2020a; Flannery-Sutherland et al., 2022). These endeavours further highlight the importance of targeting spatially restricted portions of the fossil record to avoid the effects of spatial sampling bias. In chapter three, I developed a method which provides spatially standardised subsets of the fossil record suitable for the estimation of the extinction and origination rates determining taxonomic diversity. Crucially, this approach demonstrates that diversification dynamics are geographically heterogeneous, a property evident in the context of the modern biosphere (Allen and Gillooly 2006) that gains

new meaning in a deep time macroevolutionary context. Diversity has not changed uniformly across the globe, even during apparently global events that have transformed the biosphere, and different regions with varied biogeographic histories contributed differentially to the cumulative signal of global diversity. This point is of paramount importance given how biodiversity hotspots arising under a variety of spatiotemporal contingencies have controlled long-term global patterns (Cermeño et al., 2022), and whose decimation during mass extinctions lead to catastrophic consequences on broader spatial scales.

Focusing on regional diversity trends is a critical step in palaeobiological thinking, but this is only possible subject to the availability of high-quality regional data. The method presented in chapter three needs spatially consistent sampling of fossil occurrences through time, a requirement that is not always met by the geological record and an issue that also confounds less stringent methods for deriving spatially standardised subsamples of fossil diversity (Close et al., 2020a; Antell 2022). While we may be able to estimate broad patterns of diversity which approach true changes in global taxonomic richness, it will not always be possible to confidently subdivide these patterns spatially to understand their biogeographic nuances. In a similar vein, it is not always possible to decompose ecosystem, region or continent-level estimates of richness based on entire biotas into their clade-specific components due to differential sampling quality between different taxonomic groups. While it is evident to say that the quality of the fossil record varies taxonomically (Lloyd et al., 2012), this is especially problematic when that variation negatively impacts clades which have played key roles in the construction of marine and terrestrial biodiversity hotspots during the Phanerozoic, a pertinent example being the patchy records of corals and sponges through the Triassic transformation of reef systems.

It is also necessary to consider the methodological shortcomings facing reliable estimation of past diversity. Traditional palaeontological methods remain incoherent in their estimates of diversity versus diversification and are generally less accurate than approaches which consider the fossil record from the perspective of a birth-death-sampling process (Warnock et al., 2020); they are often inadequate for the data to which they have historically been applied. Nonetheless, fossilised birth-death models and their relatives can still perform poorly when preservation rates are low or highly variable, when taxonomic representation is incomplete, or under a combination of these factors (Warnock et al., 2020). The diversification dynamics recovered in chapters three and four demonstrate that PyRate and mcmcDivE return

results that may be skewed by edge effects induced by stratigraphic imprecision in the data or the coarseness of the preservation model, while chapter four further show that some clades have fossil records which are still inadequate for use with even the most sophisticated inference methods. These issues may be overcome in the future by newer methods like the Bayesian Brownian Bridge or deep-learning techniques (Carlisle et al., 2023; Cooper, Flannery-Sutherland, Silvestro IN REVIEW). The spatiotemporal grain of fossil occurrence data will still impose hard analytical limits on the estimation of past diversity, however, and it may be the case that those limits are fast being approached in some parts of the fossil record.

6.1.2 Can the spatiotemporal controls on biodiversity be inferred?

Biodiversity displays spatially structured biogeographic patterns which contribute heterogeneously to temporally structured macroevolutionary and macroecological patterns. Spatially incomplete information on its variation and on geological proxies of past environmental conditions are major hurdles facing elucidation of diversity patterns and their drivers (Close et al., 2020a, 2020b; Jones and Eichenseer 2022), and are compounded by the issues of comparing geographically biased proxy records of 'global' environmental conditions and 'global' diversity to each other, or of the former to more reliable estimates from spatially standardised subsamples where the regional heterogeneity of diversity dynamics takes over. Integration of ecological modelling methods with palaeogeographic reconstructions and simulations from ESMs go some way towards overcoming these problems. Spatially structured controls on diversity can be readily identified in the present, for example the influence of latitudinal temperature gradients or bathymetric irradiance gradients (Gaston 2000; Valentine and Jablonski 2015; Schluter and Pennell 2017; Gagné et al., 2020). These controls can in turn be investigated in deep time contexts, subject to the availability of well curated fossil occurrence data and ESM simulations, to determine their impacts on past biodiversity across palaeogeography (Mannion et al., 2013).

It is subsequently tempting to identify certain variables like sea level (affecting marine and terrestrial habitat area) and CO_2 (affecting long term global temperature) as drivers of diversity through geological time (e.g., Cornette et al., 2002; Zaffos et al., 2017). The spatial heterogeneity of long-term diversity trends identified in chapter three, however, shows that life has not responded in a spatially uniform manner when confronted by global changes in such controls. Similarly, the response of corals to changing sea level and continental configuration through the Triassic was not smooth, but instead their reef-building boom shortly after the CPE

was predicated on a series of contingent ecological triggers that coincided with a late Carnian to early Norian peak in tropical marine shelf area. These findings raise the question of whether we can truly identify 'drivers' of diversity. There are undoubtedly environmental factors whose fluctuations contribute to consistent directional changes in diversity in some instances (Dunhill et al., 2018) but any directional change will still be a net result of taxonomically varied responses, nor do such drivers account for ecological interactions which can amplify or dampen taxonomic impacts in the face of abiotic forcings (Dunne et al., 2004). Similarly, we must be mindful of how conditions within the Earth-system created intervals of contingency that altered the trajectory of the biosphere through the Phanerozoic. Ultimately, the conceptualisation of drivers as overarching, temporally continuous controls is too simplistic to explain the complex macroevolutionary tales portrayed by the fossil record.

Despite the significant advantages offered by ESMs over proxy data from the geological record, discrepancies between environmental reconstructions and coarse spatiotemporal resolutions present methodological hurdles to inferring the drivers of diversity. In chapter four, my goal of estimating spatial trends in realised niche space for corals and sponges was challenged by the sensitivity of projected niche model estimates to the choice of palaeogeographic reconstructions and the confounding effects of sea level on the extent of accessible habitat space. Beyond these fundamental differences in palaeogeographic boundary conditions, model resolution may also affect the utility of ESMs in other palaeontological applications. While palaeogeographic digital elevation models have excellent spatial resolution at global scales, ESMs built atop these boundary conditions are typically constrained to lower resolutions by computational limitations (Hagen 2022). ESM simulation sets also suffer from limited temporal resolution as most endeavours have focused on stage-level or still broader reconstructions of long-term palaeogeography, climate, and productivity despite the palaeontological importance of short-term events like mass extinctions which bore discrepant environmental conditions, although this gap is starting to be bridged (Chiarenza et al., 2020). Consequently, ESMs can provide powerful insight into the drivers of diversity, but their deployment is not always straightforward subject to the clade or stratigraphic interval under examination, and the current availability of ESM simulations may not capture the precise spatiotemporal variations in environmental conditions that controlled the distributions of taxa and ecosystems in deep time. Careful framing of research questions within the constraints of the available data when working with these resources

6.1.3 Can we make inferences beyond the spatial extent of the fossil record?

The fossil record captures discontinuous portions of a biosphere that was shaped by its underlying geographic stage. Consequently, tools from GIS and ecological modelling provide a powerful inferential framework to probe the relationships between the biosphere and its palaeogeographic and climatic setting. The utility of GIS in palaeontology has been recognised for more than a decade, but its applications have generally been limited to relatively simple purposes like the quantification of geographic ranges from fossil occurrence data (Rayfield et al., 2005; Rode and Lieberman 2005; Stigall and Lieberman 2006; Stigall 2010; Kolis and Lieberman 2019). Instead, the biogeographic distributions of taxa in the fossil record in conjunction with the requirements of their phylogenetic histories provide a basis for inference across intervening portions of geographic space that are devoid of fossils. In chapter five I developed principled phylogeographic approaches to quantify the constraints that past topographies and climates imposed on taxonomic patterns of dispersal, along with the unseen environmental breadths encountered during those dispersals. These approaches, in concert with other spatially explicit methods to quantify taxon-environment relations, offer the potential to explore the biogeographic affinities and limits of clades beyond the bounds of their fossil records, although they will benefit from refinement to permit quantitative hypothesis testing and comparison.

From a methodological perspective, the full potential of GIS and ecological modelling tools has yet to be realised, particularly given the increasing availability of deep time simulations from ESMs and the remarkable breadth of techniques developed to investigate taxon-environment relationships in the present. A data-focused perspective, however, highlights the issue that the spatially complete views provided by ESMs are not always adequately matched by the fossil record. Spatially biased fossil data may still drive biased inferences regarding its broader biogeographic and phylogeographic patterns, as demonstrated by the conspicuous correspondence between archosauromorph tip positions and their inferred ancestral locations in chapter five, and this biasing effect conceivably extends to extraction of ESM simulation data at fossil occurrence locations more generally. Spatial sampling issues may be alleviated by calibration of taxon-environment relationships in the present day where such biases are greatly reduced but this is not always methodologically appropriate and requires strong assumptions regarding the stability of those relationships through geological time. While palaeontologists can make inferences which extend beyond the bounds of the fossil record, its spatial discontinuities remain challenging to overcome. Instead, there is a growing premium on

finding new approaches which will break our reliance on inference of the structure and properties of the past biosphere from its remnants in the geological record.

6.2 Future directions

Palaeontology has benefitted substantially from recognising the fossil record as a manifestation of a birth-death-sampling process, uniting two sets of phenomena that were traditionally treated separately: diversification responsible for shaping the biosphere (speciation and extinction) and sampling responsible for its observed fossil record (taphonomy and the anthropogenic filters imposed by palaeontological collection and description). What is missing from or at least is still largely considered separately to this sophisticated view of the fossil record are the spatial components of these processes. The genesis of a species is naturally followed by its dispersal, providing opportunities for subsequent speciation events by vicariance, while extinction is typically not a spatially uniform event but rather the diachronous result of range contraction through local extirpations. These processes are captured by dispersal-extinction-cladogenesis models which lack a sampling component, while dispersal-extinction-sampling models lack an origination process and are additionally heavily limited by the biogeographic scenarios they can capture (Silvestro et al., 2018). A more complex approach is to consider spatiotemporally variable diversification, dispersal, and sampling in a single framework, but the spatial components of its processes in this case are discretised into arbitrary bioregions rather than mapped onto empirical palaeogeographic boundary conditions (Cooper, Flannery-Sutherland, Silvestro IN REVIEW).

Alternately, palaeontologists are beginning to explore mechanistic, spatially explicit eco-evolutionary models of biodiversity (Saupe et al., 2020; Hagen et al., 2021a, 2021b; Cermeño et al., 2022; Dunne et al., 2023). These can capitalise on the wealth of data from deep time ESM simulations (Fig. 6.1) to capture intervals of environmental contingency and can be flexibly parameterised to help distinguish between different configurations of the same set of mechanistic components that can plausible generate a single observed pattern of diversity, the often-overlooked issue of equifinality (Hagen 2022). Currently, these models do not root diversification in a birth-death process but are starting to consider incomplete sampling (Dunne et al., 2023), satisfying the need for biodiversity inference without *a priori* reference to fossil data whilst still providing the critical bridge between past biosphere and remnant fossil record required for ground-truthing against empirical geological observations. To properly support

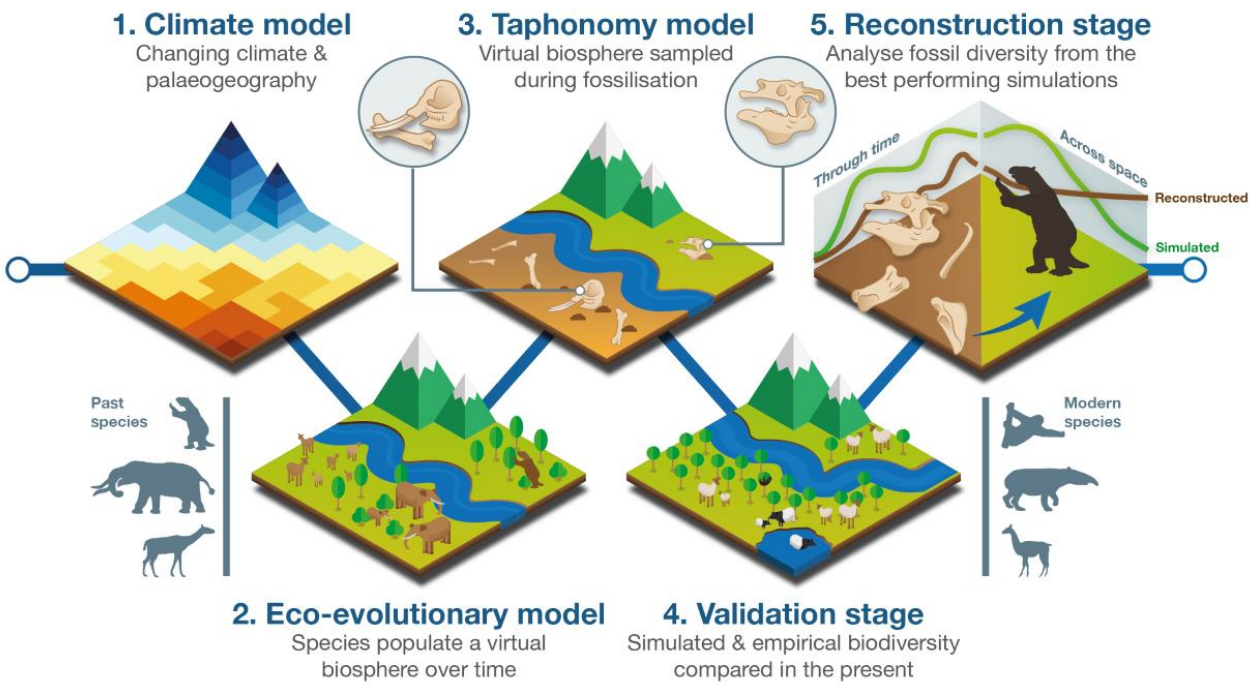


Fig. 6.1. Schematic workflow for coupling eco-evolutionary and Earth-system modelling. This workflow takes a 'forward-in-time' modelling approach where simulation outputs are validated on diversity in the present day. Melisa Morales García is thanked for designing this figure which was commissioned for a separate body of work by the author.

future advances in mechanistic, spatially explicit eco-evolutionary modelling, Earth-system simulation strategies must also be broadened to include a greater range of uncertainties in palaeogeographic boundary conditions and environmental forcings, and the disequilibria during global geological events like mass extinctions. Finally, spatially explicit eco-evolutionary models must be augmented to consider the role of ecosystem structure in determining the responses of biodiversity to Earth-system forcings, an outstanding gap between palaeobiology and ecology.

Layering of eco-evolutionary models on top of ESMs is set to yield spectacular advances in our understanding of the biosphere over variety of spatiotemporal scales and offers a tantalising framework within which to revisit the Triassic case studies tackled herein. This framework for the processes which have shaped biodiversity and its fossil record should not be viewed as a panacea, however. Instead, greatest gain will come from holistic framework which integrates insights from multiple lines of enquiry to provide a far richer view of life's history than any one technique alone can yield.

References

ADRAIN, J., WESTROP, S. (2000). An empirical assessment of taxic paleobiology. Science, 289: 110–112

AKAIKE, H. (1973). Information theory and an extension of the maximum likelihood principle. In: Petrov, B., Csáki, F. (eds). Proceeding of the Second International Symposium on Information Theory (Tsahkadsor, Armenia). Akadémia Kiadó, Budapest, Hungary, 267–281

ALLEMAND, D., TAMBUTTÉ, E., ZOCCOLA, D., TAMBUTTÉ, S. (2010). Coral calcification, cells to reefs. In: Dubinsky, Z., Stambler, N. (eds). Coral Reefs: An Ecosystem in Transition. Springer, Dordrecht, 119–150

ALLEN, A., GILLOOLY, J. (2006). Assessing latitudinal gradients in speciation rates and biodiversity at the global scale. Ecology Letters, 9: 947–954

ALLEN, B., WIGNALL, P., HILL, D., SAUPE, E., DUNHILL, A. (2020). The latitudinal diversity gradient of tetrapods across the Permo-Triassic mass extinction and recovery interval. Proceedings of the Royal Society B, 287: 20201125

ALLEN, B., CLAPHAM, M., SAUPE, E., WIGNALL, P., HILL, D. DUNHILL, A. (2023). Estimating spatial variation in origination and extinction in deep time: a case study using the Permian–Triassic marine invertebrate fossil record. Paleobiology, doi:10.1017/pab.2023.1

ALLISON, P., BRIGGS, D. (1993). Paleolatitudinal sampling bias, Phanerozoic species diversity, the end-Permian extinction. Geology, 21: 65–68

ALLMON, W., DIETL, G., HENDRICKS, J., ROSS, R. (2018). Bridging the two fossil records: Paleontology's "big data" future resides in museum collections. In: Rosenberg, G., Clary, R. (eds). Museums at the Forefront of the History and Philosophy of Geology: History Made, History in the Making. Geological Society of America Special Papers, 535: 35–44

ALROY, J. (1996). Constant extinction, constrained diversification, and uncoordinated stasis in North American mammals. Palaeogeography, Palaeoclimatology, Palaeoecology, 127: 285–311

ALROY, J. (2000). New methods for quantifying macroevolutionary patterns and processes. Paleobiology, 26: 707–733

ALROY, J. (2010a). Fair sampling of taxonomic richness and unbiased estimation of origination and extinction rates. In: Alroy, J., Hunt, G. (eds). Quantitative Methods in Paleobiology. Paleontological Society Papers, 16: 55–80

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

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

ALROY, J. (2014a). Accurate and precise estimates of origination and extinction rates. Paleobiology, 40: 374–397

ALROY, J. (2014b). A simple Bayesian method of inferring extinction. Paleobiology, 40: 584–607

ALROY, J. (2015a). A more precise speciation and extinction rate estimator. Paleobiology, 41: 633–639

ALROY, J. (2015b). A new twist on a very old binary similarity coefficient. Ecology, 96: 575–586

ALROY, J. (2018). Limits to species richness in terrestrial communities. Ecology Letters, 21: 1781–1789

ALROY, J., MARSHALL, C., BAMBACH, R., BEZUSKO, K., FOOTE, M., FÜRSICH, F., HANSEN, T., HOLLAND, S., IVANY, L., JABLONSKI, D., JACOBS, D., JONES, D., KOSNIK, LIDGARD, S., LOW, S., MILLER, A., NOVACK-GOTTSHALL, P., OLSZEWSKI, T., PATSZOWSKY, M., RAUP, D., ROY, K., SEPKOSKI, J., SOMMERS, M., WAGNER, P., WEBBER, A. (2001). Effects of sampling standardization on estimates of Phanerozoic marine diversification. Proceedings of the National Academy of Sciences of the United States of America, 98: 6261–6266

ALROY, J., ABERHAN, M., BOTTJER, D., FOOTE, M., FURSICH, F., HARRIES, P., HENDY, A., HOLLAND, S., IVANY, L., KIESSLING, W., KOSNIK, M., MARSHALL, C., MCGOWAN, A., MILLER, A., OLSZEWSKI, PATZKOWSKY, M., PETERS, S., VILLIER, L., WAGNER, P., BONUSO, BORKOW, P., BRENNEIS, B., CLAPHAM, M., FALL, L., FERGUSON, C., HANSON, V., KRUG, A., LAYOU, K., LECKEY, E., NURNBERG, S., POWERS, C., SESSA, J., SIMPSON, C., TOMASOVYCH, A., VISAGGI, C. (2008). Phanerozoic trends in the global diversity of marine invertebrates. Science, 321: 97–100

ALTEKAR, G., DWARKADAS, S. HUELSENBECK, J., RONQUIST. F. (2004). Parallel Metropolis-coupled Markov chain Monte Carlo for Bayesian phylogenetic inference. Bioinformatics, 20: 407–415

ANDREOLETTI, J., ZWAANS, A., WARNOCK, R., AGUIRRE-FERNANDEZ, G., BARIDO-SOTTANI, J., GUPTA, A., STADLER, T., MANCEAU, M. (2022). The occurrence birth–death process for combined-evidence analysis in macroevolution and epidemiology. Systematic Biology, 71: 1440–1452

ANGIOLINI, L., BALINI, M., GARZANTI, E., NICORA, A., TINTORI, A., CRASQUIN, S., MUTTONI, G. (2003). Permian climatic and paleogeographic changes in Northern Gondwana: the Khuff Formation of Interior Oman. Palaeogeography, Palaeoclimatology, Palaeoecology, 191: 269–300

ANTELL, G. (2022). divvy: Spatial Subsampling of Biodiversity Occurrence Data. R package version 0.1.1.9. https://github.com/GwenAntell/divvy

APALDETTI, C., MARTINEZ, R., CERDA, I., POL, D., ALCOBER, O. (2018). An early trend towards gigantism in Triassic sauropodomorph dinosaurs. Nature Ecology and Evolution, 2: 1227–1232

ARRHENIUS, O. (1921). Species and area. Journal of Ecology, 9: 95–99

AVARIA-LLAUTUREO, J., VENDITTI, C., RIVADENEIRA, M., INOSTROZA-MICHAEL, O., RIVERA, R., HERNANDEZ, C., CANALES-AGUIRRE, C. (2021). Historical warming consistently decreased size, dispersal and speciation rate of fish. Nature Climate Change, 11: 787–793

BAHCALL, J., PINSONNEAULT, M., BASU, S. (2001). Solar models: current epoch and time dependences, neutrinos, and helioseismological properties. The Astrophysical Journal, 555: 990–1012

BALBI, M., PETIT, E., CROCI, S., NABUCET, J., GEORGES, R., MADEC, L., ERNOUT, A. (2019). Ecological relevance of least cost path analysis: an easy implementation method for landscape urban planning. Journal of Environmental Management, 244: 61–68

BAPST, D. (2013a). When can clades be potentially resolved with morphology? PLoS ONE 8: e62312

BAPST, D. (2013b). A stochastic rate-calibrated method for time-scaling phylogenies of fossil taxa. Methods in Ecology and Evolution, 4: 724–733

BAPST, D. (2014). Assessing the effect of time-scaling methods on phylogeny-based analyses in the fossil record. Paleobiology, 40: 331–351

BAPST, D. (2012). paleotree: an R package for paleontological and phylogenetic analyses of evolution. Methods in Ecology and Evolution, 3: 803–807

BAPST, D., HOPKINS. M. (2017). Comparing cal3 and other a posteriori time-scaling approaches in a case study with the pterocephaliid trilobites. Paleobiology, 4: 49–67

BARIDO-SOTTANI, J., PETT, W., O'REILLY, J., WARNOCK, R. (2019). FossilSim: an R package for simulating fossil occurrence data under mechanistic models of preservation and recovery. Methods in Ecology and Evolution, 10: 835–840

BARIDO-SOTTANI, J., VAN TIEL, N., HOPKINS, M., WRIGHT, D., STADLER, T., WARNOCK, R. (2020). Ignoring fossil age uncertainty leads to inaccurate topology and divergence time estimates in time calibrated tree inference. Frontiers in Ecology and Evolution, 8: doi.10.3389/fevo.2020.00183

BARNOSKY, A. (2001). Distinguishing the effects of the Red Queen and Court Jester on Miocene mammal evolution in the northern Rocky Mountains. Journal of Vertebrate Paleontology, 21: 172–185

BARNOSKY, A., CARRASCO, M., DAVIS, E. (2005). The impact of the species–area relationship on estimates of paleodiversity. PloS Biology, 3: e266

BARRETT, P., HASEGAWA, Y., MANABE, M., ISAJI, S., MATSUOKA, H. (2003). Sauropod dinosaurs from the Lower Cretaceous of eastern Asia: taxonomic and biogeographical implications. Palaeontology, 45: 1197–1217

BARRETT, P., LCISCIO, L., VIGLIETTI, P., BRODERICK, T., SUAREZ, C., SHARMAN, G., JONES, A., MUNYIKWA, D., EDWARDS, S., CHAPELLE, K., DOLLMAN, K., ZONDO, M., CHOINIERE, J. (2020). The age of the Tashinga Formation (Karoo Supergroup) in the Mid-Zambezi Basin, Zimbabwe and the first phytosaur from mainland sub-Saharan Africa. Gondwana Research, 81: 445–460

BEAUCHAMP, B., GRASBY, S. (2012). Permian lysocline shoaling and ocean acidification along NW Pangea led to carbonate eradication and chert expansion. Palaeogeography, Palaeoclimatology, Palaeoecology, 350: 73–90

BELL, J., DAVY, S., JONES, T., TAYLOR, M., WEBSTER, N. (2013). Could some coral reefs become sponge reefs as our climate changes? Global Change Biology, 91: 2613–2624

BELL, J., BENNETT, H., ROVELLINI, A., WEBSTER, N. (2018). Sponges to be winners under near-future climate scenarios. BioScience, 68: 855–868

BENJAMINI, Y., HOCHBERG, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society, 57: 289–300

BENSON, R. (2018). Dinosaur macroevolution and macroecology. Annual Review of Ecology, Evolution, and Systematics, 49: 379–408

BENSON, R., BUTLER, R., LINDGREN, J. SMITH, A. (2010). Mesozoic marine tetrapod diversity: mass extinctions and temporal heterogeneity in geological megabiases affecting vertebrates. Proceedings of the Royal Society B, 277: 829–834

BENSON, R., BUTLER, R., CLOSE, R., SAUPE, E. RABOSKY, D. (2021). Biodiversity across space and time in the fossil record. Current Biology, 31: 1225–1236

BENTON, M. (1995). Diversification and extinction in the history of life. Science, 268: 52–58

BENTON, M. (1997). Models for the diversification of life. Trends in Ecology and Evolution, 12, 490–495

BENTON, M. (2012). No gap in the Middle Permian record of terrestrial vertebrates. Geology, 40: 339–342

BENTON, M. (2015a). Exploring macroevolution using modern and fossil data. Proceedings of the Royal Society B, 282: 20150569

BENTON, M. (2015b). Palaeodiversity and formation counts: redundancy or bias? Palaeontology, 58: 1003–1029

BENTON, M., DUNHILL, A., LLOYD, G., MARX, F. (2011). Assessing the quality of the fossil record: insights from vertebrates. In: McGowan, A., Smith, A. (eds). Comparing the Geological and Fossil Records. Geological Society of London Special Publication, 358: 63–94

BENTON, M., RUTA, M., DUNHILL, A., SAKAMOTO, M. (2013). The first half of tetrapod evolution, sampling proxies, and fossil record quality. Palaeogeography. Palaeoclimatology. Palaeoecology, 372: 18–41

BENTON, M., BERNARDI, M., KINSELLA, C. (2018). The Carnian Pluvial Episode and the origin of dinosaurs. Journal of the Geological Society, 175: 1019–1026

BENTON, M., WILF, P., SAUQUET, H. (2022). The Angiosperm Terrestrial Revolution and the origins of modern biodiversity. New Phytologist, 233: 2017–2035

BENTON, M., WU, F. (2022). Triassic revolution. Frontiers in Earth Science, 10: doi.10.3389/feart.2022.899541

BETTS, H., PUTTICK, M., CLARK, J., WILLIAMS, T., DONOGHUE, P., PISANI, D. (2018). Integrated genomic and fossil evidence illuminates life's early evolution and eukaryote origin. Nature Ecology and Evolution, 2: 1556–1562

BERNARDI, M., KLEIN, H., PETTI, F., EZCURRA, M. (2015). The origin and early radiation of archosauriforms: integrating the skeletal and footprint record. PLoS ONE, 10: e0128449

BERNARDI, M., GIANOLLA, P., PETTI, F., MIETTO, P., BENTON, M. (2018). Dinosaur diversification linked with the Carnian Pluvial Episode. Nature Communications, 9: 1499

BERNARDI, M., PETTI, F., BENTON, M. (2018). Tetrapod distribution and temperature rise during the Permian-Triassic mass extinction. Proceedings of the Royal Society B, 285: 20172331

BEYIN, A., HALL, J., DAY, C. (2019). A Least Cost Path Model for hominin dispersal routes out of the East African Rift region (Ethiopia) into the Levant. Journal of Archaeological Science Reports, 23: 763–772

BODDY, C., MITCHELL, E., MERDITH, A., LIU, A. (2022). Palaeolatitudinal distribution of the Ediacaran macrobiota. Journal of the Geological Society, 179: jgs2021-030

BOLEY, D., RANJAN, G., ZHANG, Z. (2011). Commute times for a directed graph using an asymmetric Laplacian. Linear Algebra and its Applications, 435: 224–242

BOUCKAERT, R. (2016). Phylogeography by diffusion on a sphere: whole world phylogeography. PeerJ, 4: e2406

BOYLE, J. (2017). GeoRange: Calculating geographic range from occurrence data. R package version 0.1.0. https://cran.r-project.org/web/packages/GeoRange/index.html

BRAYARD, A., KRUMENACKER, L., BOTTING, J., JENKS, J., BYLUND, K., FARA, E., VENNIN, E., OLIVIER, N., GOUDEMAND, N., SAUCEDE, T., CHARBONNIER, S., ROMANO, C., DOGUZHAEVA, L., THUY, B., HAUTMANN, M., STEPHEN, D., THOMAZO, C., ESCARGUEL, G. (2017). Unexpected Early Triassic marine ecosystem and the rise of the Modern evolutionary fauna. Science Advances, 3: e1602159

BRIGGS, D., LIEBERMAN, B., HENDRICKS, J., HALGEDAHL, S., JARRARD, R. (2008). Middle Cambrian Arthropods from Utah. Journal of Paleontology, 82: 238–254

BRITT, B., DALLA VECCHIA, F., CHURE, D., ENGELMANN, G., WHITING, M., SCHEETZ R. (2018). *Caelestiventus hanseni* gen. et sp. nov. extends the desert-dwelling pterosaur record back 65 million years. Nature Ecology and Evolution, 2: 1386–1392

BROCKLEHURST, N. (2015). A simulation-based examination of residual diversity estimates as a method of correcting for sampling bias. Palaeontologica Electronica, 18: 1–15

BROCKLEHURST, N. (2018). An examination of the impact of Olson's extinction on tetrapods from Texas. PeerJ, 6: e4767

BROCKLEHURST, N., DAY, M., RUBIDGE, B., FROBISCH, J. (2017). Olson's Extinction and the latitudinal biodiversity gradient of tetrapods in the Permian. Proceedings of the Royal Society B, 284: doi.10.1098/rspb.2017.0231

BROCKLEHURST, N., FROBISH, J. (2018). The definition of bioregions in palaeontological studies of diversity and biogeography affects interpretations: Palaeozoic tetrapods as a case study. Frontiers in Earth Science, 6: doi.10.3389/feart.2018.00200

BROCKLEHURST, N., DAY, M., FROBISCH, J. (2018). Accounting for differences in species frequency distributions when calculating beta diversity in the fossil record. Methods in Ecology and Evolution, 9: 1409–1420

BROENNIMANN, O., DI COLA, V., GUISAN, A. (2023). ecospat: Spatial Ecology Miscellaneous Methods. R package version 3.5. https://CRAN.R-project.org/packages/ecospat/index.html

BROWN, J., LOMOLINO, M. (2001). Concluding remarks: historical perspective and the future of island biogeography theory. Global Ecology and Biogeography, 9: 87–92

BRUSATTE, S., BENTON, M., RUTA, M., LLOYD, G. (2008). Superiority, competition, and opportunism in the evolutionary radiation of dinosaurs. Science, 321: 1485–1488

BRUSATTE, S., MONTANARI, S., YI, H., NORELL, M. (2011). Phylogenetic corrections for morphological disparity analysis: new methodology and case studies. Palaeobiology, 37: 1–22

BRUSATTE, S., BUTLER, R., NIEDZWIEDZKI, G., SULEJ, T., BRONOWICZ, R., SATKUNAS, J. (2013). First record of Mesozoic terrestrial vertebrates from Lithuania: phytosaurs (Diapsida: Archosauriformes) of probable Late Triassic age, with a review of phytosaur biogeography. Geological Magazine, 150: 110–122

BRUSATTE, S., O'CONNOR, J., JARVIS, E. (2015a). The origin and diversification of birds. Current Biology, 25: 888–898

BRUSATTE, S., BUTLER, R., MATEUS, O. STEYER, J. (2015). A new species of *Metoposaurus* from the Late Triassic of Portugal and comments on the systematics and biogeography of metoposaurid temnospondyls. Journal of Vertebrate Paleontology, 35: e912988

BUFFAN, L., JONES, L., SCOTESE, C., ZAHIROVIC, S. DOMIER, M., VARELA, S. (2023). Mind the uncertainty Global plate model choice impacts deep time palaeobiological studies. Earth ArXiv, doi.10.31223/X5GD4D

BUSH, A., MARKEY, M., MARSHALL, C. (2004). Removing bias from diversity curves: the effects of spatially organized biodiversity on sampling-standardization. Paleobiology, 30: 666–686

BUTLER, R., SULLIVAN, C., EZCURRA, M., LIU, J., LECUONA, A., SOOKIAS, R. (2014). New clade of enigmatic early archosaurs yields insights into early pseudosuchian phylogeny and the biogeography of the archosaur radiation. BMC Evolutionary Biology, 14: 128

BUTTON, D., LLOYD, G., EZCURRA, M., BUTLER, R. (2017). Mass extinctions drove increased global faunal cosmopolitanism on the supercontinent Pangaea. Nature Communications, 8: 733

CALEDE, J. (2020). Pattern and Processes of the Mammalian Turnover of the Arikareean in the Northern Rocky Mountains. Journal of Vertebrate Paleontology, 40: e1767117

CAMPOY, A., ADDAMO, A. MACHORDOM, A. MEADE, M. RIVADENIERA, C. HERNANDEZ, C. VENDITTI, C. (2020). The origin and correlated evolution of symbiosis and coloniality in scleractinian corals. Frontiers in Marine Science, 7: 461

CARLISLE, E., JANIS, C., PISANI, D., DONOGHUE, P., SILVESTRO, D. (2023). A timescale for placental mammal diversification based on Bayesian modeling of the fossil record. Current Biology, 33: 1–10

CARUTENUTO, F., TSIKARIDZE, N., ROOK, L., LORDKIPANIDZE, D., LONGO, L., CONDEMI, S., RAIA, P. (2016). Venturing out safely: the biogeography of *Homo erectus* dispersal out of Africa. Journal of Human Evolution, 95: 1–12

CASTIGLIONE, S., TESONA, G., PICCOLO, M., MELCHIONNA, M., MONDANARO, A., SERIO, C., FEBBRARO, M., RAIA, P. (2018). A new method for testing evolutionary rate variation and shifts in phenotypic evolution. Methods in Ecology and Evolution, 9: 974–983

CASEY, M., FRASER, N., KOWALEWSKI, M. (2007). Quantitative Taphonomy of a Triassic reptile: *Tanytrachelos ahynis* from the Cow Branch Formation, Dan River Basin, Solite Quarry, Virginia. Palaios, 22: 598–611

CERMEÑO, P., GARCIA-COMAS, C., POHL, A., WILLIAMS, S., BENTON, M., CHAUDHARY, C., GLAND, G., MULLER, R., RIDGWELL, A., VALLINA, S. (2022). Post-extinction recovery of the Phanerozoic oceans and biodiversity hotspots. Nature, 607: 507–511

CERNY, D., MADZIA, D., SLATER, G. (2021). Empirical and methodological challenges to the model-based inference of diversification rates in extinct clades. Systematic Biology, 71: 153–171

CERNY, D., SIMONOFF, A. (2023). Statistical evaluation of character support reveals the instability of higher-level dinosaur phylogeny. Nature Scientific Reports, 13: 9273

CHADWICK, N., MORROW, K. (2010). Competition among sessile organisms on coral reefs. In: Dubinsky, Z., Stambler, N. (eds). Coral Reefs: An Ecosystem in Transition. Springer, Dortrecht, 347–371

CHAMBERLAIN. S., BARVE, V., MCGLINN, D., OLDONI, D., DESMET, P., GEFFERT, L., RAM, K. (2023). rgbif: Interface to the Global Biodiversity Information Facility API. R package version 3.7.5. https://cran.r-project.org/web/packages/rgbif/index.html

CHANDRA, A., RAGHAVAN, P., RUZZO, W., SMOLENSKY, R., TIWARI, P. (1996). The electrical resistance of a graph captures its commute and cover times. Computational Complexity, 6: 312–340

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

CHAO, A., GOTELLI, N., HSIEH, T., SANDER, E., MA, K., COLWELL, R., ELLISON, A. (2014). Rarefaction and extrapolation with Hill numbers: a framework for sampling and estimation in species diversity studies. Ecological Monographs, 84: 45–67

CHAVES-FONNEGRA, A., RIEGL, B., ZEA, S., LOPEZ, J., SMITH, T., BRANDT, M., GILLIAM, D. (2017). Bleaching events regulate shifts from corals to excavating sponges in algae-dominated reefs. Global Change Biology, 24: 773–785

CHEN, Z. BENTON, M. (2012). The timing and pattern of biotic recovery following the end-Permian mass extinction. Nature Geoscience, 5: 375–383

CIAMPAGLIO, C., KEMP, M., MCSHEA, D. (2001). Detecting changes in morphospace occupation patterns in the fossil record: characterization and analysis of measures of disparity. Paleobiology, 27: 695–715

CISNEROS, J., ABDALA, F., ATAYMAN-GUVEN, S., RUBIDGE, B., SENSOR, A., SCHULTZ, C. (2012). Carnivorous dinocephalian from the Middle Permian of Brazil and tetrapod dispersal in Pangaea. Proceedings of the National Academy of Science of the United States of America, 109: 1584–1588

CHIARENZA, A., MANNION, P., LUNT, D., FARNSWORTH, A., JONES, L., KELLAND, S., ALLISON, P. (2019). Ecological niche modelling does not support climatically driven dinosaur diversity decline before the Cretaceous/Paleogene mass extinction. Nature Communications, 10: 1091

CHIARENZA, A., FARNSWORTH, A., MANNION, P., LUNT, D., VALDES, P., MORGAN, J., ALLISON, P. (2020). Asteroid impact, not volcanism, caused the end-Cretaceous dinosaur extinction. Proceedings of the National Academy of Sciences of the United States of America, 117: 17084–17093

CHIARENZA, A., MANNION, P., FARNSWORTH, A., CARRANO, M., VARELA, S. (2022). Climatic constraints on the biogeographic history of Mesozoic dinosaurs. Current Biology, 32: 570–585

CLARAMUNT, S. (2022). CladeDate: Calibration information generator for divergence time estimation. Methods in Ecology and Evolution, 13: 2331–2338

CLEAL, C., THOMAS, B. (2010). Botanical nomenclature and plant fossils. Taxon, 59: 261–268

CLOSE, R., BENSON, R., UPCHURCH, P., BUTLER, R. (2017). Controlling for the species-area effect supports constrained long-term Mesozoic terrestrial vertebrate diversification. Nature Communications, 8: 15381

CLOSE, R., EVERS, S., ALROY, J., BUTLER, R. (2018). How should we estimate diversity in the fossil record? Testing richness estimators using sampling-standardised discovery curves. Methods in Ecology and Evolution, 9: 1386–1400

CLOSE, R., BENSON, R., ALROY, J., CARRANO, M., CLEARY, T., DUNNE, E., MANNION, P., UHEN, M., BUTLER, R. (2020a). The apparent exponential radiation of Phanerozoic land vertebrates is an artefact of spatial sampling biases. Proceedings of the Royal Society B, 287: 1924

CLOSE, R., BENSON, R., SAUPE, E., CLAPHAM, M., BUTLER, R. (2020b). The spatial structure of Phanerozoic marine animal diversity. Science, 368: 420–424

CLOWES, C., CRAMPTON, J., BLAND, B., COLLINS, K., PREBBLE, J., RAINE, J., STROGEN, D., TEREZOW, M., WOMACK, T. (2021). The New Zealand Fossil Record File: a unique database of biological history. New Zealand Journal of Geology and Geophysics, 64: 62–71

COHEN, K., FINNERY, S., GIBBARD, P., FAN, J. (2013). The ICS International Chronostratigraphic Chart. Episodes, 36: 199–204

CONDAMINE, F., GUINOT, G., BENTON, M. CURRIE, P. (2021). Dinosaur biodiversity declined well before the asteroid impact, influenced by ecological and environmental pressures. Nature Communications, 12: 3833

CONNOR, E., MCCOY, E. (1979). The statistics and biology of the species-area relationship. American Naturalist, 113: 791–833

COOPER, R., FLANNERY-SUTHERLAND, J., SILVESTRO, D. (IN REVIEW). Estimating global biodiversity patterns through time using deep learning. Nature Communications

CORNETTE, J., LIEBERMAN, B., GOLDSTEIN, R. (2002). Documenting a significant relationship between macroevolutionary origination rates and Phanerozoic pCO_2 levels. Proceedings of the National Academy of Sciences of the United States of America, 99: 7832–7835

COUCE, E., RIDGWELL, A., HENDY, E. (2012). Environmental controls on the global distribution of shallow-water coral reefs. Journal of Biogeography, 39: 1508–1523

COX, P., BETTS, R., BUNTON, C., ESSERY, R., ROWNTREE, P., SMITH, J. (1999). The impact of new land surface physics on the GCM simulation of climate and climate sensitivity, Climate Dynamics, 15: 183–203

COX, K., WOODS, M., REIMCHEN, T. (2021). Regional heterogeneity in coral species richness and hue reveals novel global predictors of reef fish intra-family diversity. Nature Scientific Reports: 11, 18275

CRACRAFT, J. (1975). Historical biogeography and earth history: perspectives for a future synthesis. Annals of the Missouri Botanical Garden, 62: 227–250

CRISP, M., TREWICK, S., COOK, L. (2011). Hypothesis testing in biogeography. Trends in Ecology and Evolution, 26: 66–72

CSARDI, G., NEPUSZ, T. (2006). The igraph software package for complex network research. InterJournal Complex Systems, 1695

DAL CORSO, J., MIETTO, P., NEWTON, R., PANCOST, R., PRETO, N., ROGHI, G., WIGNALL, P. (2012). Discovery of a major negative ^{13}C spike in the Carnian (Late Triassic) linked to the eruption of Wrangellia flood basalts, Geology, 40: 79–82

DAL CORSO, J., GIANOLLA, P., NEWTON, R., FRANCESCHI, M., ROGHI, G., CAGGIATI, M., RAUCSIK, B., BUDAI, T., HAAS, J., PRETO, N. (2015). Carbon isotope records reveal synchronicity between carbon cycle perturbation and the "Carnian Pluvial Event" in the Tethys realm (Late Triassic). Global and Planetary Change, 127: 79–90

DAL CORSO, J., SONG, H., CALLEGARO, S., CHU, D., SUN, Y., HILTON, J., GRASBY, S., JOACHIMSKI, M., WIGNALL, P. (2022). Environmental crises at the Permian–Triassic mass extinction. Nature Reviews Earth and Environmental Science, 3: 197–214

DAL CORSO, J., BERNARDI, M., SUN, Y., SONG, H., SEYFULLAH, L., PRETO, N., GIANOLLA, P., RUFFELL, A., KUSTATSCHER, E., ROGHI, G., MERICO, A., HOHN, S., SCHMIDT, A., MARZOLI, A., NEWTON, R., WIGNALL, P., BENTON, M. (2020). Extinction and dawn of the modern world in the Carnian (Late Triassic). Science Advances, 6: eaba0099

DAVY, S., ALLEMAND, D., WEIS, V. (2012). Cell biology of cnidarian-dinoflagellate symbiosis. Microbiology and Molecular Biology Reviews, 76: 229–261

DE MARCO, P., NOBREGA, C. (2018). Evaluating collinearity effects on species distribution models: an approach based on virtual species simulation. PLoS ONE, 13: e0202403

DEWEY, J., BIRD, J. (1970). Plate tectonics and geosynclines. Tectonophysics, 10: 625–638

DICKINSON, R., HENDERSON-SELLERS, A., KENNEDY, P., WILSON, M. (1986). Biosphere–Atmosphere Transfer Scheme (BATS) for the NCAR CCM NCAR/TN-275-STR. National Center for Atmospheric Research, Boulder Colorado

DIDIER, G., ROYER-CARENZI, M., LAURIN, M. (2012). The reconstructed evolutionary process with the fossil record. Journal of Theoretical Biology, 315: 26–37

DI MARTINO, E., JACKSON, J., TAYLOR, P. JOHNSON, K. (2018). Differences in extinction rates drove modern biogeographic patterns of tropical marine biodiversity. Science Advances, 4: eaaq1508

DIETL, G. (2019). Conservation palaeobiology and the shape of things to come. Philosophical Transactions of the Royal Society B, 374: 20190294

DIJKSTRA, E. (1959). A note on two problems in connexion with graphs. Numerische Mathematik, 1: 269–271

DONOGHUE, M., DOYLE, J., GAUTHIER, J., KLUGE, A., ROWE, T. (1989). The importance of fossils in phylogeny reconstruction. Annual Review of Ecology and Systematics, 20: 431–460

DORMAN, M. (2022). nngeo: k-Nearest Neighbor Join for Spatial Data. R package version 0.4.6, https://cran.r-project.org/web/packages/nngeo/index.html

DUNHILL, A., HANNISDAL, B., BENTON, M. (2014). Disentangling rock record bias and common-cause from redundancy in the British fossil record. Nature Communications, 5: 4818

DUNHILL, A., FOSTER, W., SCIBERRAS, J., TWITCHETT, R. (2017). Impact of the Late Triassic mass extinction on functional diversity and composition of marine ecosystems. Palaeontology, 61: 133–148

DUNHILL, A., HANNISDAL, B., BROCKLEHURST, N., BENTON, M. (2017). On formation-based sampling proxies and why they should not be used to correct the fossil record. Palaeontology, 61: 119–132

DUNHILL, A., FOSTER, W., AZAELE, S., SCIBERRAS, J., TWITCHETT, R. (2018). Modelling determinants of extinction across two Mesozoic hyperthermal events. Proceedings of the Royal Society B, 285: 20180404

DUNNE, E., CLOSE, R., BUTTON, D., BROCKLEHURST, N., CASHMORE, D., LLOYD, G., BUTLER, R. (2018). Diversity change during the rise of tetrapods and the impact of the ‘Carboniferous rainforest collapse’. Proceedings of the Royal Society B, 285: 20172730

DUNNE, E., FARNSWORTH, A., GREENE, S., LUNT, D., BUTLER, R. (2021). Climatic drivers of latitudinal variation in Late Triassic tetrapod diversity. Palaeontology, 64: 101–117

DUNNE, E., FARNSWORTH, A., BENSON, R., GODOY, P., GREENE, S., VALDES, P., LUNT, D., BUTLER, R. (2023). Climatic controls on the ecological ascendancy of dinosaurs. Current Biology, 33: 206–214

DUNNE, J., WILLIAMS, R., MARTINEZ, N. (2004). Network structure and robustness of marine food webs. Marine Ecology Progress, 273: 291–302

EDDELBUETTEL, D., FRANÇOIS. R. (2011). Rcpp: Seamless R and C++ Integration. Journal of Statistical Software, 40: 1–18

EDUARDO, A., MARTINEZ, P., GOUVEIA, S., SANTOS, F., DE ARAGAO, W., MORALES-BARBERO, J., KERBER, L., LIPARINI, A. (2018). Extending the paleontology–biogeography reciprocity with SDMs: Exploring models and data in reducing fossil taxonomic uncertainty. PLoS One, 13: e0194725

ELITH, J., GRAHAM, C., ANDERSON, R., DUDIK, M., FERRIER, S., GUISAN, A., HIJMANS, R., HUETTMANN, F, LEATHWICK, J., LEHMANN, A., LI, J., LOHMANN, L., LOISELLE, B., MANION, G., MORITZ, C., NAKAMURA, M., NAKAZAWA, Y., OVERTON, J., PETERSON, T., PHILLIPS, S., RICHARDSON, K., SCACHETTI-PEREIRA, R., SCHAPIRE, R., SOBERON, J., WILLIAMS, S., WISZ M., ZIMMERMAN, N. (2006). Novel methods improve prediction of species' distributions from occurrence data. Ecography, 29: 129–151

ELITH, J., KEARNEY, M., PHILLIPS, S. (2010). The art of modelling range-shifting species. Methods in Ecology and Evolution, 1: 330–342

ELSLER, A. (2019). Macroevolution of Early Tetrapods. Unpublished PhD thesis, School of Earth Sciences, University of Bristol

EMIG, C. (2003). Proof that Lingula (Brachiopoda) is not a living-fossil, and emended diagnoses of the Family Lingulidae. Carnets de Geologie, L01: 1–8

ERICSON, J., GOLDSTEIN, R. (1980). Work space: a new approach to the analysis of energy expenditure within site catchments. In: Findlow, F., Ericson, J. (eds.). Catchment Analysis: Essays On Prehistoric Resource Space. Department of Anthropology, University of California, Los Angeles, 21–30

ETHERINGTON, T. (2016). Least-cost modelling and landscape ecology: concepts, applications, and opportunities. Current Landscape Ecology Reports, 1: 40–53

EVANS, S., KING, M. (1993). A new specimen of Protorosaurus (Reptilia: Diapsida) from the Marl Slate (late Permian) of Britain. Proceedings of the Yorkshire Geological Society, 49: 229–234

EVANS, M., SMITH, S., FLYNN, R., DONOGHUE, M. (2009). Climate, niche evolution, and diversification of the "bird-cage" evening primroses (Oenothera, sections Anogra and Kleinia). The American Naturalist, 173: 225–240

EZAKI, Y. (1997). The Permian coral *Numidiaphyllum*: new insights into anthozoan phylogeny and Triassic scleractinian origins. Palaeontology, 40: 1–14

EZAKI, Y. (2000). Palaeoecological and phylogenetic implications of a new scleractiniamorph genus from Permian sponge reefs, south China. Palaeontology, 43: 199–217

EZARD, T., PEARSON, O., AZE, T. PURVIS, A. (2012). The meaning of birth and death (in macroevolutionary birth–death models). Biology Letters, 8: 139–142

EZCURRA, M. (2010). Biogeography of Triassic tetrapods: evidence for provincialism and driven sympatric cladogenesis in the early evolution of modern tetrapod lineages. Proceedings of the Royal Society B, 277: 2547–2552

EZCURRA, M. (2016). The phylogenetic relationships of basal archosauromorphs, with an emphasis on the systematics of proterosuchian archosauriforms. PeerJ, 4: e1778

EZCURRA, M., SCHEYER, T., BUTLER, R. (2014). The origin and early evolution of Sauria: reassessing the Permian saurian fossil record and the timing of the crocodile-lizard divergence. PLoS ONE, 9: e89165

EZCURRA, M., FIORELLI, L., MARTINELLI, A., ROCHER, S., VON BACZKO, M., EZPELETA, M., TABORDA, J., HECHENLEITNER, E., TROTTEYN, M., DESOJO, J. (2017). Deep faunistic turnovers preceded the rise of dinosaurs in southwestern Pangaea. Nature Ecology and Evolution, 1: 1477–1483

EZCURRA, M., NESBITT, S., BRONZATI, M., DALLA VECCHIA, F., AGNOLIN, F., BENSON, R., EGLI, F., CABREIRA, S., EVERS, S., GENTIL, A., IRMIS, R., MARTINELLI, A., NOVAS, F., DA SILVA, L., SMITH, N., STOCKER, M., TURNER, A., LANGER, M. (2020). Enigmatic dinosaur precursors bridge the gap to the origin of Pterosauria. Nature, 588: 445–449

EZCURRA, M., JONES, A., GENTIL, A., BUTLER, R. (2021). Early Archosauromorphs: the crocodile and dinosaur precursors. In: Alderton, D., Elias, S. (eds). Encyclopedia of Geology, 2nd edition, Academic Press, United Kingdom, 175–185

FALKOWSKI, P., SCHOFIELD, O., KATZ, M., VAN DE SCHOOTBRUGGE, B., KNOLL, A. (2004). Why is the land green and the ocean red? In: Thierstein, H., Young J. (eds). Coccolithophores: From Molecular Processes to Global Impact. Springer-Verlag, Berlin Heidelberg, 429–453

FAN, J., SHEN, S., ERWIN, D., SADLER, P., MACLEOD, N., CHENG, Q., HOU, X., YANG, J., WANG, X., WANG, Y., ZHANG, H., CHEN, X., LI, G., ZHANG, Y., SHI, Y., YUAN, D., CHEN, Q., ZHANG, L., LI, C., ZHAO, Y. (2020). A high-resolution summary of Cambrian to Early Triassic marine invertebrate biodiversity. Science, 367: 272–277

FENG, X., PARK, D., LIANG, Y., PANDEY, R., PAPES, M. (2019). Collinearity in ecological niche modeling: confusions and challenges. Ecology and Evolution, 9: 10365–10376

FENTON, I., WOODHOUSE, A., AZE, T., LAZARUS, D., RENAUDIE, J., DUNHILL, A., YOUNG, J., SAUPE, E. (2021). Triton, a new species-level database of Cenozoic planktonic foraminiferal occurrences. Nature Scientific Data, 8: 160

FLANNERY-SUTHERLAND, J., RAJA, N., KOCSIS, A., KIESSLING, W. (2022a). fossilbrush: an R package for automated detection and resolution of errors in palaeontological occurrence data. Methods in Ecology and Evolution, 13: 2404–2418

FLANNERY-SUTHERLAND, J., SILVESTRO, D., BENTON, M. (2022b). Global diversity dynamics in the fossil record are regionally heterogeneous. Nature Communications, 13: 275

FLÜGEL, E. (2002). Triassic reef patterns. In: Flügel, E., Kiessling, W. (eds). Patterns of Phanerozoic Reef Crises. Society of Economic Palaeontologists and Mineralogists Special Publication, 72: 391–464

FLÜGEL, E., STANLEY, G. (1984). Reorganization, development and evolution of post-Permian reefs and reef organisms. Palaeontographica Americana, 54: 177–186

FLÜGEL, E., KIESSLING. W. (2002). Patterns of Phanerozoic reef crises. Society of Economic Palaeontologists and Mineralogists Special Publication, 72: 691–733

FOFFA, D., DUNNE, E., NESBITT, S., BUTLER, R., FRASER, N., BRUSATTE, S., FARNSWORTH, A., LUNT, D., VALDES, P., WALSH, S., BARRETT, P. (2022). *Scleromochlus* and the early evolution of Pterosauromorpha, Nature, 610: 313–318

FOOTE, M. (1994). Temporal variation in extinction risk and temporal scaling of extinction metrics. Paleobiology, 20: 424–444

FOOTE, M. (1996). On the probability of ancestors in the fossil record. Paleobiology, 22: 141–151

FOOTE, M. (2000a). Origination and extinction components of taxonomic diversity: general problems. Paleobiology, 26: 74–102

FOOTE, M. (2000b). Origination and extinction components of taxonomic diversity: Paleozoic and post-Paleozoic dynamics. Paleobiology, 26: 578–605

FOOTE, M. (2005). Pulsed origination and extinction in the marine realm. Paleobiology, 31: 6–20

FOOTE, M. (2007). Symmetric waxing and waning of marine invertebrate genera. Paleobiology, 33: 517–529

FOOTE, M., CRAMPTON, J., BEU, A., MARSHALL, B., COOPER, R., MAXWELL, P., MATCHAM, I. (2007). Rise and fall of species occupancy in Cenozoic fossil mollusks. Science, 318: 1131–1134

FOREY, P., FORTEY, R., KENRICK, P., SMITH, A. (2009). Taxonomy and fossils: a critical appraisal. Philosophical Transactions of the Royal Society of London B, 359: 649–653

FORTEY, R., COCKS, L. (1988). Biogeography and palaeogeography of the Sibumasu terrane in the Ordovician: a review. In: Hall, R., Holloway, J. (eds). Biogeography and Geological Evolution of SE Asia. Backhuys Publishers, Leiden, Netherland, 43–56

FOSTER, G., ROYER, D. LUNT, D. (2017). Future climate forcing potentially without precedent in the last 420 million years. Nature Communications, 8: 14845

FOSTER, W., HEINDEL, K., RICHOZ, S., GLIWA, J., LEHRMANN, D., BAUD, A., KOLAR-JURKOVSEK, T., ALJINOVIC, D., JURKOVSEK, B., KORN, D., MARTINDALE, R., PECKMANN, J. (2019). Suppressed competitive exclusion enabled the proliferation of Permian/Triassic boundary microbialites. Depositional Record, 6: 62–74

FOSTER, W., ALLEN, B., KITZMANN, N., MUNCHMEYER, J., RETTELBACK, T., WITTS, J., WHITTLE, R., LARINA, E., CLAPHAM, M., DUNHILL, A. (2023). How predictable are mass extinction events? Royal Society Open Science, 10: 221507

FOTH, C., EZCURRA, M., SOOKIAS, R., BRUSATTE, S., BUTLER, R. (2016). Unappreciated diversification of stem archosaurs during the Middle Triassic predated the dominance of dinosaurs. BMC Evolutionary Biology, 16: 188

FOTH, C., SOOKIAS, R., EZCURRA, M. (2021). Rapid initial morphospace expansion and delayed morphological disparity peak in the first 100 million years of the archosauromorph evolutionary radiation. Frontiers in Earth Science, 9: doi.10.3389/feart.2021.723973

FRANKOWIAK, K., RONIEWICZ, E., STOLARSKI, J. (2021). Photosymbiosis in Late Triassic scleractinian corals from the Italian Dolomites. PeerJ, 9: e11062

GAGNÉ, T., REYGONDEAU, G., JENKINS, C., SEXTON, J., BOGRAD, S., HAZEN, E., HOUTAN, K. (2020). Towards a global understanding of the drivers of marine and terrestrial biodiversity. PloS ONE, 15: e0228065

GALLAGHER, A., CREEL, S., WILSON, R., COOKE, S. (2017). Energy landscapes and the landscape of fear. Trends in Ecology and Evolution, 32: 88–96

GARDNER, J., SURYA, K., ORGAN, C. (2019). Early tetrapodomorph biogeography: controlling for fossil record bias in macroevolutionary analyses. Comptus Rendus Palevol, 18: 699–709

GASTON, K. (2000). Global patterns in biodiversity. Nature, 405: 220–227

GATTOLIN, G., PRETO, N., BREDA, A., FRANCESCHI, M. ISOTTON, M., GIANOLLA, P. (2015). Sequence stratigraphy after the demise of a high-relief carbonate platform (Carnian of the Dolomites): sea-level and climate disentangled. Palaeogeography, Palaeoclimatology, Palaeoecology, 423: 1–17

GAVRYUSHKINA, A., ZHANG, C. (2021). Total-evidence dating and the fossilized birth–death model. In: Ho, S. (ed). The Molecular Evolutionary Clock, Springer Nature, 175–193

GELMAN, A., RUBIN, D. (1992). Inference from iterative simulation using multiple sequences. Statistical Science, 7: 457–472

GJESFELD, E., SILVESTRO, D., CHANG, J., KOCH, B., FOSTER, J., ALFARO, M. (2020). A quantitative workflow for modeling diversification in material culture. PLoS ONE, 15: e0227579

GLEASON, H. (1922). On the relation of species and area. Ecology 3: 158–162

GOLONKA, J. (2007). Late Triassic and Early Jurassic palaeogeography of the world. Palaeogeography, Palaeoclimatology, Palaeoecology, 244: 297–307

GOMBIN, J., VAIDYANATHAN, R., AGAFONKIN, V. (2020). concaveman: a very fast 2D concave hull algorithm. R package version 1.1.0, https://cran.r-project.org/web/packages/concaveman/index.html

GOUGH, D. (1981). Solar interior structure and luminosity variations. Solar Physics, 74: 21–34

GOULD, S. (2002). The Structure of Evolutionary Theory. Belknap Press, Harvard

GRADSTEIN, F., OGG, J. SCHMITZ, M., OGG, G. (2020). A Geologic Timescale. Elsevier, Amsterdam

GRAHAM, R., LUNDELIUS, E. (2010). FAUNMAP II: New data for North America with a temporal extension for the Blancan, Irvingtonian and early Rancholabrean. FAUNMAP II Database, version 1.0, https://ucmp.berkeley.edu/faunmap/index.html

GREEN, S., BORUFF, B., BONNELL, T., GRUETER, C. (2020). Chimpanzees use least-cost routes to out-of-sight goals. Current Biology, 30: 4528–4533

GRENIÉ, M., BERTI, E., CARVAJAL-QUINTERO, J., SAGOUIS, A., WINTER, M. (2022). Harmonizing taxon names in biodiversity data: a review of tools, databases, and best practices. Methods in Ecology and Evolution, 14: 12–25

GRIFFIN, C., WYND, B., MUNYIKWA, D., BRODERICK, T., ZONDO, M., TOLAN, S., LANGER, M., NESBITT, S., TARUVINGA, H. (2022). Africa's oldest dinosaurs reveal early suppression of dinosaur distribution. Nature, 609: 313–319

GUILLON, W., BROWN, J. (2021). A new method for integrating ecological niche modeling with phylogenetics to estimate ancestral distributions. Systematic Biology, 70: 1033–1045

GUPTA A., MANCEAU M., VAUGHAN T., KHAMMASH M., STADLER T. (2020). The probability distribution of the reconstructed phylogenetic tree with occurrence data. Journal of Theoretical Biology, 488: 110115

GUINDAN, S., DE MAIO, N. (2021). Accounting for spatial sampling patterns in Bayesian phylogeography. Proceedings of the National Academy of Sciences of the United States of America, 118: e2105273118

GUISAN, A., ZIMMERMANN, N. (2000). Predictive habitat distribution models in ecology. Ecological Modelling, 135: 147–186

HAGEN, O., FLUCK, B., FOPP, F., CABRAL, J., HARTIG, F., PONTARP, M., RANGEL, T., PELLISSIER, L. (2021a). gen3sis: A general engine for eco-evolutionary simulations of the processes that shape earth's biodiversity. PLoS Biology, 19: e3001340

HAGEN, O., SKEELS, A., ONSTEIN, R., JETZ, W., PELLISSIER. L. (2021b). Earth history events shaped the evolution of uneven biodiversity across tropical moist forests. Proceedings of the National Academy of Sciences of the United States of America, 118: e2026347118

HAGEN, O. (2022). Coupling eco-evolutionary mechanisms with deep-time environmental dynamics to understand biodiversity patterns. Ecography, 4: e06132

HALLAM, A. (1981). Major bio-events in the Triassic and Jurassic. In: Walliser O. (ed). Global Events and Event Stratigraphy in the Phanerozoic, Springer Berlin, Heidelberg, 265–283

HANCOCK, A., LEHMBERG, E., BLACKMON, H. (2022). Phylogenetics in space: how continuous spatial structure impacts tree inference. Molecular Phylogenetics and Evolution, 173: 107505

HANNISDAL, B., PETERS, S. (2011). Phanerozoic Earth-system evolution and marine biodiversity. Science, 334: 1121–1124

HARTMAN, S., LOVELACE, D., LINZMEIER, B., MATHEWSON, P., PORTER, W. (2022). Mechanistic thermal modelling of Late Triassic terrestrial amniotes predicts biogeographic distribution. Diversity, 14: 973

HARVEY, P., MAY, R., NEE, S. (1994). Phylogenies without fossils. Evolution, 48: 523–529

HEATH, T., HUELSENBECK, J., STADLER, T. (2014). The fossilized birth–death process for coherent calibration of divergence-time estimates. Proceedings of the National Academy of Sciences of the United States of America, 111: 2957–2966

HEATH, T., WRIGHT, A., PETT, W. (2019). Combined-Evidence Analysis and the Fossilized Birth-Death Process for Stratigraphic Range Data. Github tutorial (accessed 03/04/23), https://revbayes.github.io/tutorials/fbd/#Hoehna2014b

HEY, J. (1992). Using phylogenetic trees to study speciation and extinction. Evolution, 46: 627–640

HIJMANS, R. (2019). geosphere: spherical trigonometry. R package version 1.5.10. https://cran.r-project.org/web/packages/geosphere/index.html

HIJMANS, R., GRAHAM, C. (2006). Testing the ability of climate envelope models to predict the effect of climate change on species distributions. Global Change Biology, 12: 2272–2281

HIJMANS, R., PHILLIPS, S., LEATHWICK, J., ELITH, J. (2022). dismo: Species Distribution Modeling. R package version 1.3-9. https://cran.r-project.org/web/packages/dismo/index.html

HIJMANS, R. (2023). raster: Geographic Data Analysis and Modeling. R package version 3.6.20, https://cran.r-project.org/web/packages/raster/index.html

HMICH, D., SCHNEIDER, J. W., SABER, H., VOIGT, S. EL WARTITI, M. (2006). New continental Carboniferous and Permian faunas of Morocco: implications for biostratigraphy, palaeobiogeography and palaeoclimate. In: Lucas, S., Schneider, J., Cassinis, G. (eds). Non-Marine Permian Biostratigraphy and Biochronology: An Introduction. Geological Society of London Special Publication, 265: 297–324

HSIEH, T., MA, K. CHAO, A. (2020). iNEXT: iNterpolation and EXTrapolation for species diversity. R package version 2.0.20. https://cran.r-project.org/web/packages/iNEXT/index.html

HUANG, Y., CHEN, Z., ZHAO, L., STANLEY, G., YAN, J., PEI, Y., YANG, W., HUANG, J. (2019). Restoration of reef ecosystems following the Guadalupian–Lopingian boundary mass extinction: Evidence from the Laibin area, South China. Paleogeography, Palaeoclimatology, Palaeoecology, 519: 8–22

HUNTER, A., DONOVAN, S. (2007). Field sampling bias, museum collections and completeness of the fossil record. Lethaia, 38: 305–314

HUSSON, J., PETERS, S. (2018). Nature of the sedimentary rock record and its implications for Earth-system evolution. Emerging Topics in Life Sciences, 28: 125–136

HUSSON, L., SALLES, T., LEBATARD, A., ZERATHE, S., BRAUCHER, R., NOERWIDI, S., ARIBOWO, S., MALLARD, C., CARCAILLET, J., NATAWIDJAJA, D., AUMAITRE, G., BOURLES, D., KEDDADOUCHE, K. (2022). Javanese *Homo erectus* on the move in SE Asia circa 1.8 Ma. Nature Scientific Reports, 12: 19012

IPCC CLIMATE CHANGE (2022). Impacts, Adaptation, and Vulnerability. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge

JABLONSKI, D. (2008). Extinction and the spatial dynamics of biodiversity. Proceedings of the National Academy of Sciences of the United States of America, 12: 11528–11535

JACKSON, J. (1977). Competition on marine hard substrata: the adaptive significance of solitary and colonial strategies. American Naturalist, 111: 743–767

JAGIELSKA, N., BRUSATTE, S. (2021). Pterosaurs. Current Biology, 31: 973–992

JETZ, W., THOMAS, G., HARTMANN, K., MOONERS, A. (2012). The global diversity of birds in space and time. Nature, 491: 444–448

JIANG, H., YUAN, J., CHEN, Y., OGG, J., YAN, J. (2019). Synchronous onset of the Mid-Carnian Pluvial Episode in the East and West Tethys: conodont evidence from Hanwang, Sichuan, South China. Palaeogeography, Palaeoclimatology, Palaeoecology, 520: 173–180

JOACHIMSKI, M., LAI, X., SHEN, S., JIANG, H., LUO, G., CHEN, B., CHEN, J., SUN, Y. (2012). Climate warming in the latest Permian and the Permian-Triassic mass extinction. Geology, 40: 195–198

JONES, L., MANNION, P., FARNSWORTH, A., VALDES, P., KELLAND, S., ALLISON, P. (2019). Coupling of palaeontological and neontological reef coral data improves forecasts of biodiversity responses under global climatic change. Royal Society Open Science, 6: 182111

JONES, L., DEAN, C., MANNION, P., FARNSWORTH, A., ALLISON, P. (2021). Spatial sampling heterogeneity limits the detectability of deep time latitudinal biodiversity gradients. Proceedings of the Royal Society B, 288: 20202762

JONES, L., EICHENSEER, K. (2022). Uneven spatial sampling distorts reconstructions of Phanerozoic seawater temperature. Geology, 50: 238–242

JONES, L., MANNION, P., FARNSWORTH, A., BRAGG, F., LUNT, D. (2022). Climatic and tectonic drivers shaped the tropical distribution of coral reefs. Nature Communications, 13: 3120

JONES, L., GEARTY, W., ALLEN, B., EICHENSEER, K., DEAN, C., GALVÁN S., KOUVARI, M., DUNNE, E., GODOY, P., NICHOLL, C., FLANNERY-SUTHERLAND, J., DILLON, E., CHIARENZA, A. (2023). palaeoverse: a community-driven R package to support palaeobiological analyses. Methods in Ecology and Evolution, doi.10.1111/2041-210X.14099

KAMMERER, C., NESBITT, S., FLYNN, J., RANIVOHARIMANANA, L., WYSS, A. (2020). A tiny ornithodiran archosaur from the Triassic of Madagascar and the role of miniaturization in dinosaur and pterosaur ancestry. Proceedings of the National Academy of Sciences of the United States of America, 117: 17932–17936

KAMMERER, C., VIGLIETTI, P., BUTLER, E., BOTHA, J. (2023). Rapid turnover of top predators in African terrestrial faunas around the Permian-Triassic mass extinction. Current Biology, 33: 2283–2290

KASS, R., RAFTERY, A. 1995). Bayes Factors. Journal of the American Statistical Association, 90: 773–795

KEARNEY, M., WINTLE, B., PORTER, W. (2010). Correlative and mechanistic models of species distribution provide congruent forecasts under climate change. Conservation Letters, 3: 203–213

KENDALL, D. (1948). On the generalised "birth-and-death" process. The Annals of Mathematical Statistics, 19: 1–15

KIDWELL, S., FLESSA, K. (1995). The quality of the fossil record: populations, species and communities. Annual Review in Ecology and Systematics, 26: 269–299

KIDWELL, S. (2015). Biology in the Anthropocene: Challenges and insights from young fossil records. Proceedings of the National Academy of Sciences of the United States of America, 112: 4922–4929

KIESSLING, W. (2009). Reef expansion during the Triassic: spread of photosymbiosis balancing climatic cooling. Palaeogeography, Palaeoclimatology, Palaeoecology, 290: 11–19

KIESSLING, W. (2010). Geologic and biologic controls on the evolution of reefs. Annual Review of Ecology and Evolution, 40: 173–192

KIESSLING, W., ABERHAN, M., BRENNEIS, B., WAGNER, P. (2007). Extinction trajectories of benthic organisms across the Triassic-Jurassic boundary. Palaeogeography, Palaeoclimatology, Palaeoecology, 224: 201–222

KIESSLING, W., SIMPSON, C., FOOTE, M. (2010). Reefs as cradles of evolution and sources of biodiversity in the Phanerozoic. Science, 327: 196–198

KIESSLING, W., SIMPSON, C. (2010). On the potential for ocean acidification to be a general cause of ancient reef crises. Global Change Biology, 17: 56–67

KIESSLING, A. KOCSIS, A. (2015). Biodiversity dynamics and environmental occupancy of fossil azooxanthellate and zooxanthellate scleractinian corals. Paleobiology, 41: 1–13

KIESSLING, W., RAJA, N., RODEN, V., TURVEY, S., SAUPE, E. (2019). Addressing priority questions of conservation science with palaeontological data. Philosophical Transactions of the Royal Society B, 374: 1788

KIESSLING, W., KRAUSE, M. (2022). PARED - An Online Database of Phanerozoic Reefs. https://www.paleo-reefs.pal.uni-erlangen.de

KING, B., RUCKLIN, M. (2020). Tip dating with fossil sites and stratigraphic sequences. PeerJ, 8: 32617191

KNOLL, A., FOLLOWS, M. (2016). A bottom-up perspective on ecosystem change in Mesozoic oceans. Proceedings of the Royal Society B, 283: 20161755

KOCH, B., SILVESTRO, D., FOSTER, J. (2021). The evolutionary dynamics of cultural change (as told through the birth and brutal, blackened death of metal music). SocArXiv, doi.10.31235/osf.io/659bt

KOCSIS, Á. (2021). icosa: global triangular and penta-hexagonal grids based on tessellated icosahedra. R package version 0.10.1. https://cran.r-project.org/web/packages/icosa/index.html

KOCSIS, Á., RAJA, N. (2019). chronosphere: Earth-system history variables. R package version 0.5.0. https://cran.r-project.org/web/packages/chronosphere/index.html

KOCSIS, Á., REDDIN, C., KIESSLING, W. (2018a). The biogeographical imprint of mass extinctions. Proceedings of the Royal Society B, 285: 1878

KOCSIS, Á., REDDIN, C., KIESSLING, W. (2018b). The stability of coastal benthic biogeography over the last 10 million years. Global Ecology and Biogeography, 27: 1106–1120

KOCSIS, Á., REDDIN, C., ALROY, J., KIESSLING, W. (2019). The R package divDyn for quantifying diversity dynamics using fossil sampling data. Methods in Ecology and Evolution, 10, 735–743

KOCSIS, A., REDDIN, C., SCOTESE, C., VALDES, P., KIESSLING, W. (2021). Increase in marine provinciality over the last 250 million years governed more by climate change than plate tectonics. Proceedings of the Royal Society B, 288: 20211342

KOCSIS, Á., SCOTESE, C. (2021). Mapping paleocoastlines and continental flooding during the Phanerozoic. Earth-Science Reviews, 213: 103463

KOEN, E., BOWMAN, J., WALPOLE, A. (2012). The effect of cost surface parameterization on landscape resistance estimates. Molecular Ecology Resources, 12: 686–696

KOLIS, K., LIEBERMAN, B. (2019). Using GIS to examine biogeographic and macroevolutionary patterns in some late Paleozoic cephalopods from the North American Midcontinent Sea. PeerJ, 7: e6910

KOLODZIEJ, B., SALAMON, K., MORYCOWA, E., SZULC, J., LABAJ, M. (2018). Platy corals from the Middle Triassic of Upper Silesia, Poland: implications for photosymbiosis in the first scleractinians. Palaeogeography, Palaeoclimatology, Palaeoecology, 490: 533–545

KOPPERUD, B., LIDGARD, S., LIOW, L. (2018). Text-mined fossil biodiversity dynamics using machine learning. Proceedings of the Royal Society B, 286: 20190022.

KOPPERUD, B., MAGEE, A., HOHNA, S. (2023). Rapidly changing speciation and extinction rates can be inferred in spite of nonidentifiability. Proceedings of the National Academy of Sciences of the United States of America, 120: e2208851120

KRUG, A., JABLONSKI, D. (2012). Long-term origination rates are reset only at mass extinctions. Geology, 40: 731–734

KUHNERT, D., STADLER, T., VAUGHAN, T., DRUMMOND, A. (2016). Phylodynamics with migration: a computational framework to quantify population structure from genomic data. Molecular Biology and Evolution, 33: 2102–2116

LANDWEHRS, J., FEULNER, G., PETRI, S., SAMES, B., WAGREICH, M. (2021a). Investigating Mesozoic climate trends and sensitivities with a large ensemble of climate model simulations. Paleoceanography, Paleoclimatology, 36: e2020PA004134

LANDWEHRS, J., FEULNER, G., PETRI, S., SAMES, B., WAGREICH, M. (2021). Data from Climate Model Ensemble Simulations for the Mesozoic Climate Evolution. GFZ Data Services, 2021-02-04, https://doi.org/10.5880/PIK.2020.009

LARMET, V. (2022). cppRouting: Algorithms for Routing and Solving the Traffic Assignment Problem. R package version 3.1, https://cran.r-project.org/web/packages/cppRouting/index.html

LAZARUS, D., WEINKAUF, M., DIVER, P. (2012). Pacman profiling: a simple procedure to identify stratigraphic outliers in high-density deep-sea microfossil data. Paleobiology, 38: 144–161

LEGG, D., SUTTON, M., EDGECOMBE, G. (2013). Arthropod fossil data increase congruence of morphological and molecular phylogenies. Nature Communications, 4: 2485

LEHTONEN, S., SILVESTRO, D. KARGER, D., SCOTESE, C., TUOMISTO, H., KESSLER, M., PENA, C., WAHLBERG, N., ANTONELLI, A. (2017). Environmentally driven extinction and opportunistic origination explain fern diversification patterns. Nature Scientific Reports, 7: 4831

LEMMON, A., LEMMON, E. (2008). A likelihood framework for estimating phylogeographic history on a continuous landscape. Systematic biology, 57: 544–561

LIEBERMAN, B. (2003). Paleobiogeography: the relevance of fossils to biogeography. Annual review of Ecology, Evolution and Systematics, 34: 51–69

LINDNER, A., CAIRNS, S., CUNNINGHAM, C. (2008). From offshore to onshore: multiple origins of shallow-water corals from deep-sea ancestors. PLoS One, 3: e2429

LIOW, L., QUENTAL, T., MARSHALL, C. (2010a). When can decreasing diversification rates be detected with molecular phylogenies and the fossil record? Systematic Biology, 59: 646–659

LIOW, L., SKAUG, H., ERGON, T., SCHWEDER, T. (2010). Global occurrence trajectories of microfossils: environmental volatility and the rise and fall of individual species, Paleobiology, 36: 224–252

LIU, J., ANGIELCZYK, K., ABDALA, F. (2021). Permo-Triassic tetrapods and their climate implications. Global and Planetary Change, 205: 103618

LLOYD, G. (2012). A refined modelling approach to assess the influence of sampling on palaeobiodiversity curves: new support for declining Cretaceous dinosaur richness. Biology Letters, 8: 123–126

LLOYD, G., YOUNG, J., SMITH, A. (2012). Taxonomic structure of the fossil record is shaped by sampling bias. Systematic Biology, 61: 80–89

LLOYD, G., SLATER, G. (2021). A total-group phylogenetic metatree for Cetacea and the importance of fossil data in diversification analyses. Systematic Biology, 70: 922–939

LOMOLINO, M. (2000). Ecology's most general, yet protean pattern: The species-area relationship. Journal of Biogeography, 27: 17–26

LOPEZ-ANTONANZAZ, R., MTITCHELL, J., SOMOES, T., CONDAMINE, F., AGUILEE, R., PELAEZ-CAMPOMANES, P., RENAUD, S., ROLLAND, J., DONOGHUE, P. (2022). Integrative phylogenetics: tools for palaeontologists to explore the tree of life. Biology (Basel), 11: 1185

LOUCA, S., PENNELL, M. (2020). Extant timetrees are consistent with a myriad of diversification histories. Nature, 580: 502–505

LOUCA, S., PENNELL, M. (2021). Why extinction estimates from extant phylogenies are so often zero. Current Biology, 31: 3168–3173

LOVELACE, D., HARTMAN, S., MATHEWSON, P, LINZMEIER, B., PORTER, W. (2020). Modeling dragons: using linked mechanistic physiological and microclimate models to explore environmental, physiological, and morphological constraints on the early evolution of dinosaurs. PLoS ONE, 15: e0223872

LUNT, D., FARNSWORTH, A., LOPTSON, C., FOSTER, G, MARKWICK, P., O'BRIEN, C., PANCOST, R., ROBINSON, S., WROBEL, N. (2016). Palaeogeographic controls on climate and proxy interpretation. Climates of the Past, 12: 1181–1198

LYSTER, S., WHITTAKER, A., ALLISON, P., LUNT, D., FARNSWORTH, A. (2020). Predicting sediment discharges and erosion rates in deep time – examples from the late Cretaceous North American continent. Basin Research, 32: 1547–1573

MADIN, J., ANDERSON, K., ANDREASEN, M., BRIDGE, T., CAIRNS, S., CONNOLLY, S., DARLING, E., DIAZ, M., FALSTER, D., FRANKLIN, E., GATES, R., HARMER, A., HOOGENBOOM, M., HUANG, D., KEITH, S., KOSNIK, M., KUO, C., LOUGH, J., LOVELOCK, C., LUIZ, O., MARTINELLI, J., MIZEREK, T., PANDOLFI, J., POCHON, X., PRATCHETT, M., PUTNAM, H., ROBERTS, T., STAT, M., WALLACE, C., WIDMAN, E., BAIRD, A. (2016). The Coral Trait Database, a curated

database of trait information for coral species from the global oceans. Nature Scientific Data, 3: 160017

MAGALLON, S., SANDERSON, D. (2001). Absolute diversification rates in angiosperm clades. Evolution, 55: 1762–1780

MACARTHUR, R., WILSON, E. (1967). The Theory of Island Biogeography. Princeton University Press, Princeton

MACPHERSON, A., LOUCA, S., MCLAUGHLIN, A., JOY, J., PENNELL, M. (2021). Unifying phylogenetic birth–death models in epidemiology and macroevolution. Systematic Biology, 71: 172–189

MANGERUD, G., PATERSON, N., RIDING, J. (2019). The temporal and spatial distribution of Triassic dinoflagellate cysts. Review of Palaeobotany and Palynology, 261: 53–66

MANNION, P., UPCHURCH, P., BENSON, R., GOSWAMI, A. (2013). The latitudinal biodiversity gradient through deep time. Trends in Ecology and Evolution, 29: 42–50

MANNION, P., BENSON, R., CARRANO, M., TENNANT, J., JUDD, J., BUTLER, R. (2015). Climate constrains the evolutionary history and biodiversity of crocodylians. Nature Communications, 6: 8438.

MARKWICK, P. (2019). Palaeogeography in exploration. Geological Magazine, 156: 366–407

MARKWICK, P., VALDES, P. (2004). Palaeo-digital elevation models for use as boundary conditions in coupled ocean–atmosphere GCM experiments: a Maastrichtian (late Cretaceous) example. Palaeogeography, Palaeoclimatology, Palaeoecology, 213: 37–63

MARSHALL, C. (1990). Confidence intervals on stratigraphic ranges. Paleobiology, 16; 1–10

MARSHALL, C. (1994). Confidence intervals on stratigraphic ranges: partial relaxation of the assumption of randomly distributed fossil horizons. Paleobiology, 20: 459–469

MARSHALL, C. (1995). Confidence intervals on stratigraphic ranges with non-random distributions of fossil horizons. Paleobiology, 23: 165 – 173

MARSHALL, C. (2017). Five palaeobiological laws needed to understand the evolution of the living biota. Nature Ecology and Evolution, 1: 165

MARSHALL, C. (2019). Using the fossil record to evaluate timetree timescales. Frontiers in Genetics, 10: 1049

MARSHALL, C., FINNEGAN, S., CLITES, E., HOLROYD, P., BONUSO, N., CORTEZ, C., DAVIS, E., DIETL, G., DRUCKENMILLER, P., ENG, R., GARCIA, C., ESTES-SMARGIASSI, K., HENDY, A., HOLLIS, K., LITTLE, H., NESBITT, E., ROOPNARINE, P., SKIBINSKI, L., VENDETTI, J., WHITE, D. (2018). Quantifying the dark data in museum fossil collections as palaeontology undergoes a second digital revolution. Biology Letters, 14: 20180431

MARSHALL, J., LAKIN, J., TROTH, I., WALLACE-JOHNSON, S. (2021). UV-B radiation was the Devonian-Carboniferous boundary terrestrial extinction kill mechanism. Science Advances, 6: eaba0768

MARSOLA, J., FERREIRA, G., LANGER, M., BUTTON, D., BUTLER, R. (2018). Increases in sampling support the southern Gondwanan hypothesis for the origin of dinosaurs. Palaeontology, 62: 473–482

MARTINDALE, R. BERELSON, R., CORSETTI, F., BOTTJER, D., WEST, A. (2012). Constraining carbonate chemistry at a potential ocean acidification event (the Triassic–Jurassic boundary) using the presence of corals and coral reefs in the fossil record. Palaeogeography, Palaeoclimatology, Palaeoecology, 350: 114–123

MARTINDALE, R., FOSTER, W., VELLEDITS, F. (2019). The survival, recovery, and diversification of metazoan reef ecosystems following the end-Permian mass extinction event. Palaeogeography, Palaeoclimatology, Palaeoecology, 513: 100–115

MARTINEZ-PEREZ, C., PLASENCIA, P., CASCALES-MINANA, B., MAZZA, M., BOTELLA, H. (2014). New insights into the diversity dynamics of Triassic conodonts. Historical Biology, 26: 591–602

MATZKE, N. (2013). Probabilistic historical biogeography: new models for founder-event speciation, imperfect detection, and fossils allow improved accuracy and model-testing. Frontiers of Biogeography, 5: 242–248

MATZKE, N., WRIGHT, A. (2016). Inferring node dates from tip dates in fossil Canidae: the importance of tree priors. Biology Letters, 12: 20160328

MAZAHERI-JOHARI, M., GIANOLLA, P., MATHER, T., FRIELING, J., CHU, D., DAL CORSO, J. (2021). Mercury deposition in Western Tethys during the Carnian Pluvial Episode (Late Triassic). Nature Scientific Reports, 11: 17339

MCEWEN, J., WALLIS, C., DOCHERTY, M. (2022). Machine learning assisted Bayesian model comparison: learnt harmonic mean estimator. aXiv, 2111.12720v2

MCGHEE JR., G., SHEEHAN, P., BOTTJER, D., DROSER, M. (2004). Ecological ranking of Phanerozoic biodiversity crises: ecological and taxonomic severities are decoupled. Palaeogeography, Palaeoclimatology, Palaeoecology, 4: 289–297

MCGOWAN, A., SMITH A. (2007). The shape of the Phanerozoic diversity curve. how much can be predicted from the sedimentary rock record of Western Europe? Palaeontology, 50: 765–777

MCGOWAN, A., SMITH, A. (2008). Are global Phanerozoic marine diversity curves truly global? A study of the relationship between regional rock records and global Phanerozoic marine diversity. Paleobiology, 34: 80–103

MCGOWAN, A., SMITH A. (2011). Comparing the geological and fossil records: implications for biodiversity studies. Geological Society of London Special Publications, 358

MCHARG, I. (1967). Where should highways go? Landscape Architecture, 57: 179–181

MESEGUER, A., LOBO, J., REE, R., BEERLING, D., SANMARTIN, I. (2015). Integrating fossils, phylogenies, and niche models into biogeography to reveal ancient evolutionary history: the case of *Hypericum* (Hypericaceae). Systematic Biology, 64: 215–232

METCALFE, I., CROWLEY, J. (2021). Upper Permian and Lower Triassic conodonts, high-precision U-Pb zircon ages and the Permian-Triassic boundary in the Malay Peninsula. Journal of Asian Earth Sciences, 199: 104403

MEYER, A., WIENS, J. (2017). Estimating diversification rates for higher taxa: BAMM can give problematic estimates of rates and rate shifts. Evolution, 72: 39–53

MICHELS, E., COTTENIE, K., NEYS, K., DE GELAS, K., COPPIN, P., DE MEESTER, L. (2001). Geographical and genetic distances among zooplankton populations in a set of interconnected ponds: a plea for using GIS modelling of the effective geographical distance. Molecular Ecology, 10: 1929–1938

MILLS, B., DONNADIEU, Y., GODDERIS, Y. (2021). Spatial continuous integration of Phanerozoic global biogeochemistry and climate. Gondwana Research, 100: 73–86

MILLS, M., KRAUSE, A., JARVIS, I., CRAMER, B. (2023). Evolution of atmospheric O2 through the Phanerozoic, revisited. Annual Review of Earth and Planetary Sciences, 51: 253–276

MITCHELL, J., ETIENNE, R., RABOSKY, D. (2019). Inferring diversification rate variation from phylogenies with fossils. Systematic Biology, 68: 1–18

MIYASHITA, T., COATES, M., FARRAR, R., LARSON, P., MANNING, P., WOGELIUS, R., EDWARDS, N., ANNE, J., BERGMANN, U., PALMER, A., CURRIE, P. (2019). Hagfish from the Cretaceous Tethys Sea and a reconciliation of the morphological–molecular conflict in early vertebrate phylogeny. Proceedings of the National Academy of Sciences of the United States of America, 116: 2146–2151

MOHARREK, F., TAYLOR, P., SILVESTRO, D., JENKINS, H., GORDON, D., WAESCHENBACK, A. (2022). Diversification dynamics of cheilostome Bryozoa based on a Bayesian analysis of the fossil record. Palaeontology, 65: e12586

MONTOYA, M., GRIESEL, A., LEVERMANN, A., MIGNOT, J., HOFMANN, M., GANOPOLSKI, A., RAHMSTORF, S. (2005). The Earth-system model of intermediate complexity CLIMBER-3α. Part I: description and performance for present-day conditions. Climate Dynamics, 25: 237–263

MORALES-BARBERO, J., VEGA-ALVAREZ, J. (2018). Input matters matter: bioclimatic consistency to map more reliable species distribution models. Methods in Ecology and Evolution, 10: 212–224

MORLON, H., PARSON, T., PLOTKIN, J. (2011). Reconciling molecular phylogenies with the fossil record. Proceedings of the National Academy of Sciences of the United States of America, 108: 16327–16332

MORLON, H. (2014). Phylogenetic approaches for studying diversification. Ecology Letters, 17: 508–525

MORLON, H., ROBIN, S., HARTIG, F. (2022). Studying speciation and extinction dynamics from phylogenies: addressing identifiability issues. Trends in Ecology and Evolution, 37: 497–506

MUELLER, T., FAGAN, W. (2008). Search and navigation in dynamic environments: from individual behaviours to population distributions. Oikos, 117: 654–664

MULLER, R., ROEST, W., ROYER, J., GAHAGAN, L., SCLATER, J. (1997). Digital isochrons of the world's ocean floor. Journal of Geophysical Research, 102: 3211–3214

MULLER, R., GARCIA, M. (2020). A paraphyletic ‘Silesauridae' as an alternative hypothesis for the initial radiation of ornithischian dinosaurs. Biology Letters, 16: 20200417

MULLER, R, GARCIA, M. (2023). A new silesaurid from Carnian beds of Brazil fills a gap in the radiation of avian line archosaurs. Nature Scientific Reports, 13: 4981

MUSCENTE, A., SCHIFFBAUER, J., BROCE, J., LAFLAMME, M., O’DONNELL, K., BOAG, T., MYEYER, M., HAWKINS, A., MCNAMARA, M., MCKENZIE, L., STANLEY, G., HINMAN, N., HOFMANN, X. (2017). Exceptionally preserved fossil assemblages through geologic time and space. Gondwana Research, 48: 164–188

MUSCENTE, A., PRABHU, A., ZHING, H., ELEISH, A., MEYER, M., FOX, P., HAZEN, R., KNOLL, A. (2018). Quantifying ecological impacts of mass extinctions with network analysis of fossil communities. Proceedings of the National Academy of the United States of America, 115: 517–522

NEAL, R. (1994). Contribution to the discussion of “approximate Bayesian inference with the weighted likelihood bootstrap” by Newton MA, Raftery AE. Journal of the Royal Statistical Society Series A (Methodological), 56: 41–42

NEE, S., MOOERS, A., HARVEY, P. (1992). Tempo and mode of evolution revealed from molecular phylogenies. Proceedings of the National Academy of Sciences of the United States of America, 89: 8322–8326

NEE, S., HOLMES, E., MAY, R., HARVEY, P. (1994a). Extinction rates can be estimated from molecular phylogenies. Philosophical Transactions of the Royal Society of London B, 344: 77–82

NEE, S., MAY, R., HARVEY, P. (1994b). The reconstructed evolutionary process. Philosophical Transactions of the Royal Society of London B, 344: 305–311

NESBITT, S. (2011). The early evolution of archosaurs: relationships and the origin of major clades. Bulletin of the American Museum of Natural History, 352: 1–292

NEWELL, N. (1959). The Nature of the Fossil Record. Proceedings of the American Philosophical Society, 103: 264–285

NIX, H. (1986). A biogeographic analysis of Australian elapid snakes. In: Longmore, R. (ed). Atlas of Elapid Snakes of Australia. Australian Flora and Fauna Series Number 7. Australian Government Publishing Service: Canberra, 4–15

NORMAN, K., CHAMBERLAIN, S., BOETTIGER, C. (2020). taxadb: A high-performance local taxonomic database interface. Methods in Ecology and Evolution, 11: 1153–1159

NORMAN, D., BARON, M., GARCIA, M., MULLER, R. (2022). Taxonomic, palaeobiological and evolutionary implications of a phylogenetic hypothesis for Ornithischia (Archosauria: Dinosauria). Zoological Journal of the Linnean Society, 196: 1273–1309

NOVACK-GOTSHALL, P. (2016). Love, not war, drove the Mesozoic marine revolution. Proceedings of the National Academy of Sciences of the United States of America, 113: 14471–14473

NOW COMMUNITY (2020). New and Old Worlds Database of Fossil Mammals (NOW). https://nowdatabase.org/now/database/

O'DONOVAN, C., MEADE, A., VENDITTI, C. (2018). Dinosaurs reveal the geographical signature of an evolutionary radiation. Nature Ecology and Evolution, 2: 452–458

O'REILLY, J., DOS REIS, M., DONOGHUE, P. (2015). Dating tips for divergence-time estimation. Trends in Genetics, 31: 637–650

OLROYD, S., SIDOR, C. (2017). A review of the Guadalupian (middle Permian) global tetrapod fossil record. Earth-Science Reviews, 171: 583–597

OLSEN, P., SHA, J., FANG, Y., CHANG, C., WHITESIDE, J., KINNEY, S., SUES, H., KENT, D., SCHALLER, M., VAJDA, V. (2022). Arctic ice and the ecological rise of the dinosaurs. Science Advances, 8: eabo6342

OYSTON, J., WILKINSON, M., RUTA, M., WILLS, M. (2022). Molecular phylogenies map to biogeography better than morphological ones. Communications Biology, 5: 521

PADGHAM, M. (2019). dodgr: an R package for network flow aggregation. Transport Findings, 2: doi.10.32866/6945

PAGEL, M., MEADE, A., BARKER, D. (2004). Bayesian estimation of ancestral character states on phylogenies. Systematic Biology, 53: 673–684

PARADIS, E., SCHLIEP, K. (2019). ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. Bioinformatics, 35: 526–528

PARDO, J., HUTTENLOCKER, A., MARCOT, J. (2008). Stratocladistics and evaluation of evolutionary modes in the fossil record: an example from the ammonite genus *Semiformiceras*. Palaeontology, 51: 767–773

PARRISH, J. (1993). Climate of the supercontinent Pangea. The Journal of Geology, 101: 215–233

PARRISH, J., ZIEGLER, A., SCOTESE, C. (1982). Rainfall patterns and the distribution of coals and evaporites in the Mesozoic and Cenozoic. Palaeogeography, Palaeoclimatology, Palaeoecology, 40: 67–101

PAYNE, J., LEHRMANN, D., WEI, J., KNOLL, A. (2006). The pattern and timing of biotic recovery from the End-Permian extinction on the Great Bank of Guizhou, Guizhou Province, China. Palaios, 21: 63–85

PAYNE, J. VAN DE SCHOOTBRUGGE, B. (2007). Triassic life in the oceans: links between planktonic and benthic recovery and radiation. In: Falkowski, P., Knoll, A. (eds). Evolution of primary producers in the sea. Academic Press, Amsterdam, 165–189

PEBESMA, E. (2018). Simple features for R: standardized support for spatial vector data. The R Journal, 10: 439–446

PEBESMA, E., BIVAND, R. (2005). Classes and methods for spatial data in R. R News, 5: 9–13

PEBESMA, E. (2018). Simple Features for R: Standardized Support for Spatial Vector Data. The R Journal, 10: 439-446,

PEBESMA, E., BIVAND, R. (2023). Spatial data science: with applications in R. Chapman and Hall/CRC

PENN-CLARKE, C., HARPER, D. (2021). Early–Middle Devonian brachiopod provincialism and bioregionalization at high latitudes: A case study from southwestern Gondwana. Geological Society of America Bulletin, 133: 819–836

PERCIVAL, L., RUHL, M., HESSELBO, S., JENKYNS, H., MATHER, T., WHITESIDE, J. (2017). Mercury evidence for pulsed volcanism during the end-Triassic mass extinction. Proceedings of the National Academy of Sciences of the United States of America, 114: 7929–7934

PETERS, R. (1991). A critique for ecology. Cambridge University Press, Cambridge

PETERS, S. (2005). Geologic constraints on the macroevolutionary history of marine animals. Proceedings of the National Academy of Sciences of the United States of America, 102: 12326–12331.

PETERS, S. (2006). Genus extinction, origination, and the durations of sedimentary hiatuses. Paleobiology, 32: 387–407

PETERS, S., HEIM, N. (2012). Stratigraphic distribution of marine fossils in North America. Geology, 39: 259–262

PETERS, S., ZHANG, C., LIVNY, M., RE., C. (2014). A machine reading system for assembling synthetic paleontological databases. PloS ONE, 9: e113523

PETERS, S., MCCLENNEN, M. (2016). The Paleobiology Database application programming interface. Paleobiology, 42: 1–7

PETERSON, A., SOBERÓN, J., PEARSON, R., ANDERSON, R., MARTÍNEZ-MEYER, E., NAKAMURA, M. ARAÚJO, M. (2011). Ecological niches and geographic distributions. Princeton University Press, Princeton

PETERSON, A., ANAMZA, T. (2015). Ecological niches and present and historical geographic distributions of species: a 15-year review of frameworks, results, pitfalls, and promises. Folia Zoologica, 64: 207–218

PHILLIPS, S., ANDERSON, R., SCHAPIRE, R. (2006). Maximum entropy modelling of species geographic distributions. Ecological Modelling, 190: 231–259

PHILLIPS, S., ANDERSON, R., DUDÍK, M., SCHAPIRE, R., BLAIR, M. (2017). Opening the black box: an open-source release of MaxEnt. Ecography, 40: 887–893

PILOWSKY, J., COLWELL, R., RAHBEK, C., FORDHAM, D. (2022). Process-explicit models reveal the structure and dynamics of biodiversity patterns. Science Advances, 8, eabj2271

PIMIENTO, C., ANTONELLI. A. (2022). Integrating deep-time palaeontology in conservation prioritisation. Frontiers in Ecology and Evolution, 10: 959364

PINTO, N., KEITT, T. (2009). Beyond the least-cost path: evaluating corridor redundancy using a graph-theoretic approach. Landscape Ecology, 24: 253–266

PLAISANCE, L., CALEY, M., BRAINARD, R., KNOWLTON, N. (2011). The diversity of coral reefs: what are we missing? PloS ONE, 6: e25026

PLOTNICK, R., WAGNER, P. (2006). Round up the usual suspects: common genera in the fossil record and the nature of wastebasket taxa. Paleobiology, 32: 126–146

PLUMMER, M., BEST, N., COWLES, K., VINES, K. (2006). coda: Convergence diagnosis and output analysis for MCMC. R News, 6: 7–11

POWELL, M. (2007). Geographic range and genus longevity of late Paleozoic brachiopods. Paleobiology, 33: 530–546

POWELL, M., MOORE, B., SMITH, T. (2015). Origination, extinction, invasion, and extirpation components of the brachiopod latitudinal biodiversity gradient through the Phanerozoic Eon. Paleobiology, 41: 330–341

PRADELLI, L., LEARDI, J., EZCURRA, M. (2021). Body size disparity of the archosauromorph reptiles during the first 90 million years of their evolution. Ameghiniana, 59: 47–77

PRETO, N., KUSTATSCHER, E., WIGNALL, P. (2010). Triassic climates – state of the art and perspectives. Palaeogeography, Palaeoclimatology, Palaeoecology, 290: 1–10

PRETO, N., WILLEMS, H., GUAIUMI, C., WESTPHAL, H. (2013). Onset of significant pelagic carbonate accumulation after the Carnian Pluvial Event (CPE) in the western Tethys. Facies, 59: 891–914

PROTHERO, D. (2015). Garbage in, garbage out: the effect of immature taxonomy on database compilations of North American fossil mammals. New Mexico Museum of Natural History and Science Bulletin, 68: 257–264

PULLIAM, H. (2002). On the relationship between niche and distribution. Ecology Letters, 3: 349–361

PYRON, R. (2011). Divergence time estimation using fossils as terminal taxa and the origins of Lissamphibia. Systematic Biology, 60: 466–481

QUENTAL, T., MARSHALL, C. (2009). Extinction during evolutionary radiations: reconciling the fossil record with molecular phylogenies. Evolution, 63: 3158–3167

QUENTAL, T., MARSHALL, C. (2010). Diversity dynamics: molecular phylogenies need the fossil record. Trends in Ecology and Evolution, 25: 434–441

QUIGG, A., FINKEL, Z., IRWIN, A., ROSENTHAL, Y., HO, T., REINFELDER, J., SCHOFIELD, O., MOREL, F., FALKOWSKI, P. (2003). The evolutionary inheritance of elemental stoichiometry in marine phytoplankton. Nature, 425: 291–294

R CORE TEAM (2022). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

RABOSKY, D. (2010). Extinction rates should not be estimated from molecular phylogenies. Evolution, 64: 1816–1824

RABOSKY, D. (2014). Automatic detection of key innovations, rate shifts, and diversity-dependence on phylogenetic trees. PloS One, 9: e89543

RABOSKY, D., BENSON, R. (2021). Ecological and biogeographic drivers of biodiversity cannot be resolved using clade age-richness data. Nature Communications, 12: 2945

RAFTERY, A., NEWTON, M., SATAGOPAN, J., KRIVITSKY, P. (2007). Estimating the integrated likelihood via posterior simulation using the harmonic mean identity. In: Bernardo, J., Bayarri, M., Berger, J., Daid, A., Heckerman, D., Smith, A., West, M. (eds). Bayesian Statistics. Oxford University Press, Oxford, 1–145

RAJA, N., DUNNE, E., MATIWANE, A., MING KHAN, T., NATSCHER, P., GHILARDI, A., CHATTOPADHYAY, D. (2022). Colonial history and global economics distort our understanding of deep-time biodiversity. Nature Ecology and Evolution, 6: 145–154

RAJA, N., DIMITRIJEVIĆ, D., KRAUSE, M., KIESSLING, W. (2022). Ancient Reef Traits, a database of trait information for reef-building organisms over the Phanerozoic. Nature Scientific Data, 9: 425

RAJA, N., PANDOLFI, J., KIESSLING, W. (2023). Modularity explains large-scale reef booms in Earth's history. Facies, 69: 15

RAMBAUT, A., DRUMMOND, A., XIE, D., BAELE, G., SUCHARD, M. (2018). Posterior summarisation in Bayesian phylogenetics using Tracer 1.7. Systematic Biology, 67: 901–904

RANGEL, T., EDWARDS, N., HOLDEN, P., DINIZ-FILHO, J., GOSLING, W., COELHO, M., CASSEMIRO, F., RAHBEK, C., COLWELL, R. (2018). Modeling the ecology and evolution of biodiversity: biogeographical cradles, museums, and graves. Science, 361, 361: eaar5452

RAUP, D. (1972). Taxonomic diversity during the Phanerozoic. Science, 177: 1065–1071

RAUP, D. (1975). Taxonomic diversity estimation using rarefaction. Paleobiology, 1: 333–342

RAUP, D. (1976). Species diversity in the Phanerozoic: an interpretation. Paleobiology, 2: 289–297

RAUP, D. (1985). Mathematical models of cladogenesis. Paleobiology, 11: 42–52

RAUP, D., GOULD, S., SCHOPF, T., SIMBERLOFF, D. (1973). Stochastic models of phylogeny and the evolution of diversity. The Journal of Geology, 81: 525–542

RAYFIELD, E., BARRETT, P., McDONNELL, R., WILLIS, K. (2005). A Geographical Information System (GIS) study of Triassic vertebrate biochronology. Geological Magazine, 142: 327–354

REVELL, L. (2012). phytools: An R package for phylogenetic comparative biology (and other things). Methods in Ecology and Evolution, 3: 217–223

RIBEIRO, B., VELAZCO, S., GUIDONI-MARTINS, K., TESSAROLO, G., JARDIM, L., BACHMAN, S., LOYOLA, R. (2022). bdc: A toolkit for standardizing, integrating and cleaning biodiversity data. Methods in Ecology and Evolution, 13: 1421–1428

RICKETTS, T. (2001). The matrix matters: effective isolation in fragmented landscapes. American Naturalist, 158: 87–99

RIDGEWELL, A. ZEEBE, R. (2005). The role of the global carbonate cycle in the regulation and evolution of the Earth-system. Earth and Planetary Science Letters, 234: 299–315

RIDGWELL, A., HARGREAVES, J., EDWARDS, N., ANNAN, J., LENTON, T., MARSH, R., YOOL, A., WATSON, A. (2007). Marine geochemical data assimilation in an

efficient Earth-system model of global biogeochemical cycling. Biogeosciences, 4: 87–104

RIES, J. (2009). Review: the effects of secular variation in seawater Mg/Ca on marine biocalcification. Biogeosciences Discussions, 6: 7325–7452

RODE, A., LIEBERMAN, B. (2005). Integrating evolution and biogeography: a case study involving Devonian crustaceans. Journal of Paleontology, 79: 267–276

ROJAS, A., CALATAYUD, J., KOWALEWSKI, M., NEUMAN, M., ROSVALL, M. (2021). A multiscale view of the Phanerozoic fossil record reveals the three major biotic transitions. Communications Biology, 4: 309

ROMANO, M., BERNARDI, M., PETTI, F., RUBIDGE, B., HANCOX, J., BENTON, M. (2020). Early Triassic terrestrial tetrapod fauna: a review. Earth Science Reviews, 210: 10331

RONQUIST, F., KLOPFSTEIN, S., VILHELMSEN, L., SCHULMEISTER, S., MURRAY, D., RASNITSYN, A. (2012). A total-evidence approach to dating with fossils, applied to the early radiation of the Hymenoptera. Systematic Biology, 61: 973–999

RONQUIST, F., HUELSENBECK, J. (2003). MRBAYES 3: Bayesian phylogenetic inference under mixed models. Bioinformatics, 19: 1572–1574

ROOPNARINE, P., ANGIELCZYK, J., WANG, S., HERTOG, R. (2006). Trophic network models explain instability of Early Triassic terrestrial communities. Proceedings of the Royal Society B, 274, 2077–2086

ROOPNARINE, P., ANGIELCZYK, K. (2011). The evolutionary palaeoecology of species and the tragedy of the commons. Biology Letters, 8: 147–150

ROSCHER, M., SCHNEIDER, J. (2006). Permo-Carboniferous climate: Early Pennsylvanian to Late Permian climate development of central Europe in a regional and global context. Geological Society of London Special Publication, 265: 95–136

RUDWICK, M. (1997). Georges Cuvier, Fossil Bones, and Geological Catastrophes. University of Chicago Press, Chicago

SAHNEY, S., BENTON, M., FALCON-LANG, H. (2010). Rainforest collapse triggered Carboniferous tetrapod diversification in Euramerica. Geology, 38: 1079–1082

SAITO, R., KAIHO, K., OBA, M., TONG, J., CHEN, Z., TAKAHASHI, S., CHEN, J., TIAN, L., BISWAS, R. (2016). Secular changes in environmental stresses and eukaryotes during the Early Triassic to the early Middle Triassic. Palaeogeography, Palaeoclimatology, Palaeoecology, 451: 35–45

SAKAMOTO, M., VENDITTI, C., BENTON, M. (2016). 'Residual diversity estimates' do not correct for sampling bias in palaeodiversity data. Methods in Ecology and Evolution, 8: 453–459

SALLES, T., HUSSON L., REY, P., MALLARD, C., ZAHIROVIC, S., BOGGIANI, B., COLTICE, N., ARNOULD M. (2023). Hundred million years of landscape dynamics from catchment to global scale. Science, 379: 918–923

SANDERSON, M., DONOGHUE, M. (1996). Reconstructing shifts in diversification rates on phylogenetic trees. Trends in Ecology and Evolution, 11: 15–20

SAUPE, E., FARNSWORTH, A., LUNT, D., FIELD, D. (2019). Climatic shifts drove major contractions in avian latitudinal distributions throughout the Cenozoic. Proceedings of the National Academy of Sciences of the United States of America, 116: 12895–12900

SAUPE, E., QIAO, H., DONNADIEU, Y., FARNSWORTH, A., KENNEDY-ASSER, A., LADANT, J., LUNT, POHL, A., VALDES, P., FINNEGAN, S. (2020). Extinction intensity during Ordovician and Cenozoic glaciations explained by cooling and palaeogeography. Nature Geoscience, 13: 65–70

SCHLUTER, D., PENNELL, M. (2017). Speciation gradients and the distribution of biodiversity. Nature, 546: 48–55

SCOTESE, C. (1976). A continental drift 'flip book'. Computers and Geology, 2: 13–116

SCOTESE, C., BAKER, C. (1975). Continental drift reconstructions and animation. Journal of Geological Education, 23: 167–171

SCRUTTON, C. (1993). New kilbuchophyllid corals from the Ordovician of the Southern Uplands, Scotland. Courier Forschungs-Institut Senckenberg, 164: 153–158

SCRUTTON, C., CLARKSON E. (1991). A new scleractinian-like coral from the Ordovician of the Southern Uplands, Scotland. Palaeontology, 34: 179–194

SCRUTTON., C., JERAM, A., ARMSTRONG, H. (1997). Kilbuchophyllid corals from the Ordovician (Caradoc) of Pomeroy, Co. Tyrone: implications for coral phylogeny and for movement on the Southern Uplands Fault. Earth and Environmental Transactions of the Royal Society of Edinburgh, 88: 117–126

SEELEY, H. (1887). On the classification of the fossil animals commonly named Dinosauria. Proceedings of the Royal Society of London, 43: 165–171

SEIBLITZ, I., CAPEL, K., STOLARSKI, J., QUEK, Z., HUANG, D., KITAHARA, M. (2020). The earliest diverging extant scleractinian corals recovered by mitochondrial genomes. Nature Scientific Reports, 10: 20714

SELLWOOD, B., VALDES, P. (2006). Mesozoic climates: general circulation models and the rock record. Sedimentary Geology, 190: 269–287

SEPKOSKI, J. (1976). Species diversity in the Phanerozoic: species-area effects. Paleobiology, 2: 298–303

SEPKOSKI, J. (1978). A kinetic model of Phanerozoic taxonomic diversity I. analysis of marine orders. Paleobiology, 4: 223–251

SEPKOSKI, J. (1979). A kinetic model of Phanerozoic taxonomic diversity II. early Phanerozoic families and multiple equilibria. Paleobiology, 5: 222–251

SEPKOSKI, J. (1981). A factor analytic description of the Phanerozoic marine fossil record. Paleobiology, 7: 36–53

SEPKOSKI, J. (1984). A kinetic model of Phanerozoic taxonomic diversity III: post Palaeozoic families and mass extinctions. Paleobiology, 10: 246–267

SEPKOSKI, J. (1993). Ten years in the library: new data confirm palaeontological patterns. Paleobiology, 19: 43–51

SEPKOSKI, J., BAMBACH, R., RAUP, D., VALENTINE, J. (1981). Phanerozoic marine diversity and the fossil record. Nature, 293: 435–437

SEPKOSKI, J., RAUP, D. (1981). Periodicity in marine mass extinctions. In: Elliot, D. (ed). Dynamics of extinction. John Wiley and Sons, New York, 3–36

SETON, M., MÜLLER, R., ZAHIROVIC, S., WILLIAMS, S., WRIGHT, N., CANNON, J., WHITTAKER, J., MATTHEWS, K., MCGIRR, R. (2020). A global dataset of present-

day oceanic crustal age and seafloor spreading parameters. Geochemistry, Geophysics, Geosystems, 21: e2020GC009214

SAGOO, N., VALDES, P., FLECKER, R., GREGOIRE, L. (2013). The Early Eocene equable climate problem: can perturbations of climate model parameters identify possible solutions? Philosophical Transactions of the Royal Society A, 371: 20130123

SETON, M., WILLIAMS, S., DOMEIER, M., COLLINS, A., SIGLOCK, K. (2023). Deconstructing plate tectonic reconstructions. Nature Reviews Earth and Environment, 4: 185–204

SHARONI, S., HALEVY, I. (2023). Rates of seafloor and continental weathering govern Phanerozoic marine phosphate levels. Nature Geosciences, 16: 75–81

SHAW, J., BRIGGS, D. HULL, P. (2021). Fossilization potential of marine assemblages and environments. Geology, 49: 258–262

SHEEHAN, P. (1985). Reefs are not so different – they follow the evolutionary pattern of level-bottom communities. Geology, 13: 46–49

SIDOR, C., O'KEEFE, F., DAMIANI, R., STEYER, J., SMITH, R., LARSSON, H., SERENO, P., IDE, O., MAGA, A. (2005). Permian tetrapods from the Sahara show climate-controlled endemism in Pangaea. Nature, 434: 886–889

SIDOR, C., VILHENA, D., ANGIELCZYK, K., HUTTENLOCKER, A., NESBITT, S., PEECOOK, B., STEYER, J., SMITH, R., TSUJI, L. (2013). Provincialization of terrestrial faunas following the end-Permian mass extinction. Proceedings of the National Academy of Sciences of the United States of America, 110: 8129–8133

SIGNOR, P., LIPPS, J. (1982). Sampling bias, gradual extinction patterns, and catastrophes in the fossil record. In: Silver, L., Schultz, P. (eds). Geological Implications of Impacts of Large Asteroids, Comets on the Earth. Geological Society of America Special Publication, 190: 291–296

SILVESTRO, D., SCHNITZLER, J., ZIZKA, G. (2011). A Bayesian framework to estimate diversification rates and their variation through time and space. BMC Evolutionary Biology, 11: 311

SILVESTRO, D., SALAMIN, N., SCHNITZLER, J. (2014a). PyRate: a new program to estimate speciation and extinction rates from incomplete fossil data. Methods in Ecology and Evolution, 5, 1126–1131

SILVESTRO, D., SCHNITZLER, J., LI, L., ANTONELLI, A., SALAMIN, N. (2014b). Bayesian estimation of speciation and extinction from incomplete fossil occurrence data. Systematic Biology, 63: 349–367

SILVESTRO, D., CASCALES-MINANA, B., BACON, C., ANTONELLI, A. (2015). Revisiting the origin and diversification of vascular plants through a comprehensive Bayesian analysis of the fossil record. New Phytologist, 207: 425–436

SILVESTRO, D., ANTONELLI., A., SALAMIN, N., QUENTAL, T. (2015). The role of clade competition in the diversification of North American canids. Proceedings of the National Academy of Sciences of the United States of America, 112: 8684–8689

SILVESTRO, D., ZIZKA, A., BACON, C., CASCALES-MINANA, B., SALAMIN, N., ANTONELLI, A. (2016). Fossil biogeography: a new model to infer dispersal, extinction and sampling from palaeontological data. Philosophical Transactions of the Royal Society B: Biological Sciences, 371: 20150225

SILVESTRO, D., PIRES, M., QUENTAL, T., SALAMIN, N. (2017). Bayesian estimation of multiple clade competition from fossil data. Evolutionary Ecology Research, 18: 41–59

SILVESTRO, D., WARNOCK, R., GAVRYUSHKINA, A., STADLER, T. (2018). Closing the gap between palaeontological and neontological speciation and extinction rate estimates. Nature Communications, 9: 5237

SILVESTRO D., SALAMIN N., ANTONELLI A., MEYER X. (2019). Improved estimation of macroevolutionary rates from fossil data using a Bayesian framework. Paleobiology, 45: 546–570

SILVESTRO, D., BACON, C., DING, W., ZHANG, Q., DONOGHUE, P., ANTONELLI, A., XING, Y. (2021). Fossil data support a pre-Cretaceous origin of flowering plants. Nature Ecology and Evolution, 5: 449–457

SIMMS M., RUFFELL A. (1989). Synchroneity of climatic change and extinctions in the late Triassic. Geology, 17: 265–268

SIMMS, M., RUFFELL, A., WIGNALL, P. (2015). The Carnian humid episode of the Late Triassic: a review. Geological Magazine, 153: 271–284

SIMÕES, T., KAMMERER, C., CALDWELL, M., PIERCE, S. (2022). Successive climate crises in the deep past drove the early evolution and radiation of reptiles. Science Advances, 8: eabq1898

SIMPSON, C., KIESSLING, W. MEWIS, H., BARON-SZABO, R., MULLER, J. (2011). Evolutionary diversification of reef corals: a comparison of the molecular and fossil records. Evolution, 65: 3274–3284

SINGH, S., ELSLER, A., STUBBS, T., BOND, R., RAYFIELD, E., BENTON, M. (2021). Niche partitioning shaped herbivore macroevolution through the early Mesozoic. Nature Communications, 12: 2796

SMITH, A. (1988). Late Palaeozoic biogeography of East Asia and palaeontological constraints on plate tectonic reconstructions. Philosophical Transactions of the Royal Society A, 326: 189–227

SMITH, A. (2001). Large-scale heterogeneity of the fossil record: implications for Phanerozoic biodiversity studies. Philosophical Transactions of the Royal Society of London Series B, 356: 351–367

SMITH, A., BRIDEN, J., DREWRY, G. (1973). Phanerozoic world maps. In: Hughes, N. (ed). Organisms and Continents Through Time. Special Papers in Palaeontology, 12: 1-44

SMITH, C., LAVILLE, T., FARA, E., ESCARGUEL, G., OLIVIER, N., VENNIN, E., GOUDEMAND, N., BYLUND, K., JENKS, J., STEPHEN, D., HAUTMANN, M., CHARBONNIER, S., KRUMENACKER, L., BRAYARD, A. (2021). Exceptional fossil assemblages confirm the existence of complex Early Triassic ecosystems during the early Spathian. Nature Scientific Reports, 11: 19657

SOBERÓN, J., PETERSON, A. (2005). Interpretation of models of fundamental ecological niches and species' distributional areas. Biodiversity Informatics, 2: 1–10

SOLOW, A. (2003). Estimation of stratigraphic ranges when fossil finds are not randomly distributed. Paleobiology, 29: 181–185

SOLOW, A., ROBERTS, D. (2003). A nonparametric test for extinction based on a sighting record. Ecology, 84: 1329–1332

SONG, H., WIGNALL, P., DUNHILL, A. (2018). Decoupled taxonomic and ecological recoveries from the Permo-Triassic extinction. Science Advances, 4: eaat5091

SONG, H., HUANG, S., JIA, E., DAI, X., WIGNALL, P., DUNHILL, A. (2020). Flat latitudinal diversity gradient caused by the Permian–Triassic mass extinction. Proceedings of the National Academy of Sciences of the United States of America, 117: 17578–17583

SONG, H., HUANG, S., JIA, E., DAI, X., WIGNALL, P., DUNHILL, H. (2020). Flat latitudinal diversity gradient caused by the Permian–Triassic mass extinction. Proceedings of the National Academy of Sciences of the United States of America, 117: 17578–17583

SOLÓRZANO, A., NÚÑEZ-FLORES, M., INOSTROZA-MICHAEL, O., HERNÁNDEZ, C. (2020). Biotic and abiotic factors driving the diversification dynamics of Crocodylia. Palaeontology, 63: 415–429

SPIRIDONOV, A., LOVEJOY. S. (2022). Life rather than climate influences diversity at scales greater than 40 million years. Nature, 607: 307–312

STADLER, T. (2009). On incomplete sampling under birth–death models and connections to the sampling-based coalescent. Journal of Theoretical Biology, 261: 58–66

STADLER, T. (2010). Sampling-through-time in birth–death trees. Journal of Theoretical Biology, 267: 396–404

STADLER, T. (2013). Recovering speciation and extinction dynamics based on phylogenies. Journal of Evolutionary Biology, 26: 1203–1219

STADLER, T., KÜHNERT, D., BONHOEFFER, S., DRUMMOND, A. (2013). Birth–death skyline plot reveals temporal changes of epidemic spread in HIV and hepatitis c virus (HCV). Proceedings of the National Academy of Sciences of the United States of America, 110: 228–233

STADLER T., GAVRYUSHKINA, A., WARNOCK, R., DRUMMOND, A., HEATH, T. (2018). The fossilized birth-death model for the analysis of stratigraphic range data under different speciation modes. Journal of Theoretical Biology, 447: 41–55

STANLEY, G., HARDIE, L. (1998). Secular oscillations in carbonate mineralogy of reef-building and sediment-producing organisms driven by tectonically forced shifts in seawater chemistry. Paleogeography, Paleoclimatology, Palaeoecology, 144: 3–19

STANLEY, G., HARDIE, L. (1999). Hypercalcification: paleontology links plate tectonics and geochemistry to sedimentology. GSA Today, 9: 1–7

STANLEY, G., HELMLE, K. (2010). Middle Triassic coral growth bands and their implication for photosymbiosis. Palaios, 25: 754–763

STANLEY, G., VAN DE SCHOOTBRUGGE, B., (2018). The evolution of the coral–algal symbiosis and coral bleaching in the geologic past. In: Oppen, M., Lough, J. (eds). Coral Bleaching: Patterns, Processes, Causes and Consequences. Springer, Cambridge, 9–26

STARRFELT, J., LIOW, L. (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, 371: 20150219

STEPHENSON, P., STENGEL, C. (2020). An inventory of biodiversity data sources for conservation monitoring. PloS ONE, 15: e0242923

STIGALL, A. (2010). Using GIS to assess the biogeographic impact of species invasions on native brachiopods during the Richmondian invasion in the type-Cincinnatian (Late Ordovician, Cincinnati region). Palaeontologica Electronica, 13: 13.1.5A

STIGALL, A., LIEBERMAN, B. (2006). Quantitative palaeobiogeography: GIS, phylogenetic biogeographical analysis, and conservation insights. Journal of Biogeography, 33: 2051–2060

STRAUSS, D., SADLER, P. (1987). Confidence intervals for the ends of local taxon ranges. Technical Report 158. Department of Statistics, University of California, California.

STRAUSS, D., SADLER, P. (1989). Classical confidence intervals and Bayesian probability estimates for ends of local taxon ranges. Mathematical Geology, 21: 411–427

STUBBS, T., PIERCE, S., ELSLER, A., ANDERSON, P., RAYFIELD, E., BENTON, M. (2021). Ecological opportunity and the rise and fall of crocodylomorph evolutionary innovation. Proceedings of the Royal Society B, 288: 20210069

SUN, Y., JOACHIMSKI, M., WIGNALL, P., YAN, C., CHEN, Y., JIANG, H., WANG, L., LAI, X. (2012). Lethally hot temperatures during the Early Triassic greenhouse. Science, 338: 366–370

SUN, Y., ORCHARD, M., KOCSIS, A., JOACHIMSKI, M. (2020). Carnian–Norian (Late Triassic) climate change: evidence from conodont oxygen isotope thermometry with implications for reef development and Wrangellian tectonics. Earth and Planetary Science Letters, 534: 116082

TANNER, L. (2017). Climates of the Late Triassic: perspectives, proxies and problems. In: Tanner, H. (ed). The Late Triassic World, Springer, Cham, 59–90

TAYLOR, P., FAHRIG, L., HENEIN, K., MERRIAN, G. (1993). Connectivity is a vital element of landscape structure. Oikos, 68: 571–573

TAYLOR, R. (2002). A new bivalved arthropod from the Early Cambrian Sirius Passet fauna, North Greenland. Palaeontology, 45: 97–123

TEICHERT, S., STEINBAUER, M., KIESSLING, W. (2020). A possible link between coral reef success, crustose coralline algae and the evolution of herbivory. Nature Scientific Reports, 10: 17748

TENNANT, J., CHIARENZA, A., BARON, M. (2018). How has our knowledge of dinosaur diversity through geologic time changed through research history? PeerJ, 6: e4417

THIBODEAU, A., RITTERBUSH, K., YAGER, J., WEST, A., IBARRA, Y., BOTTJER, D., BERELSON, W., BERGQUIST, B., CORSETTI, F. (2016). Mercury anomalies and the timing of biotic recovery following the end-Triassic mass extinction. Nature Communications, 7: 11147

TOMLIN, C. (1990). Geographic information systems and cartographic modelling. Englewood Cliffs, Prentice-Hall, Portland

TORSVIK, T., COCKS, R. (2019). The integration of palaeomagnetism, the geological record and mantle tomography in the location of ancient continents. Geological Magazine, 156: 242–260

TOSTI, F., MASTANDREA, A. GUIDO, A., DEMASI, F. RUSSO, F., RIDING, R. (2014). Biogeochemical and redox record of mid–late Triassic reef evolution in the Italian Dolomites. Palaeogeography, Palaeoclimatology, Palaeoecology, 399: 52–66

TURNER, A., MILES, R. (1971). The GCARS System: a computer-assisted method of regional route location. Highway Research Record, 348: 1–15

UHEN, M., BARNOSKY, A., BILLS, B., BLOIS, J., CARRANO, M., CARRASCO, M., ERICKSON, G., ERONEN, J., FORTELIUS, M., GRAHAM., GRIMM, E., O'LEARY, M., MAST, A., PIEL., W., POLLY, P., SAILA, L. (2013). From card catalogs to computers: databases in vertebrate paleontology. Journal of Vertebrate Paleontology, 33: 13–28

URBAN, D., KEITT, T. (2001). Landscape connectivity: a graph theoretic perspective. Ecology, 82: 1205–1218

VALAVI, R., GUILLERA-ARROITA, G., MAHOZ-MAONFORT, J., ELITH, J. (2021). Predictive performance of presence-only species distribution models: a benchmark study with reproducible code. Ecological Monographs, 92: e01486

VALDES, P., ARMSTRONG, E., BADGER, M., BRADSHAW, C., BRAGG, F., CRUCIFIX, M., DAVIES-BARNARD, T., DAY, J., FARNSWORTH, A., GORDON, C., HOPCROFT, P., KENNEDY, A., LORD, N., LUNT, D., MARZOCCHI, A., PARRY, L., POPE, V., ROBERTS, W., STONE, E., TOURTE, G., WILLIAM, J. (2017). The BRIDGE HadCM3 family of climate models: HadCM3@Bristol v1.0. Geoscientific Model Development, 10: 3715–3743

VALDES, P., SCOTESE, C., LUNT, D. (2021). Deep ocean temperatures through time. Climates of the Past, 17: 1483–1506

VALENTINE, J. (1969). Patterns of taxonomic and ecological structure of the shelf benthos during Phanerozoic time. Palaeontology, 12: 684–709

VALENTINE, J., JABLONSKI, D. (2015). A twofold role for global energy gradients in marine biodiversity trends. Journal of Biogeography, 42: 997–1005

VAN DE LOO, M. (2014). The stringdist package for approximate string matching. The R Journal, 6: 111–122

VAN DAM, M., MATZKE, N. (2016). Evaluating the influence of connectivity and distance on biogeographical patterns in the south-western deserts of North America. Journal of Biogeography, 43: 1514–1532

VAN LEUSEN, M. (1993). Cartographic modelling in a cell-based GIS. In: Andresen, J., Madsen, T., Scollar, I. (eds). Computing the Past. Computer Applications and Quantitative Methods in Archaeology. Aarhus University Press, Aarhus, 105–124

VARELA, S., LOBO, J., HORTAL, J. (2011). Using species distribution models in paleobiogeography: a matter of data, predictors and concepts. Palaeogeography, Palaeoclimatology, Palaeoecology, 310: 451–463

VENDITTI, C., MEADE, A., PAGEL, M. (2011). Multiple routes to mammalian diversity. Nature, 479:393–396

VERARD, C. (2019). Plate tectonic modelling: review and perspectives. Geological Magazine, 156, 208–241

VÉRARD, C., HOCHARD, C., BAUMGARTNER, P., STAMPFLI, G. (2015). Geodynamic evolution of the Earth over the Phanerozoic: plate tectonic activity and palaeo-climatic indicators. Journal of Palaeogeography, 4: 167–188

VERHAGEN, P., NUNINGER, L., GROENHUIJZEN, M. (2019). Modelling of pathways and movement networks in archaeology: an overview of current approaches. In: Verhagen, P., Joyce, J., Groenhuijzen, M. (eds). Finding the Limits of the Limes: Modelling Demography, Economy and Transport on the Edge of the Roman Empire. Springer Cham, 217–249

VERMEIJ, G. (1987). Evolution and escalation: patterns of marine life. Princeton University Press, Princeton

VILHENA, D., SMITH, A. (2013). Spatial bias in the marine fossil record. PLoS ONE, 8: e74470

VITO, M., MUGGEO, R. (2008). segmented: an R Package to fit regression models with broken-line relationships. R News, 8: 20–25

WAGNER, P., HENDY, A., KIESSLING, W. (2007). The effects of taxonomic standardization on sampling-standardized estimates of historical diversity. Proceedings of the Royal Society B, 274: 439–444

WANG, I., SAVAGE, W., SHAFFER, B. (2009). Landscape genetics and least-cost path analysis reveal unexpected dispersal routes in the California tiger salamander (*Ambystoma californiense*). Molecular Ecology, 18: 1365–1374

WANG, X., YAO, L., LIN, W. (2018). Permian rugose corals of the world. In: Lucas, S., Shen S. (Eds.). The Permian Timescale. Geological Society of London Special Publications, 450: 165–184

WARNOCK, R., HEATH, T., STADLER, T. (2020). Assessing the impact of incomplete species sampling on estimates of speciation and extinction rates. Paleobiology, 46: 137–157

WARNTZ, W. (1957). Transportation, social physics, and the law of refraction. The Professional Geographer, 9: 2–7

WARNTZ, W. (1965). A note on surfaces and paths and applications to geographical problems. Michigan Inter-University Community of Mathematical Geographers, Ann Arbor

WARREN, B., SIMBERLOFF, D., RICKLEFS, R., AGUILEE, R., CONDAMINE, F., GRAVEL, D., MORLON, H., MOUQUET, N., ROSINDELL, J., CASQUET, J., CONTI, E., CORNAULT, J., FERNANDEZ-PALACIOS, J., HENGL, T., NORDER, S., RIJSDIJK, K., SANMARTIN, I., STRASBERG, D., TRIANTIS, K., VALENTE, L., WHITTAKER, R., GILLESPIE, R., EMERSON, B., THEBAUD, C. (2015). Islands as model systems in ecology and evolution: prospects fifty years after MacArthur-Wilson. Ecology Letters, 18: 200–217

WHITESIDE, J., GROGAN, D., OLSEN, P., KENT, D. (2011). Climatically driven biogeographic provinces of Late Triassic tropical Pangea. Proceedings of the National Academy of Sciences of the United States of America, 108: 8972–8977

WHITESIDE, J., WARD, P. (2011). Ammonoid diversity and disparity track episodes of chaotic carbon cycling during the early Mesozoic. Geology, 39: 99–102

WIENS, J. (2009). Paleontology, genomics, and combined-data phylogenetics: can molecular data improve phylogeny estimation for fossil taxa? Systematic Biology, 58: 87–99

WIGNALL, P., SUN, Y., BOND, D., IZON, G., NETWON, R., VEDRINE, S., WIDDOWSON, M., ALI, J., LAI, X., JIANG, H., COPE, H., BOTTRELL, S. (2009). Volcanism, mass extinction, and carbon isotope fluctuations in the Middle Permian of China. Science, 324: 1179–1182

WIGNALL, P., ATKINSON, J. (2020). A two-phase end-Triassic mass extinction. Earth-Science Reviews, 208: 103282

WILLIAMS, J., GRIMM, E., BLOIS, J., CHARLES, D., DAVIS, E., GORING, S., GRAHAM, R., SMITH, A., ANDERSON, M., ARROYO-CABRALES, J., ASHWORTH, A., BETANCOURT, J., BILLS, B., BOOTH, R., BUCKLAND, P., CURRY, B., GIESECKE, T., HAUSMANN, S., JACKSON, S., LATORRE, C. (2018). The Neotoma Paleoecology Database: a multi-proxy, international community-curated data resource. Quaternary Research, 89: 156–177

WINKLER, W. (1990). String Comparator Metrics and Enhanced Decision Rules in the Fellegi-Sunter Model of Record Linkage. Proceedings of the Section on Survey Research Methods, American Statistical Association, 354–359

WISNIEWSKI, A., LLOYD, G., SLATER, G. (2022). Extant species fail to estimate ancestral geographical ranges at older nodes in primate phylogeny. Proceedings of the Royal Society B, 289: 20212535

WOOD, H., MATZKE, N., GILLESPIE, R., GRISWOLD, C. (2013). Treating fossils as terminal taxa in divergence time estimation reveals ancient vicariance patterns in the palpimanoid spiders. Systematic Biology, 62: 264–284

WRIGHT, N., ZAHIROVIC, S., MULLER, R., SETON, M. (2013). Towards community-driven paleogeographic reconstructions: integrating open-access paleogeographic and paleobiology data with plate tectonics. Biogeosciences, 10: 1529–1541

WRIGHT, A., BAPST, D., BARIDO-SOTTANI, J., WARNOCK, R. (2022). Integrating fossil observations into phylogenetics using the fossilized birth–death model. Annual Review of Ecology, Evolution and Systematics, 53: 251–273

YANG, Z. (1994). Maximum likelihood phylogenetic estimation from DNA sequences with variable rates over sites: approximate methods. Journal of Molecular Evolution, 39: 306–314

YANG, Z., RANNALA, B. (1997). Bayesian phylogenetic inference using DNA sequences: a Markov Chain Monte Carlo method. Molecular Biology and Evolution, 17: 717–724

YE, S., PETERS, S. (2023). Bedrock geological map predictions for Phanerozoic fossil occurrences. Paleobiology, doi.10.1017/pab.2022.46

YESSON, C., CULHAM, A. (2006). Phyloclimatic modeling: combining phylogenetics and bioclimatic modeling. Systematic Biology, 55: 785–802

ZAFFOS, A. (2017). velociraptr. R package version 1.0. https://cran.r-project.org/web/packages/velociraptr/index.html

ZAFFOS, A., FINNEGAN, S., PETERS, S. (2017). Plate tectonic regulation of global marine animal diversity. Proceedings of the National Academy of Sciences of the United States of America, 114: 5653–5658

ZEEBE, R., WESTBROEK, P. (2003). A simple model for the CaCO3 saturation state of the ocean: The “Strangelove,” the “Neritan,” and the “Cretan” Ocean. Geochemistry, Geophysics, Geosystems, 4: 1–26

ZHANG, C., KLOPFSTEIN, S., HEATH, T., RONQUIST, F. (2016). Total-evidence dating under the fossilized birth-death process. Systematic Biology, 65: 228–249

ZHOU, G., WEI, H., FU, S. (2019). A fast and simple algorithm for calculating flow accumulation matrices from raster digital elevation. Frontiers in Earth Science, 13: 317–326

ZIZKA, A., SILVESTRO, D., ANDERMANN, T., AZEVEDO, J., RITTER, C., EDLER, D., FAROOQ, H., HERDEAN, A., ARIZO, M., SCHARN, R., SVANTESSON, S., WENGSTROM, N., ZIZKA, V., ANTONELLI, A. (2019). CoordinateCleaner: Standardized cleaning of occurrence records from biological collection databases. Methods in Ecology and Evolution, 10: 744–751

Supplementary Tables and Figures

Figures

Tables

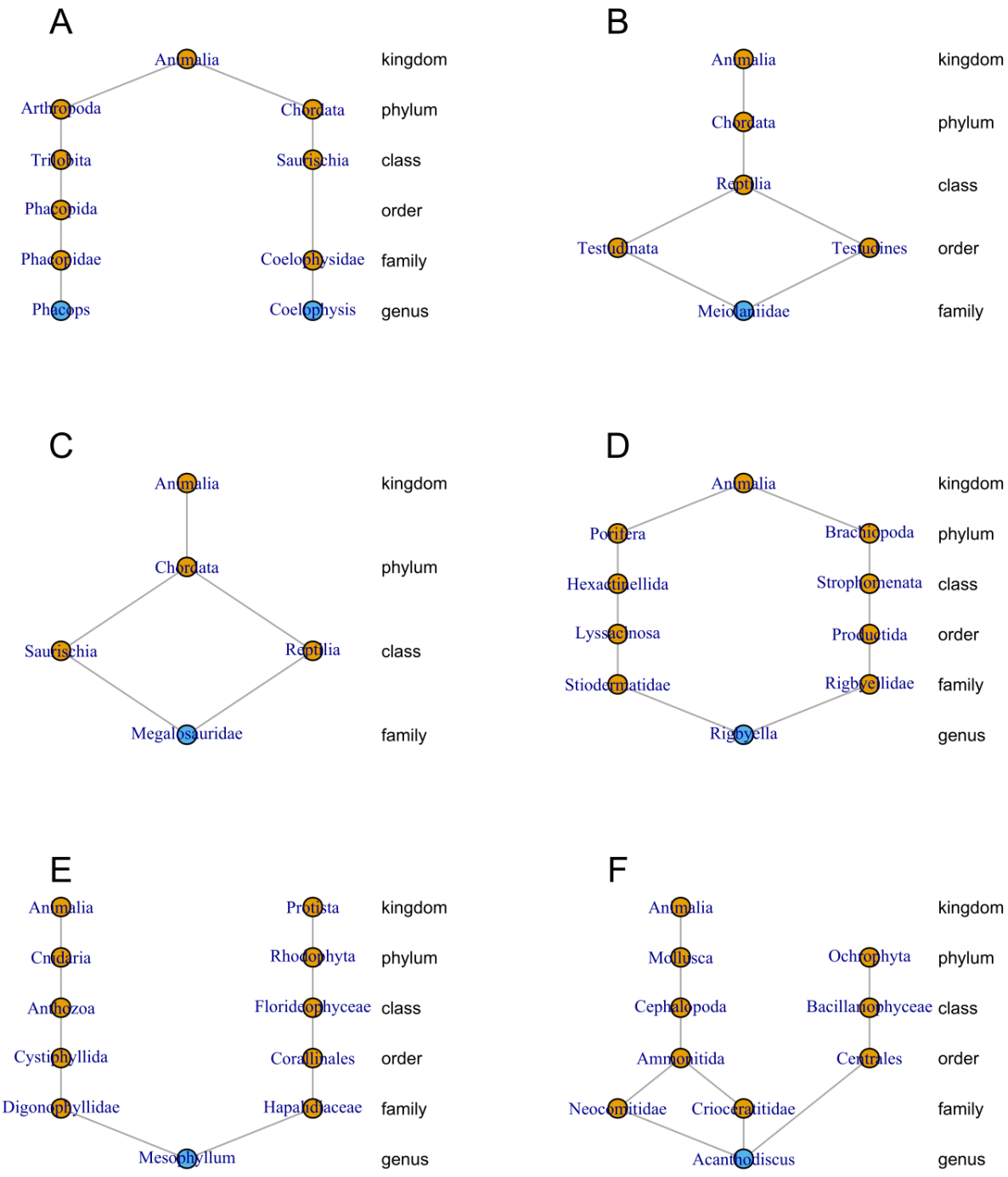


Fig. S1. Examples of valid and inconsistent classifications in the PBDB. **(A).** The relative relationships of *Coelophysis* (Dinosauria) and *Phacops* (Trilobita), displaying correct classifications. **(B, C).** Inconsistent classification where the distance to the next shared taxon is small and will be resolved to the more frequently used scheme. **(D).** Inconsistent classification of a homonym where the distance to the next shared taxon is large. **(E).** Inconsistent classification of a homonym where there is no shared taxon. **(F).** Example of a highly inconsistent classification, displaying a combined case of a homonym and a conflict in higher taxonomy.

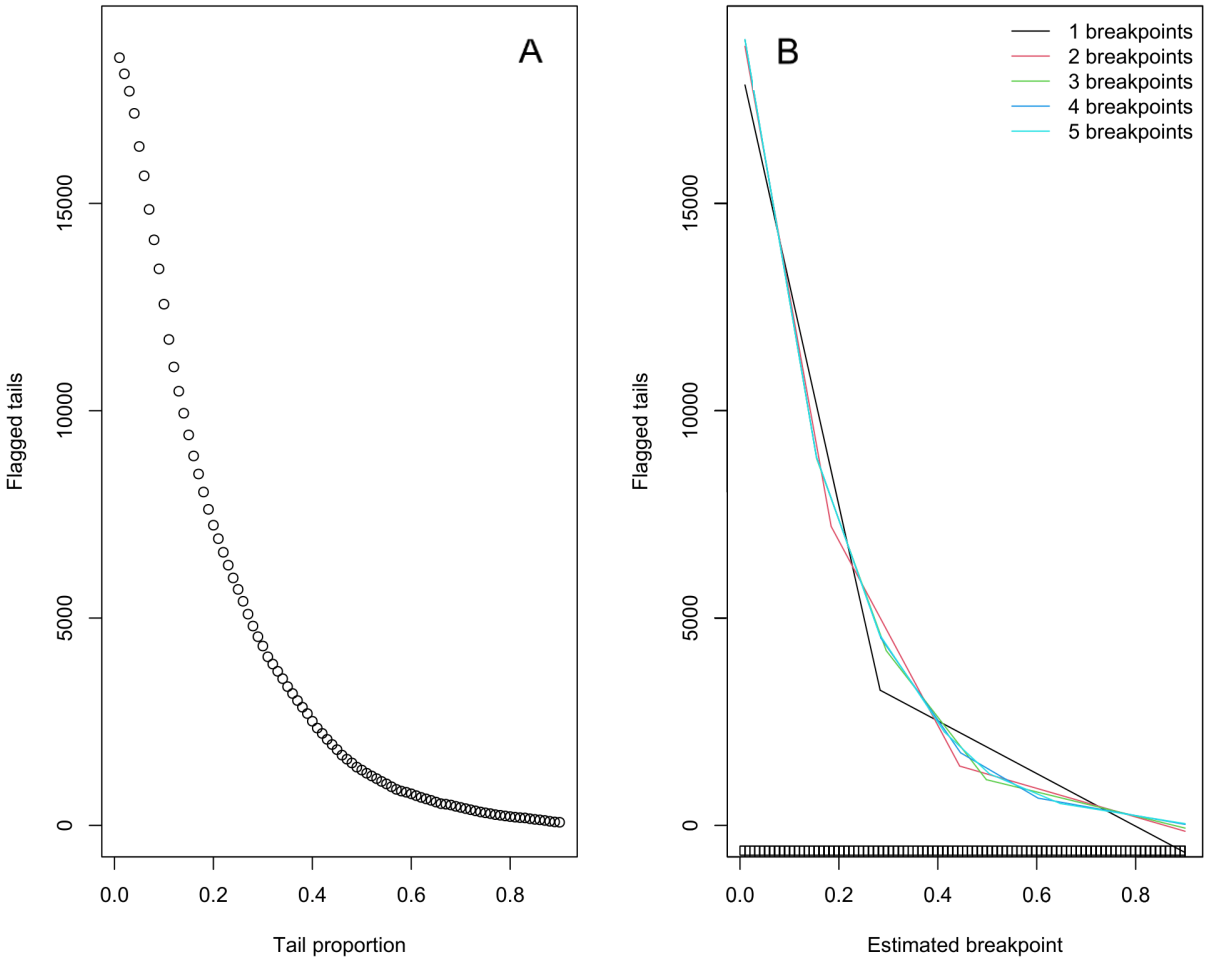


Fig. S2. Sensitivity testing of Pacmacro tail flagging. Flagging was tested under different tail stratigraphic proportions with a constant tail density proportion of 5%. **(A).** The relationship between increasing tail proportion and the number of flagged genera. **(B).** Linear models with increasing numbers of breakpoints fitted to this relationship.

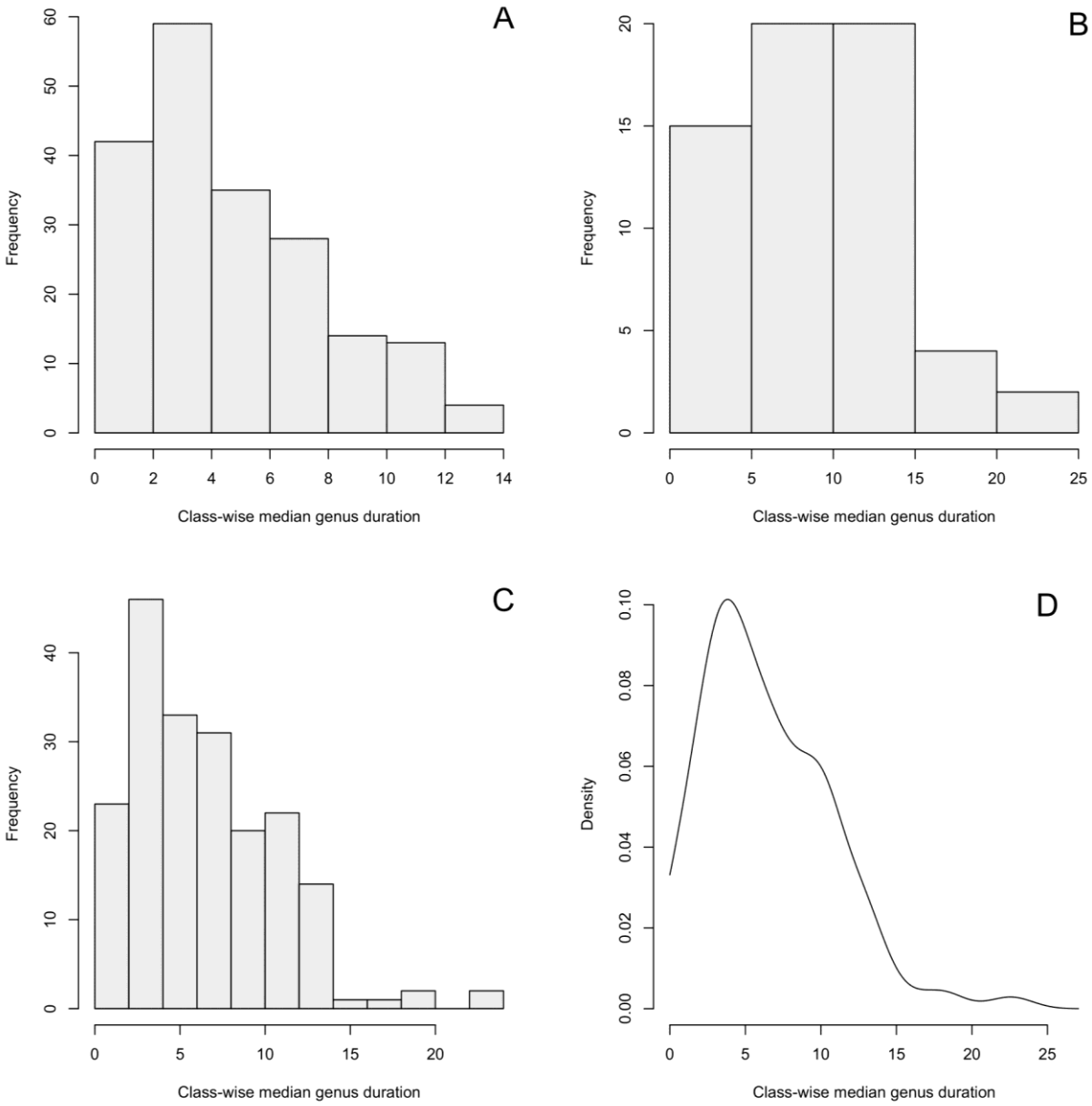


Fig. S3. Class-wise median genus durations in the Sepkoski Compendium and PBDB. (A). Durations calculated from the PBDB. **(B).** Durations calculated from the Sepkoski Compendium. **(C).** Durations calculated from a composite dataset where Sepkoski ranges were used primarily and supplemented with taxa only represented in the PBDB. **(D).** Density distribution of the composite dataset, suggesting a relaxed default of 15 Ma for interpeak thresholding.

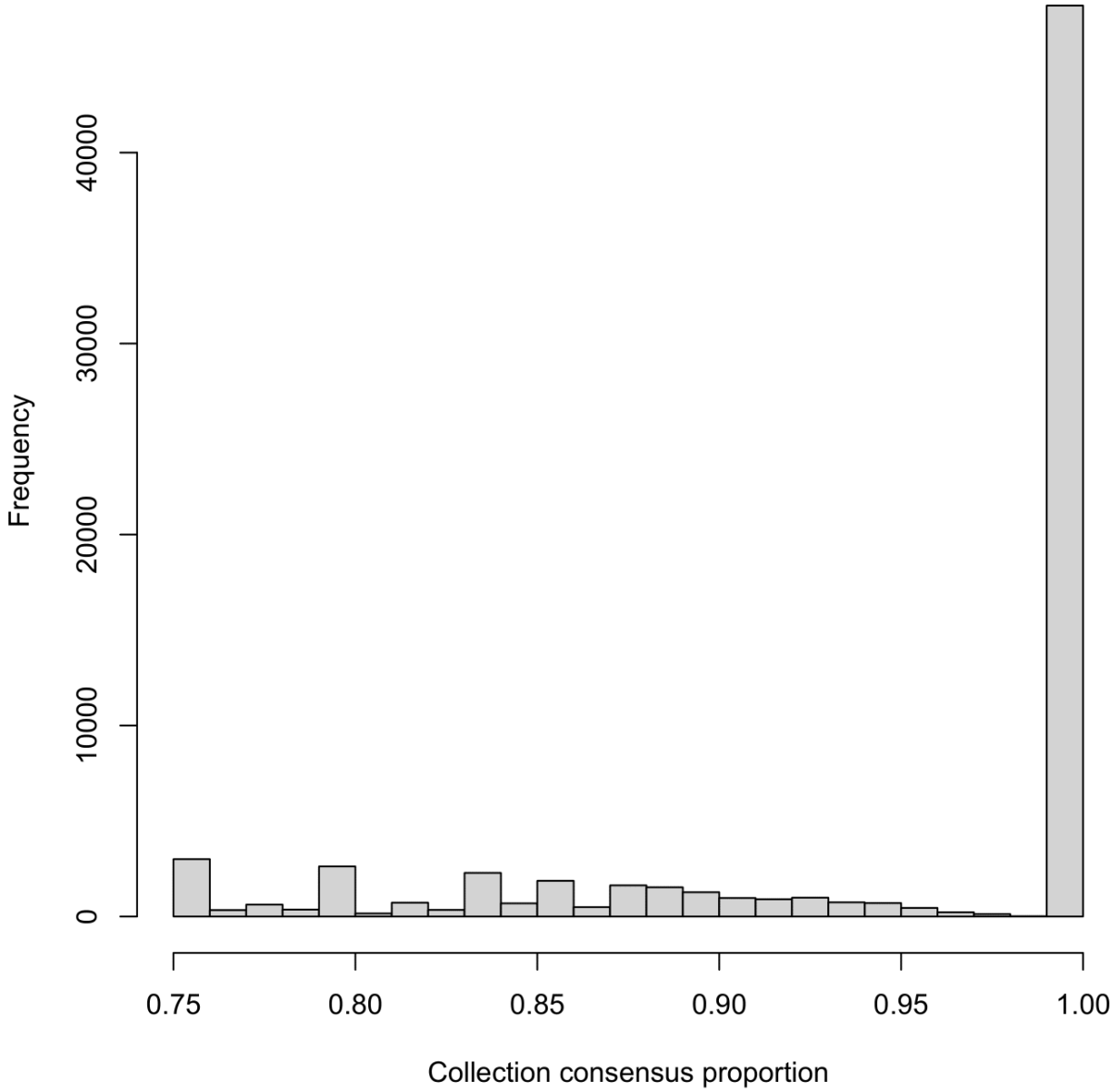


Fig. S4. Distribution of taxon range consensus proportions from collection age revision. Collection ages were revised in the PBDB collection age revision using the Sepkoski Compendium.

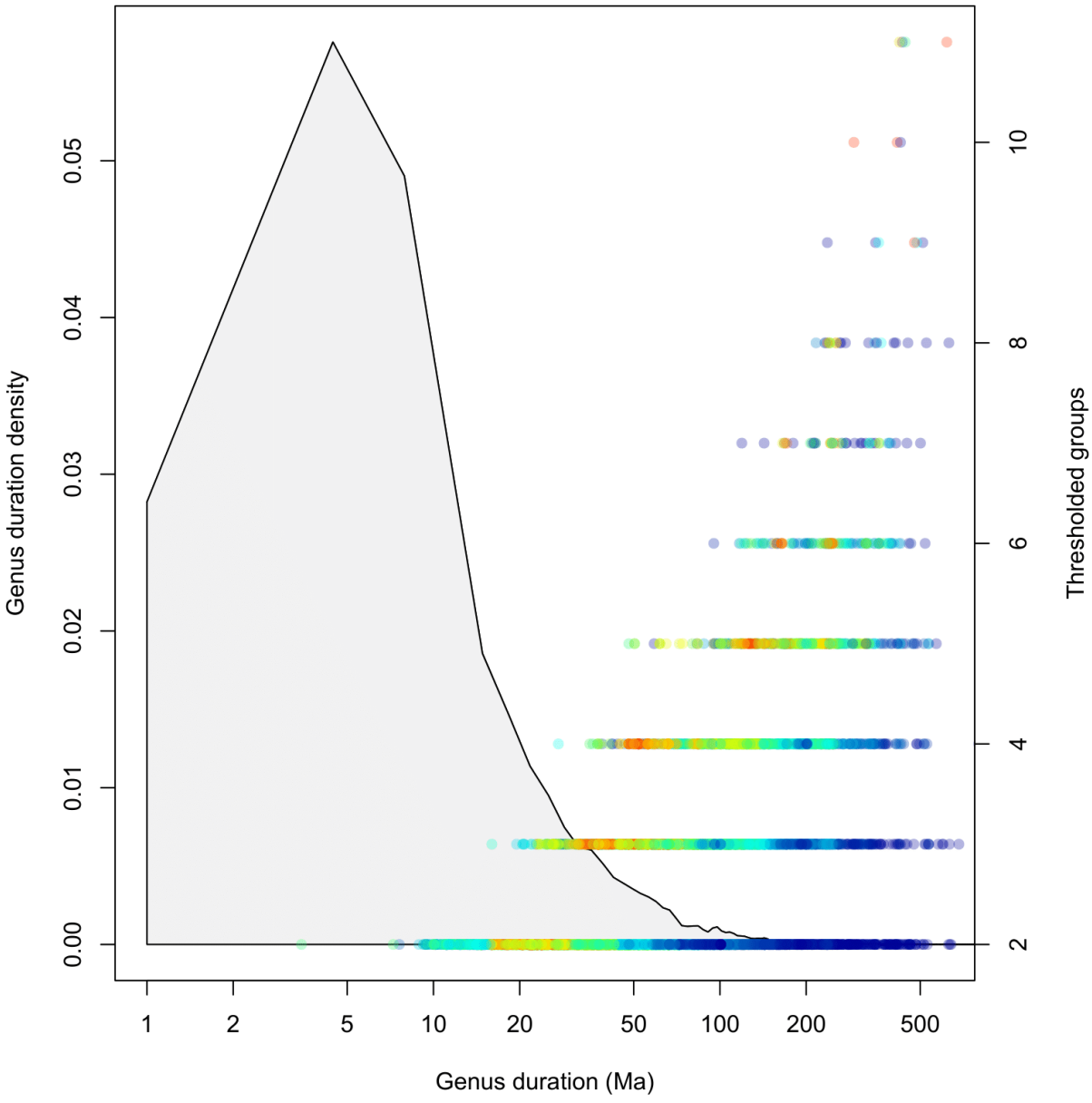


Fig. S5. Original genus durations versus number of splits induced by interpeak thresholding. In the latter, warmer colours indicate a greater density of points.

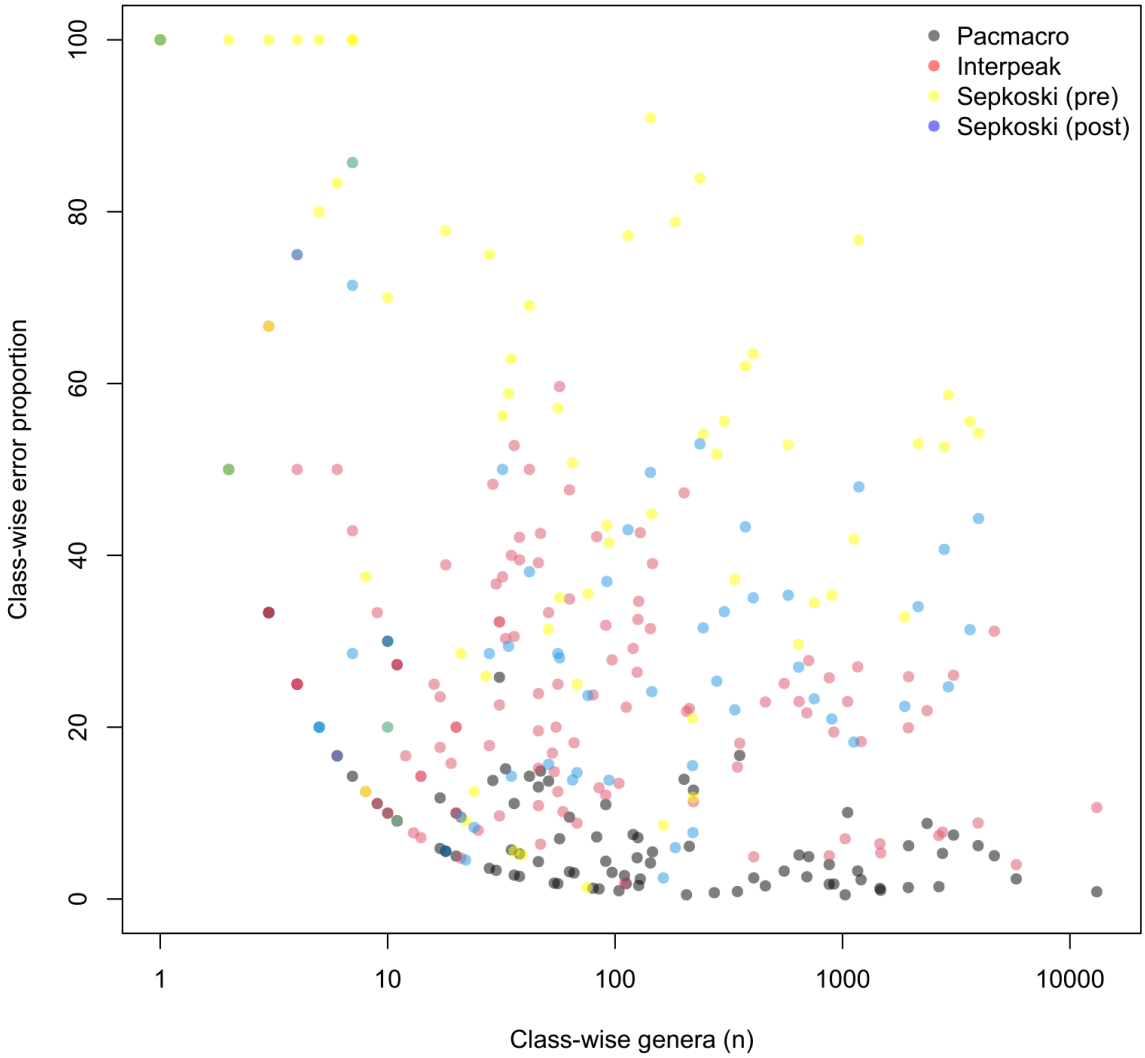


Fig. S6. The relationship between class-wise error proportion and class-wise genus count. Tallies were made each class in the PBDB, for each error flagging method.

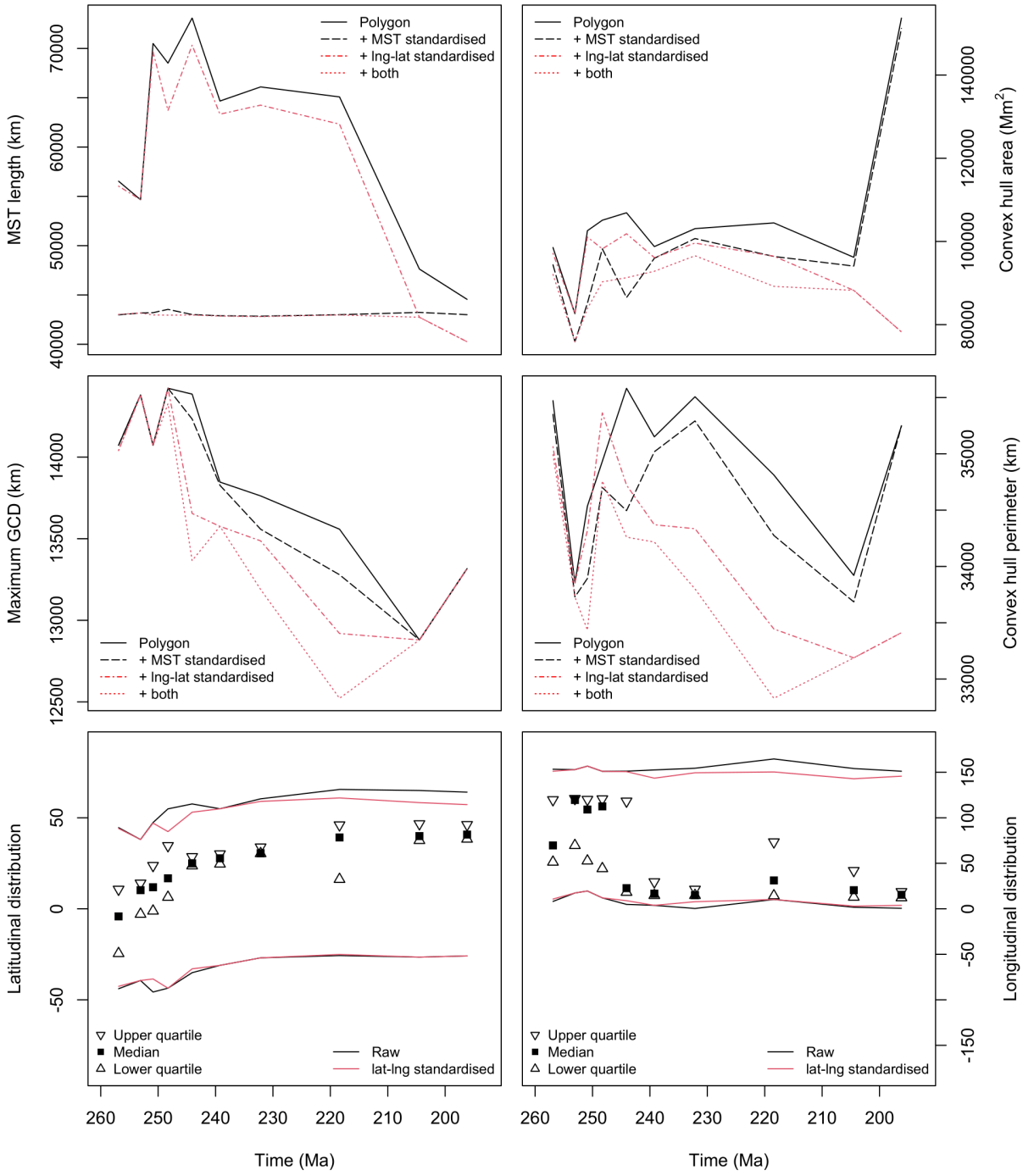


Fig. S7. Spatial properties of the Circumtethys region during spatial standardisation. Metrics are calculated at stage level, aside for the Early Jurassic. MST = minimum spanning tree.

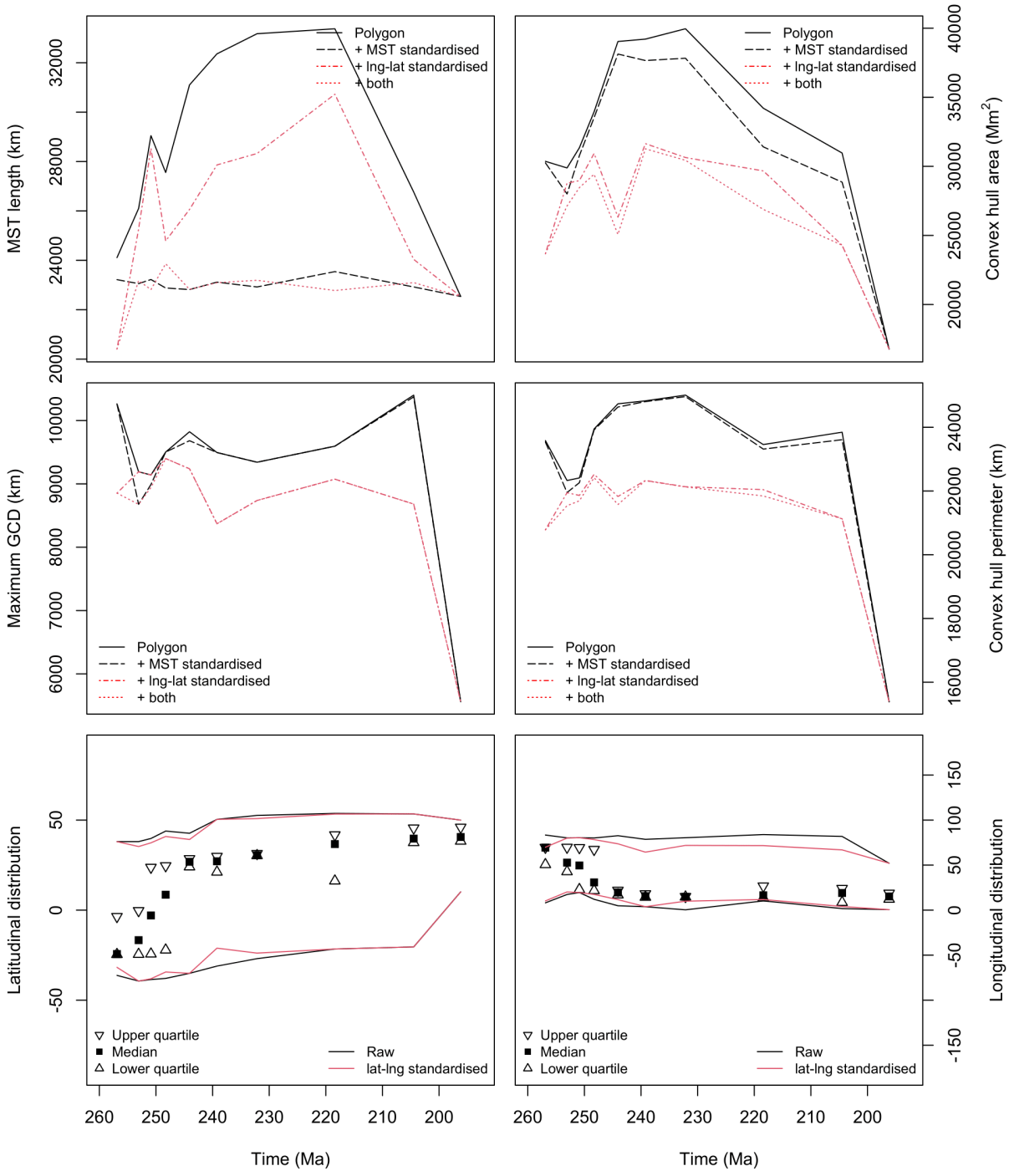


Fig. S8. Spatial properties of the West Circumtethys region during spatial standardisation. Metrics are calculated at stage level, aside for the Early Jurassic. MST = minimum spanning tree.

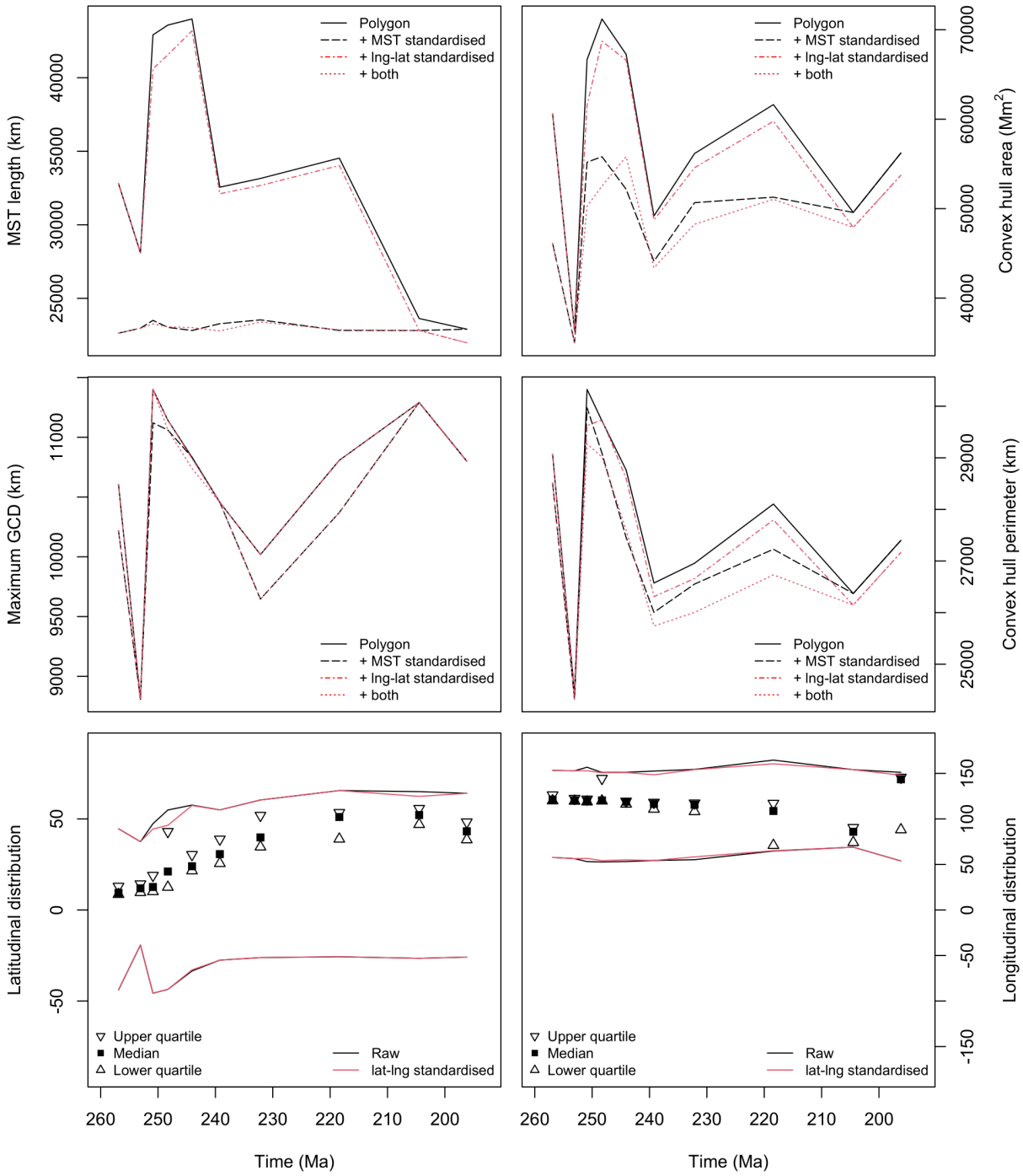


Fig. S9. Spatial properties of the East Circumtethys region during spatial standardisation. Metrics are calculated at stage level, aside for the Early Jurassic. MST = minimum spanning tree.

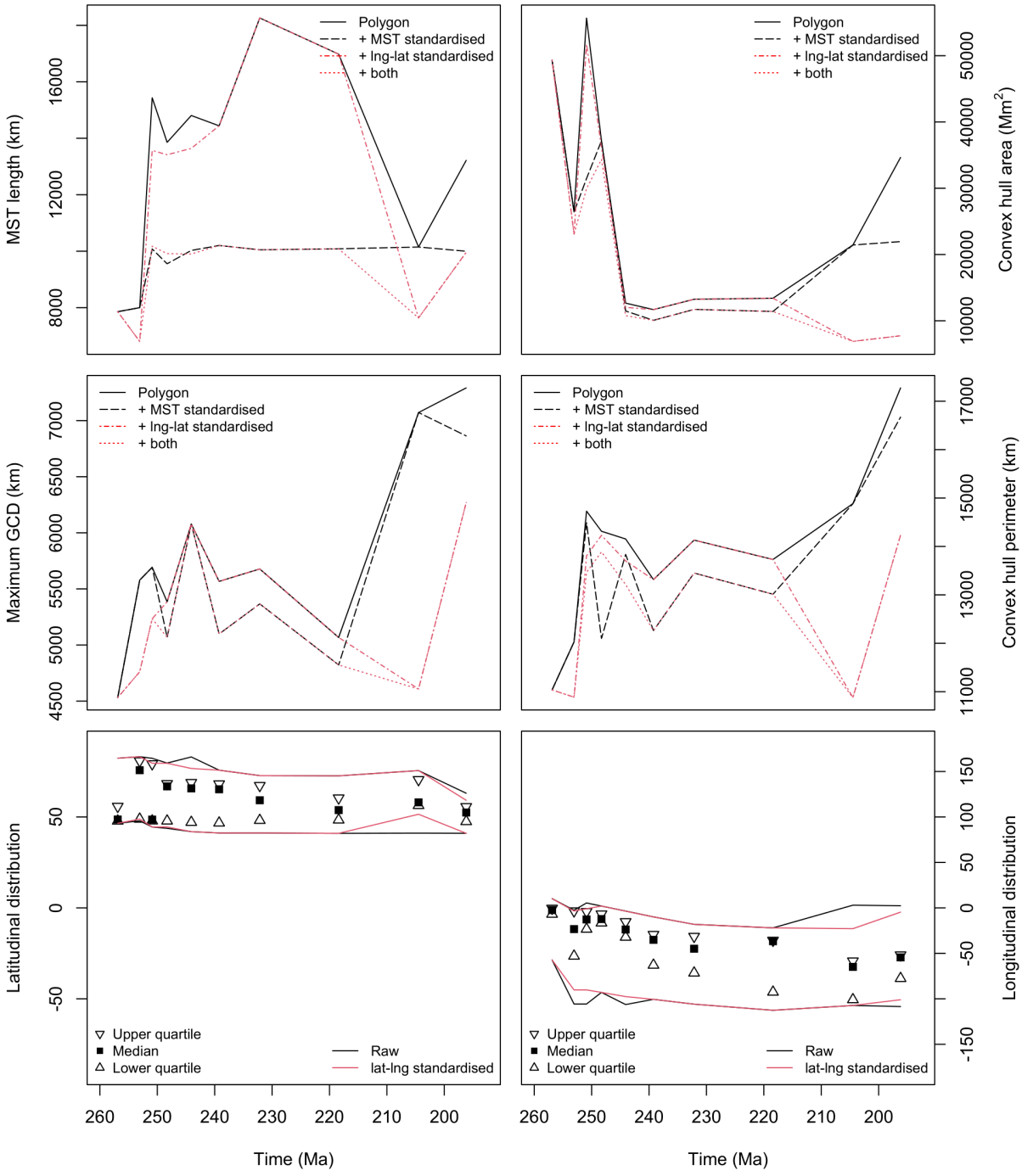


Fig. S10. Spatial properties of the Boreal region during spatial standardisation. Metrics are calculated at stage level, aside for the Early Jurassic. MST = minimum spanning tree.

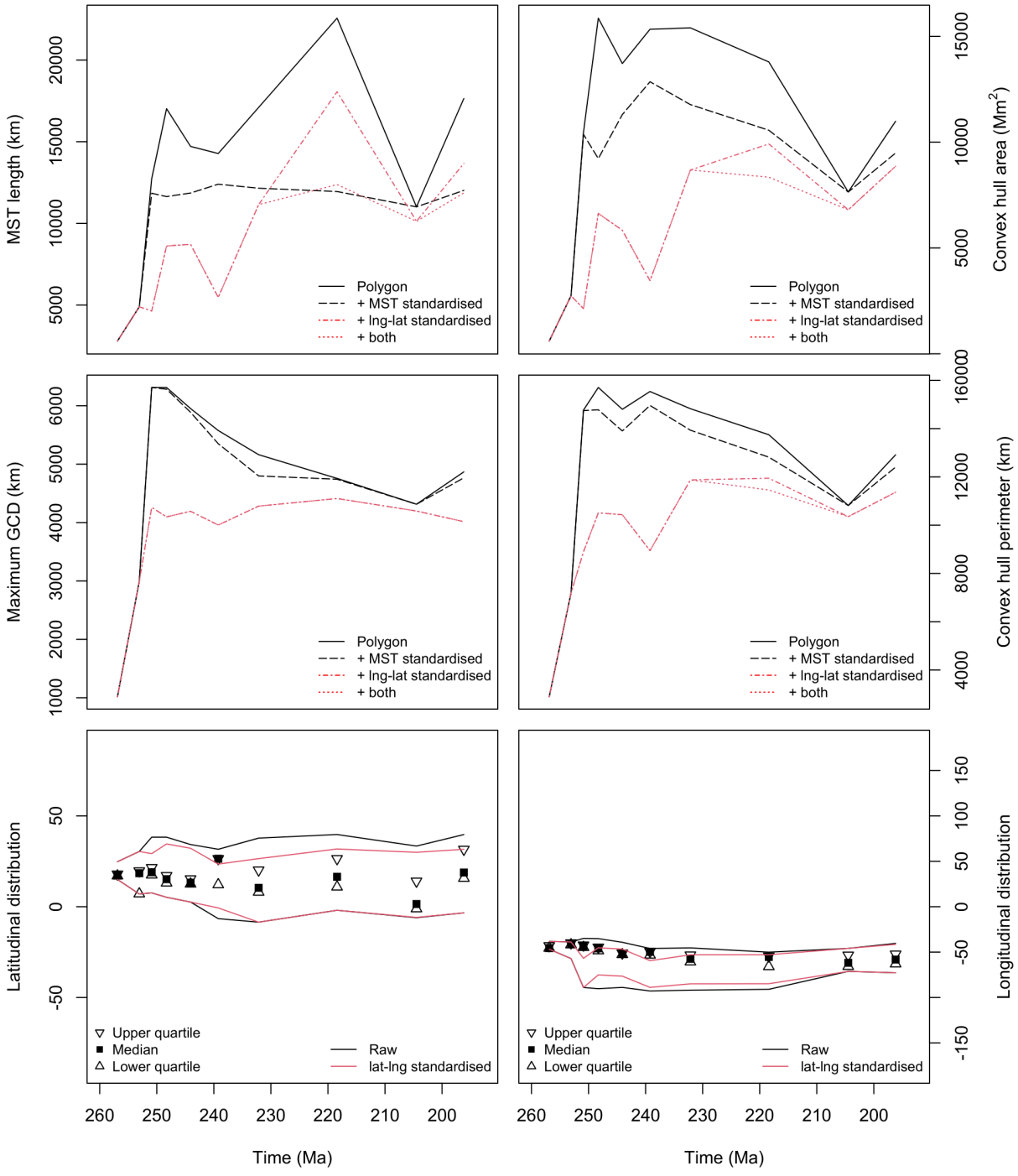


Fig. S11. Spatial properties of the North Panthalassic region during spatial standardisation. Metrics are calculated at stage level, aside for the Early Jurassic. MST = minimum spanning tree.

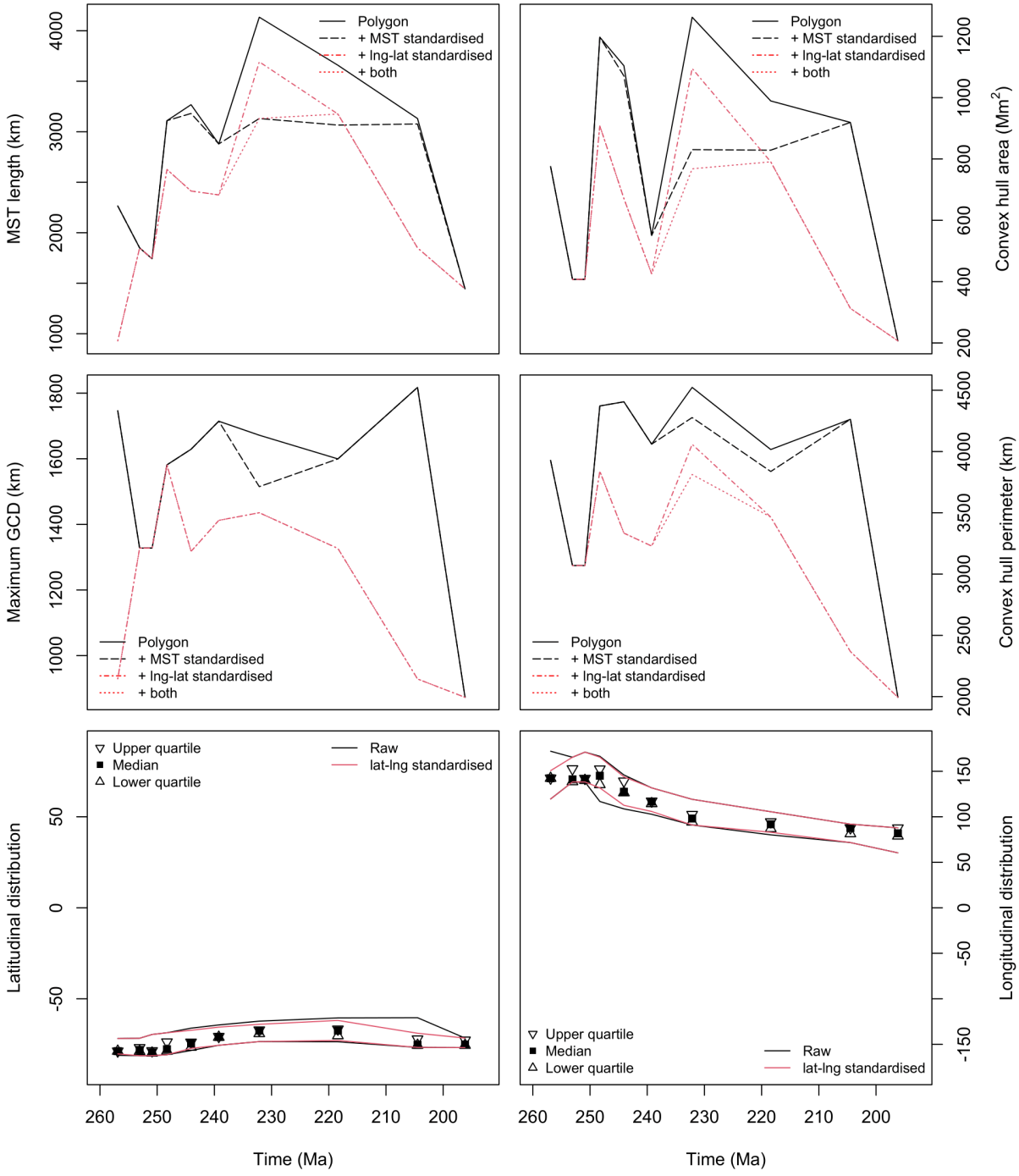


Fig. S12. Spatial properties of the Tangaroan region during spatial standardisation. Metrics are calculated at stage level, aside for the Early Jurassic. MST = minimum spanning tree.

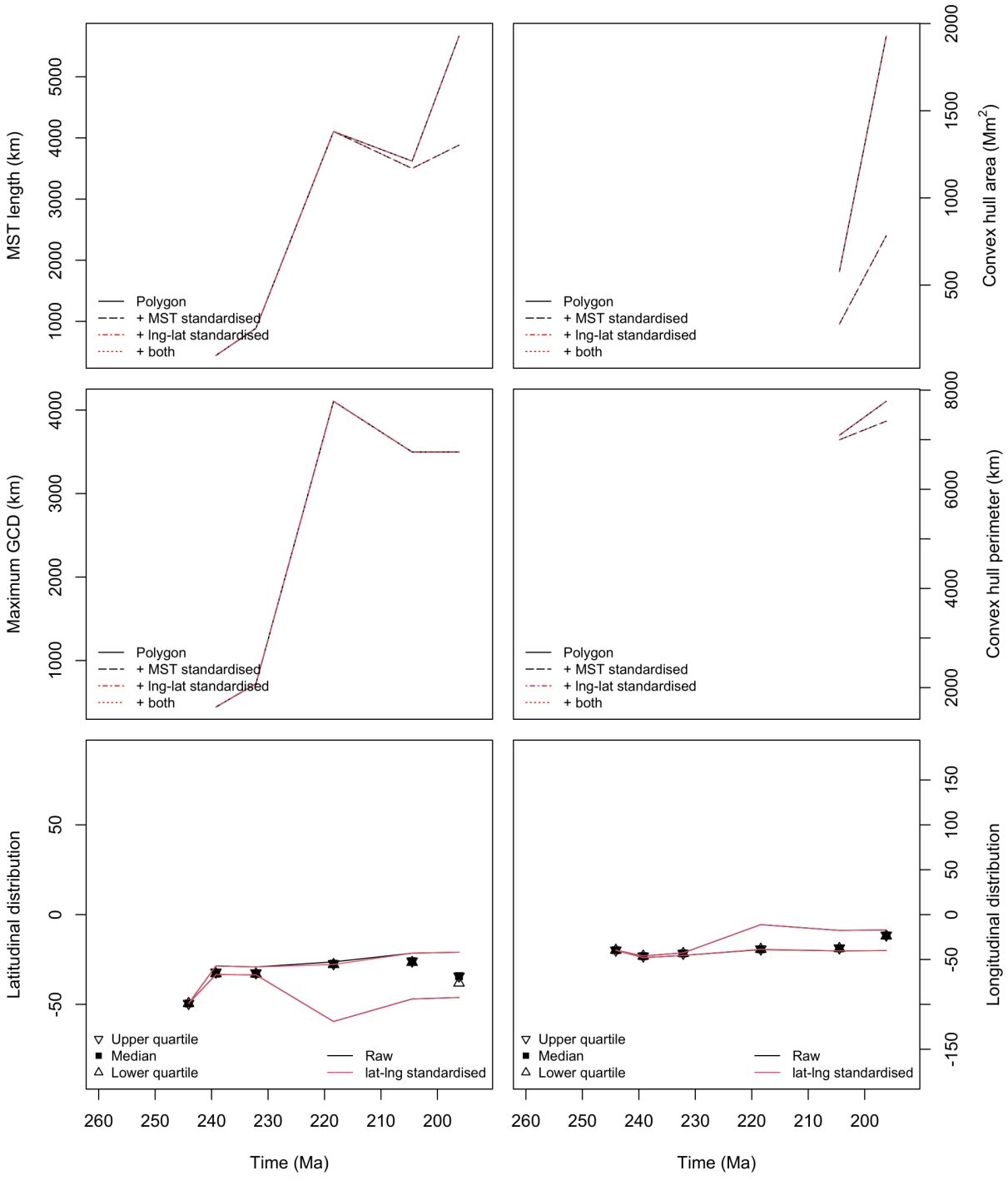


Fig. S13. Spatial properties of the South Panthalassic region during spatial standardisation. Metrics are calculated at stage level, aside for the Early Jurassic. MST = minimum spanning tree.

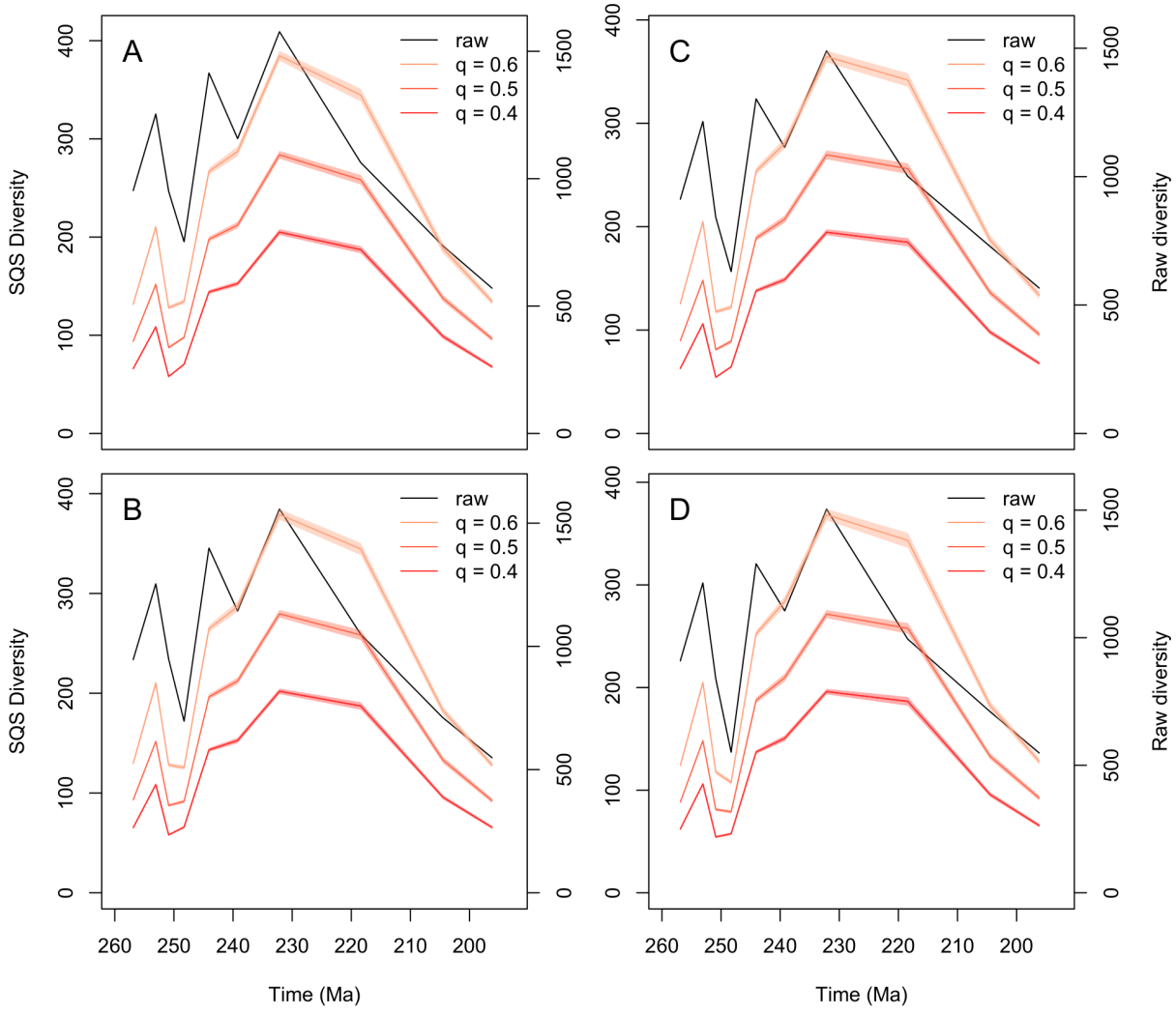


Fig. S14. Stage-wise sampling standardised diversity for the Circumtethys region. Diversity calculated using shareholder quorum subsampling at several quorum (q) levels. **(A).** Unstandardised. **(B).** MST standardisation. **(C).** Longitude-latitude standardisation. **(D).** MST + longitude latitude standardisation.

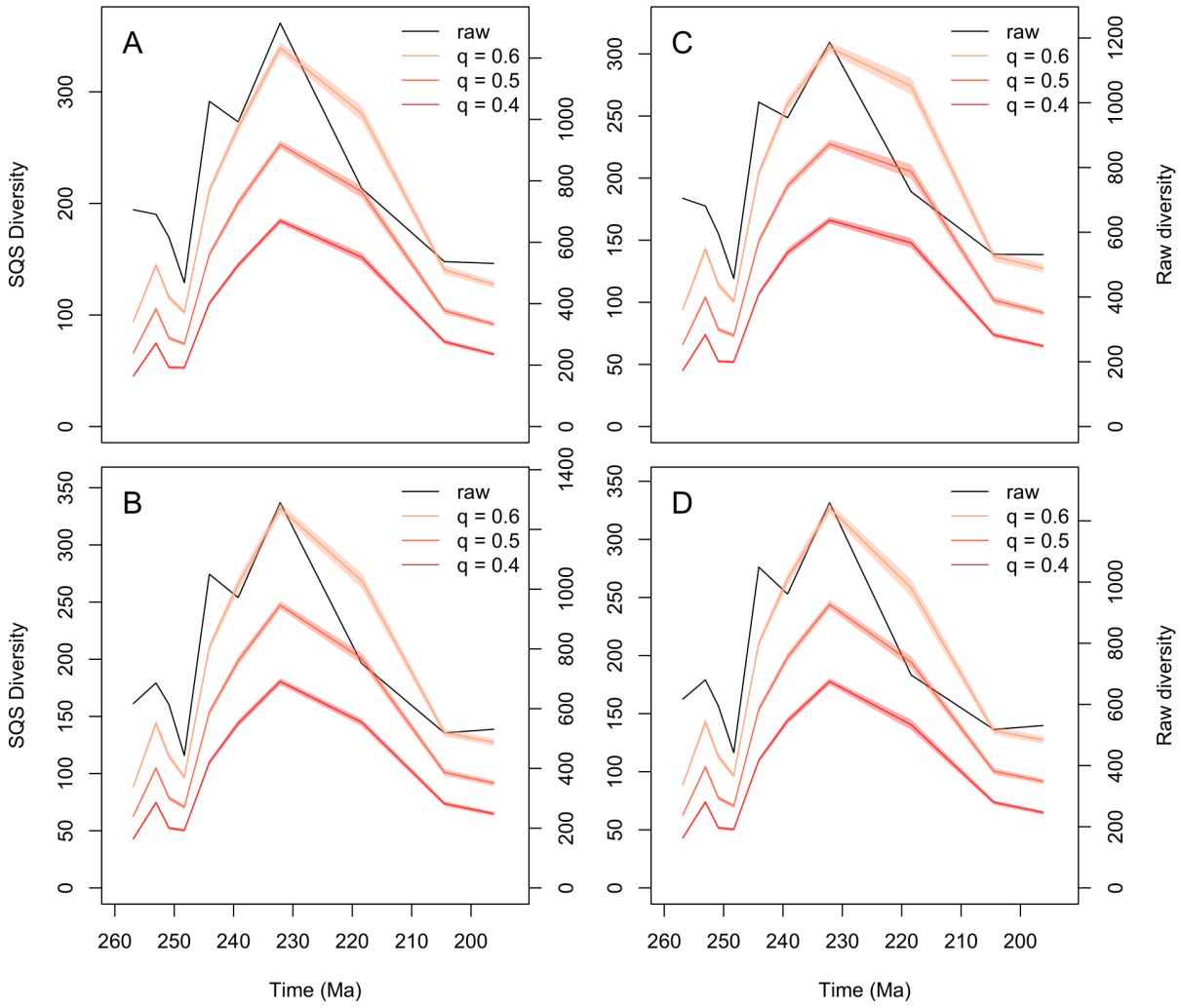


Fig. S15. Stage-wise sampling standardised diversity for the West Circumtethys region. Diversity calculated using shareholder quorum subsampling at several quorum (q) levels. **(A).** Unstandardised. **(B).** MST standardisation. **(C).** Longitude-latitude standardisation. **(D).** MST + longitude latitude standardisation.

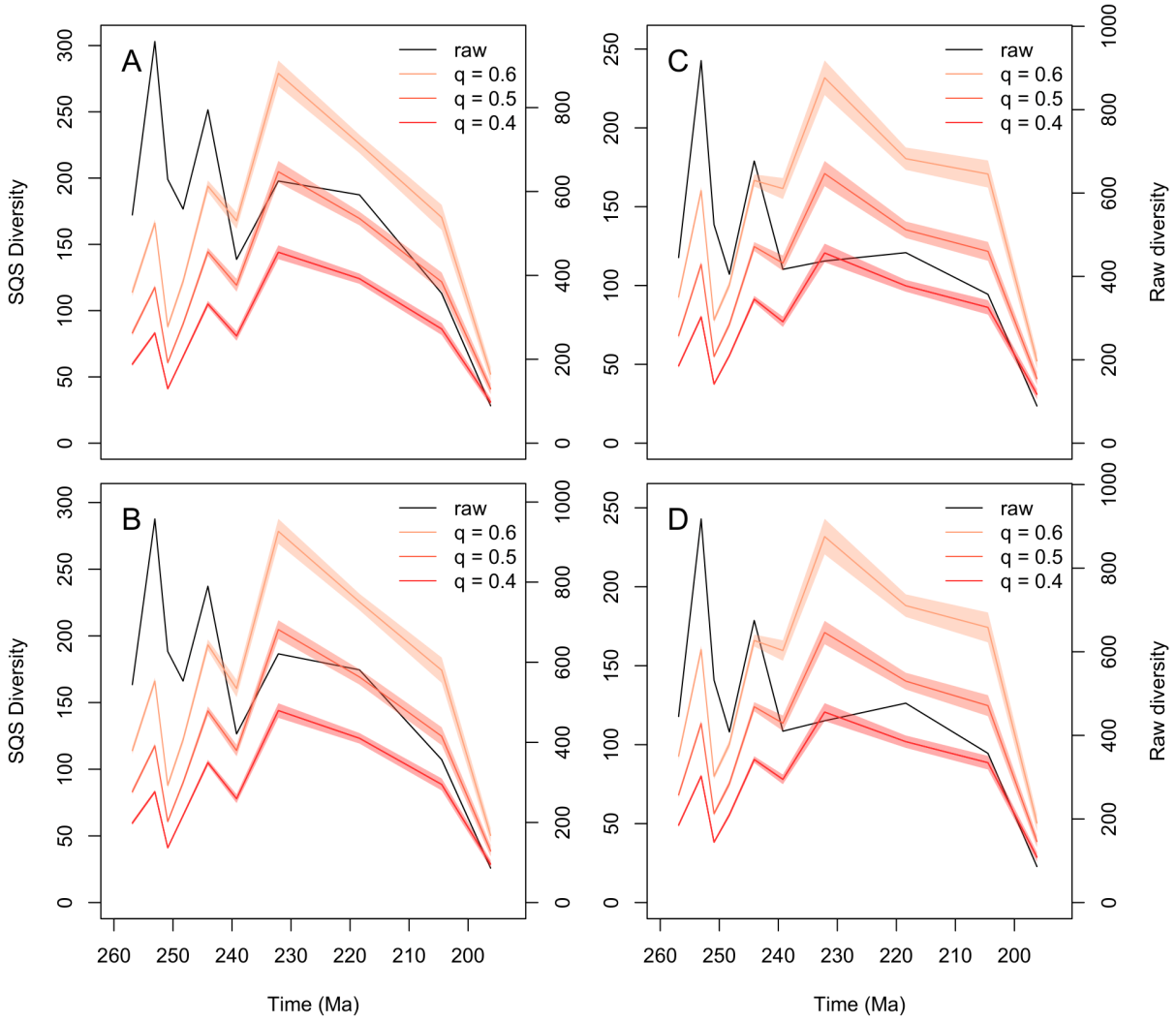


Fig. S16. Stage-wise sampling standardised diversity for the East Circumtethys region. Diversity calculated using shareholder quorum subsampling at several quorum (q) levels. **(A).** Unstandardised. **(B).** MST standardisation. **(C).** Longitude-latitude standardisation. **(D).** MST + longitude latitude standardisation.

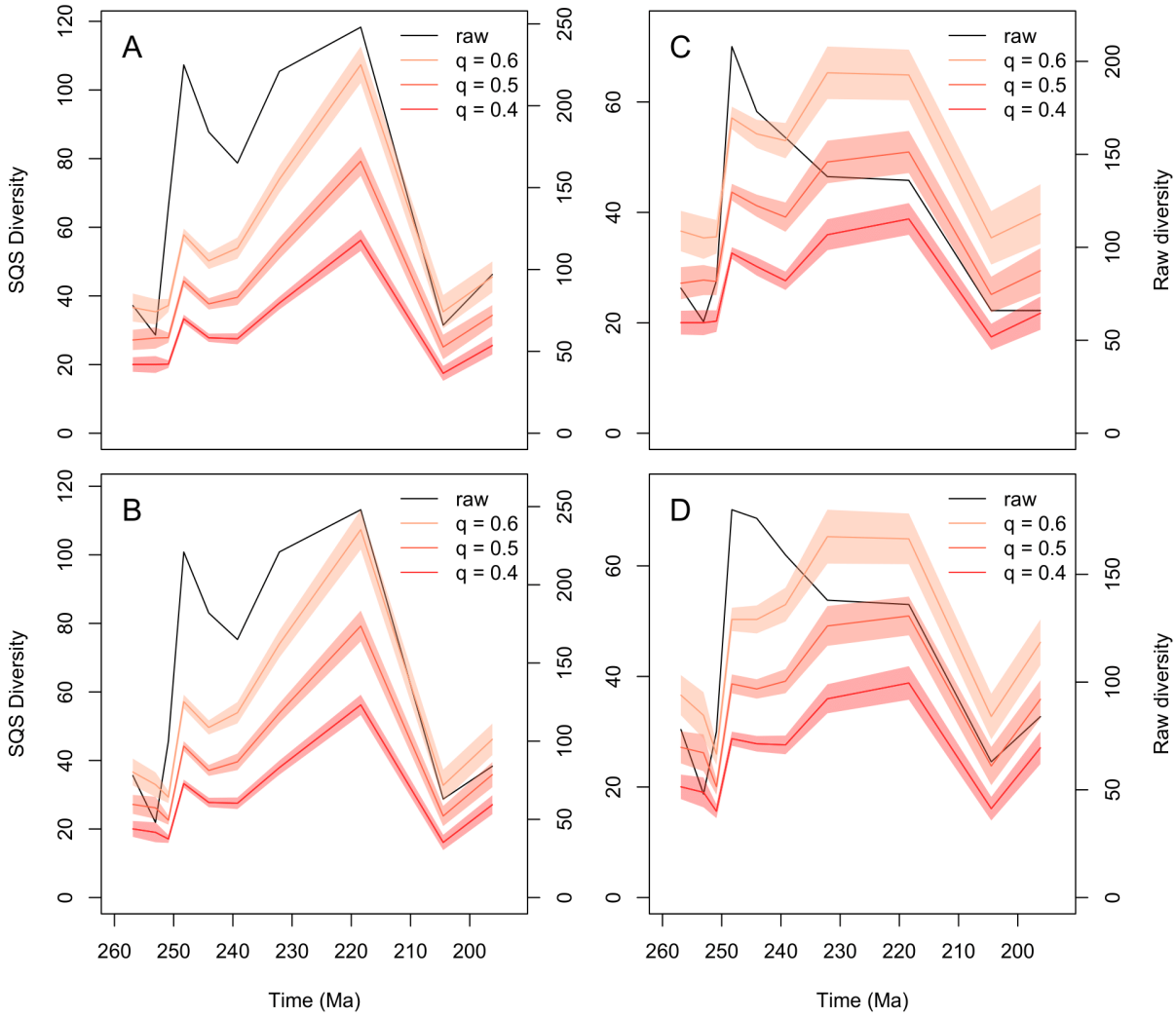


Fig. S17. Stage-wise sampling standardised diversity for the Boreal region. Diversity calculated using shareholder quorum subsampling at several quorum (q) levels. **(A).** Unstandardised. **(B).** MST standardisation. **(C).** Longitude-latitude standardisation. **(D).** MST + longitude latitude standardisation.

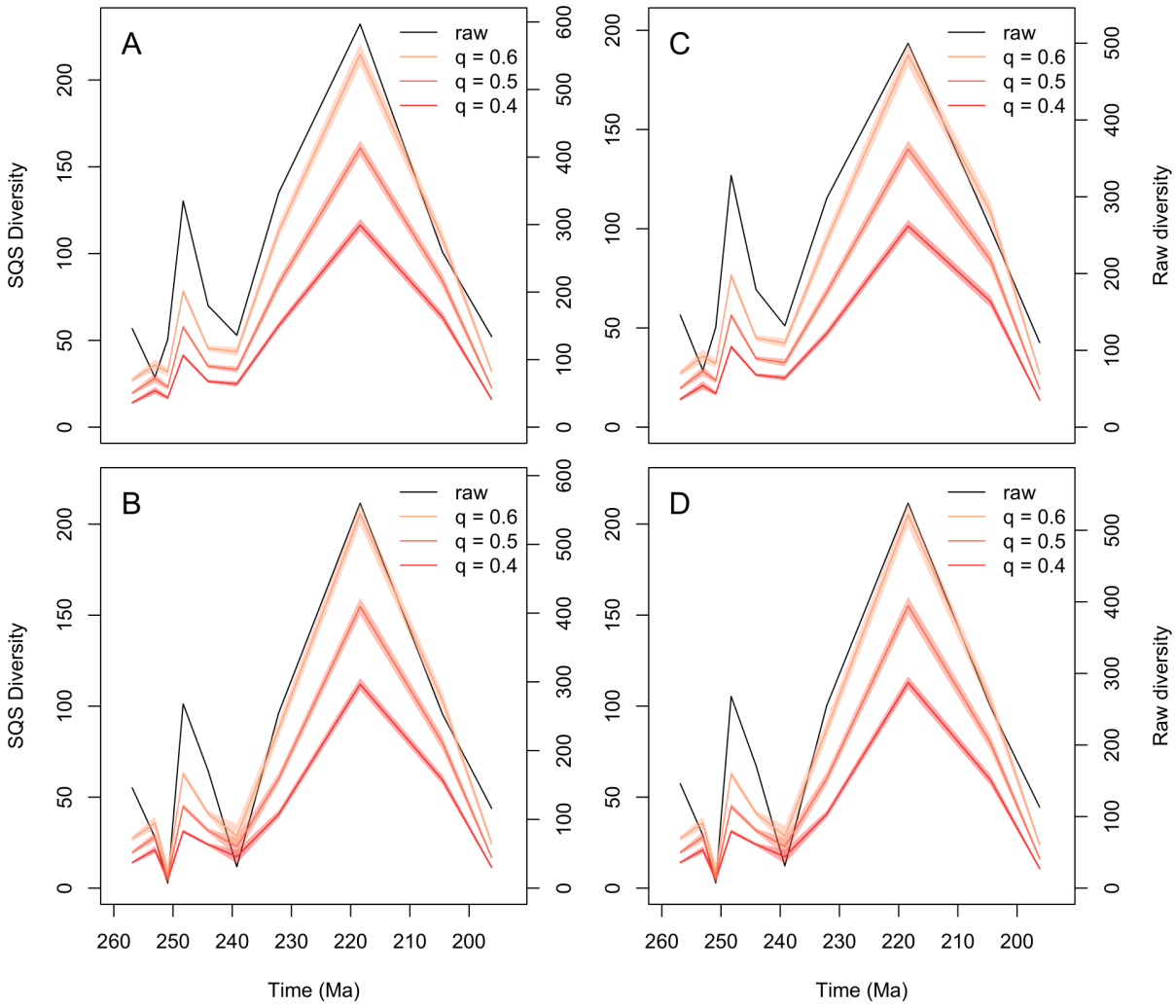


Fig. S18. Stage-wise sampling standardised diversity for the North Panthalassic region. Diversity calculated using shareholder quorum subsampling at several quorum (q) levels. **(A).** Unstandardised. **(B).** MST standardisation. **(C).** Longitude-latitude standardisation. **(D).** MST + longitude latitude standardisation.

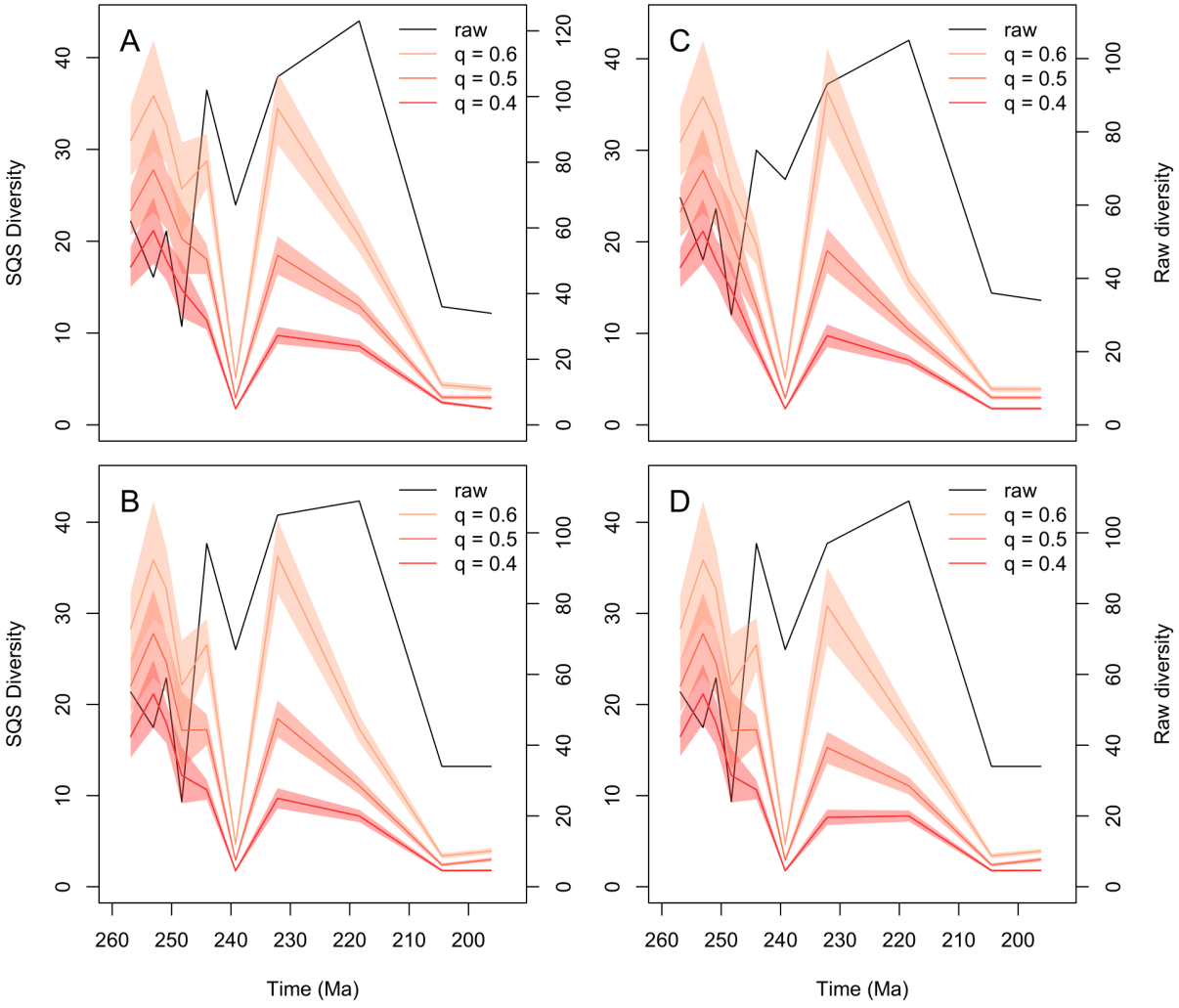


Fig. S19. Stage-wise sampling standardised diversity for the Tangaroan region. Diversity calculated using shareholder quorum subsampling at several quorum (q) levels. **(A).** Unstandardised. **(B).** MST standardisation. **(C).** Longitude-latitude standardisation. **(D).** MST + longitude latitude standardisation.

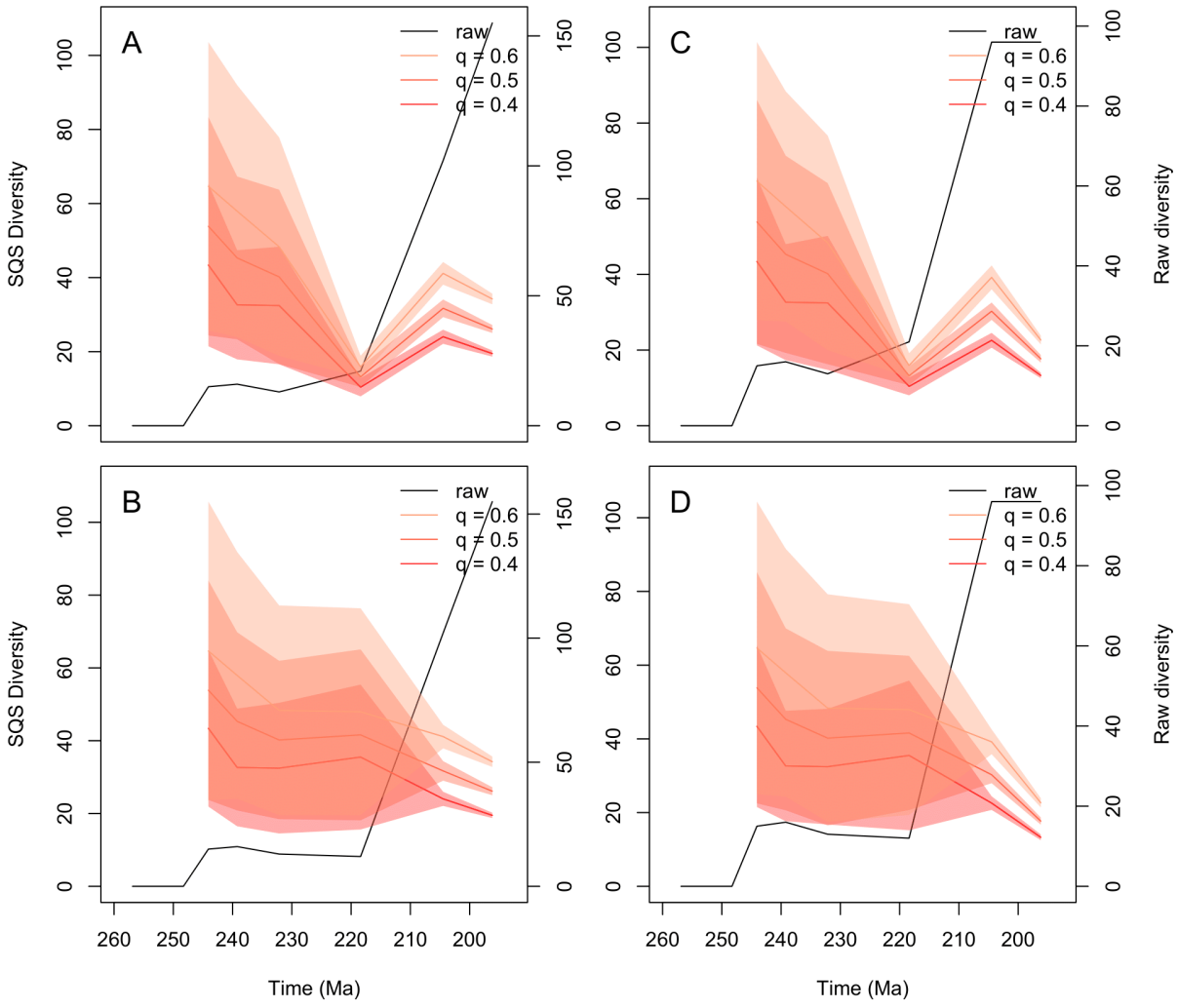


Fig. S20. Stage-wise sampling standardised diversity for the South Panthalassic region. Diversity calculated using shareholder quorum subsampling at several quorum (q) levels. **(A).** Unstandardised. **(B).** MST standardisation. **(C).** Longitude-latitude standardisation. **(D).** MST + longitude latitude standardisation.

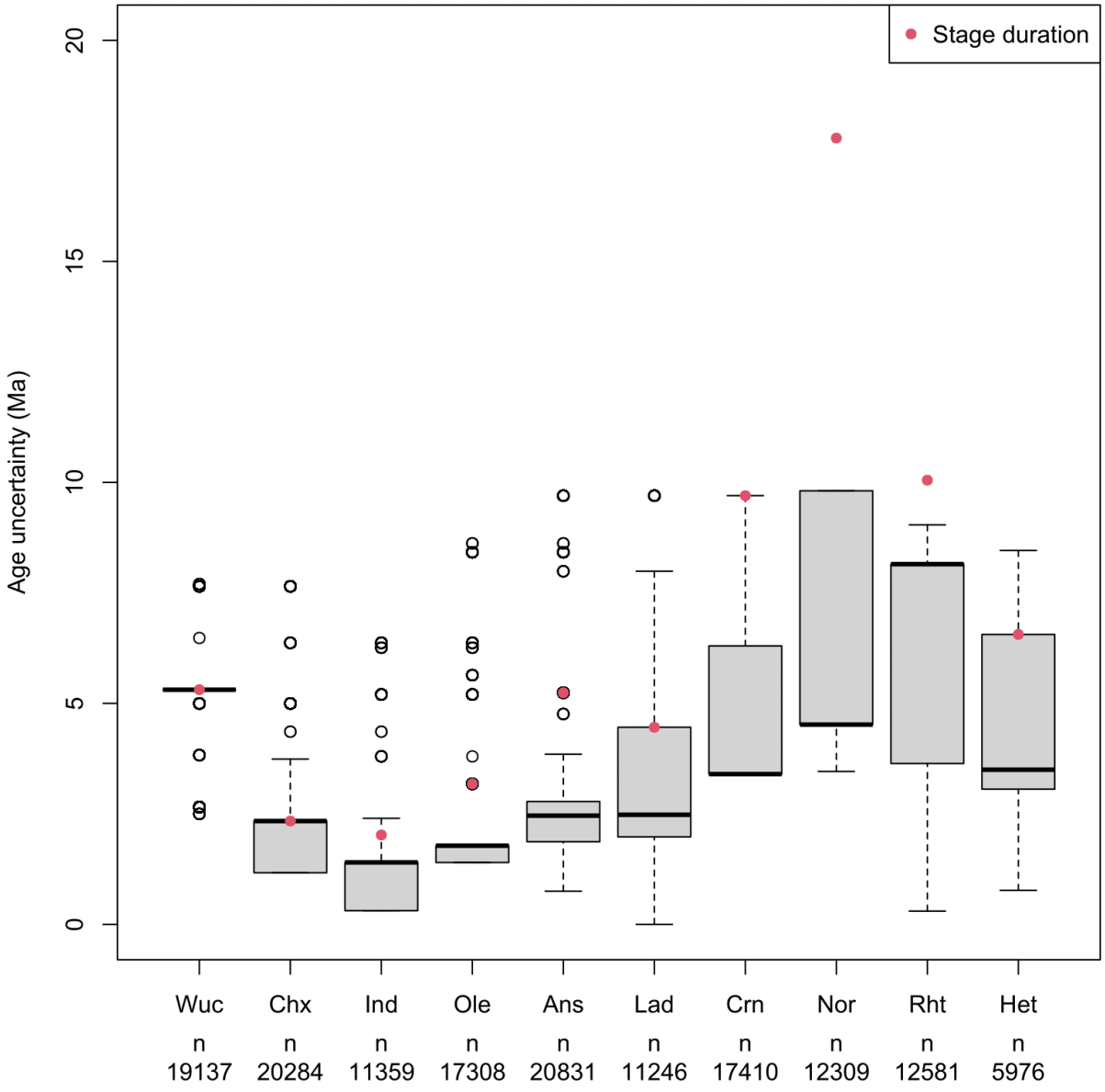


Fig. S21. Age uncertainty distributions in the stratigraphically revised fossil occurrence data. Relative to their parent bin duration, fossil occurrence ages are generally more precise, highlighting the improved age resolution in our dataset. Boxplots display the range, 1st and 3rd quartiles and median for each set of fossil occurrence age uncertainties. WUC = Wuchiapingian, CHX = Changhsingian, IND = Induan, OLE = Olenekian, ANS = Anisian, LAD = Ladinian, CRN = Carnian, NOR = Norian, RHT = Rhaetian, HET = Hettangian.

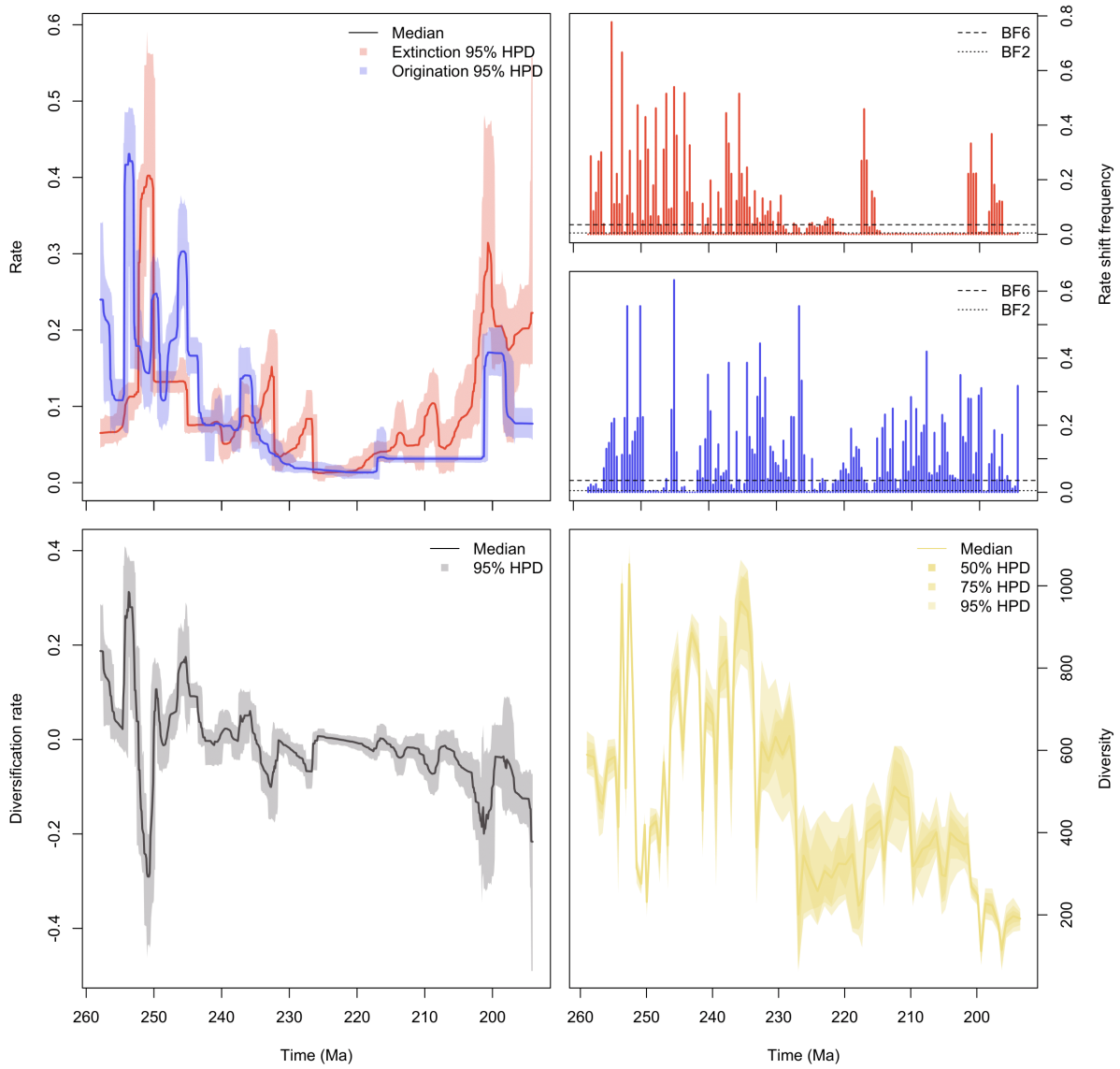


Fig. S22. Probabilistic diversity dynamics for Circumtethys (unstandardised). Diversity dynamics recovered from the regional fossil occurrence data by PyRate. BF = Bayes Factor. HPD = Highest Posterior density

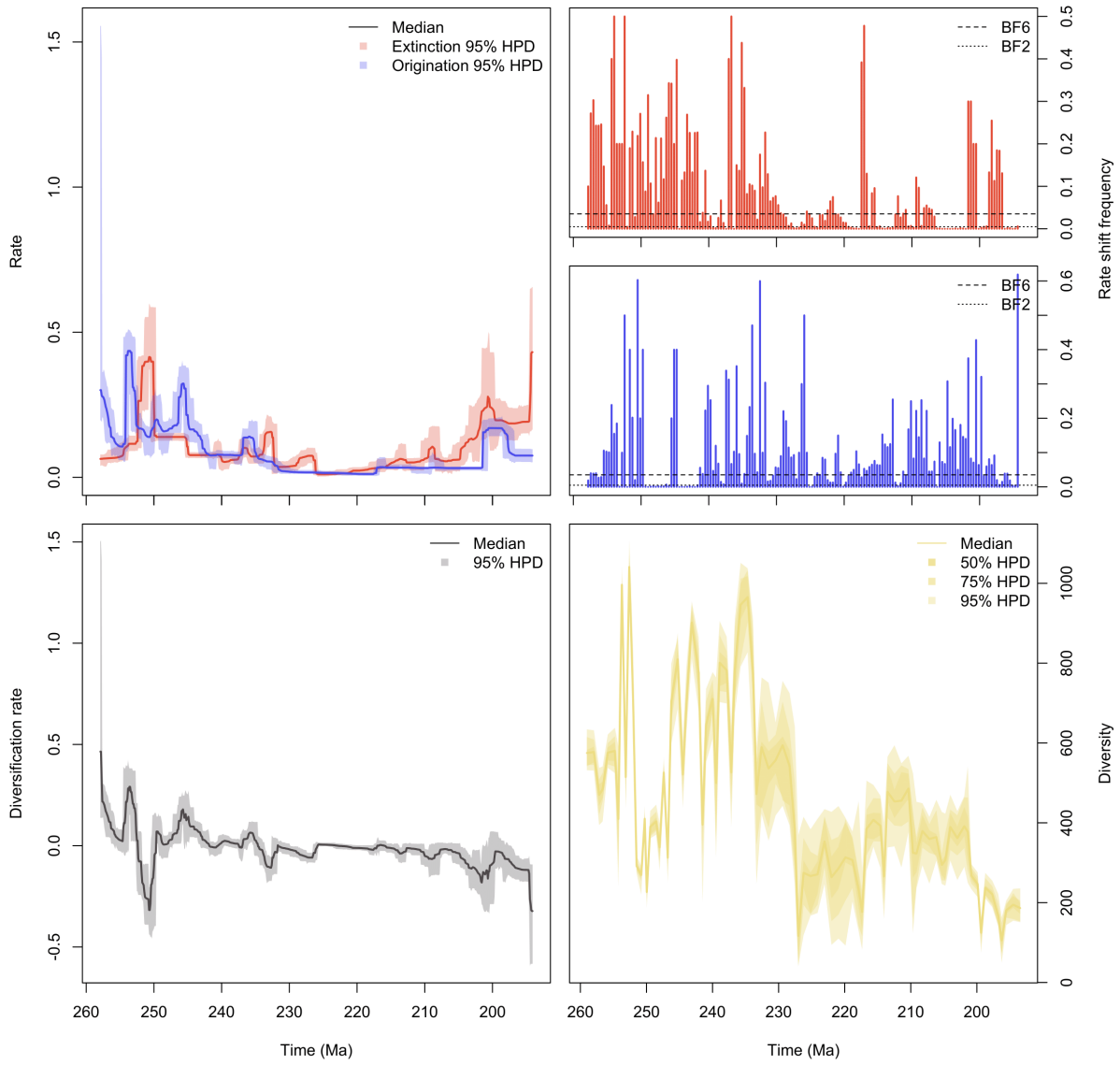


Fig. S23. Probabilistic diversity dynamics for Circumtethys (MST standardised). Diversity dynamics recovered from the regional fossil occurrence data by PyRate. BF = Bayes Factor. HPD = Highest Posterior density

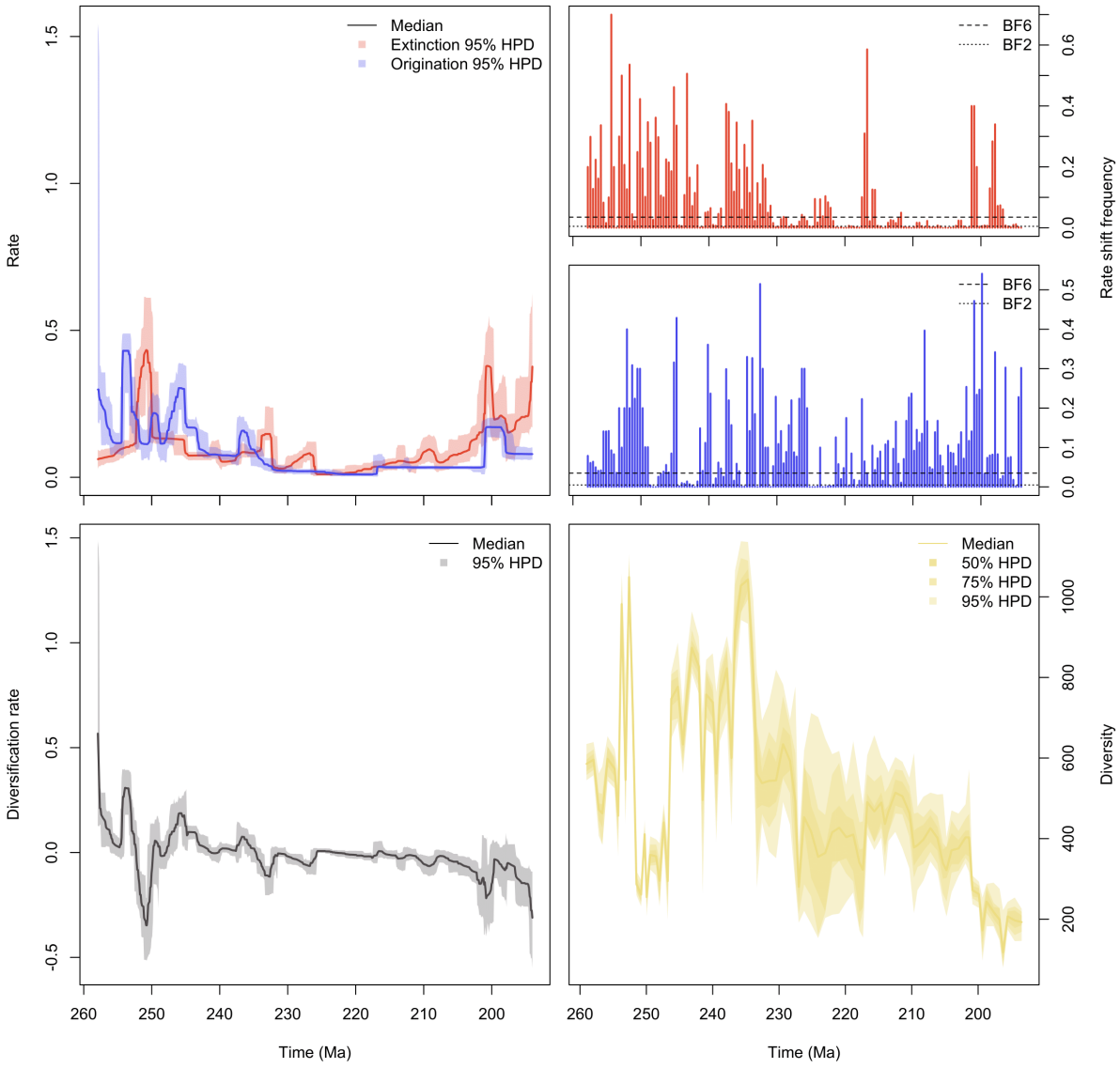


Fig. S24. Probabilistic diversity dynamics for Circumtethys (MST + lng-lat standardised). Diversity dynamics recovered from the regional fossil occurrence data by PyRate. BF = Bayes Factor. HPD = Highest Posterior density

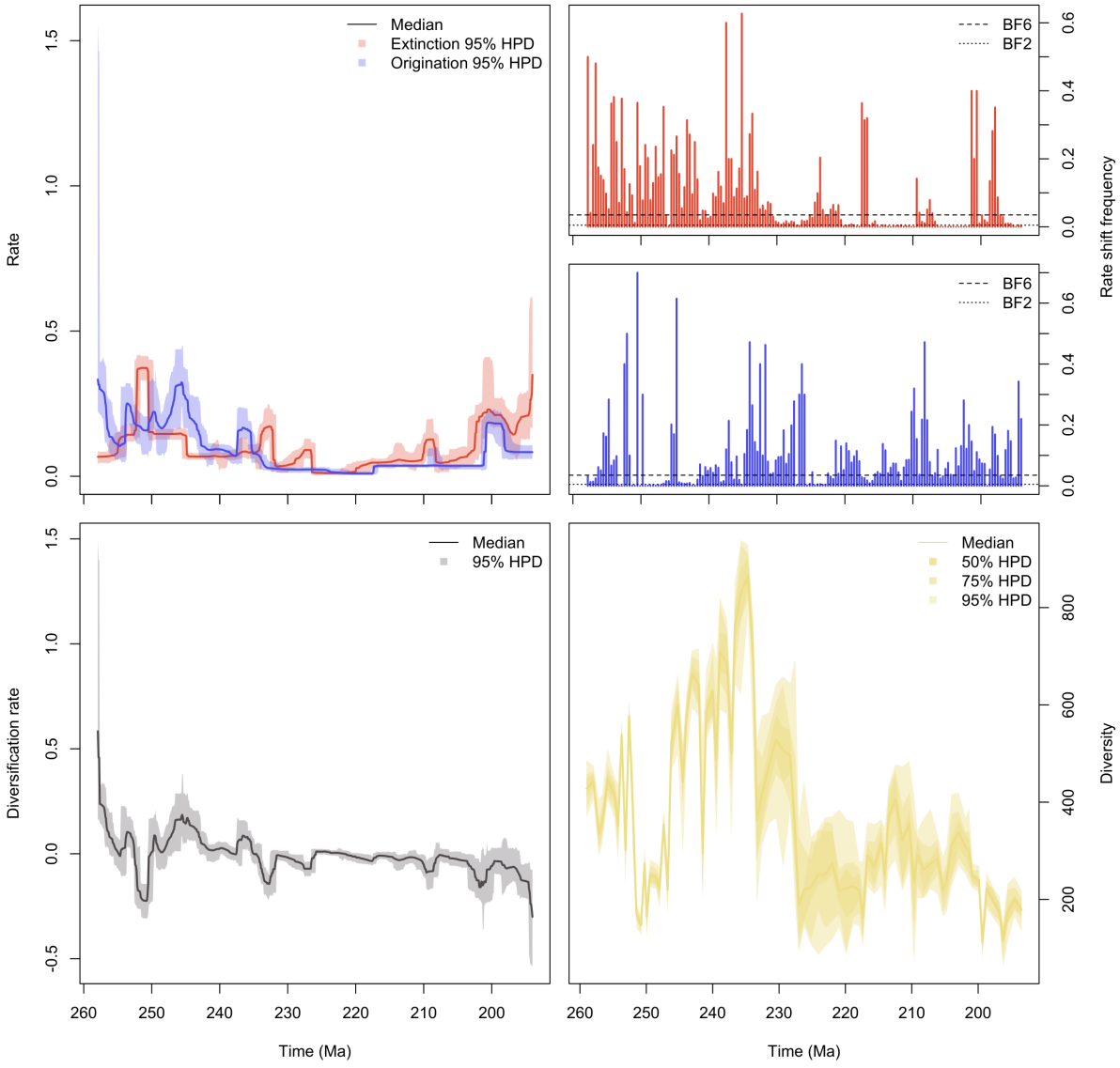


Fig. S25. Probabilistic diversity dynamics for West Circumtethys (unstandardised). Diversity dynamics recovered from the regional fossil occurrence data by PyRate. BF = Bayes Factor. HPD = Highest Posterior density

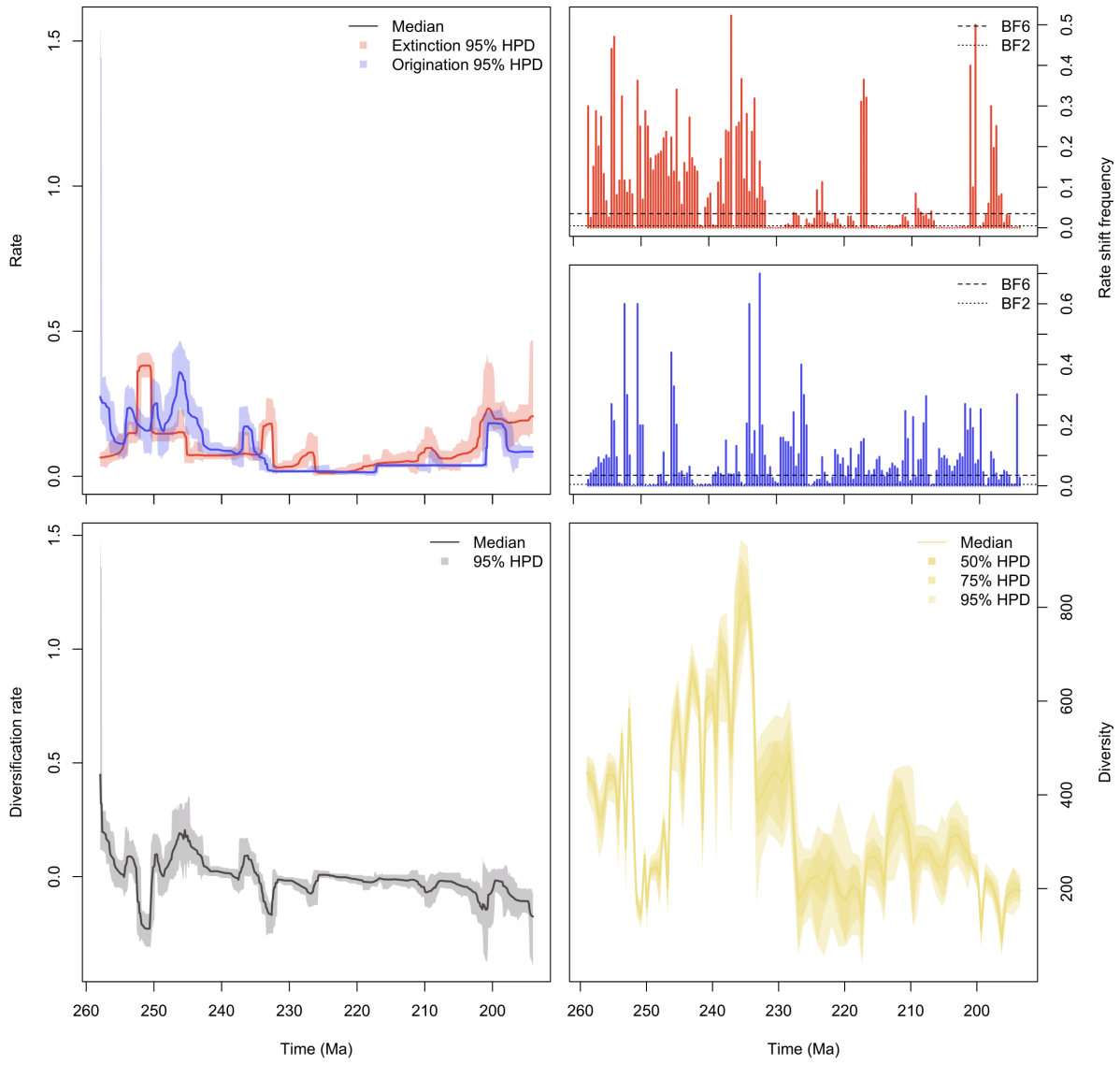


Fig. S26. Probabilistic diversity dynamics for West Circumtethys (MST standardised). Diversity dynamics recovered from the regional fossil occurrence data by PyRate. BF = Bayes Factor. HPD = Highest Posterior density

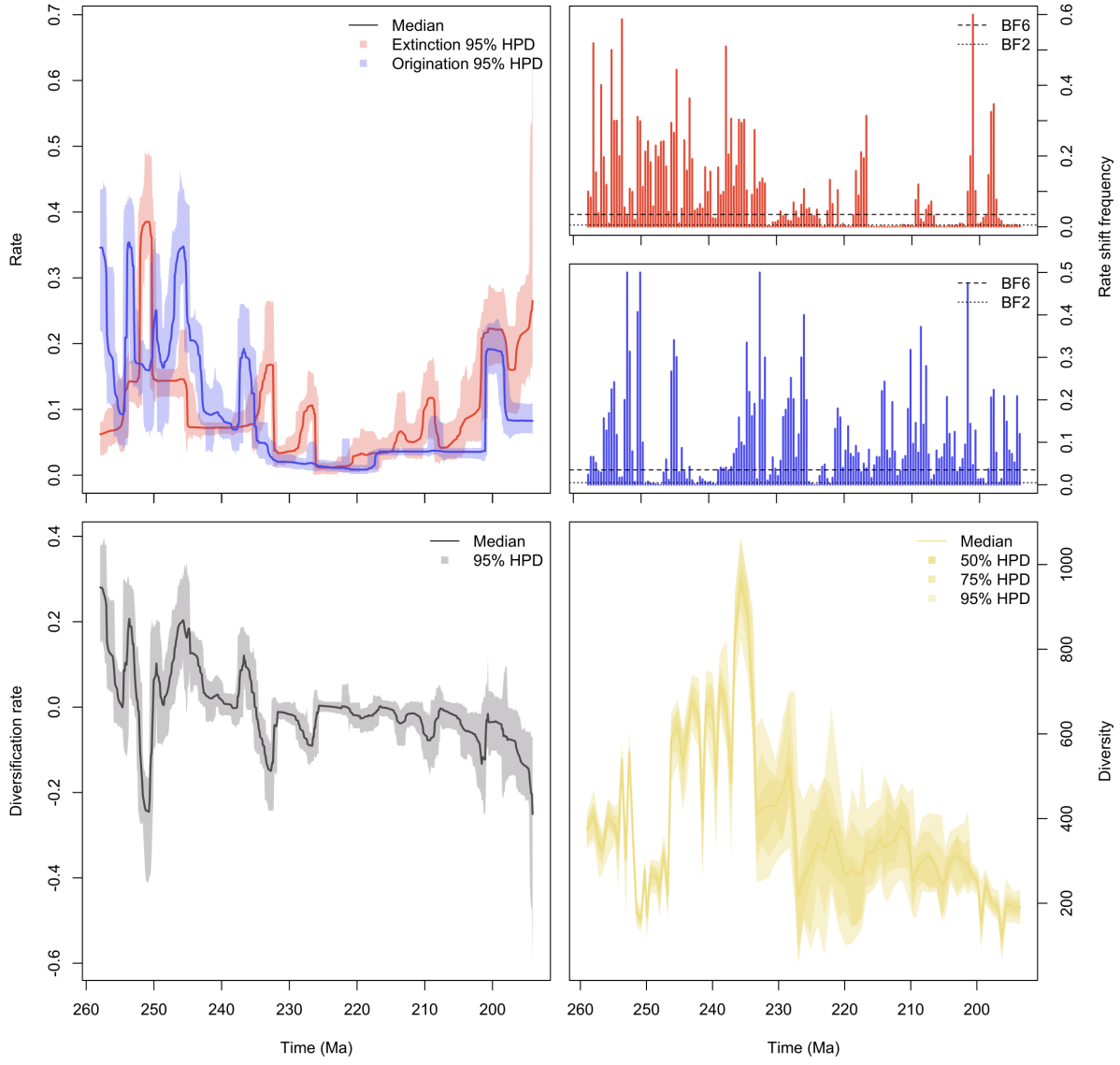


Fig. S27. Probabilistic diversity dynamics for West Circumtethys (MST + lng-lat standardised). Diversity dynamics recovered from the regional fossil occurrence data by PyRate. BF = Bayes Factor. HPD = Highest Posterior density

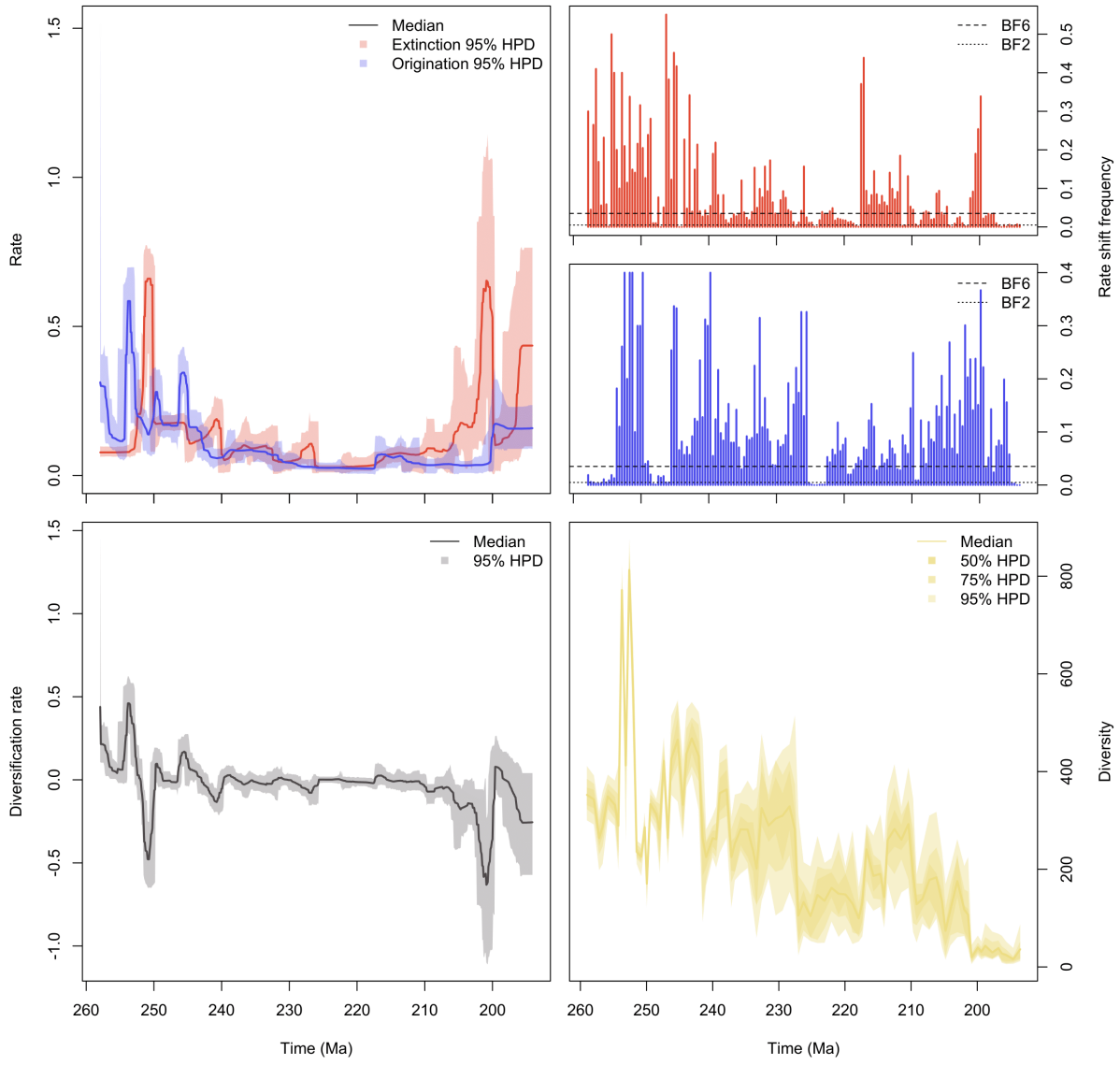


Fig. S28. Probabilistic diversity dynamics for East Circumtethys (unstandardised). Diversity dynamics recovered from the regional fossil occurrence data by PyRate. BF = Bayes Factor. HPD = Highest Posterior density

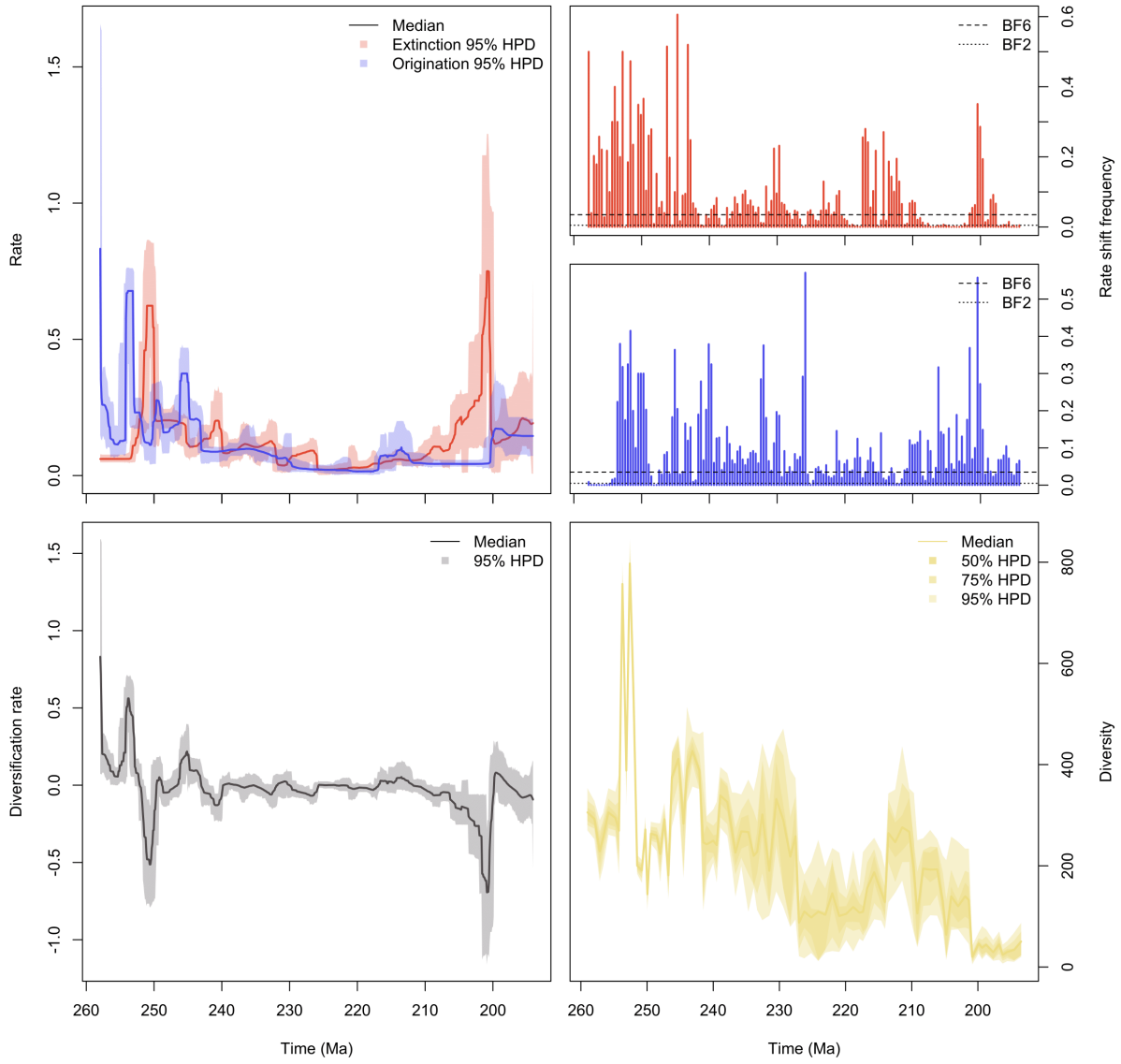


Fig. S29. Probabilistic diversity dynamics for East Circumtethys (MST standardised). Diversity dynamics recovered from the regional fossil occurrence data by PyRate. BF = Bayes Factor. HPD = Highest Posterior density

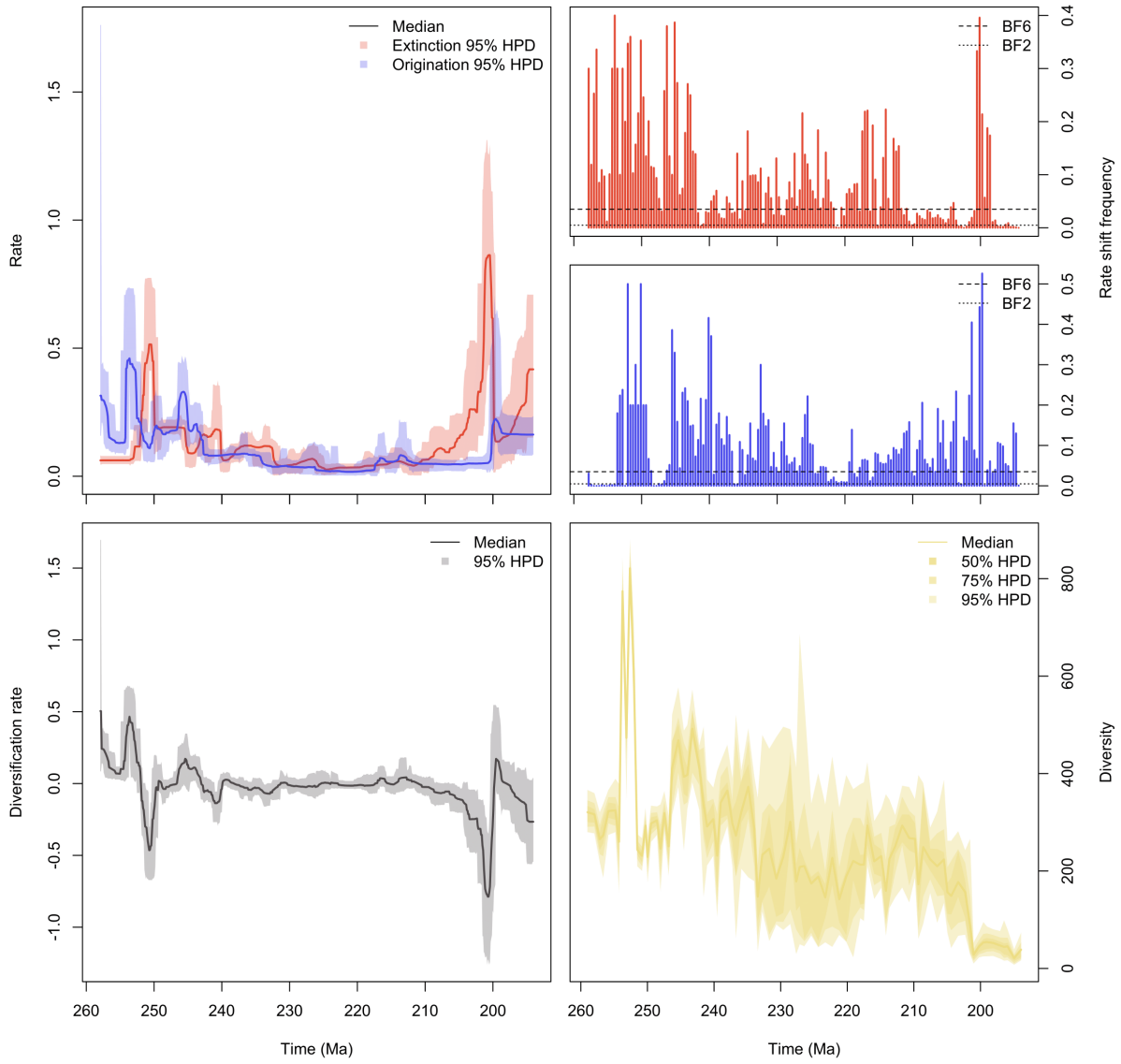


Fig. S30. Probabilistic diversity dynamics for East Circumtethys (MST + lng-lat standardised). Diversity dynamics recovered from the regional fossil occurrence data by PyRate. BF = Bayes Factor. HPD = Highest Posterior density

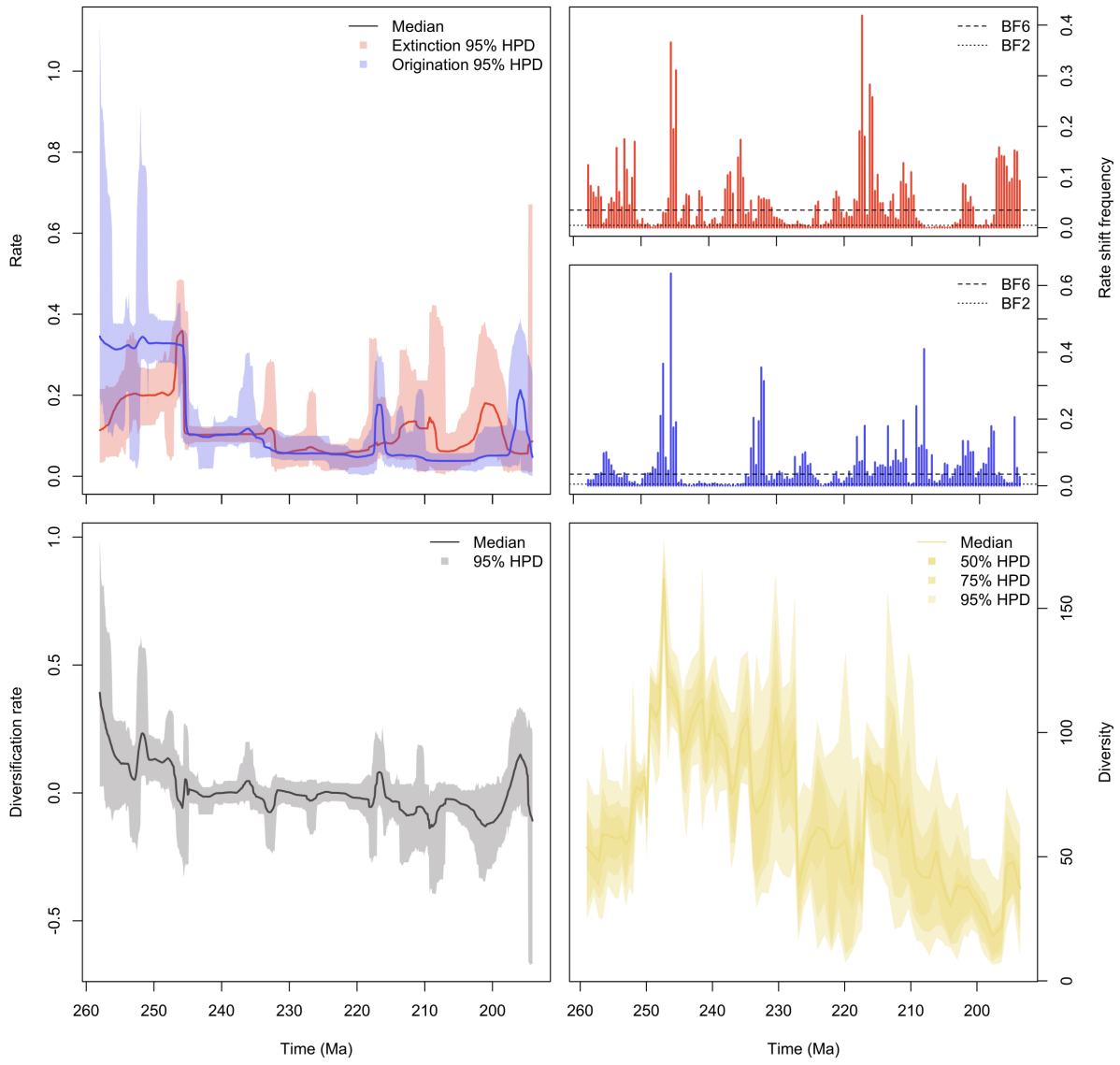


Fig. S31. Probabilistic diversity dynamics for Boreal (unstandardised). Diversity dynamics recovered from the regional fossil occurrence data by PyRate. BF = Bayes Factor. HPD = Highest Posterior density

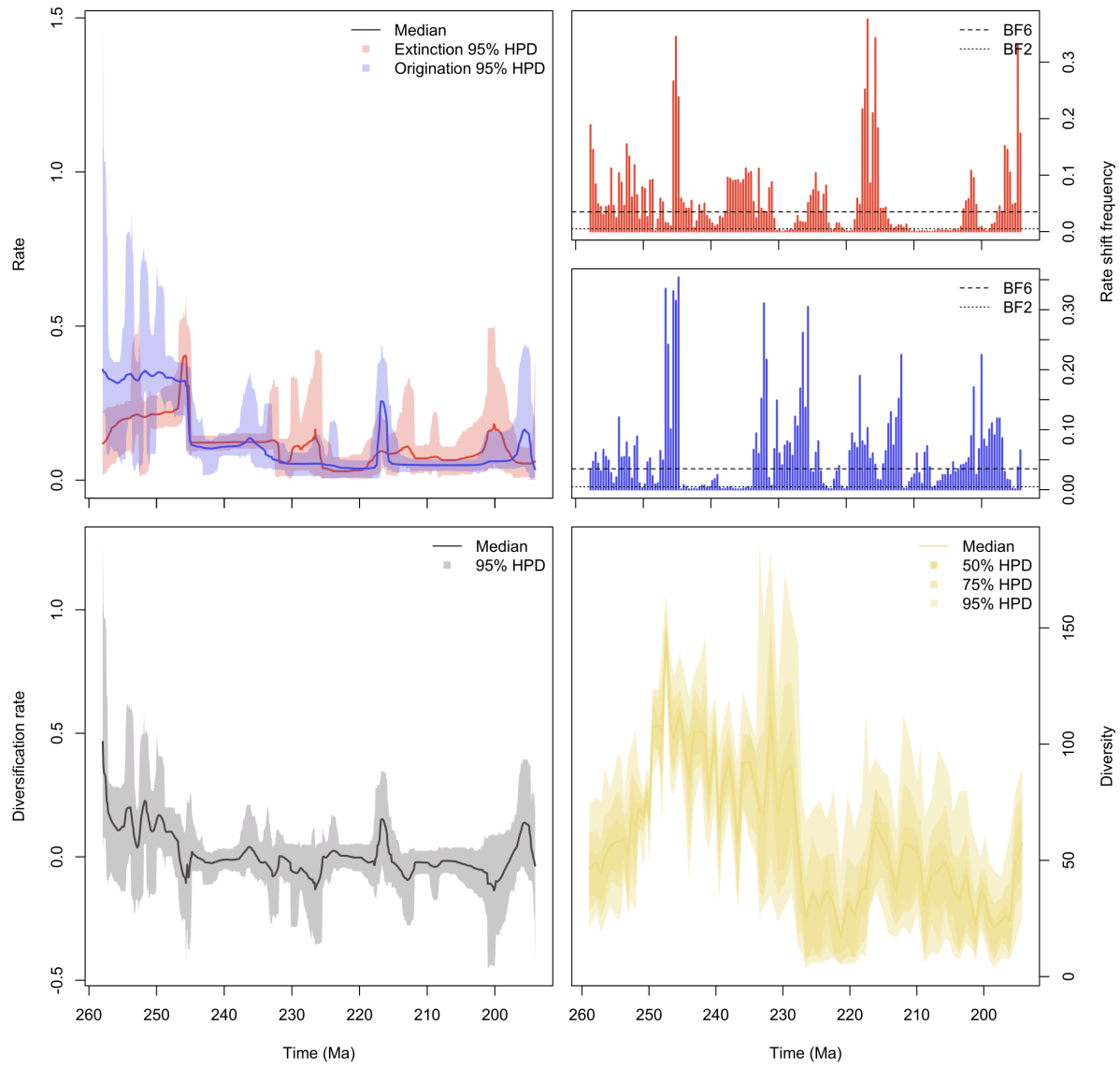


Fig. S32. Probabilistic diversity dynamics for Boreal (MST standardised). Diversity dynamics recovered from the regional fossil occurrence data by PyRate. BF = Bayes Factor. HPD = Highest Posterior density

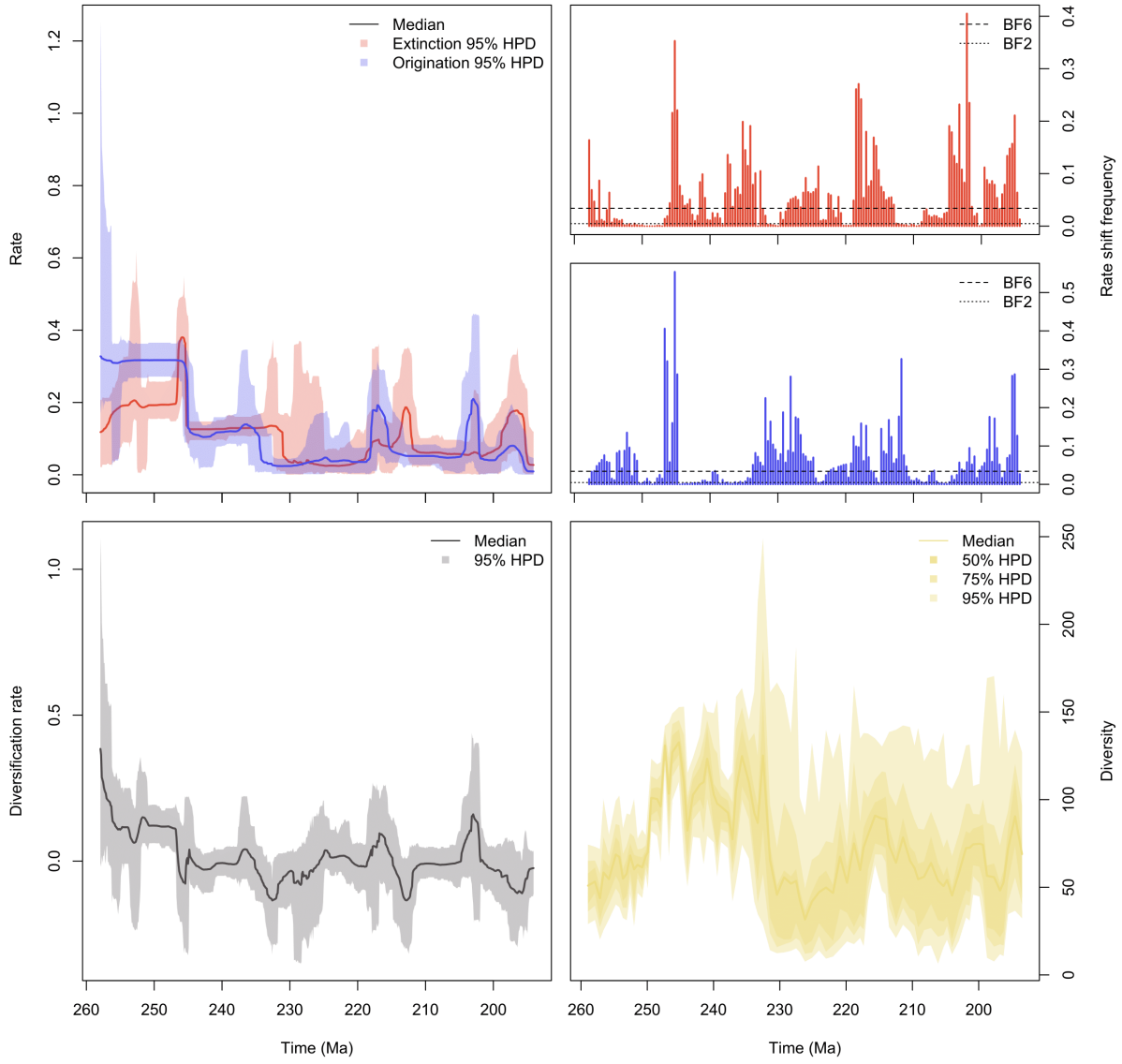


Fig. S33. Probabilistic diversity dynamics for Boreal (MST + lng-lat standardised). Diversity dynamics recovered from the regional fossil occurrence data by PyRate. BF = Bayes Factor. HPD = Highest Posterior density

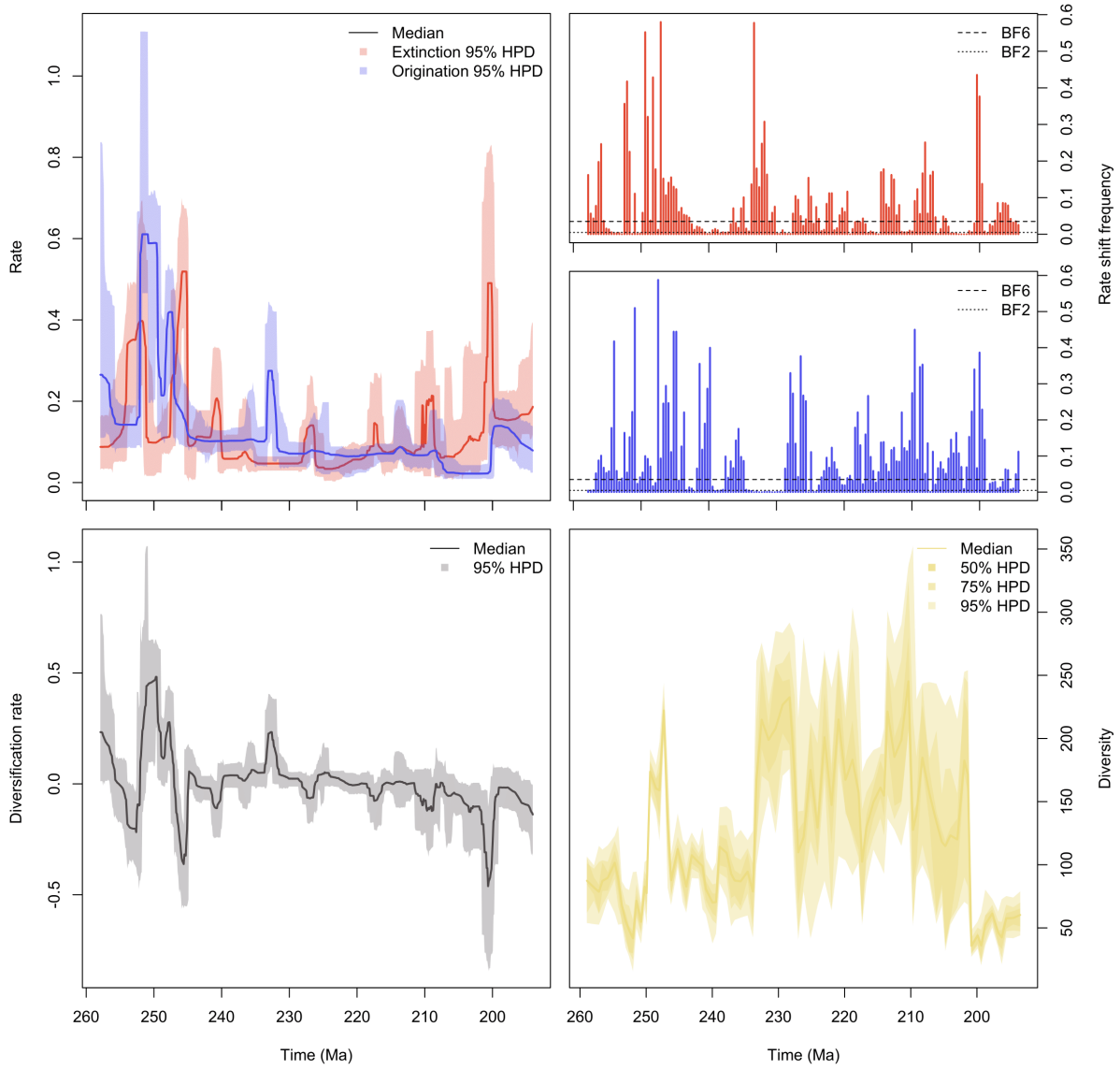


Fig. S34. Probabilistic diversity dynamics for North Panthalassic (unstandardised). Diversity dynamics recovered from the regional fossil occurrence data by PyRate. BF = Bayes Factor. HPD = Highest Posterior density

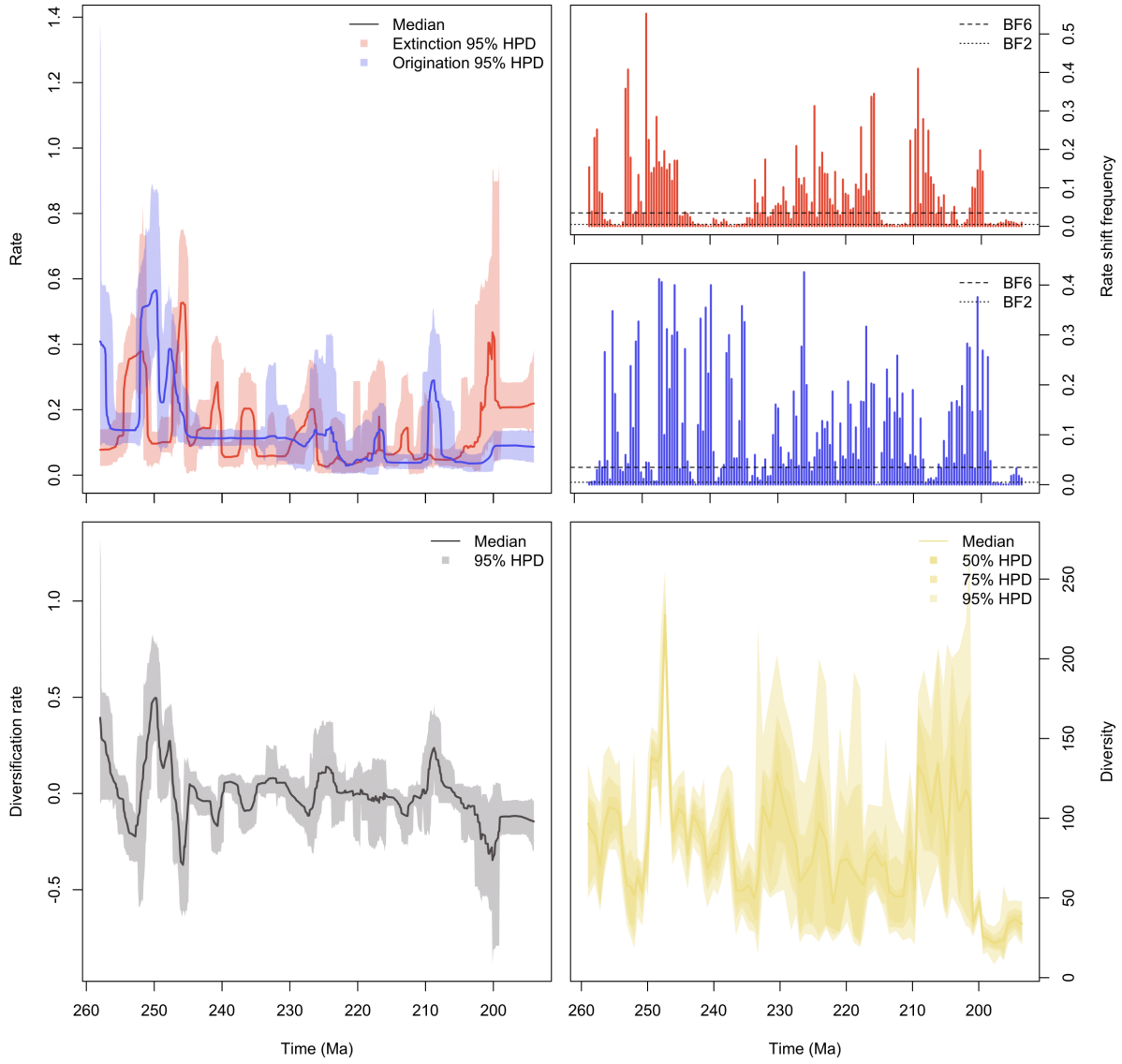


Fig. S35. Probabilistic diversity dynamics for North Panthalassic (MST standardised). Diversity dynamics recovered from the regional fossil occurrence data by PyRate. BF = Bayes Factor. HPD = Highest Posterior density

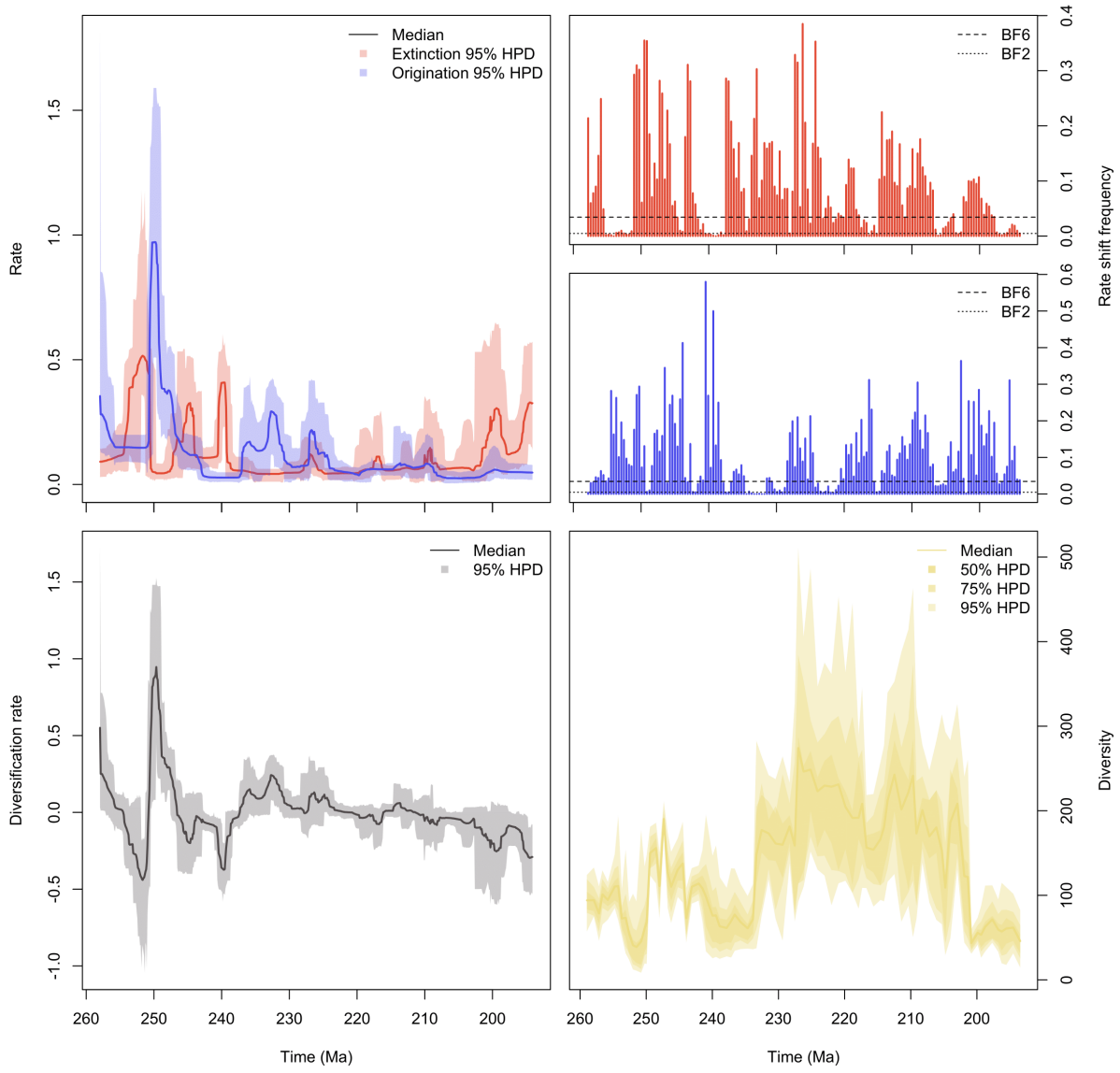


Fig. S36. Probabilistic diversity dynamics for North Panthalassic (MST + lng-lat standardised). Diversity dynamics recovered from the regional fossil occurrence data by PyRate. BF = Bayes Factor. HPD = Highest Posterior density

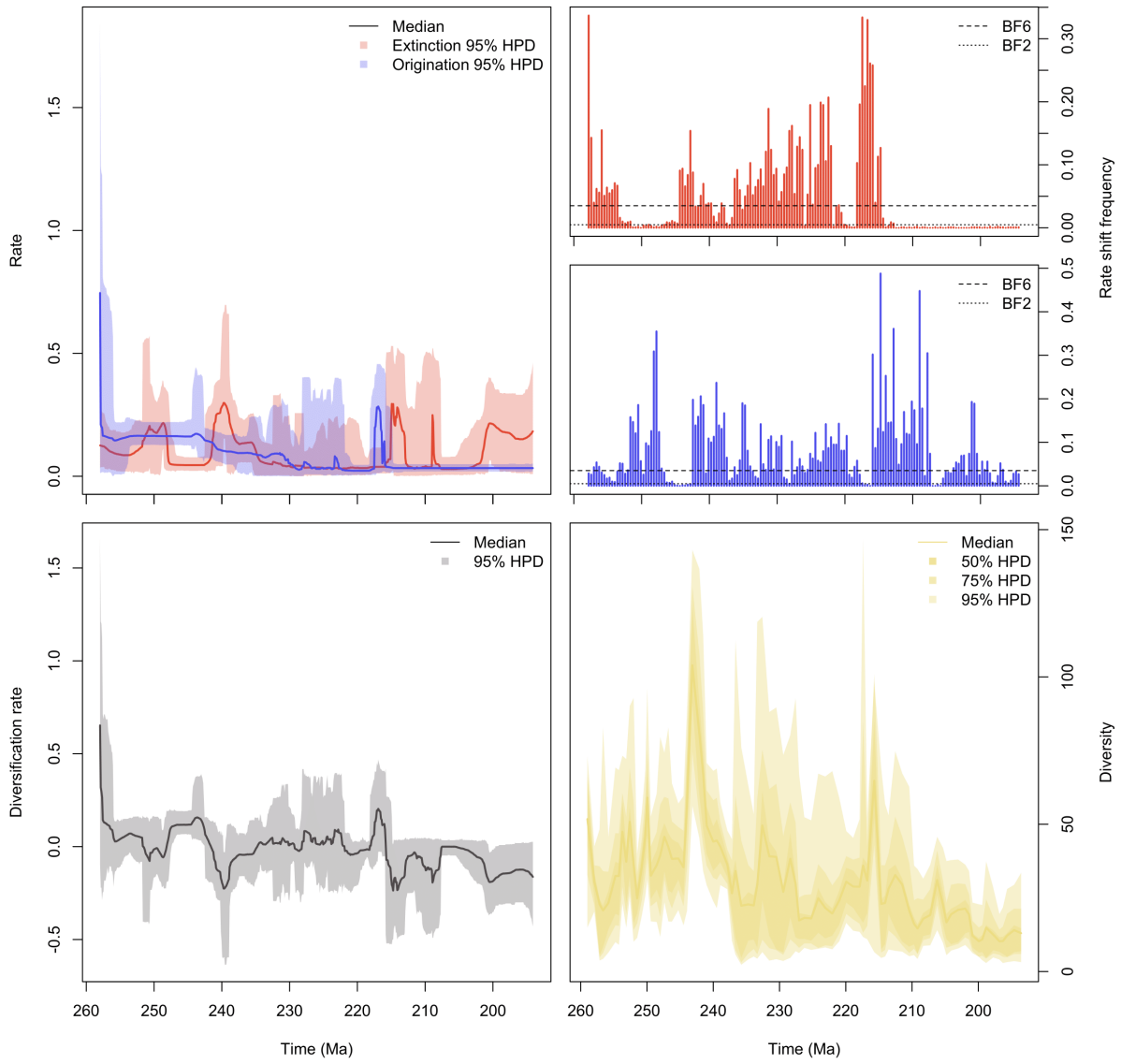


Fig. S37. Probabilistic diversity dynamics for Tangaroan (unstandardised). Diversity dynamics recovered from the regional fossil occurrence data by PyRate. BF = Bayes Factor. HPD = Highest Posterior density

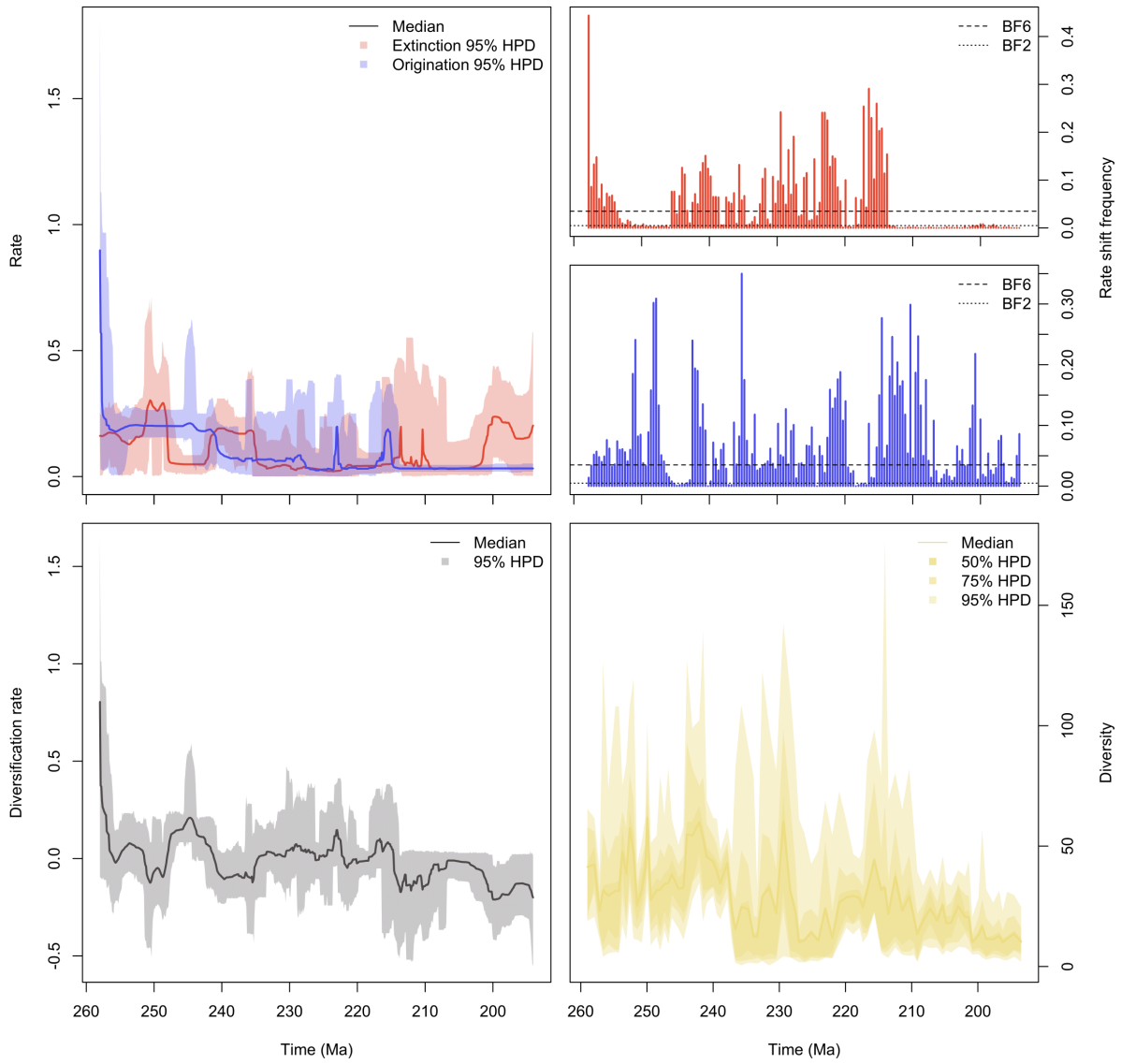


Fig. S38. Probabilistic diversity dynamics for Tangaroan (MST standardised). Diversity dynamics recovered from the regional fossil occurrence data by PyRate. BF = Bayes Factor. HPD = Highest Posterior density

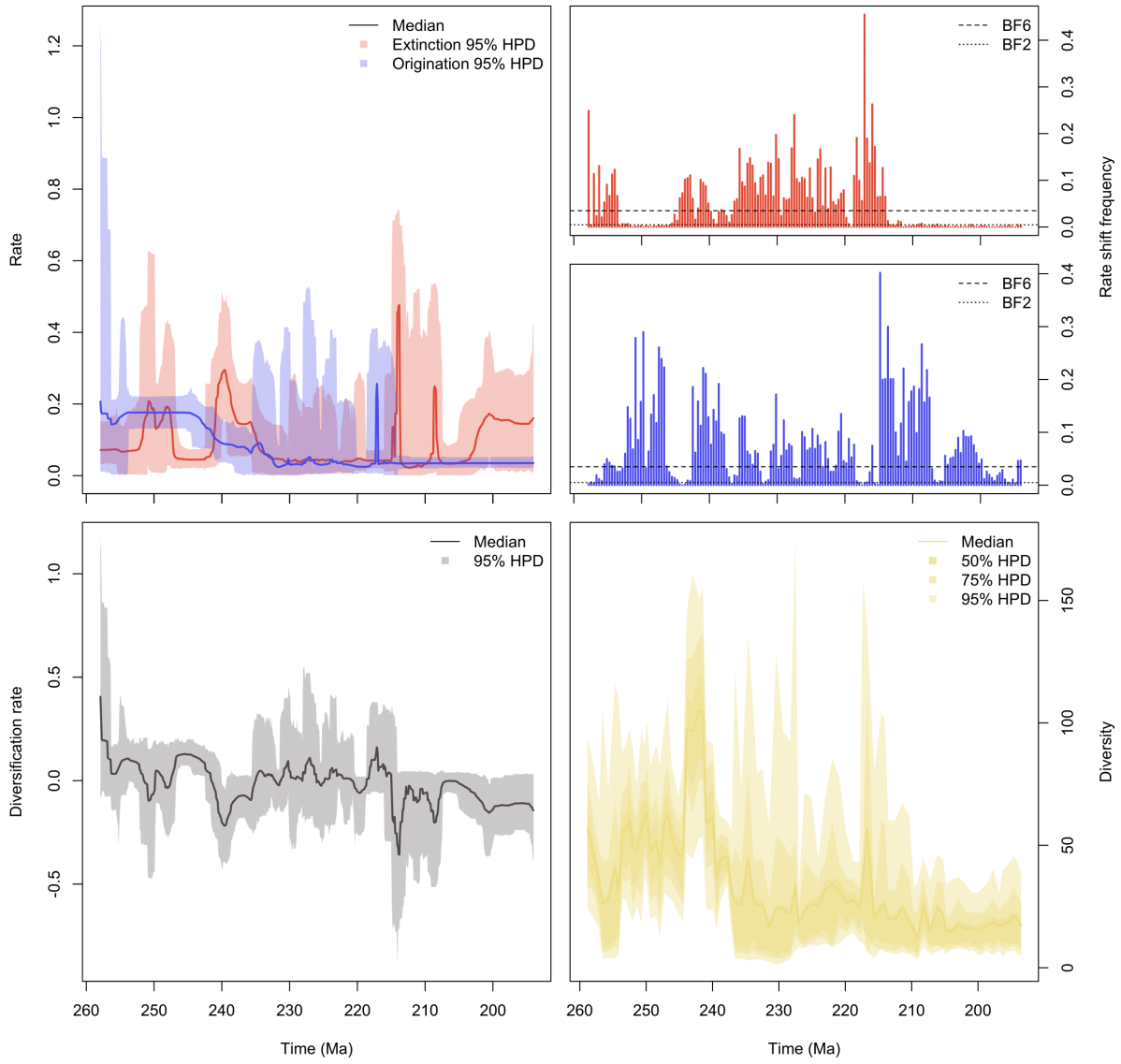


Fig. S39. Probabilistic diversity dynamics for Tangaroan (MST + lng-lat standardised). Diversity dynamics recovered from the regional fossil occurrence data by PyRate. BF = Bayes Factor. HPD = Highest Posterior density

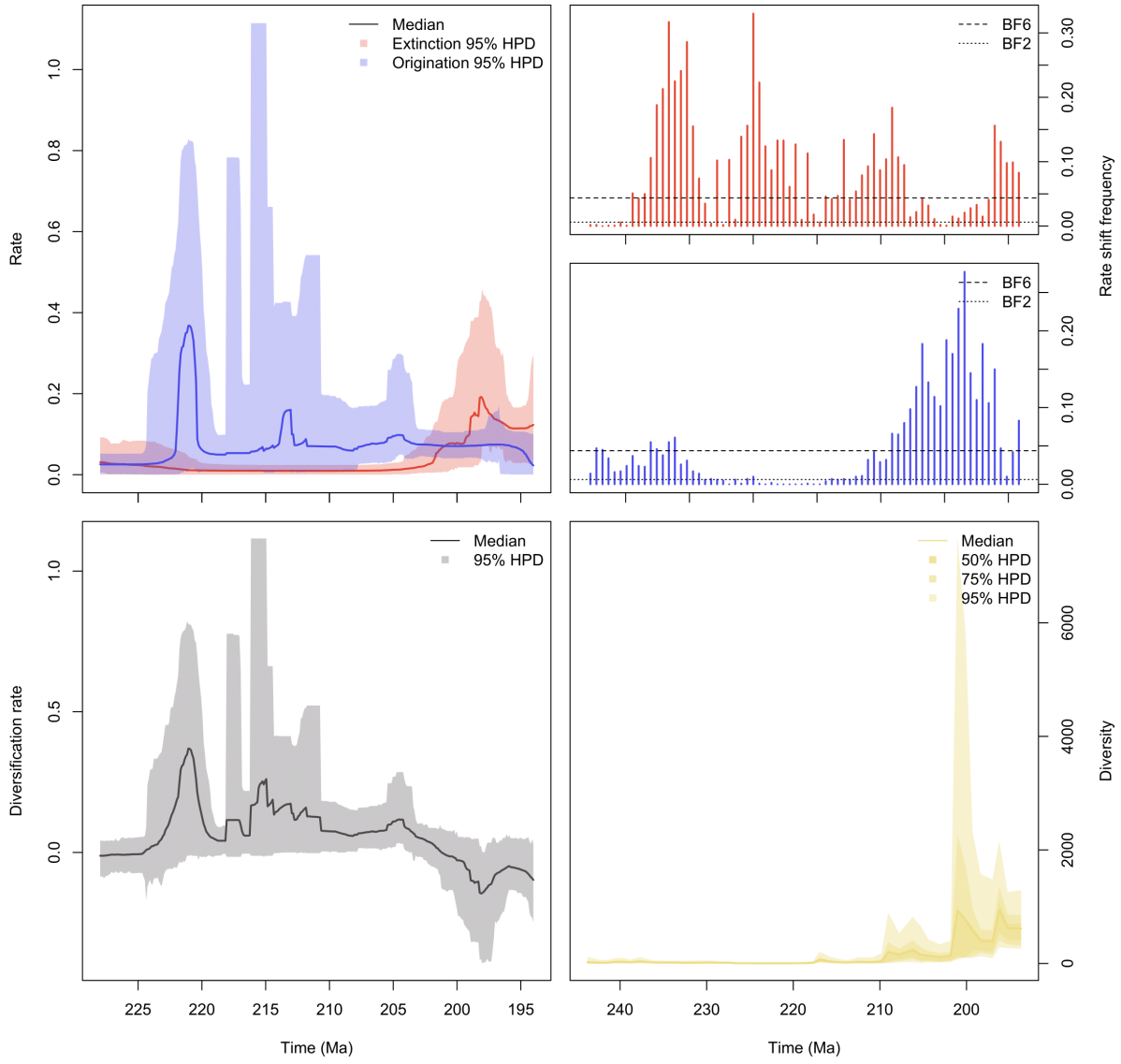


Fig. S40. Probabilistic diversity dynamics for South Panthalassic (unstandardised). Diversity dynamics recovered from the regional fossil occurrence data by PyRate. BF = Bayes Factor. HPD = Highest Posterior density

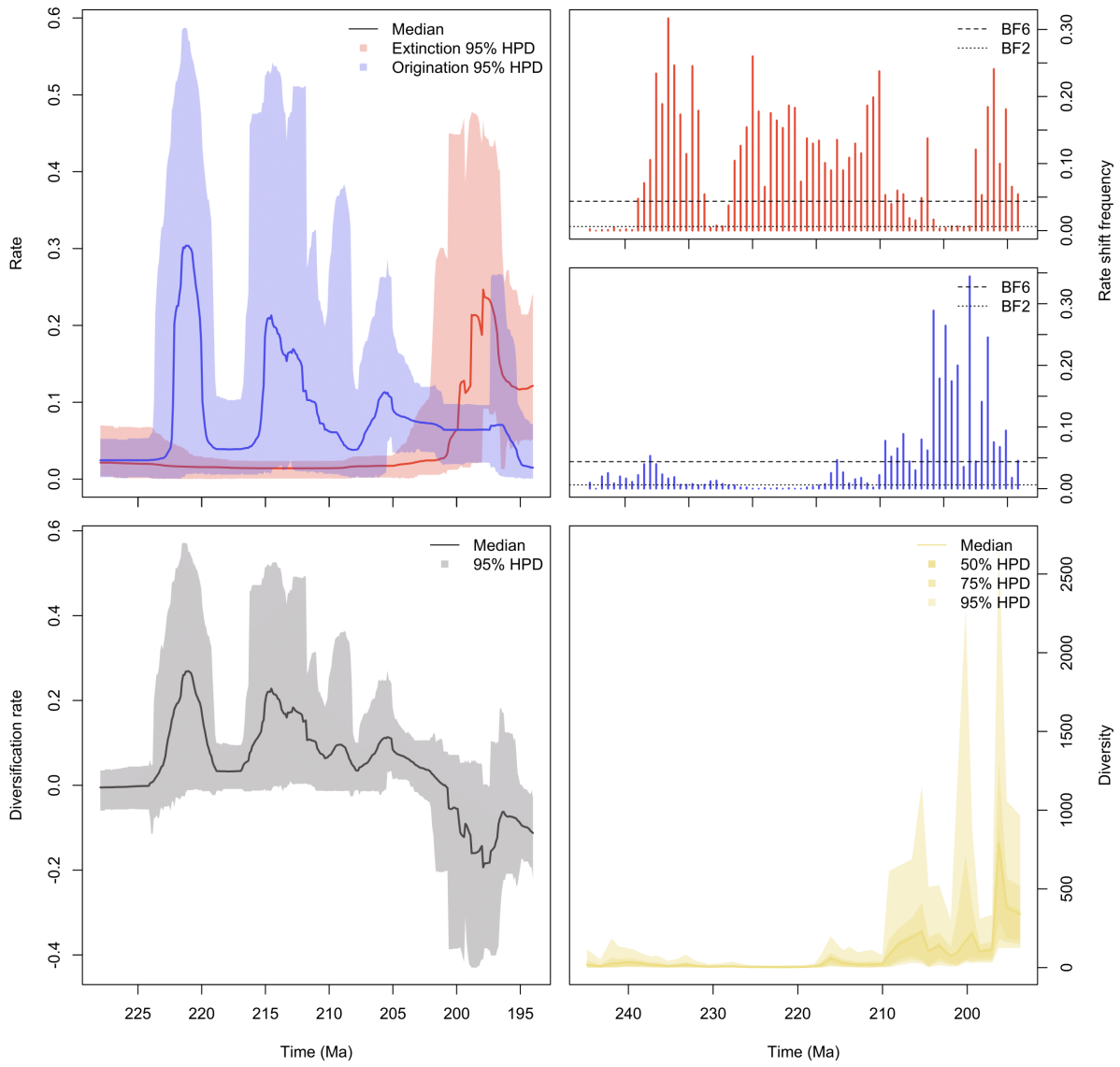


Fig. S41. Probabilistic diversity dynamics for South Panthalassic (MST standardised). Diversity dynamics recovered from the regional fossil occurrence data by PyRate. BF = Bayes Factor. HPD = Highest Posterior density

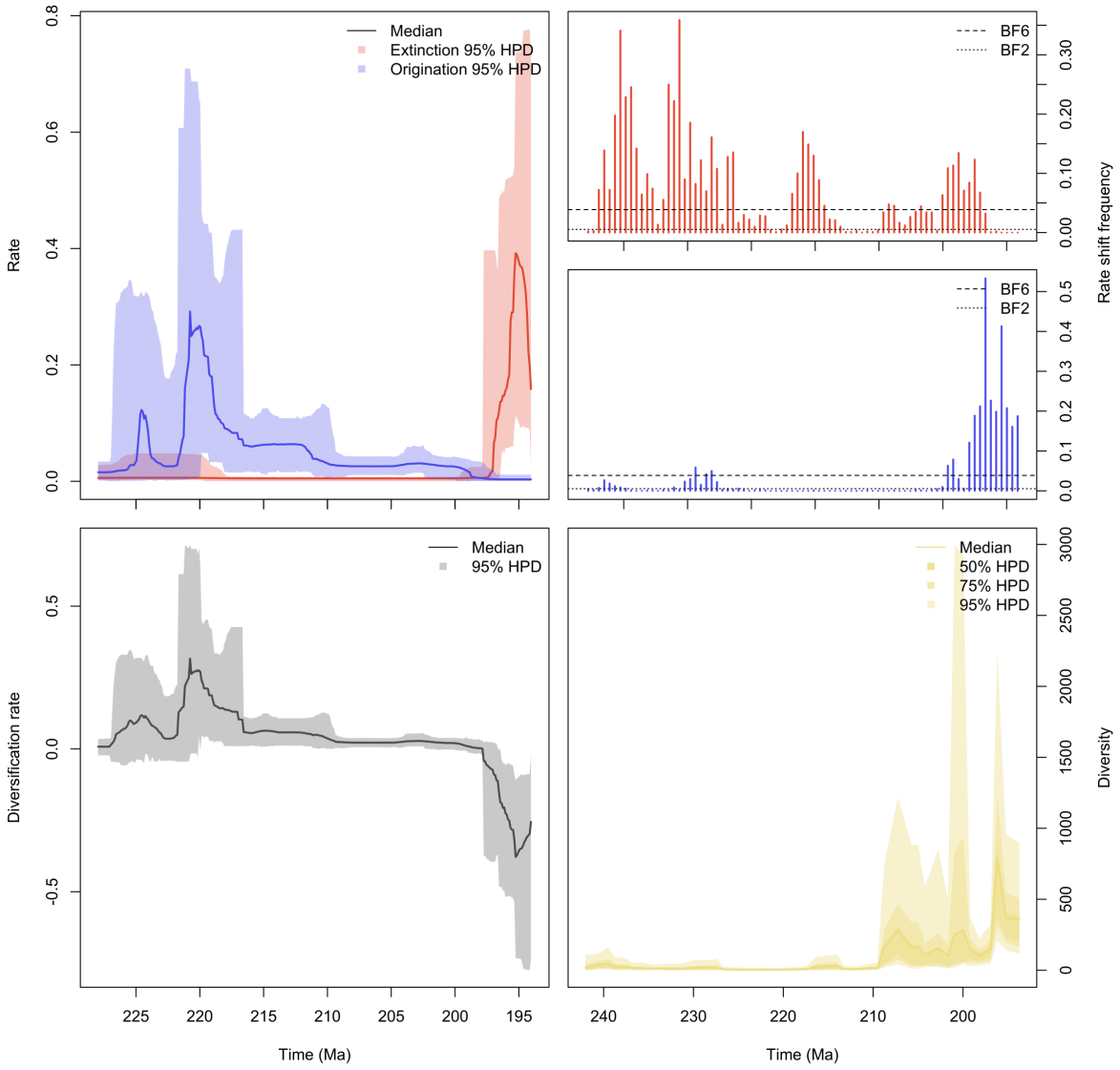


Fig. S42. Probabilistic diversity dynamics for South Panthalassic (MST + lng-lat standardised). Diversity dynamics recovered from the regional fossil occurrence data by PyRate. BF = Bayes Factor. HPD = Highest Posterior density

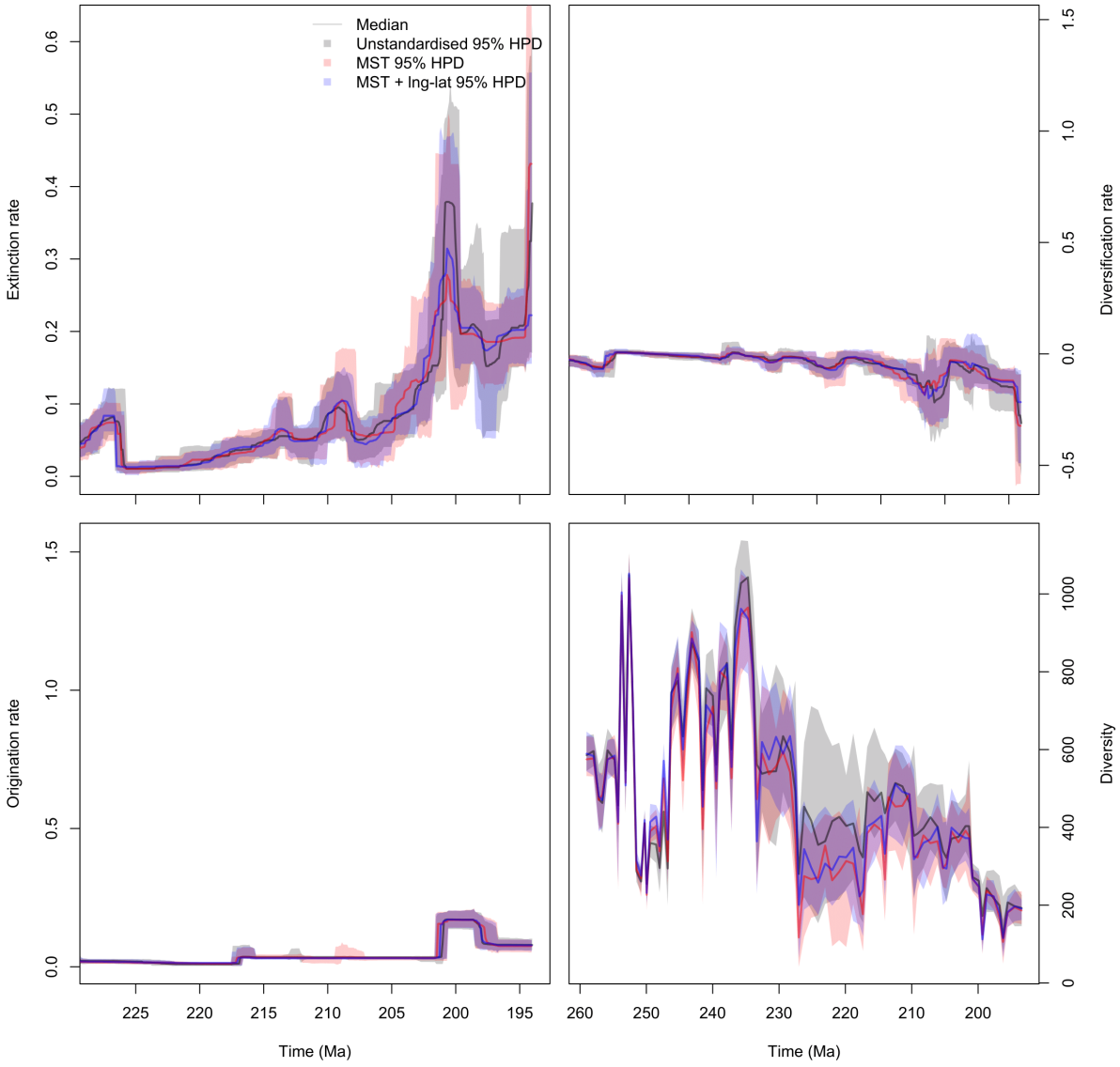


Fig. S43. Impact of spatial standardisation on diversity dynamics (Circumtethys). Comparison of diversity dynamics from unstandardised, MST-standardised, and MST + lnglat standardised datasets.

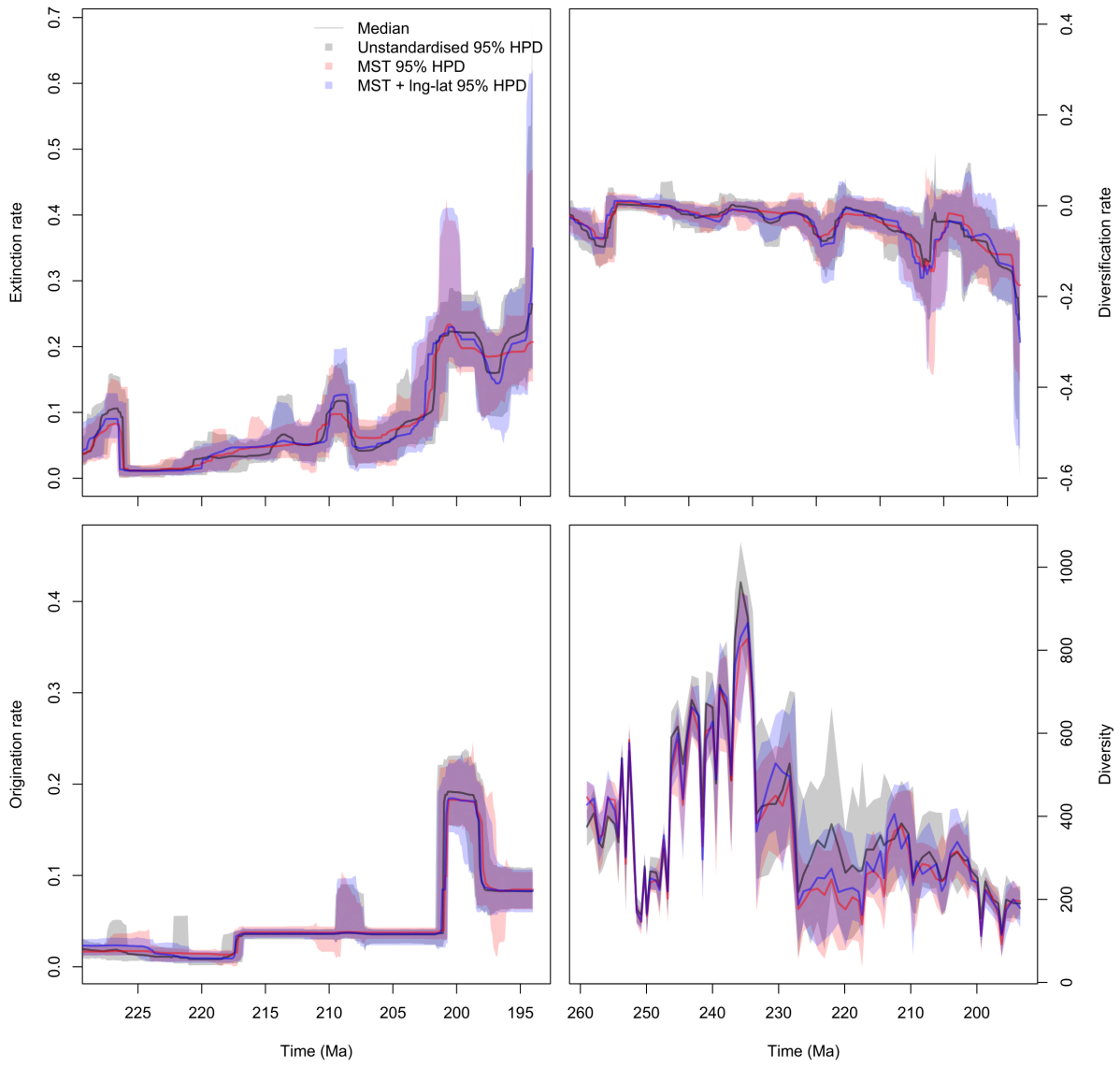


Fig. S44. Impact of spatial standardisation on diversity dynamics (West Circumtethys). Comparison of diversity dynamics from unstandardised, MST-standardised, and MST + lnglat standardised datasets.

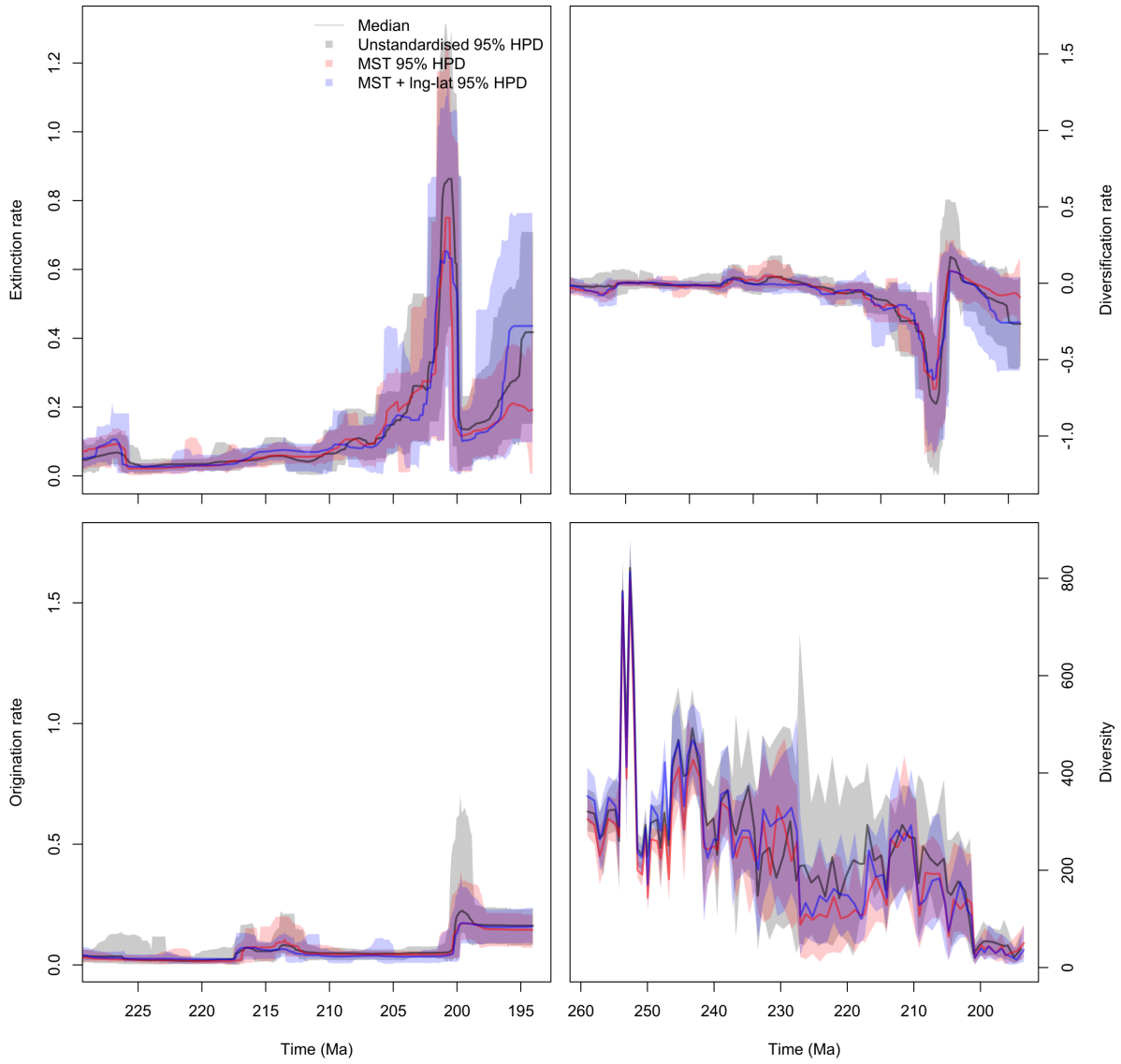


Fig. S45. Impact of spatial standardisation on diversity dynamics (East Circumtethys). Comparison of diversity dynamics from unstandardised, MST-standardised, and MST + lnglat standardised datasets.

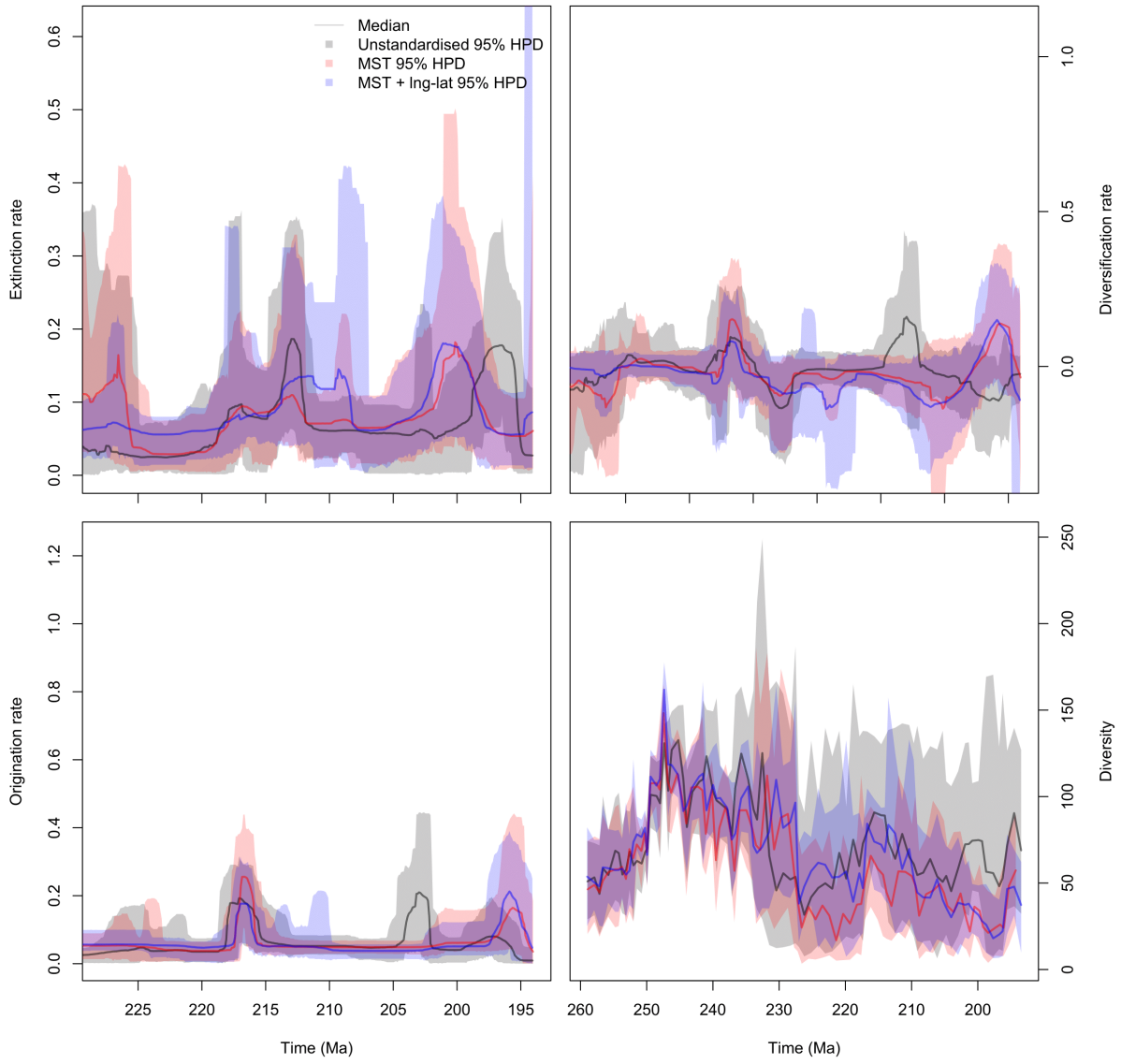


Fig. S46. Impact of spatial standardisation on diversity dynamics (Boreal). Comparison of diversity dynamics from unstandardised, MST-standardised, and MST + lnglat standardised datasets.

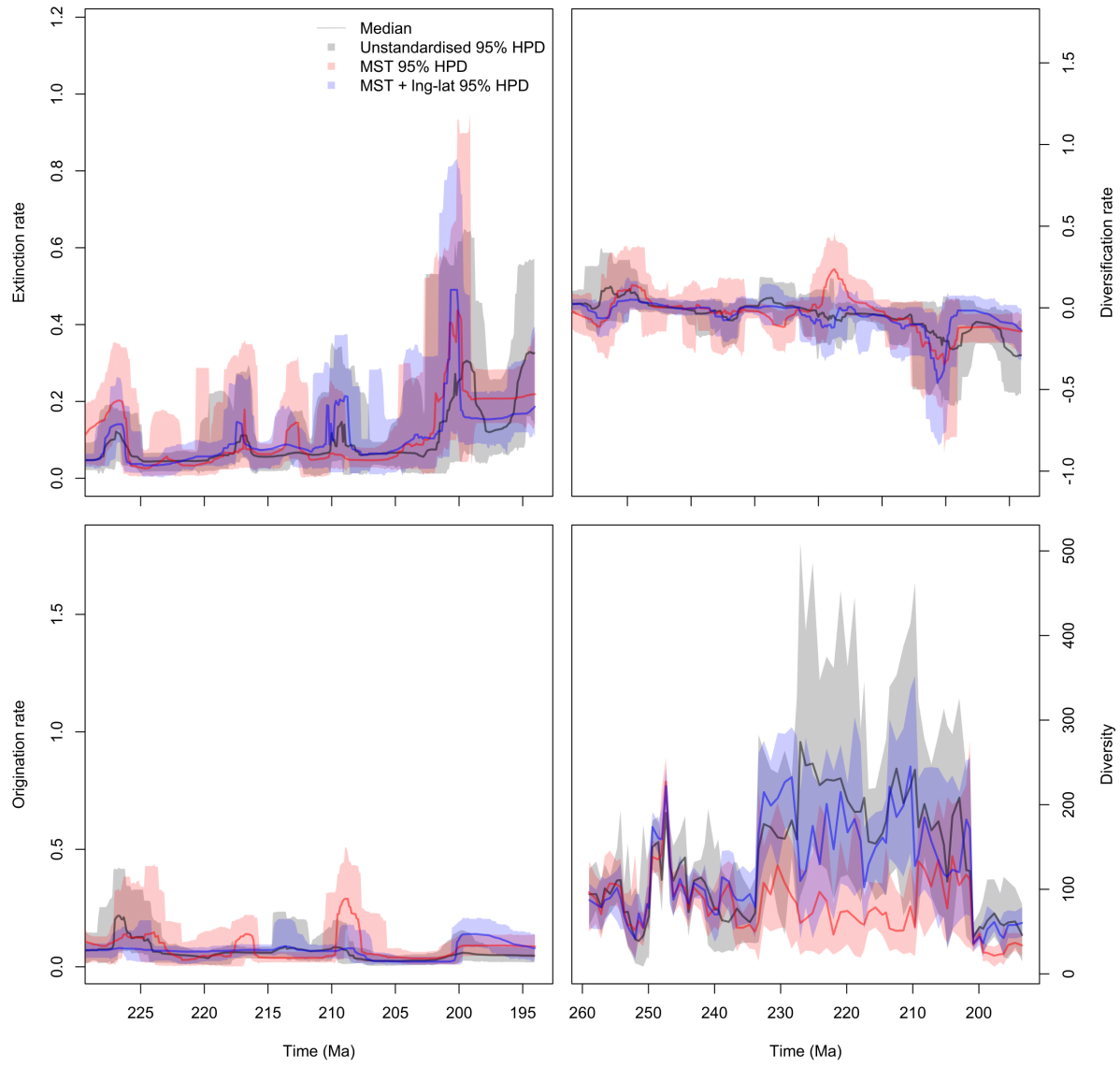


Fig. S47. Impact of spatial standardisation on diversity dynamics (North Panthalassic). Comparison of diversity dynamics from unstandardised, MST-standardised, and MST + lnglat standardised datasets.

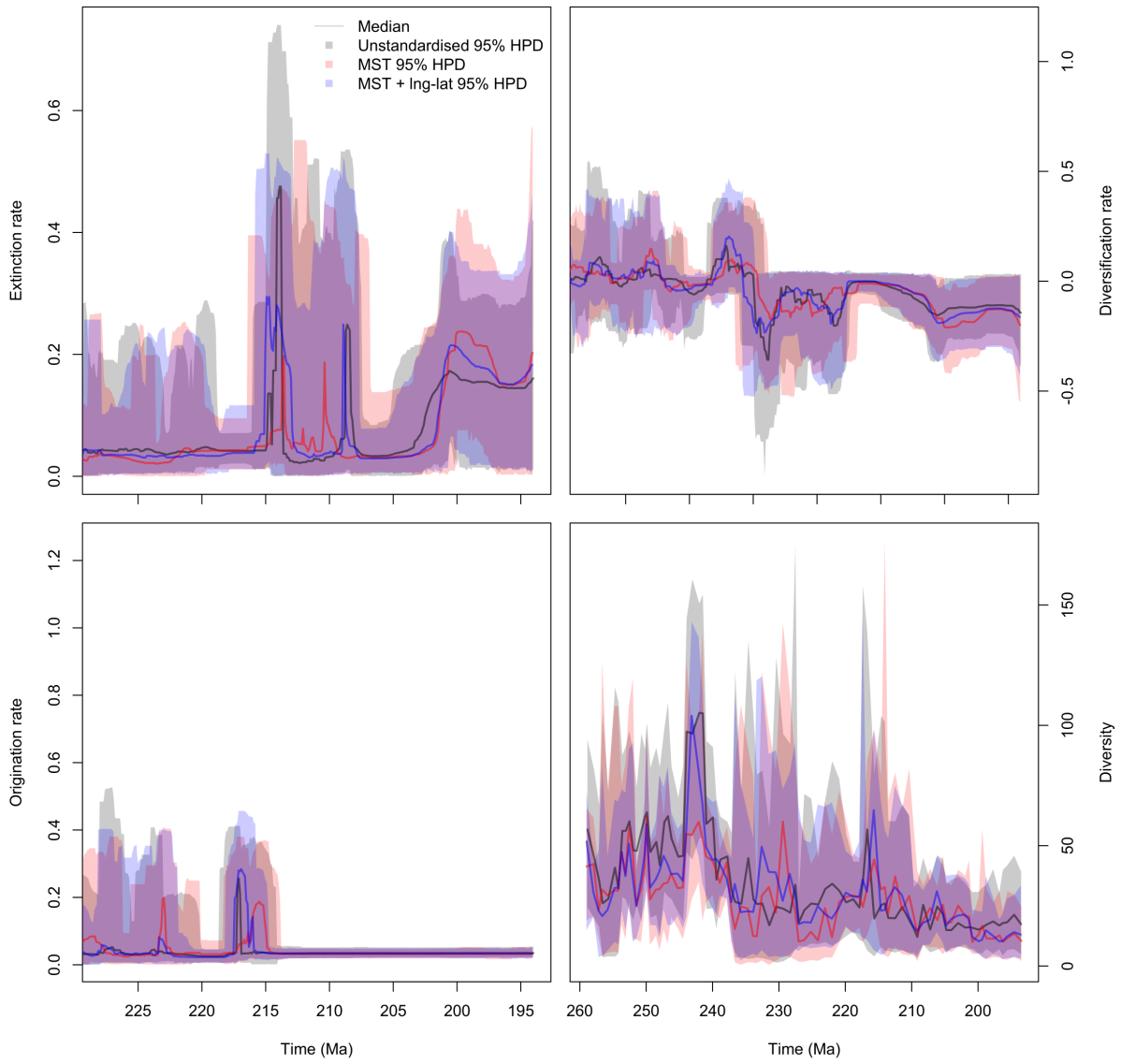


Fig. S48. Impact of spatial standardisation on diversity dynamics (Tangaroan). Comparison of diversity dynamics from unstandardised, MST-standardised, and MST + lnglat standardised datasets.

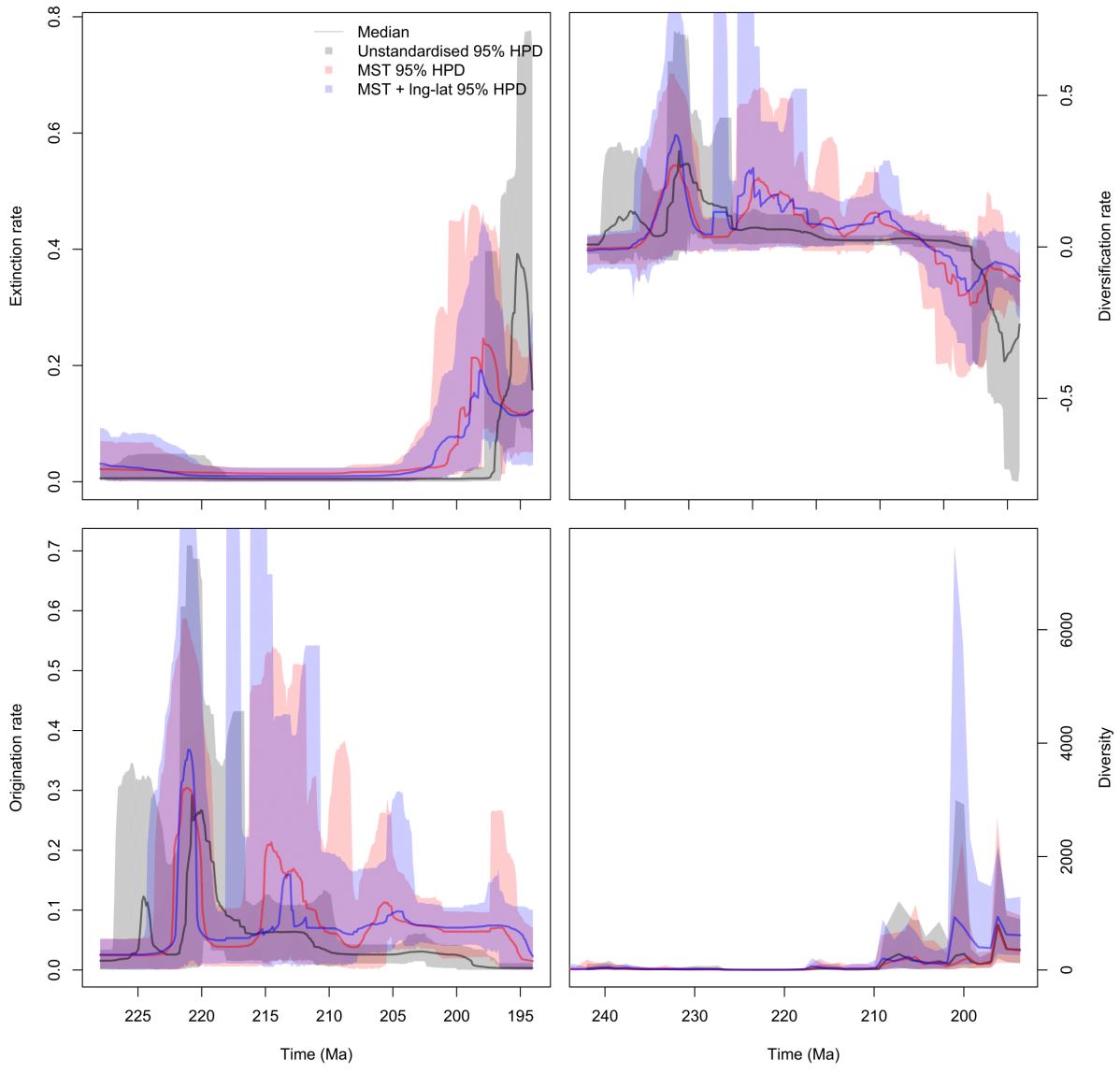


Fig. S49. Impact of spatial standardisation on diversity dynamics (South Panthalassic). Comparison of diversity dynamics from unstandardised, MST-standardised, and MST + lnglat standardised datasets.

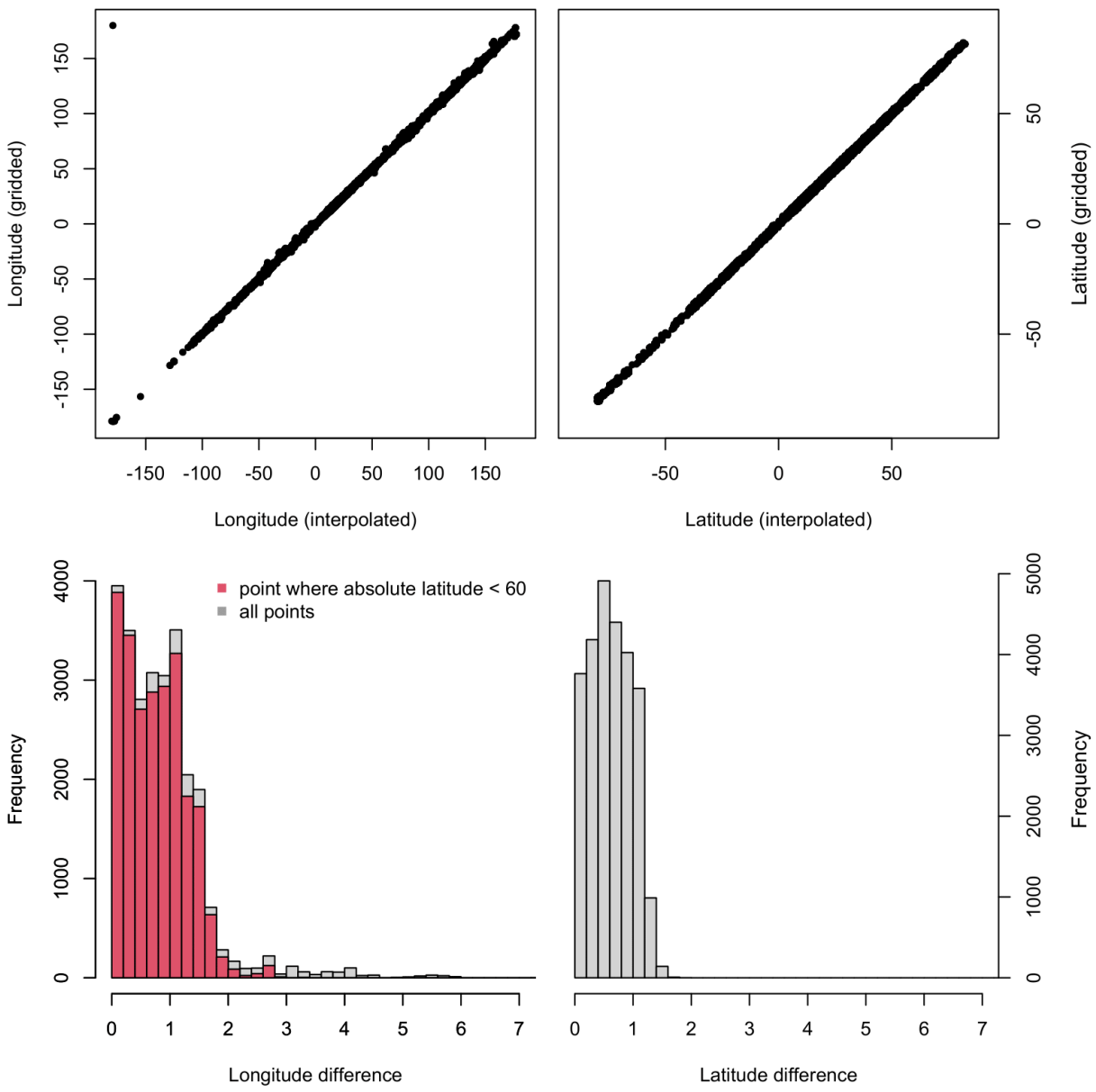


Fig. S50. Effect of rotation method on Getech palaeocoordinates. Comparison of diversity dynamics from unstandardised, MST-standardised, and MST + lnglat standardised datasets.

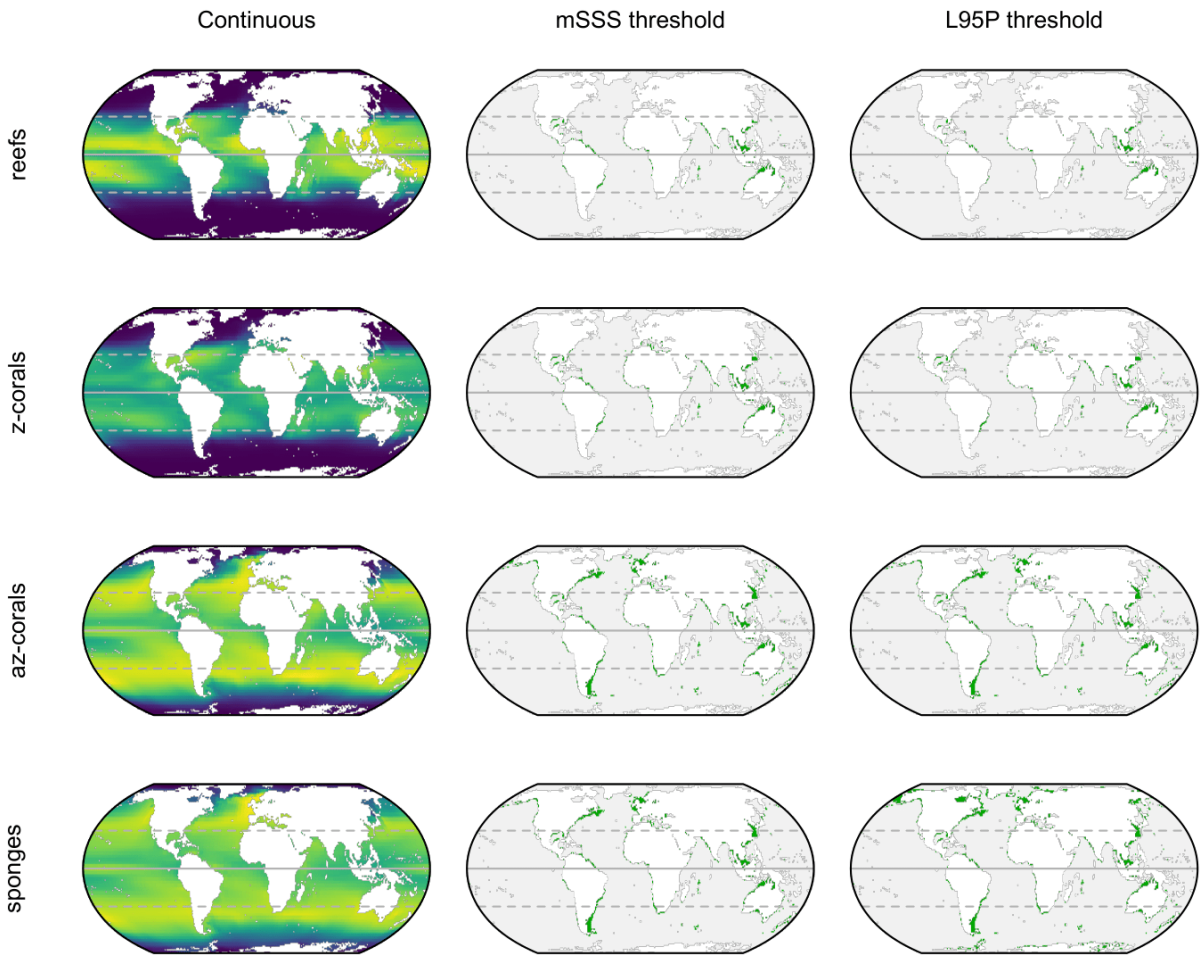


Fig. S51. Niche model projections in the present day. The continuous maps display the probability of habitability (brighter = more probable). The other maps display habitable area after masking to the photic zone and thresholding by the sum of maximum sensitivity and specificity (mSSS) and the threshold that excludes the lowest 5% of occurrence points (L95P). Horizontal lines demarcate the equator and tropics.

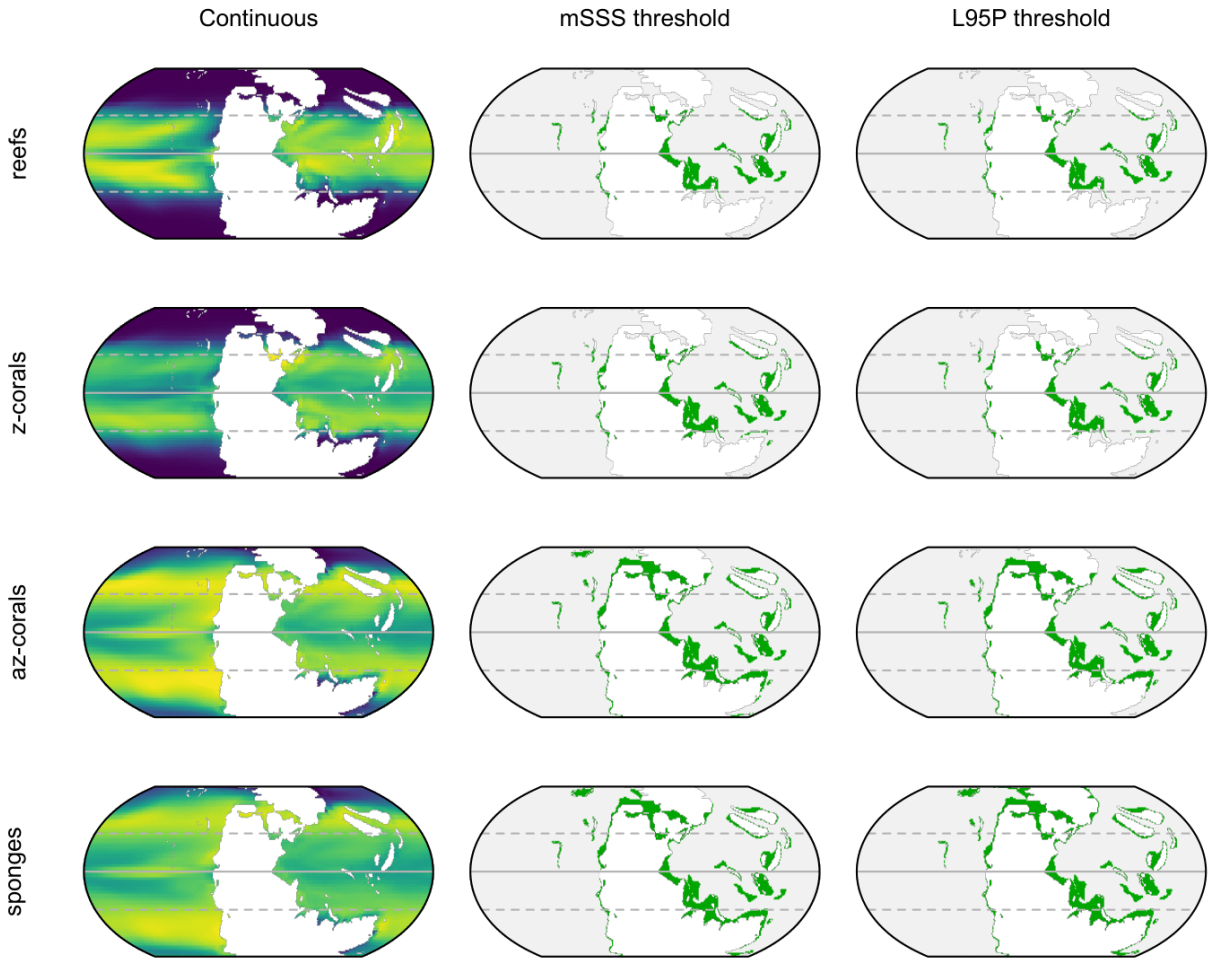


Fig. S52. Niche model projections for the Wuchiapingian (Getech + HadCM3L). The continuous maps display the probability of habitability (brighter = more probable). The other maps display habitable area after masking to the photic zone and thresholding by the sum of maximum sensitivity and specificity (mSSS) and the threshold that excludes the lowest 5% of occurrence points (L95P). Horizontal lines demarcate the equator and tropics.

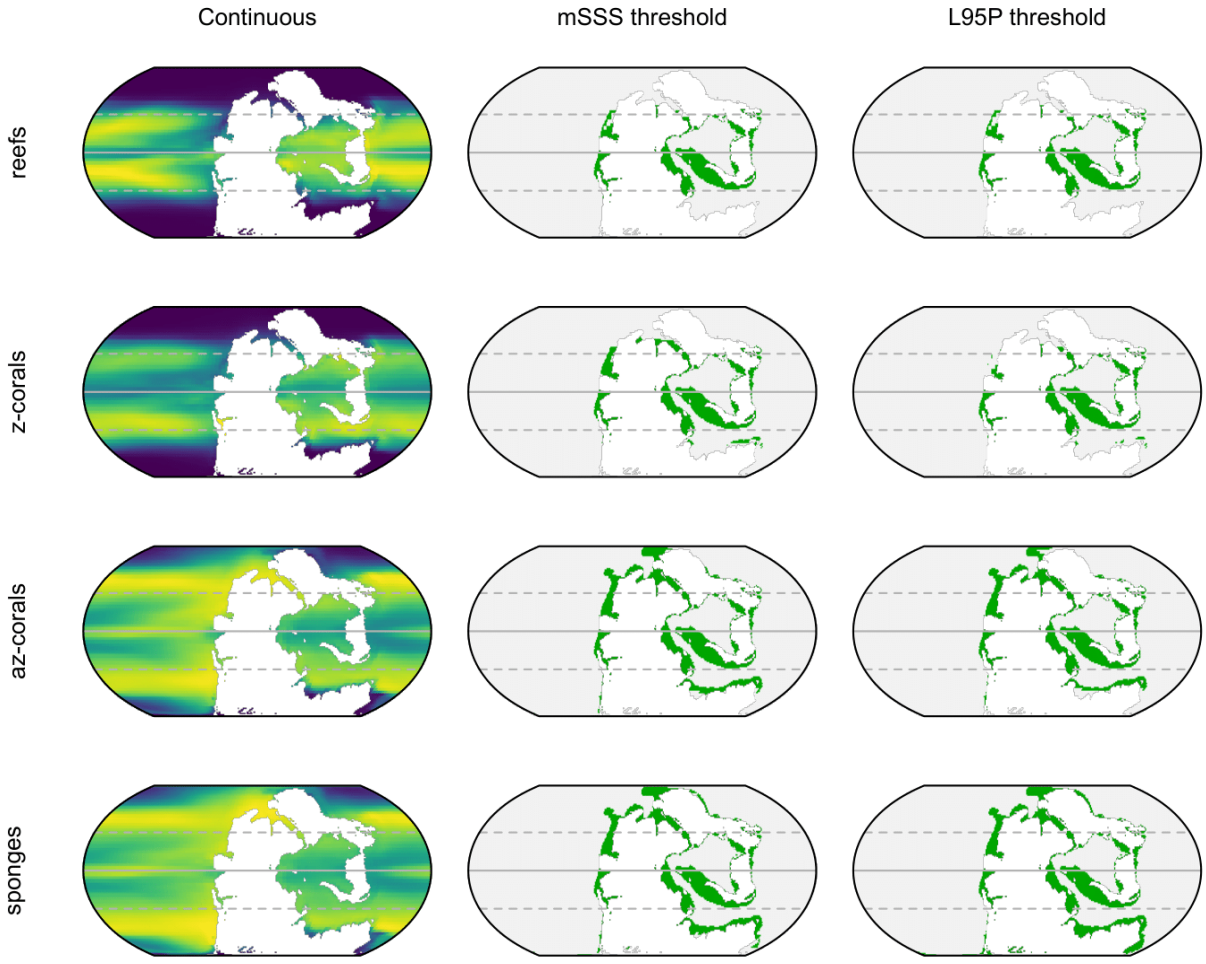


Fig. S53. Niche model projections for the Wuchiapingian (PALEOMAP + HadCM3L). The continuous maps display the probability of habitability (brighter = more probable). The other maps display habitable area after masking to the photic zone and thresholding by the sum of maximum sensitivity and specificity (mSSS) and the threshold that excludes the lowest 5% of occurrence points (L95P). Horizontal lines demarcate the equator and tropics.

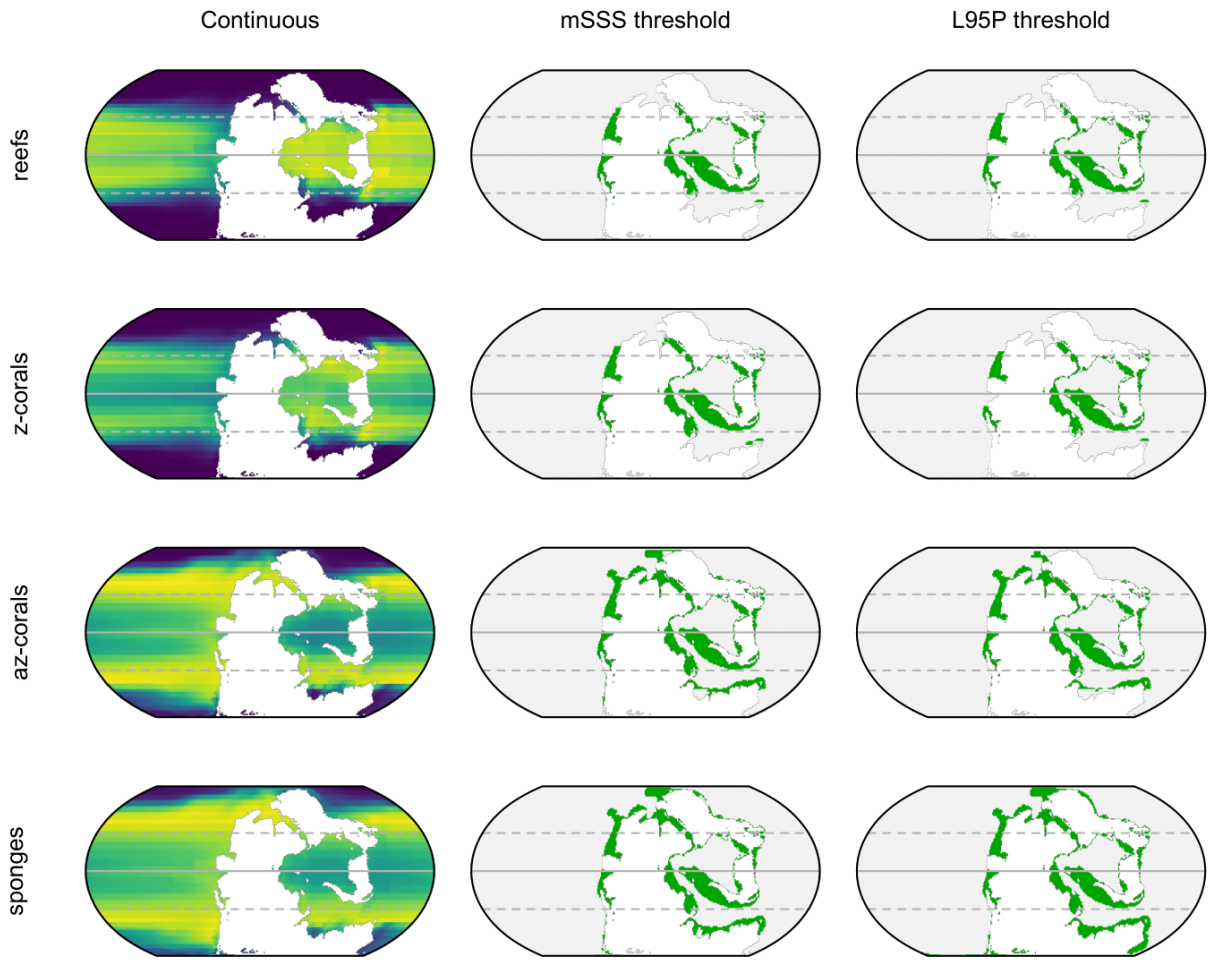


Fig. S54. Niche model projections for the Wuchiapingian (PALEOMAP + CLIMBER-3*α*). The continuous maps display the probability of habitability (brighter = more probable). The other maps display habitable area after masking to the photic zone and thresholding by the sum of maximum sensitivity and specificity (mSSS) and the threshold that excludes the lowest 5% of occurrence points (L95P). Horizontal lines demarcate the equator and tropics.

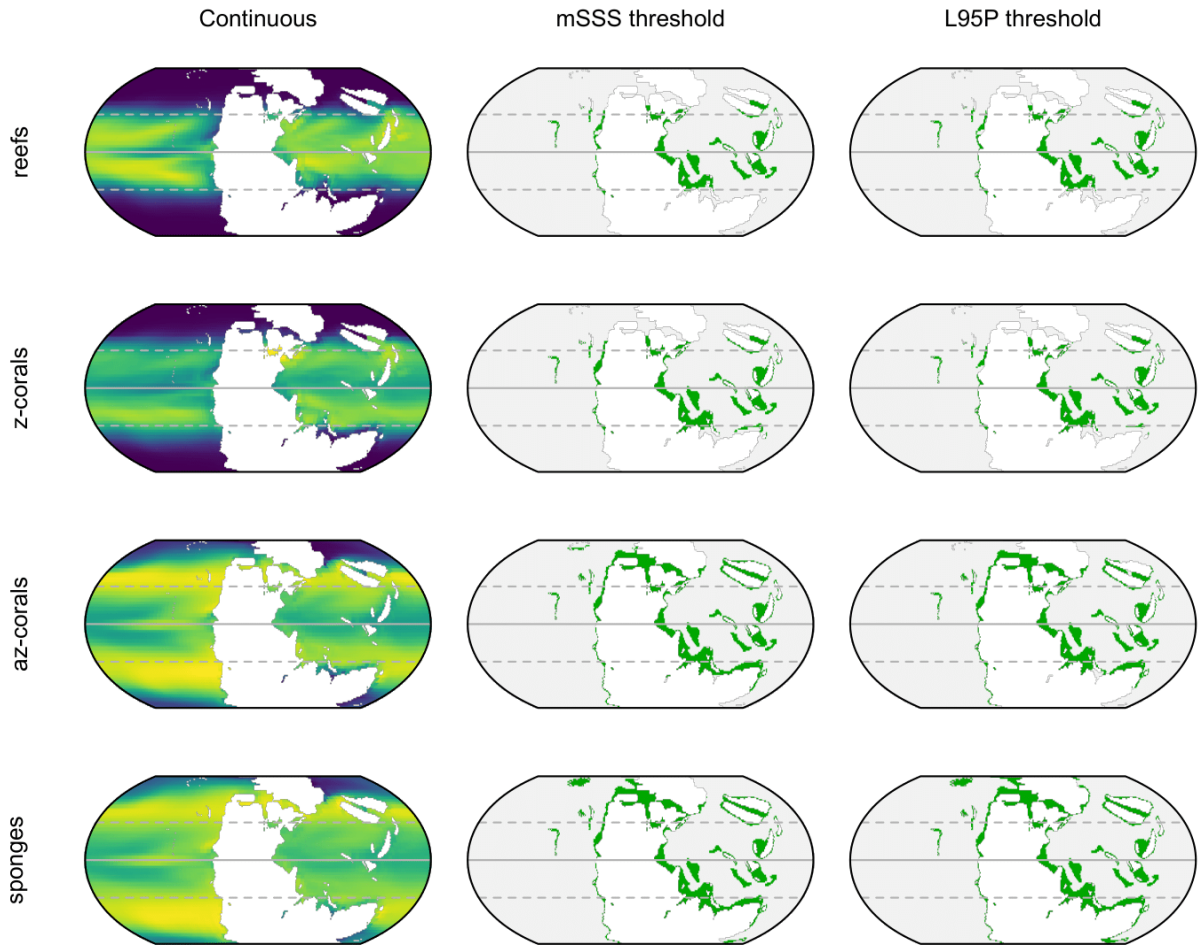


Fig. S55. Niche model projections for the Changhsingian (Getech + HadCM3L). The continuous maps display the probability of habitability (brighter = more probable). The other maps display habitable area after masking to the photic zone and thresholding by the sum of maximum sensitivity and specificity (mSSS) and the threshold that excludes the lowest 5% of occurrence points (L95P). Horizontal lines demarcate the equator and tropics.

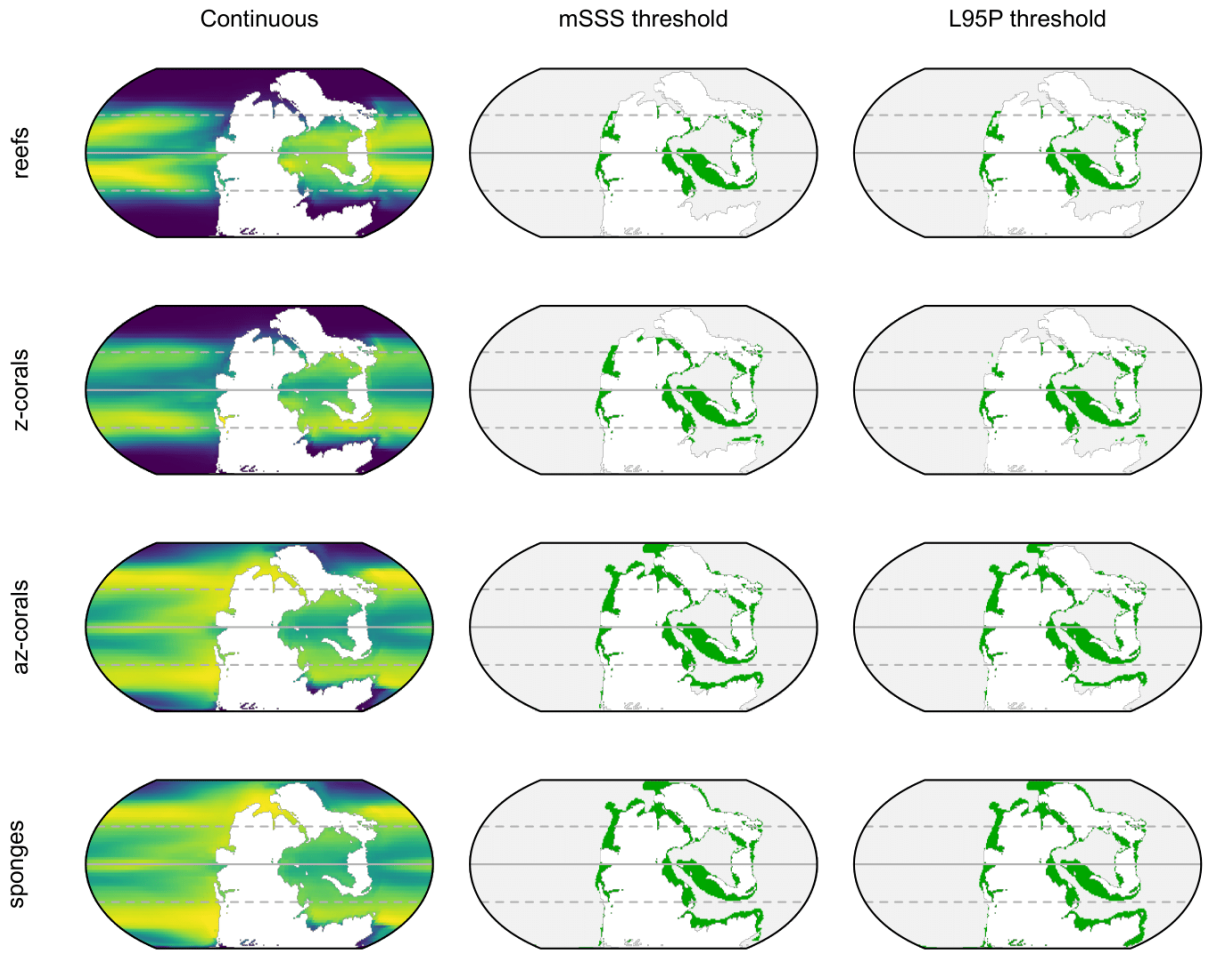


Fig. S56. Niche model projections for the Changhsingian (PALEOMAP + HadCM3L). The continuous maps display the probability of habitability (brighter = more probable). The other maps display habitable area after masking to the photic zone and thresholding by the sum of maximum sensitivity and specificity (mSSS) and the threshold that excludes the lowest 5% of occurrence points (L95P). Horizontal lines demarcate the equator and tropics.

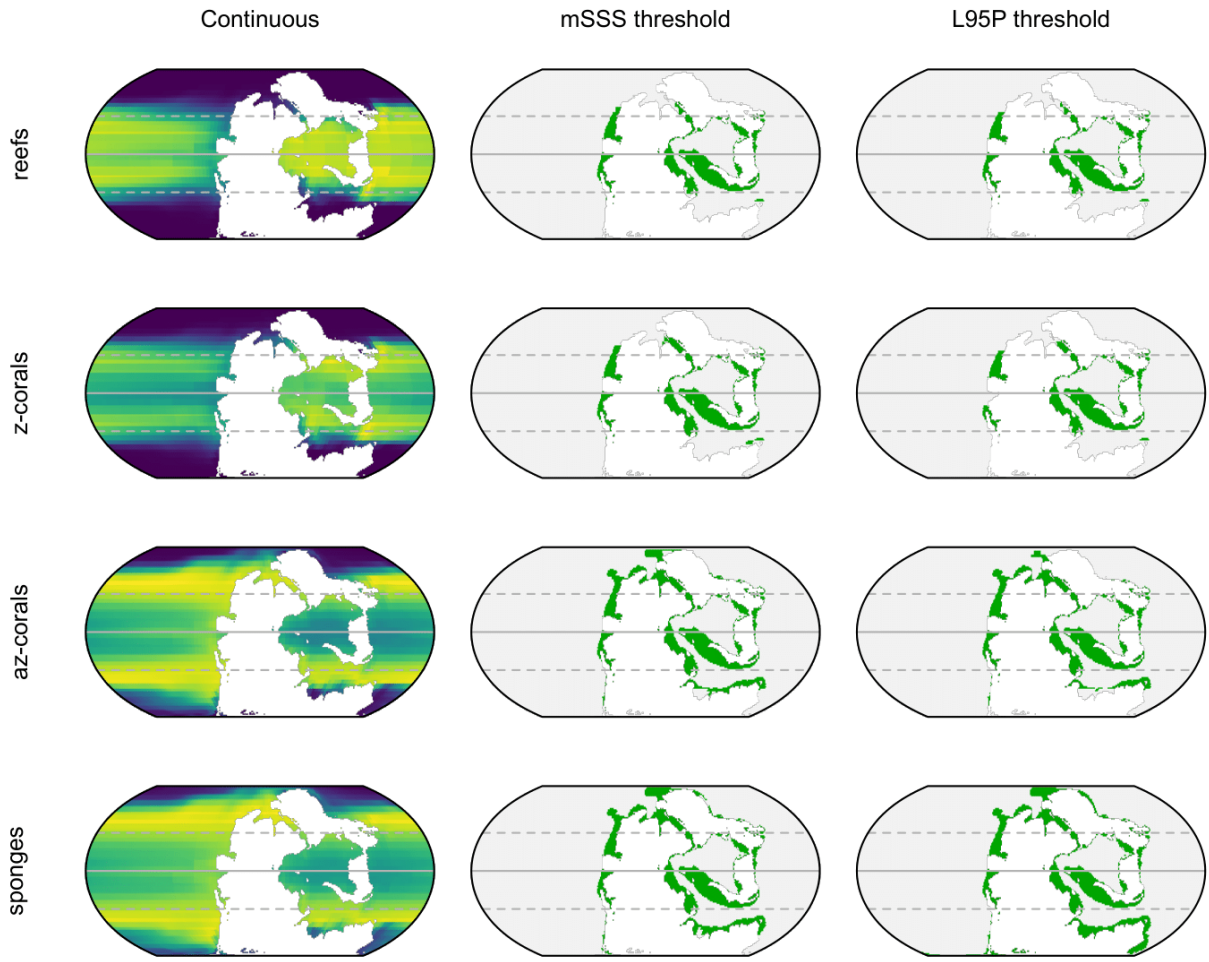


Fig. S57. Niche model projections for the Changhsingian (PALEOMAP + CLIMBER-3α). The continuous maps display the probability of habitability (brighter = more probable). The other maps display habitable area after masking to the photic zone and thresholding by the sum of maximum sensitivity and specificity (mSSS) and the threshold that excludes the lowest 5% of occurrence points (L95P). Horizontal lines demarcate the equator and tropics.

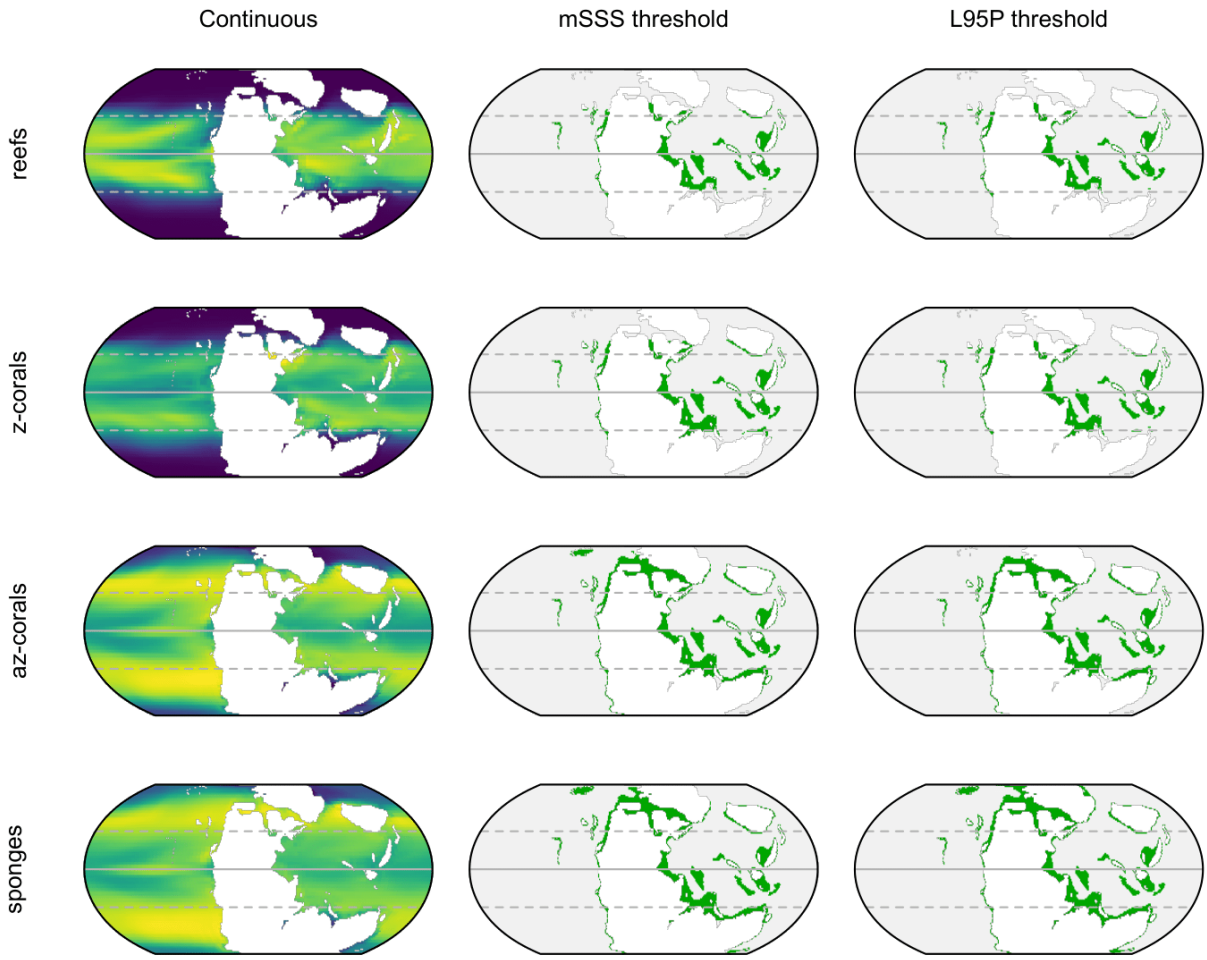


Fig. S58. Niche model projections for the Induan (Getech + HadCM3L). The continuous maps display the probability of habitability (brighter = more probable). The other maps display habitable area after masking to the photic zone and thresholding by the sum of maximum sensitivity and specificity (mSSS) and the threshold that excludes the lowest 5% of occurrence points (L95P). Horizontal lines demarcate the equator and tropics.

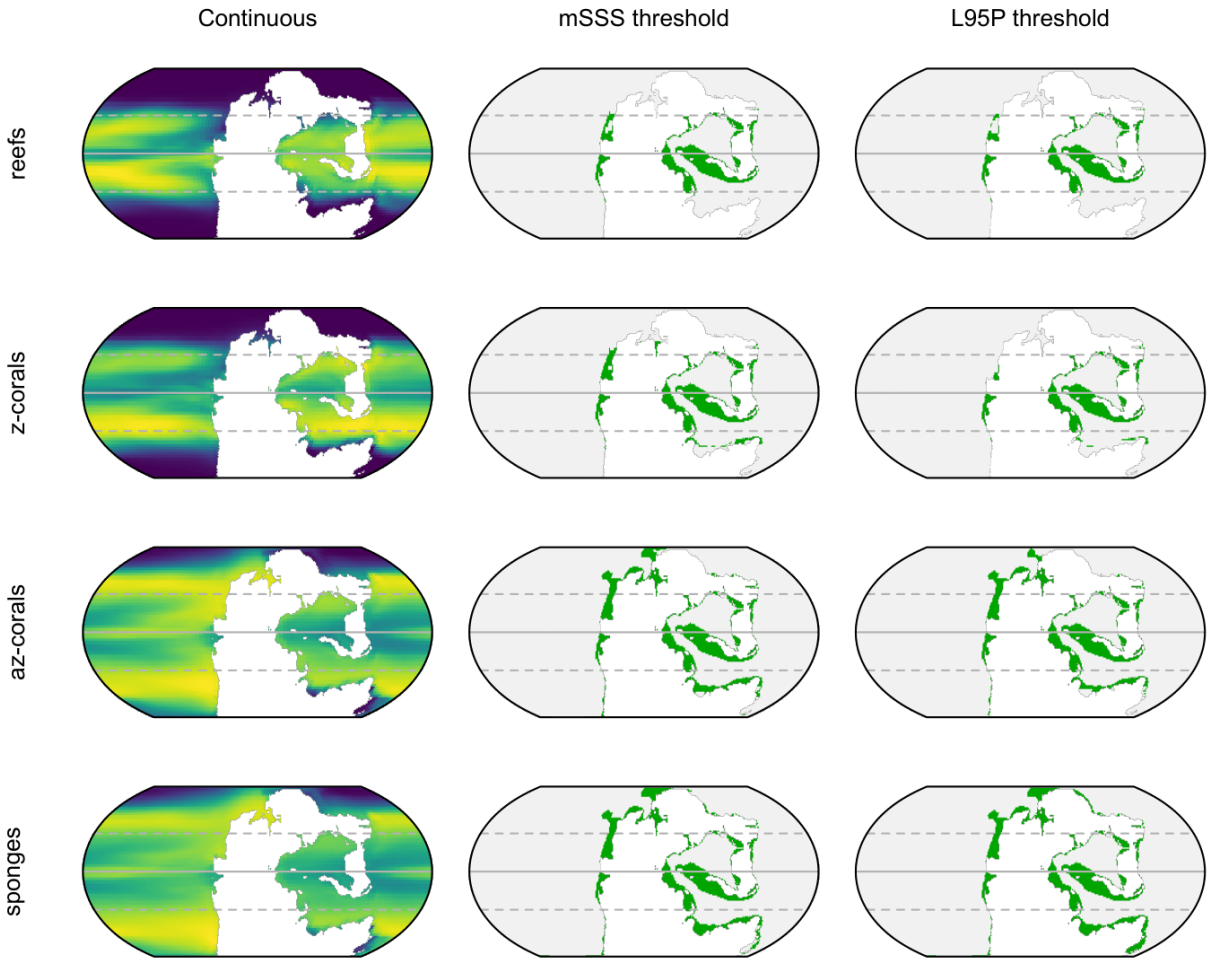


Fig. S59. Niche model projections for the Induan (PALEOMAP + HadCM3L). The continuous maps display the probability of habitability (brighter = more probable). The other maps display habitable area after masking to the photic zone and thresholding by the sum of maximum sensitivity and specificity (mSSS) and the threshold that excludes the lowest 5% of occurrence points (L95P). Horizontal lines demarcate the equator and tropics.

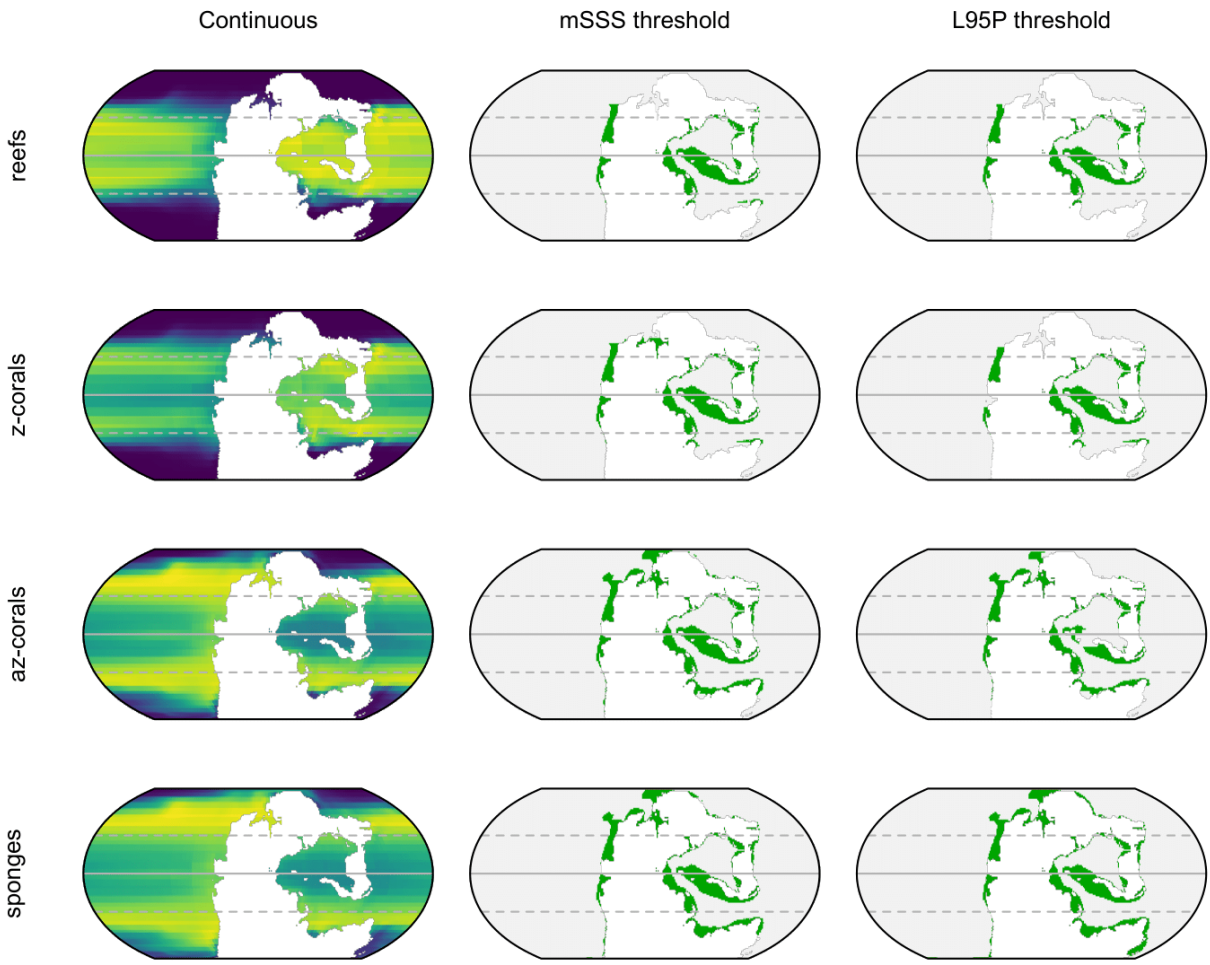


Fig. S60. Niche model projections for the Induan (PALEOMAP + CLIMBER-3*α*). The continuous maps display the probability of habitability (brighter = more probable). The other maps display habitable area after masking to the photic zone and thresholding by the sum of maximum sensitivity and specificity (mSSS) and the threshold that excludes the lowest 5% of occurrence points (L95P). Horizontal lines demarcate the equator and tropics.

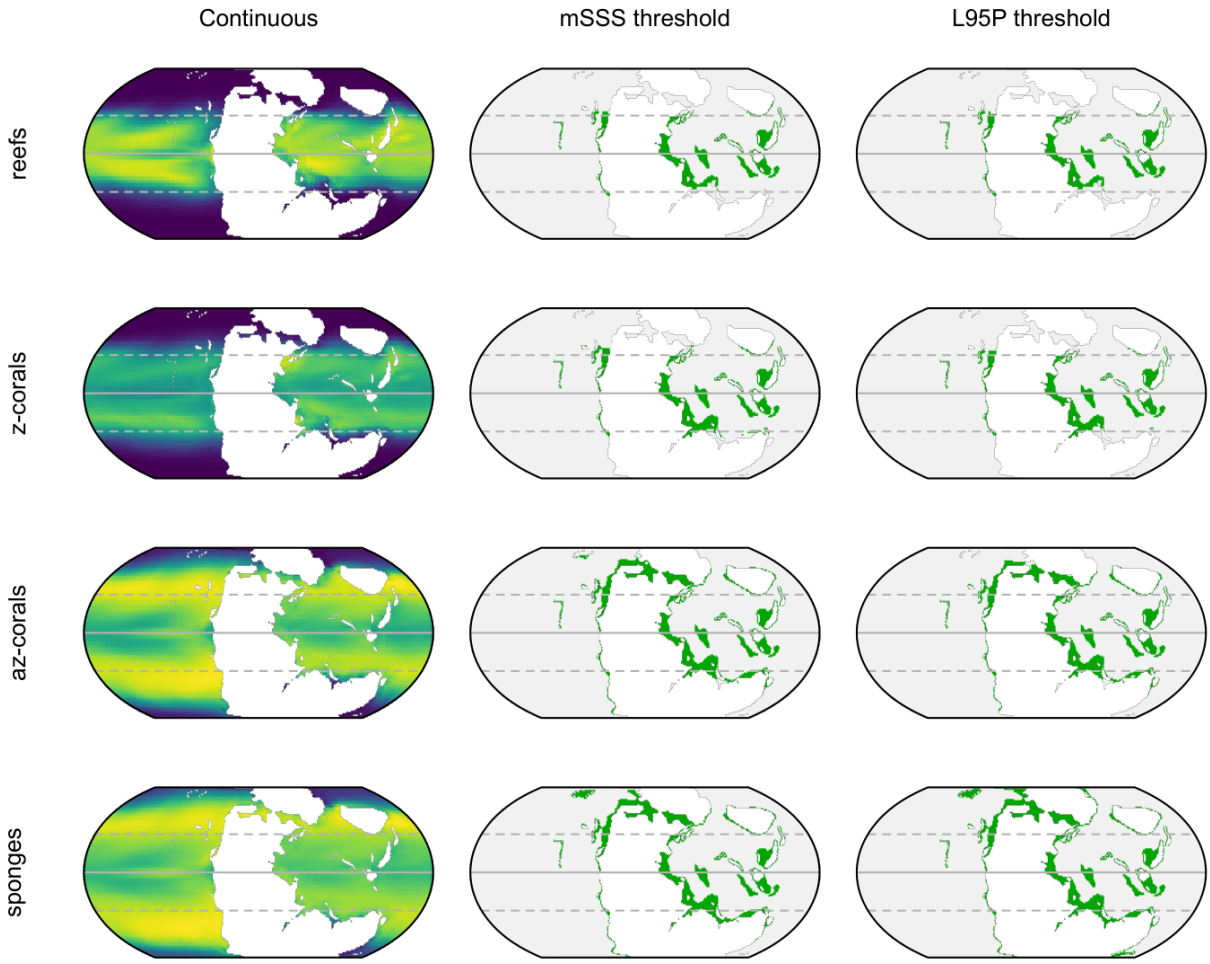


Fig. S61. Niche model projections for the Olenekian (Getech + HadCM3L). The continuous maps display the probability of habitability (brighter = more probable). The other maps display habitable area after masking to the photic zone and thresholding by the sum of maximum sensitivity and specificity (mSSS) and the threshold that excludes the lowest 5% of occurrence points (L95P). Horizontal lines demarcate the equator and tropics.

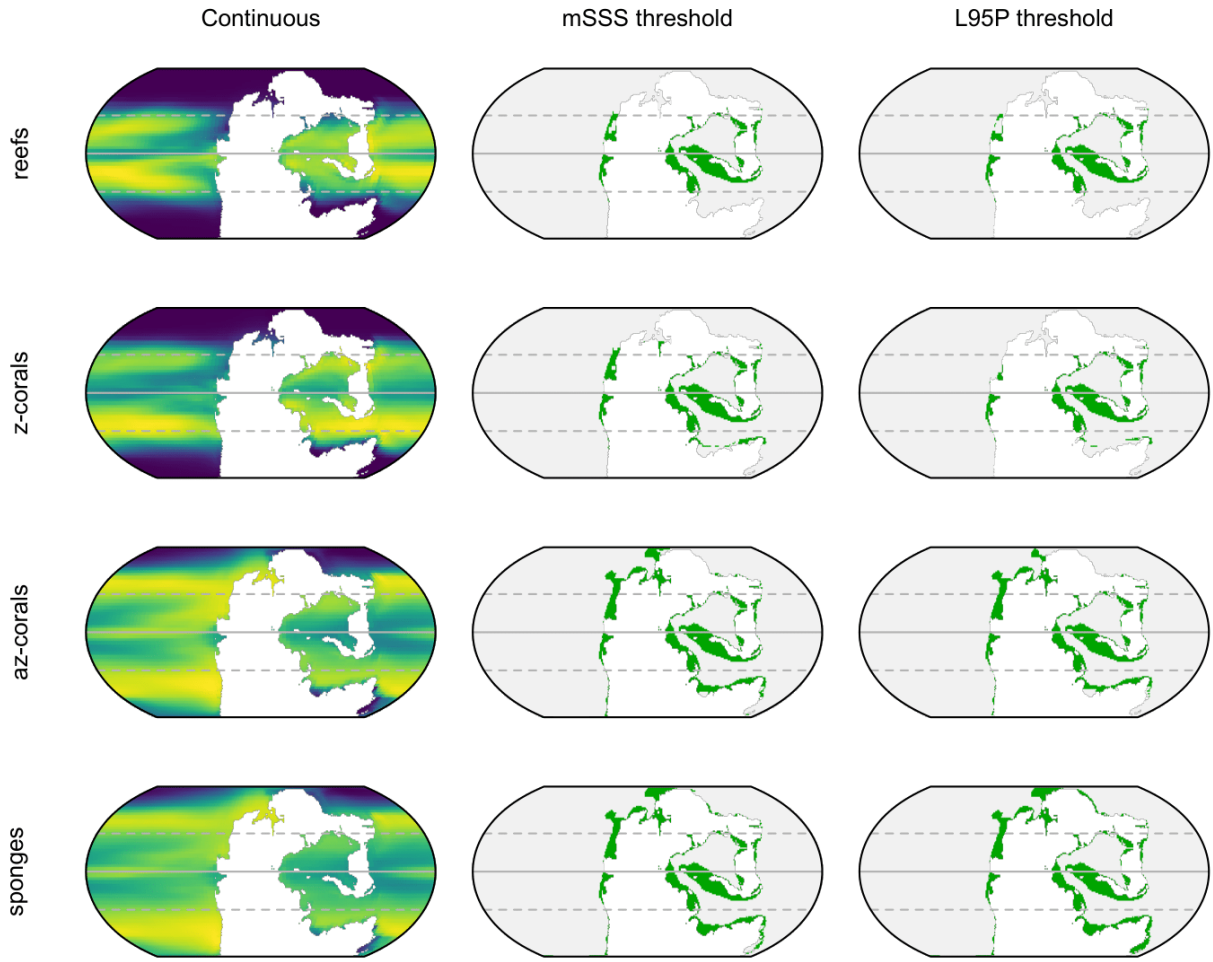


Fig. S62. Niche model projections for the Olenekian (PALEOMAP + HadCM3L). The continuous maps display the probability of habitability (brighter = more probable). The other maps display habitable area after masking to the photic zone and thresholding by the sum of maximum sensitivity and specificity (mSSS) and the threshold that excludes the lowest 5% of occurrence points (L95P). Horizontal lines demarcate the equator and tropics.

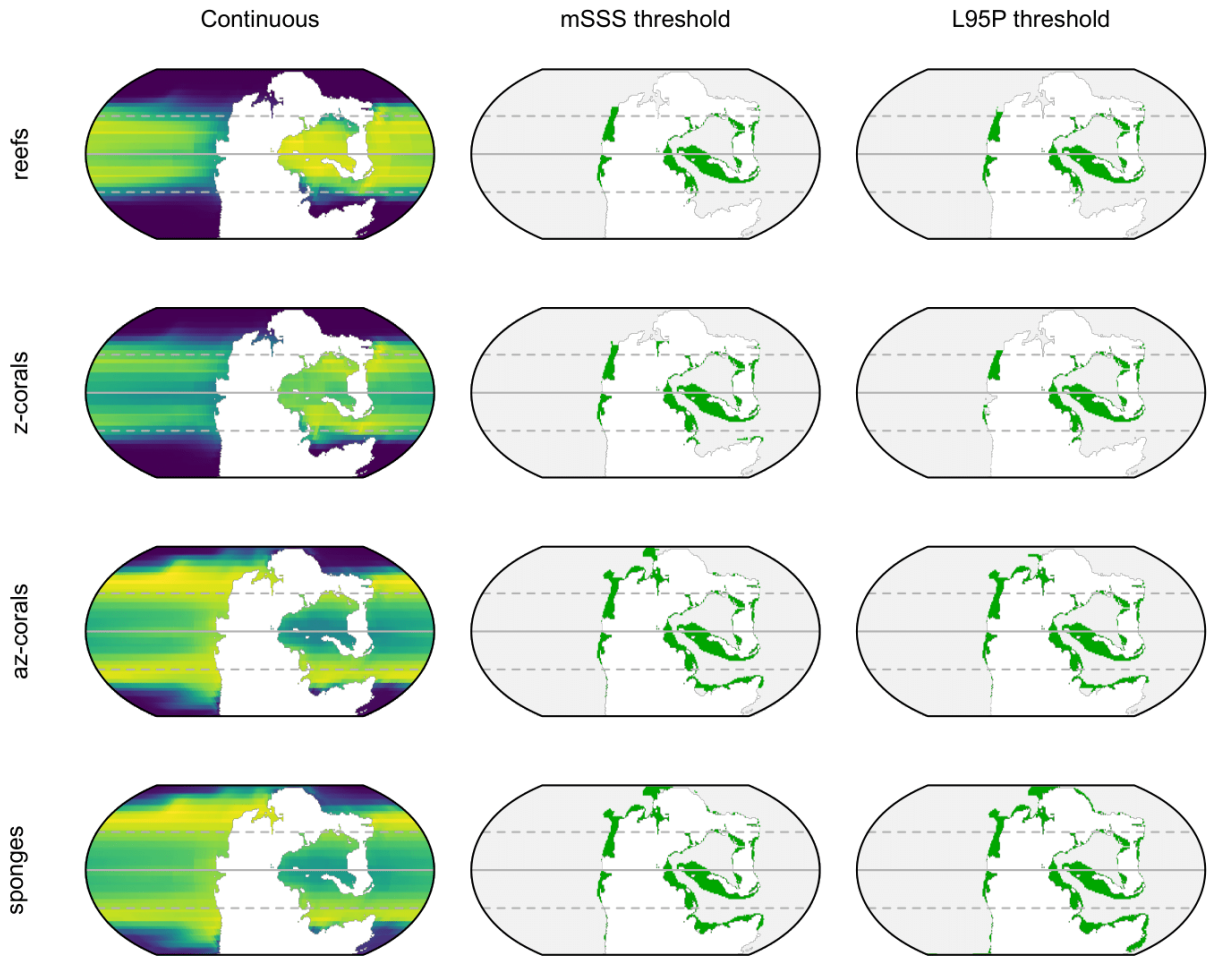


Fig. S63. Niche model projections for the Olenekian (PALEOMAP + CLIMBER-3*α*). The continuous maps display the probability of habitability (brighter = more probable). The other maps display habitable area after masking to the photic zone and thresholding by the sum of maximum sensitivity and specificity (mSSS) and the threshold that excludes the lowest 5% of occurrence points (L95P). Horizontal lines demarcate the equator and tropics.

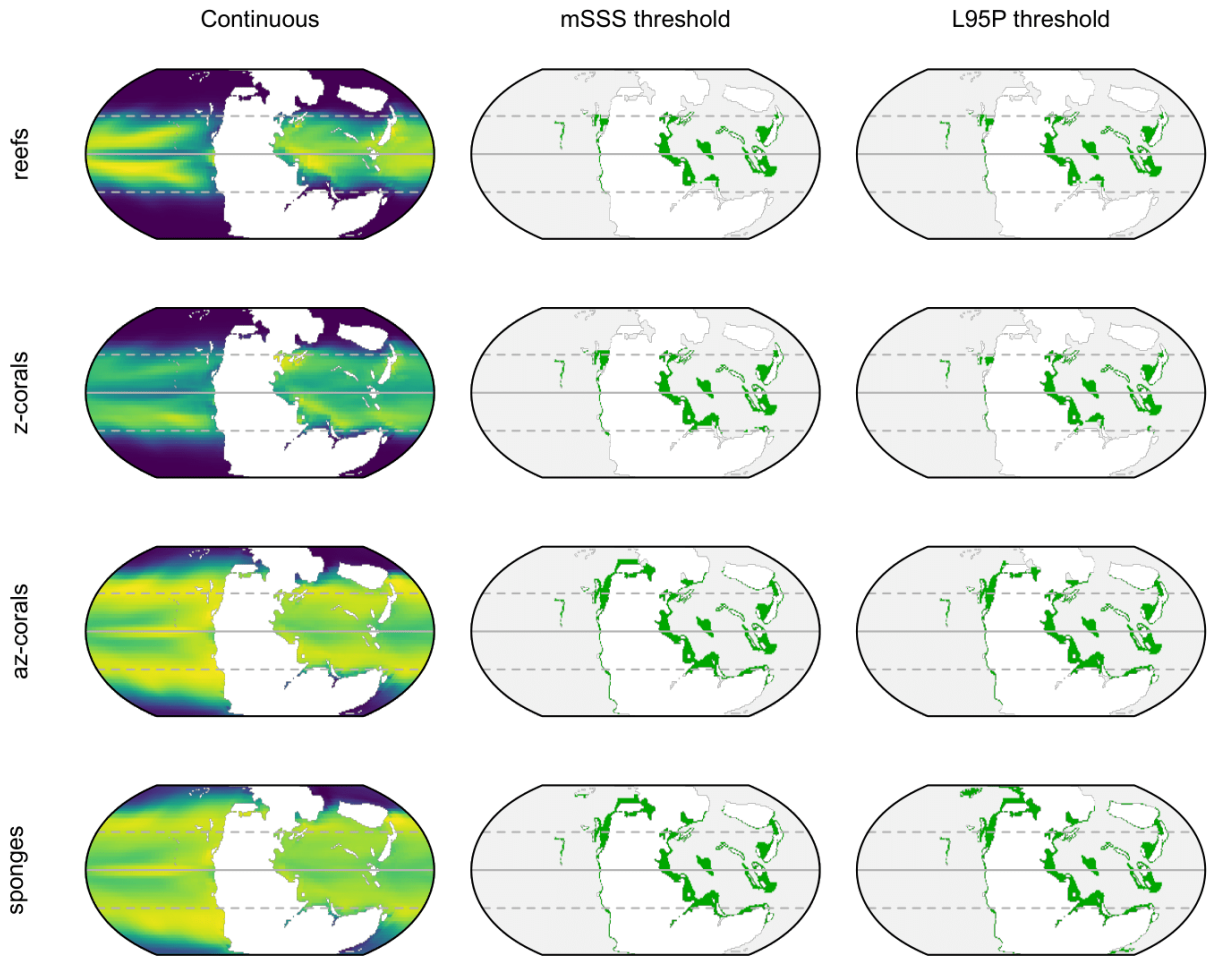


Fig. S64. Niche model projections for the Anisian (Getech + HadCM3L). The continuous maps display the probability of habitability (brighter = more probable). The other maps display habitable area after masking to the photic zone and thresholding by the sum of maximum sensitivity and specificity (mSSS) and the threshold that excludes the lowest 5% of occurrence points (L95P). Horizontal lines demarcate the equator and tropics.

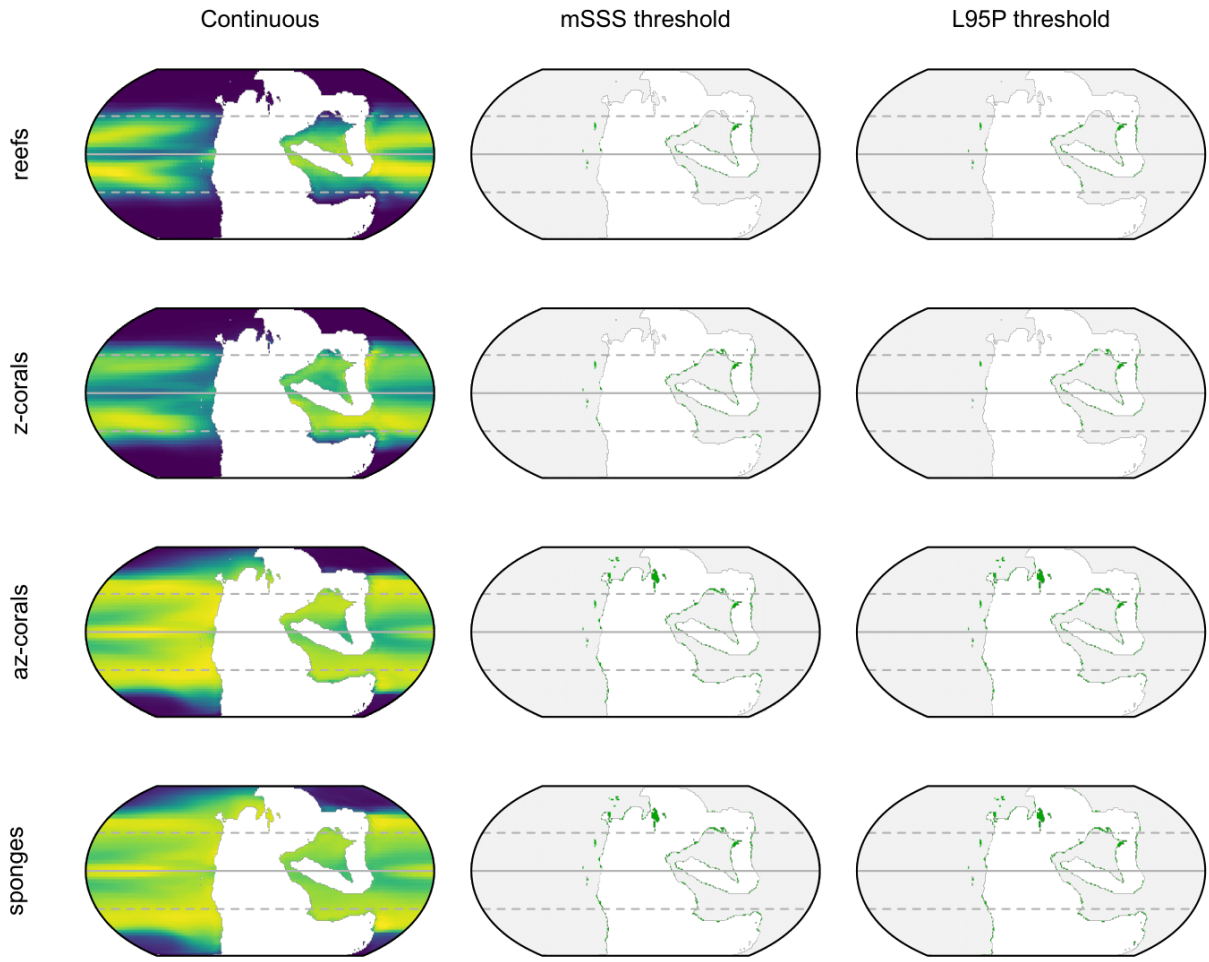


Fig. S65. Niche model projections for the Anisian (PALEOMAP + HadCM3L). The continuous maps display the probability of habitability (brighter = more probable). The other maps display habitable area after masking to the photic zone and thresholding by the sum of maximum sensitivity and specificity (mSSS) and the threshold that excludes the lowest 5% of occurrence points (L95P). Horizontal lines demarcate the equator and tropics.

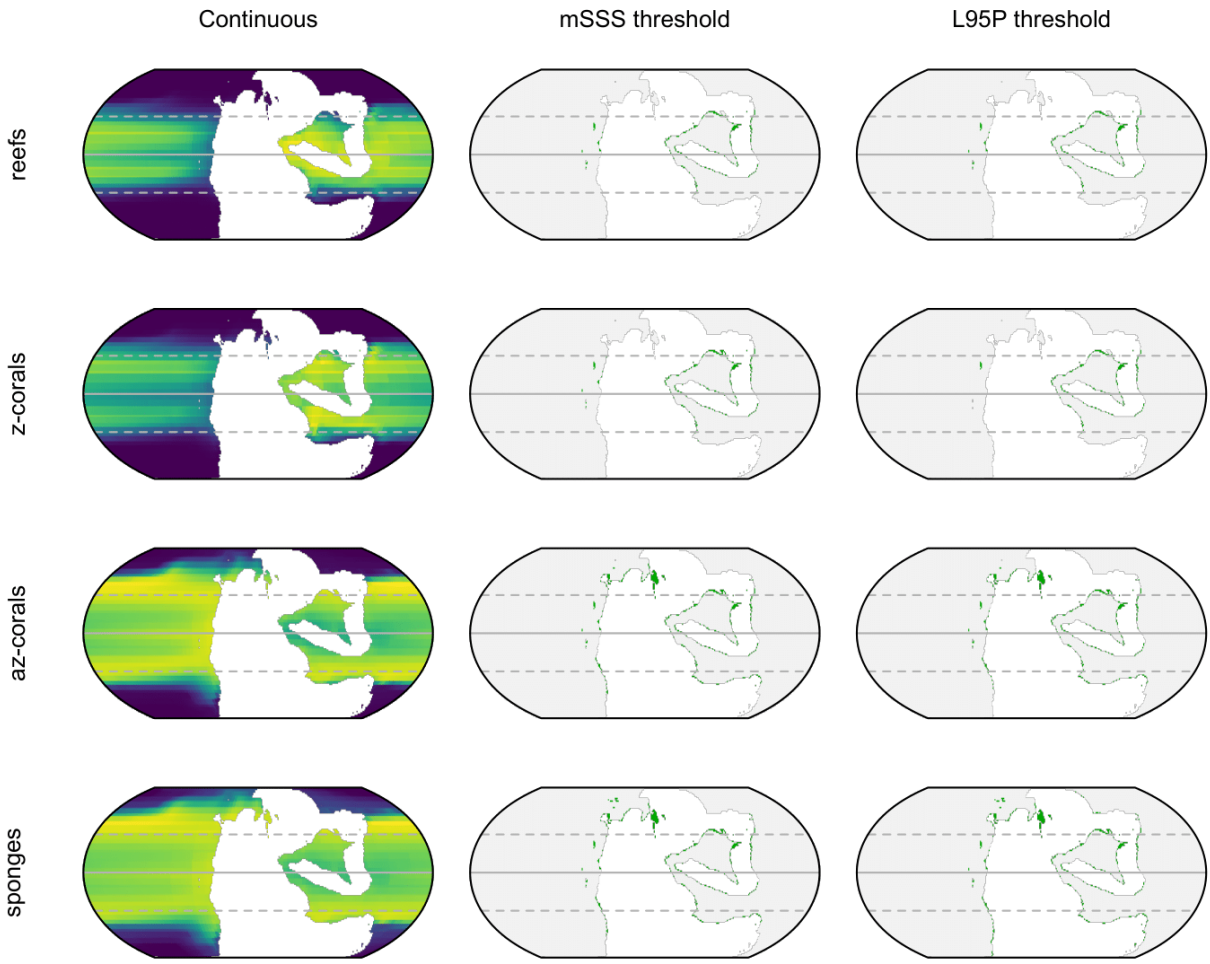


Fig. S66. Niche model projections for the Anisian (PALEOMAP + CLIMBER-3α). The continuous maps display the probability of habitability (brighter = more probable). The other maps display habitable area after masking to the photic zone and thresholding by the sum of maximum sensitivity and specificity (mSSS) and the threshold that excludes the lowest 5% of occurrence points (L95P). Horizontal lines demarcate the equator and tropics.

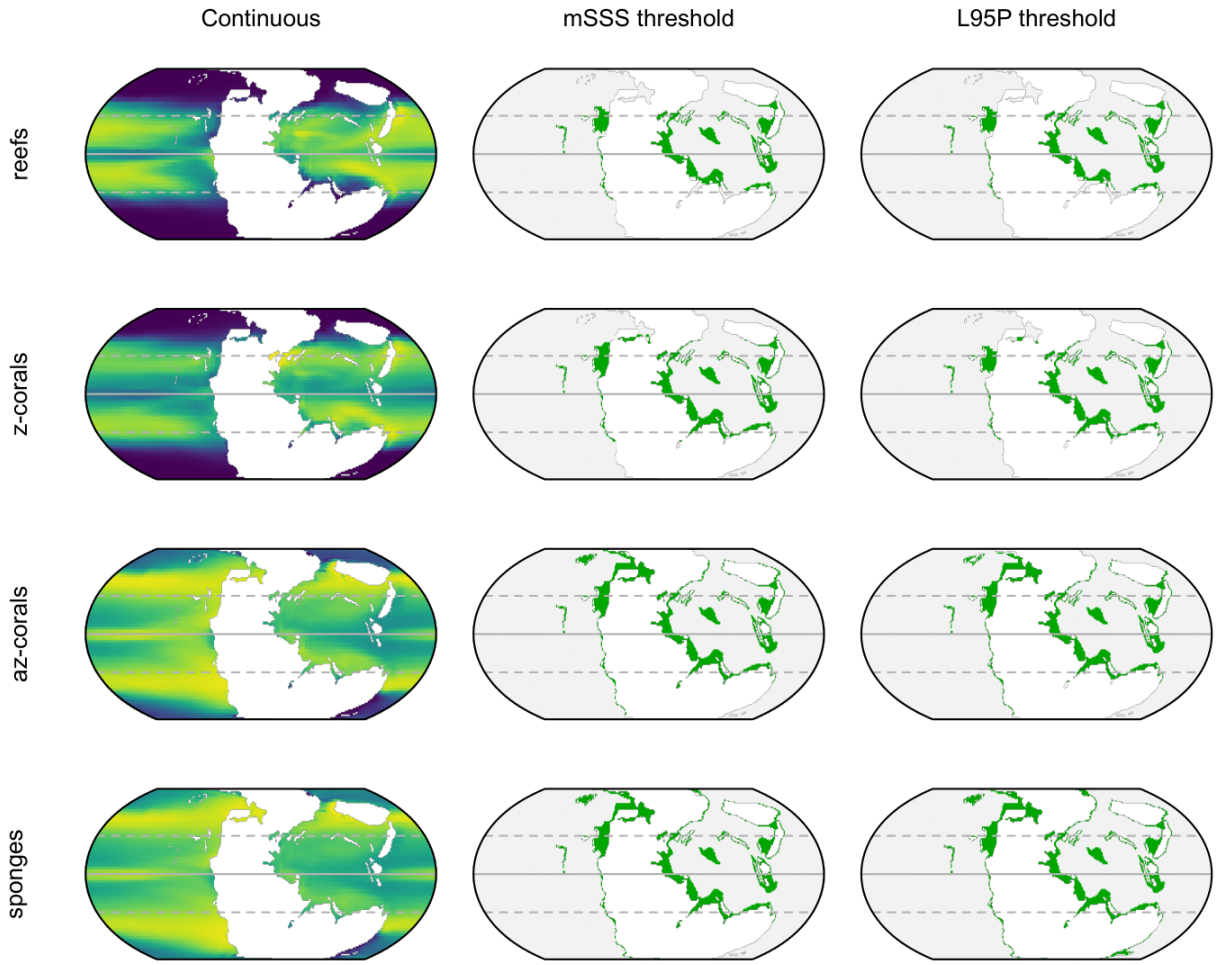


Fig. S67. Niche model projections for the Ladinian (Getech + HadCM3L). The continuous maps display the probability of habitability (brighter = more probable). The other maps display habitable area after masking to the photic zone and thresholding by the sum of maximum sensitivity and specificity (mSSS) and the threshold that excludes the lowest 5% of occurrence points (L95P). Horizontal lines demarcate the equator and tropics.

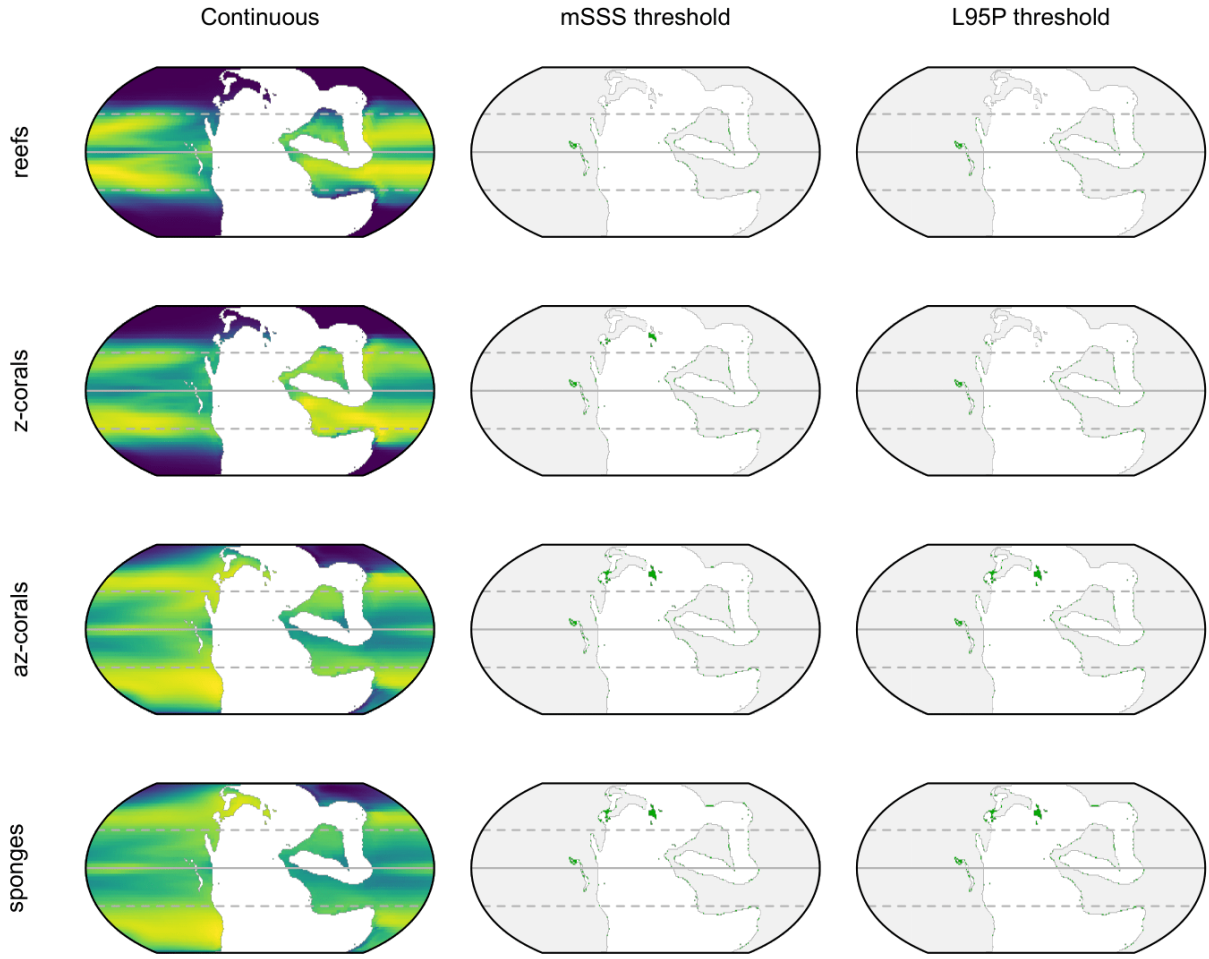


Fig. S68. Niche model projections for the Ladinian (PALEOMAP + HadCM3L). The continuous maps display the probability of habitability (brighter = more probable). The other maps display habitable area after masking to the photic zone and thresholding by the sum of maximum sensitivity and specificity (mSSS) and the threshold that excludes the lowest 5% of occurrence points (L95P). Horizontal lines demarcate the equator and tropics.

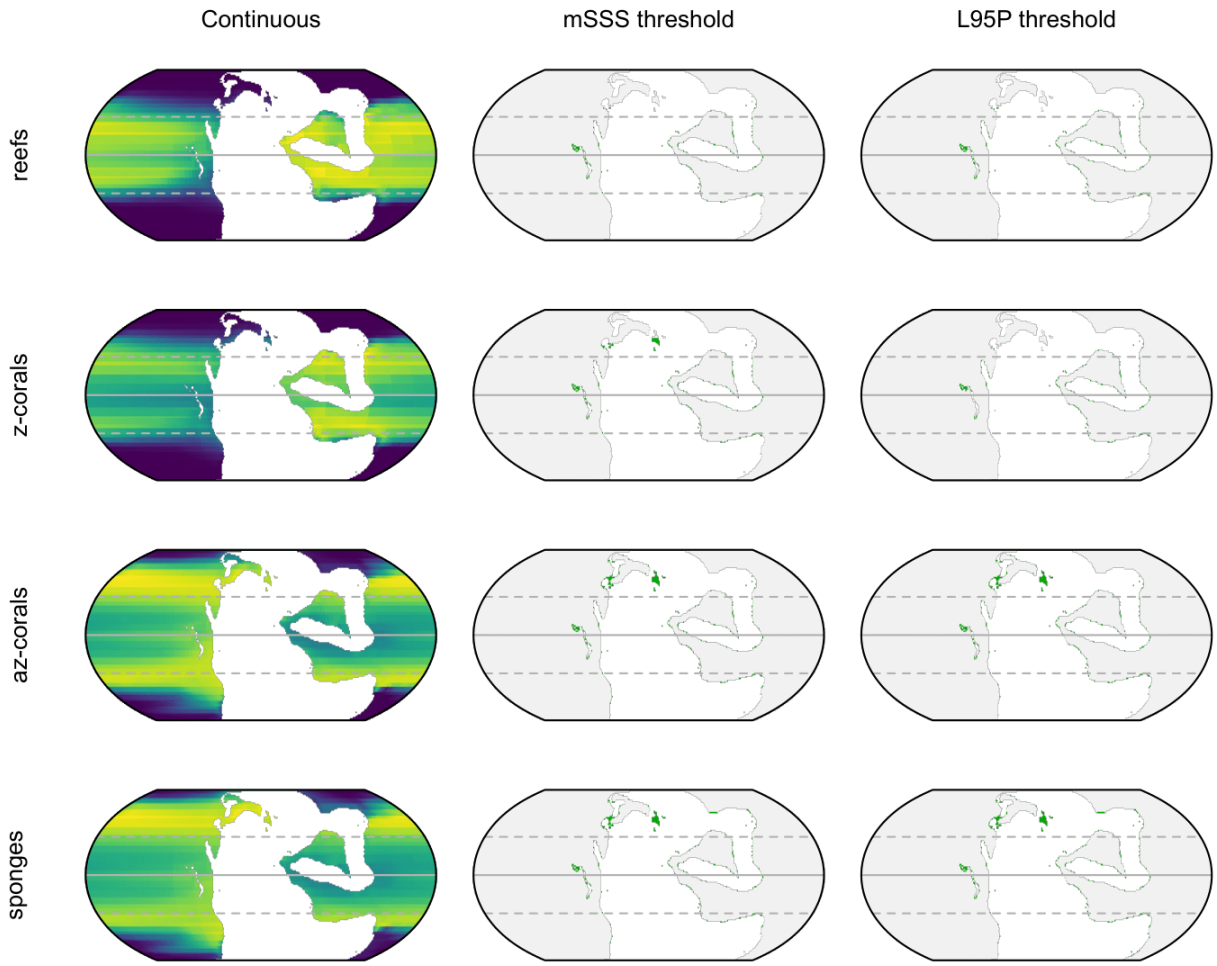


Fig. S69. Niche model projections for the Ladinian (PALEOMAP + CLIMBER-3α). The continuous maps display the probability of habitability (brighter = more probable). The other maps display habitable area after masking to the photic zone and thresholding by the sum of maximum sensitivity and specificity (mSSS) and the threshold that excludes the lowest 5% of occurrence points (L95P). Horizontal lines demarcate the equator and tropics.

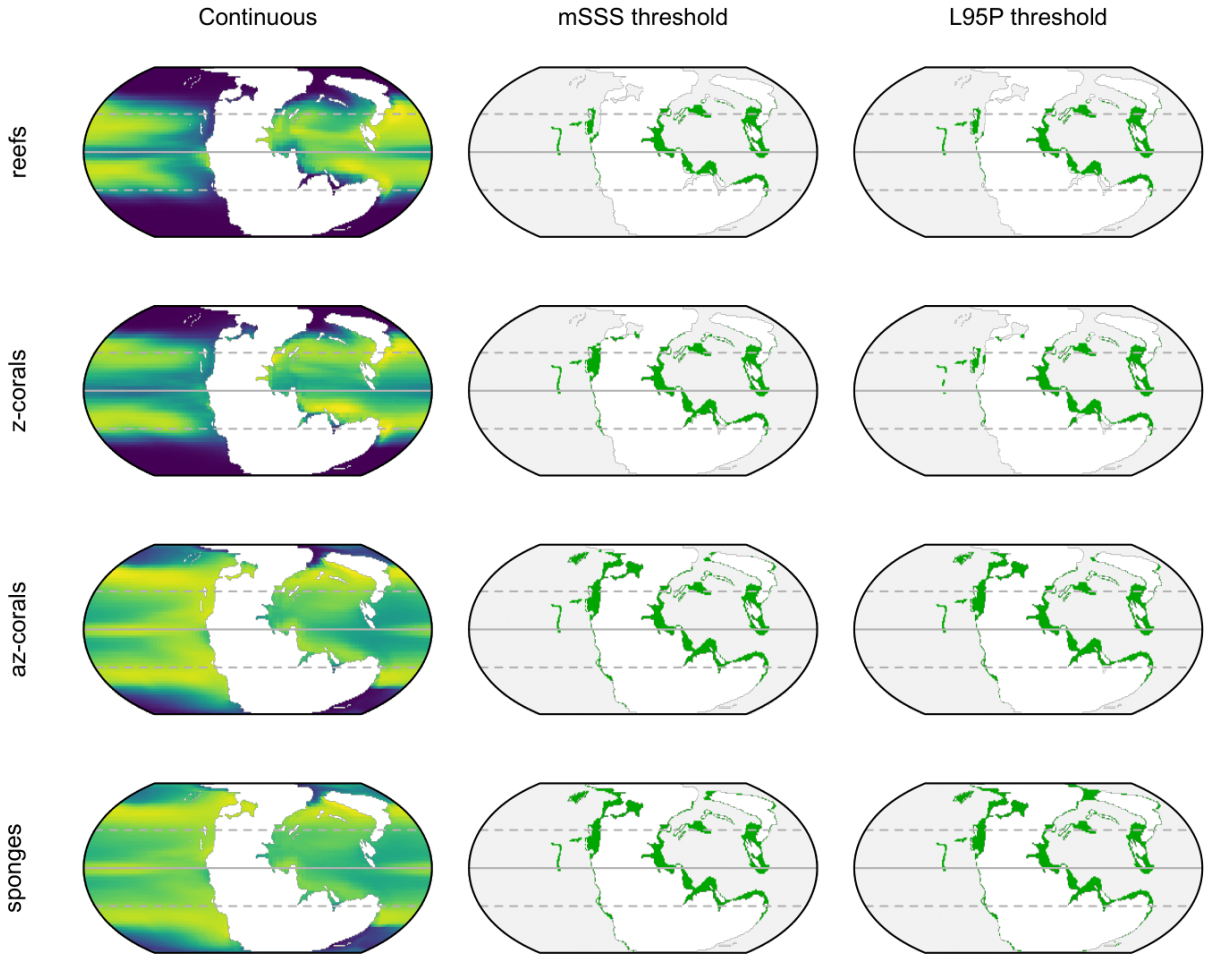


Fig. S70. Niche model projections for the early Carnian (Getech + HadCM3L). The continuous maps display the probability of habitability (brighter = more probable). The other maps display habitable area after masking to the photic zone and thresholding by the sum of maximum sensitivity and specificity (mSSS) and the threshold that excludes the lowest 5% of occurrence points (L95P). Horizontal lines demarcate the equator and tropics.

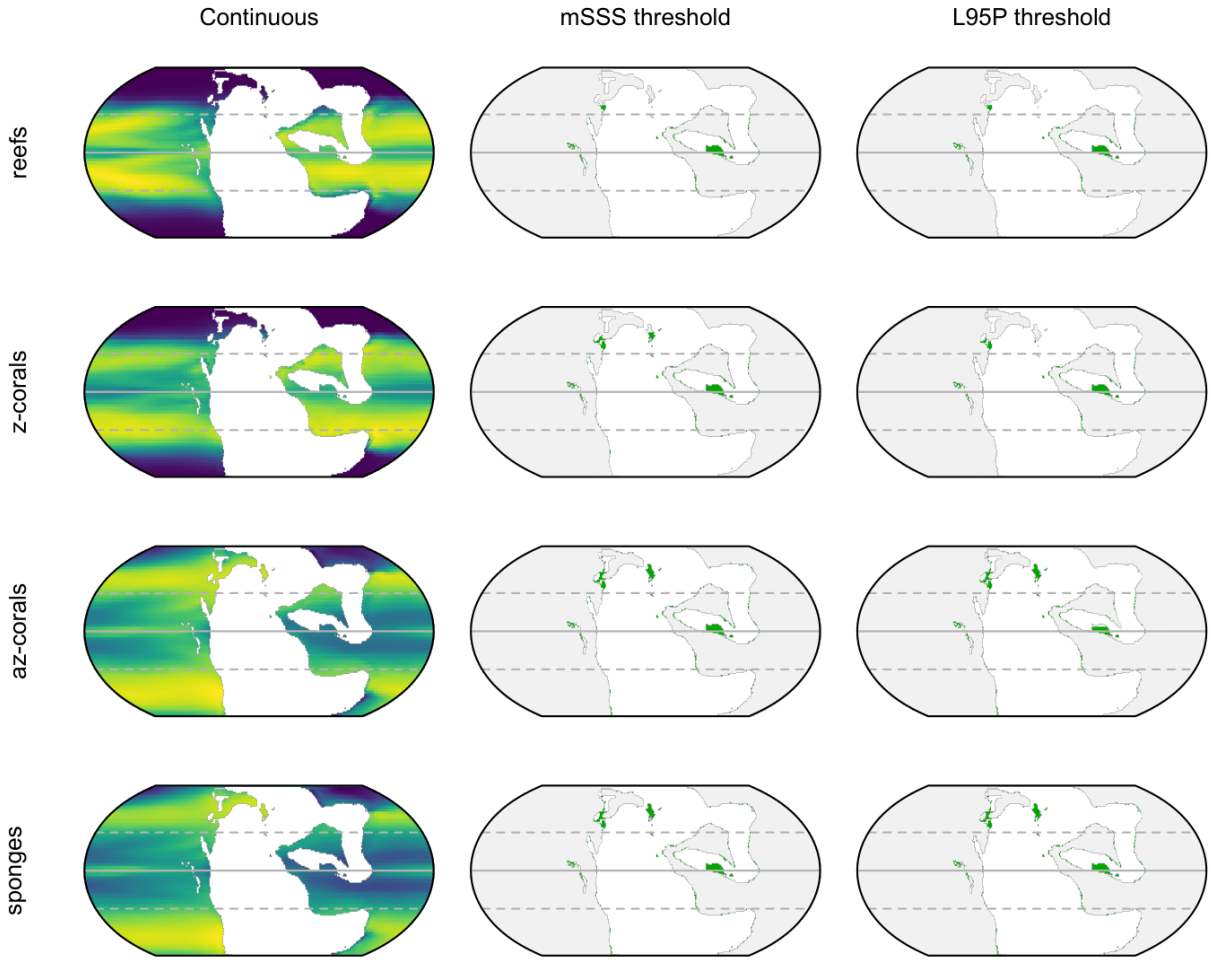


Fig. S71. Niche model projections for the early Carnian (PALEOMAP + HadCM3L). The continuous maps display the probability of habitability (brighter = more probable). The other maps display habitable area after masking to the photic zone and thresholding by the sum of maximum sensitivity and specificity (mSSS) and the threshold that excludes the lowest 5% of occurrence points (L95P). Horizontal lines demarcate the equator and tropics.

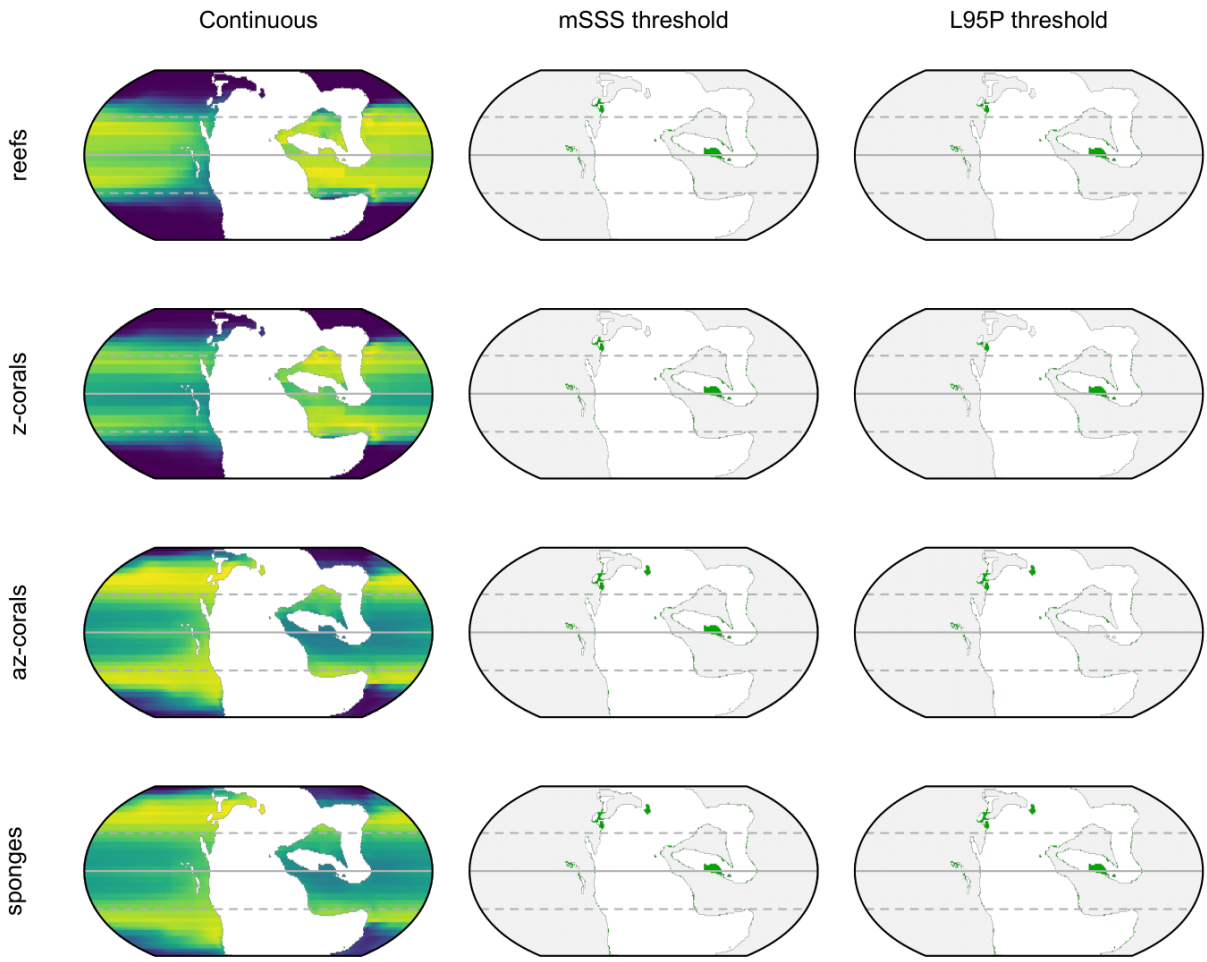


Fig. S72. Niche model projections for the early Carnian (PALEOMAP + CLIMBER-3α). The continuous maps display the probability of habitability (brighter = more probable). The other maps display habitable area after masking to the photic zone and thresholding by the sum of maximum sensitivity and specificity (mSSS) and the threshold that excludes the lowest 5% of occurrence points (L95P). Horizontal lines demarcate the equator and tropics.

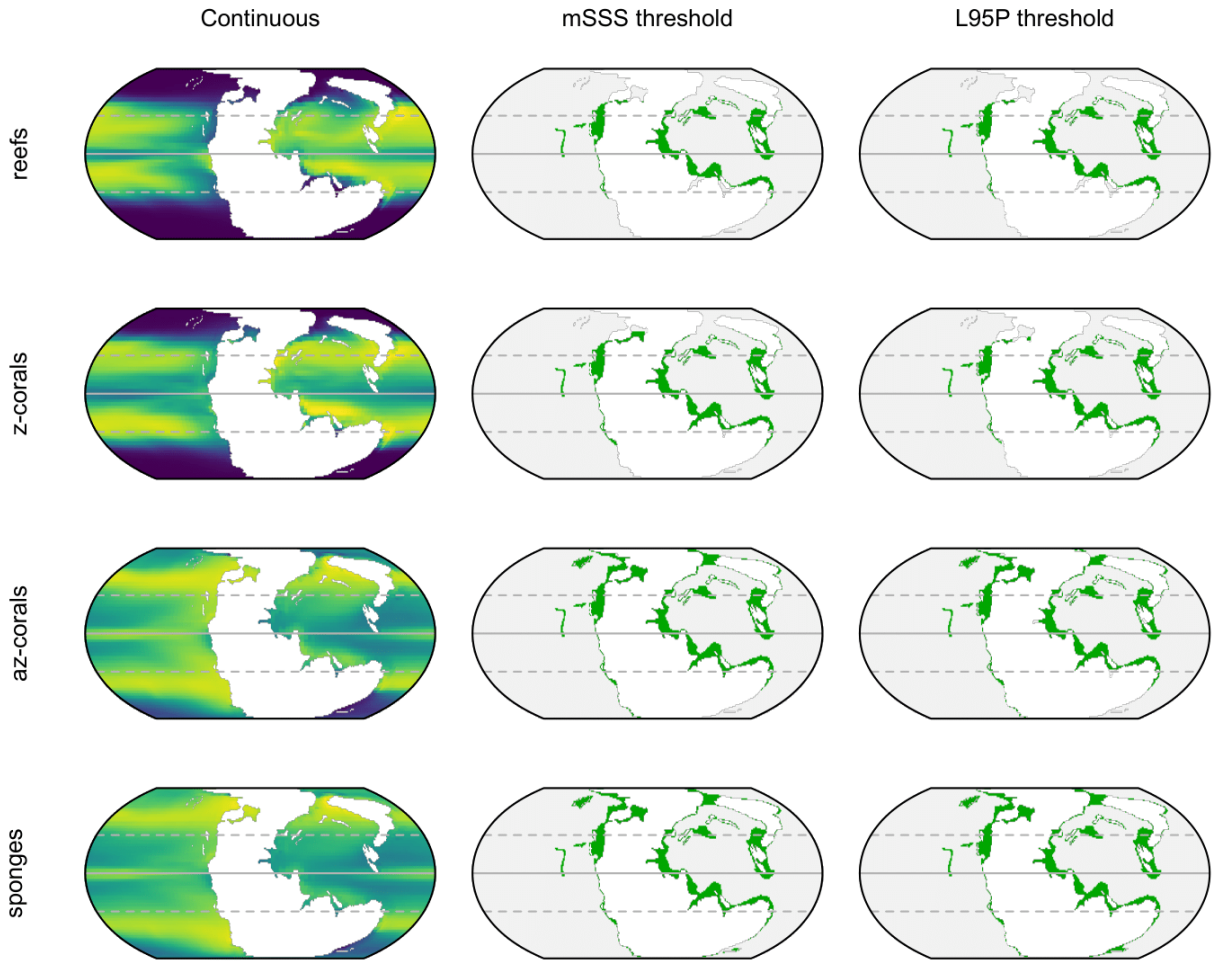


Fig. S73. Niche model projections for the late Carnian (Getech + HadCM3L). The continuous maps display the probability of habitability (brighter = more probable). The other maps display habitable area after masking to the photic zone and thresholding by the sum of maximum sensitivity and specificity (mSSS) and the threshold that excludes the lowest 5% of occurrence points (L95P). Horizontal lines demarcate the equator and tropics.

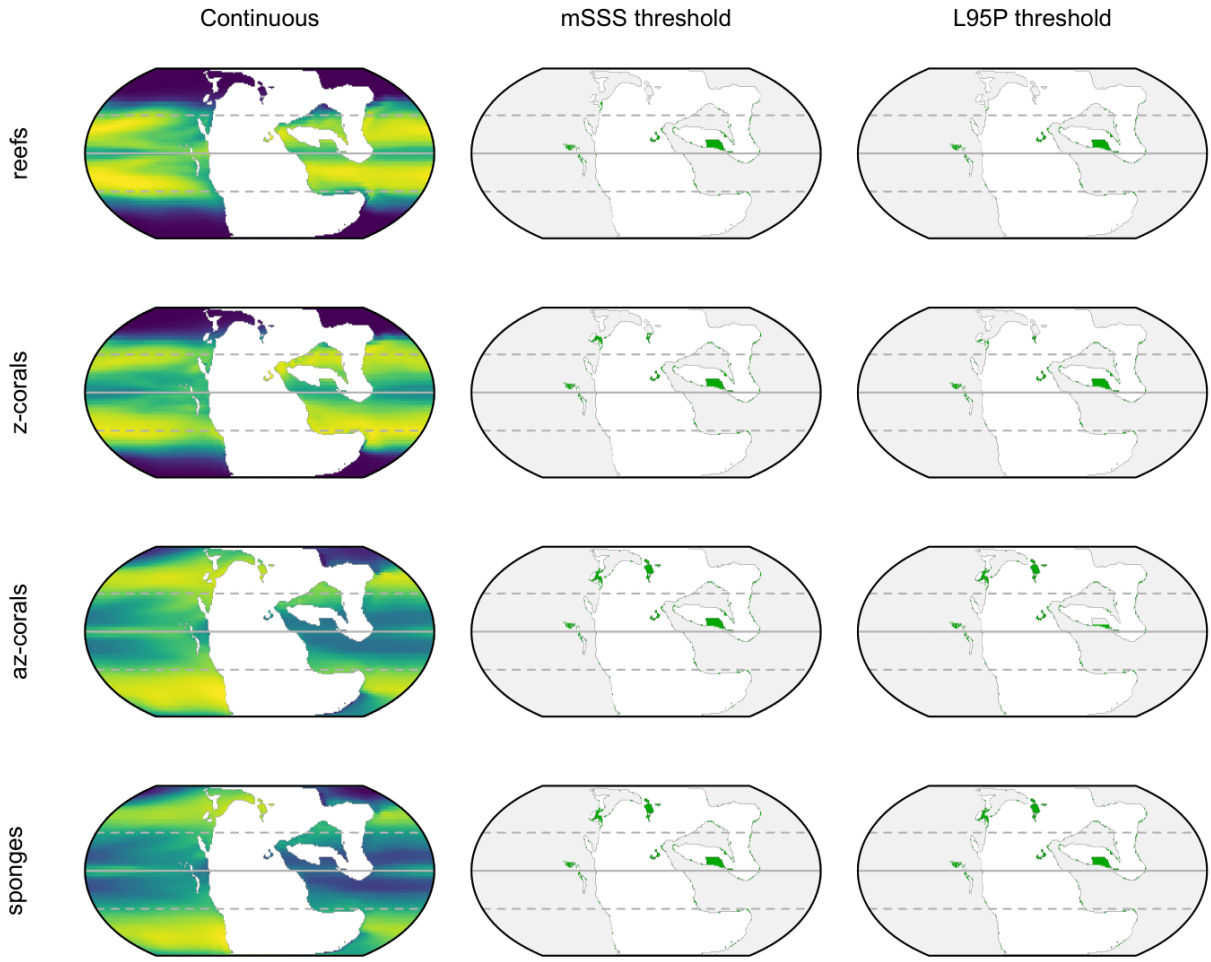


Fig. S74. Niche model projections for the late Carnian (PALEOMAP + HadCM3L). The continuous maps display the probability of habitability (brighter = more probable). The other maps display habitable area after masking to the photic zone and thresholding by the sum of maximum sensitivity and specificity (mSSS) and the threshold that excludes the lowest 5% of occurrence points (L95P). Horizontal lines demarcate the equator and tropics.

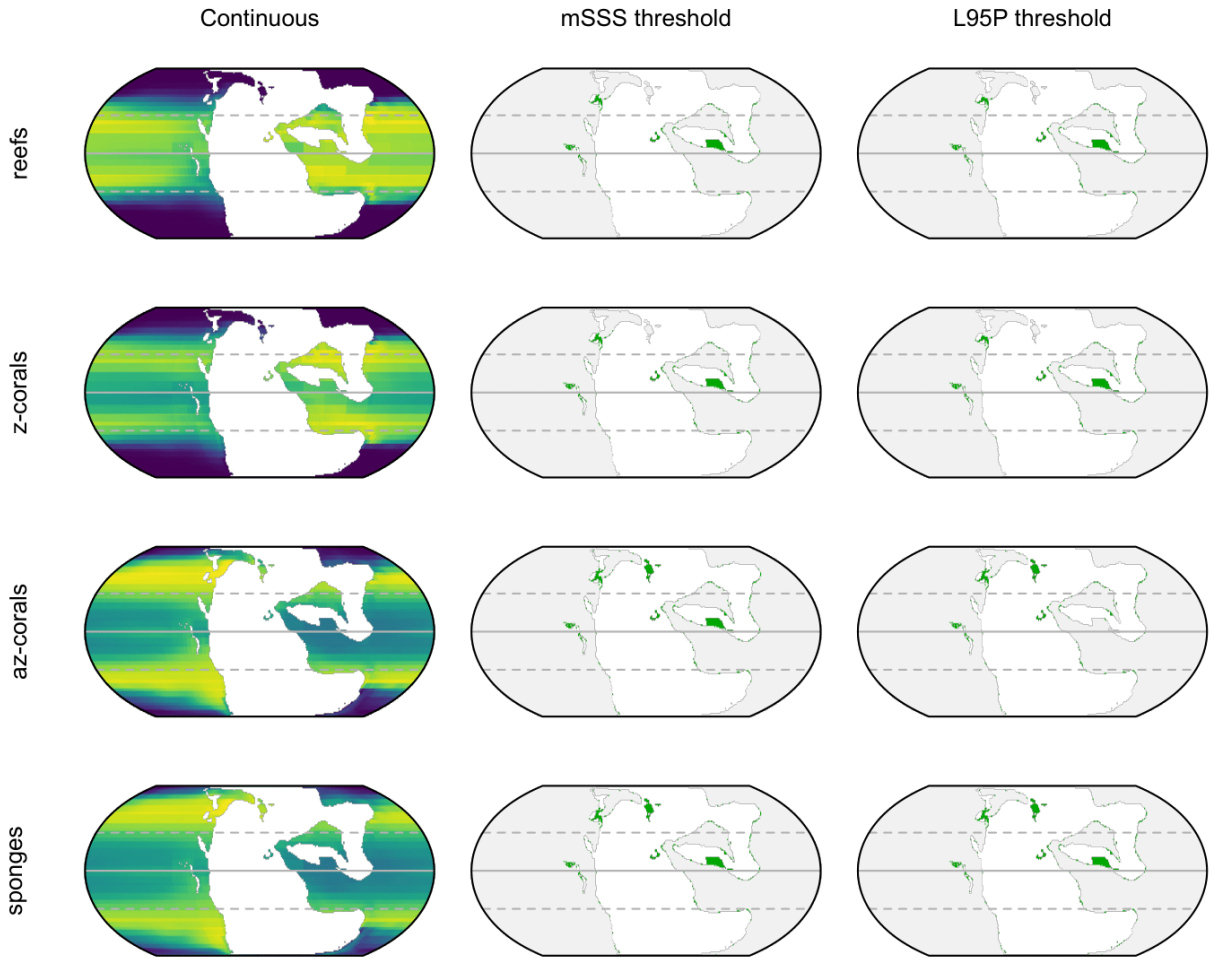


Fig. S75. Niche model projections for the late Carnian (PALEOMAP + CLIMBER-3*α*). The continuous maps display the probability of habitability (brighter = more probable). The other maps display habitable area after masking to the photic zone and thresholding by the sum of maximum sensitivity and specificity (mSSS) and the threshold that excludes the lowest 5% of occurrence points (L95P). Horizontal lines demarcate the equator and tropics.

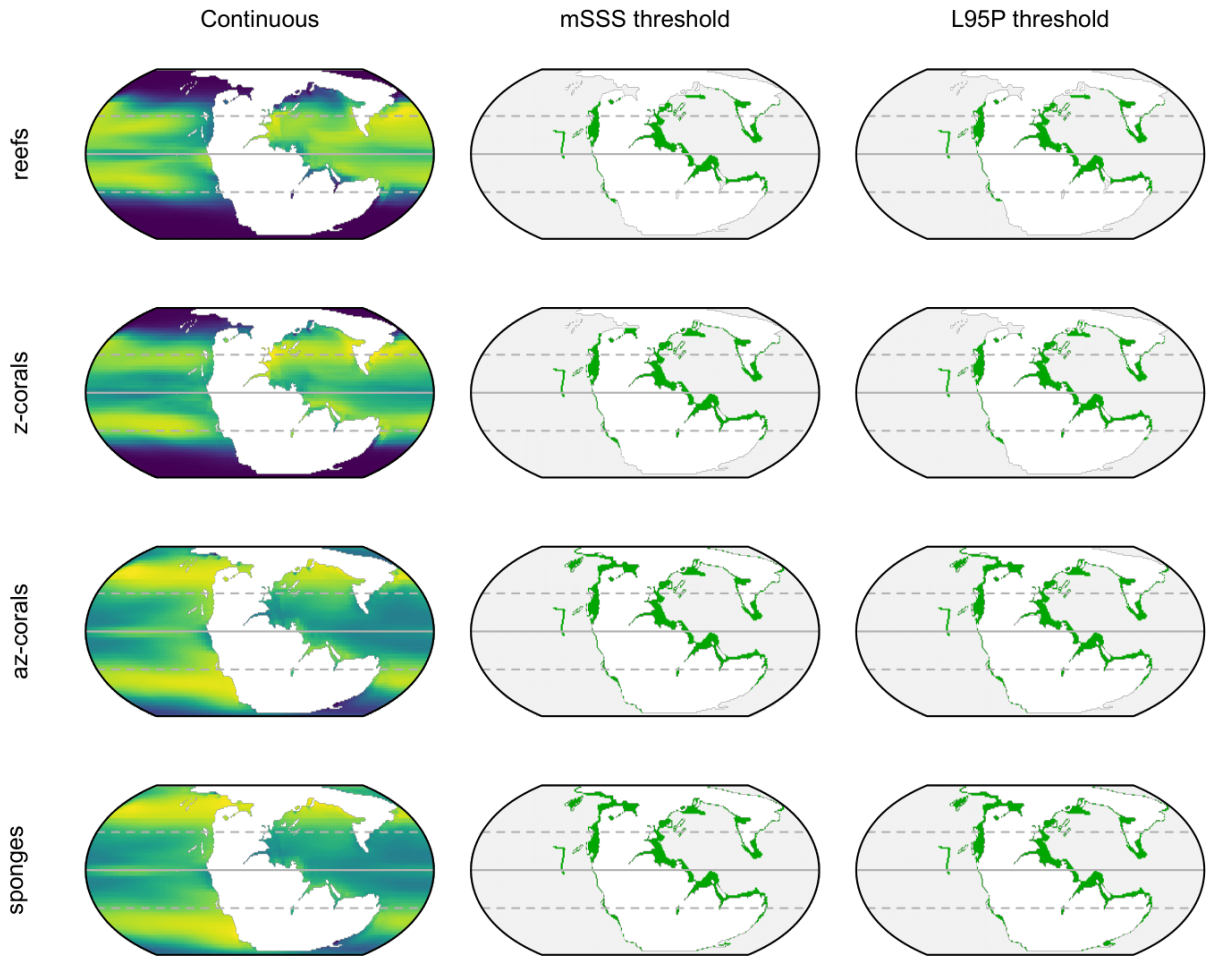


Fig. S76. Niche model projections for the early Norian (Getech + HadCM3L). The continuous maps display the probability of habitability (brighter = more probable). The other maps display habitable area after masking to the photic zone and thresholding by the sum of maximum sensitivity and specificity (mSSS) and the threshold that excludes the lowest 5% of occurrence points (L95P). Horizontal lines demarcate the equator and tropics.

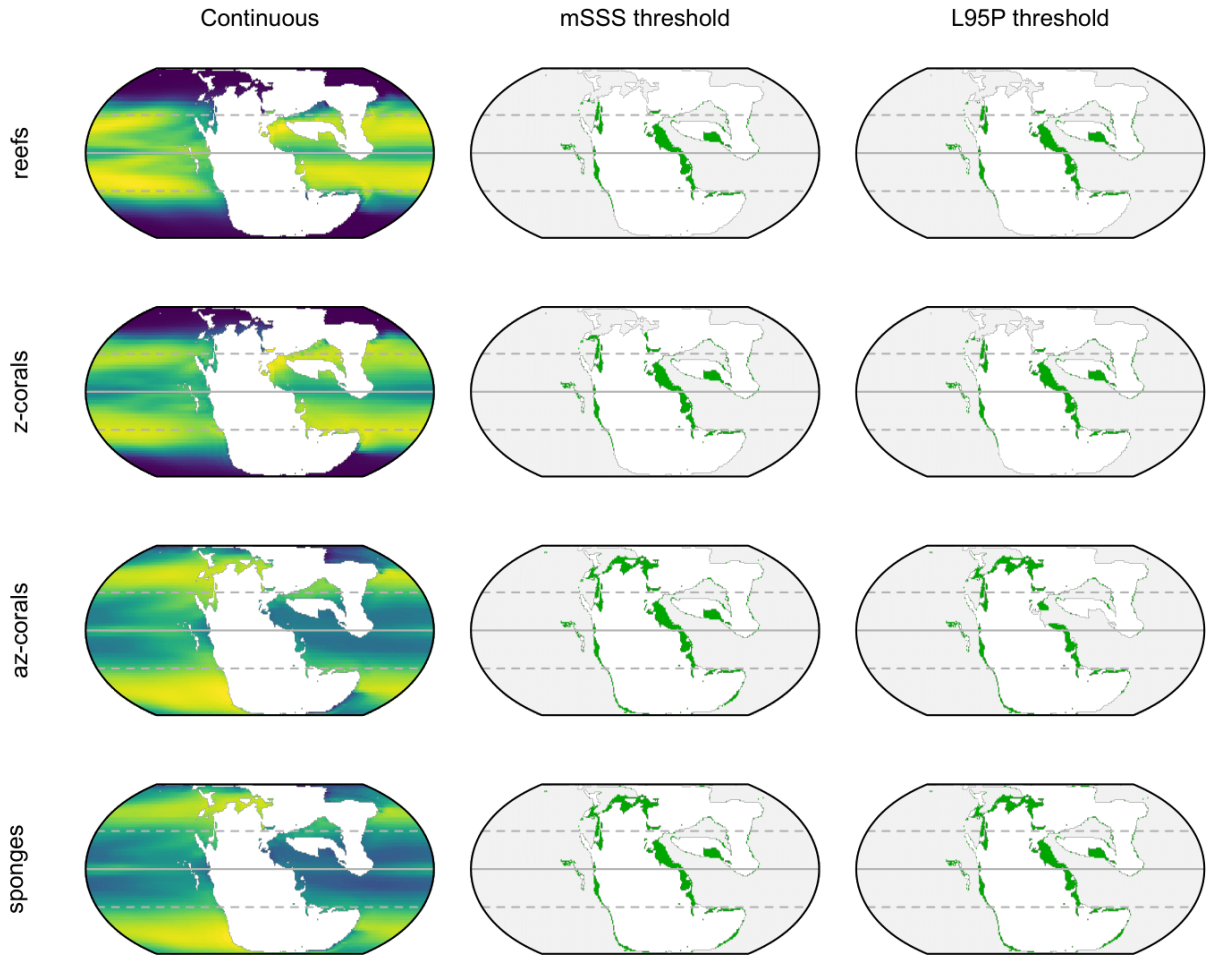


Fig. S77. Niche model projections for the early Norian (PALEOMAP + HadCM3L). The continuous maps display the probability of habitability (brighter = more probable). The other maps display habitable area after masking to the photic zone and thresholding by the sum of maximum sensitivity and specificity (mSSS) and the threshold that excludes the lowest 5% of occurrence points (L95P). Horizontal lines demarcate the equator and tropics.

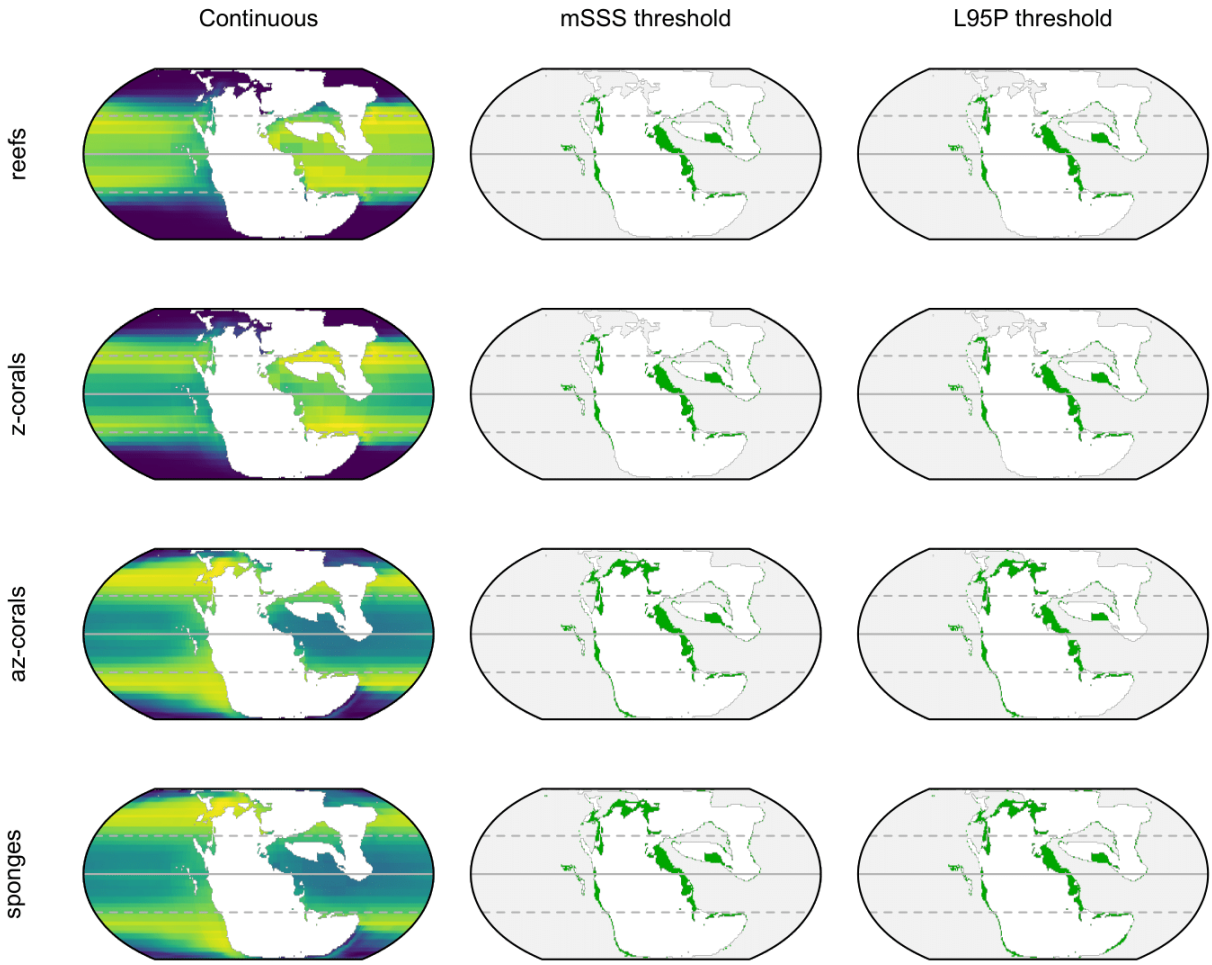


Fig. S78. Niche model projections for the early Norian (PALEOMAP + CLIMBER-3*α*). The continuous maps display the probability of habitability (brighter = more probable). The other maps display habitable area after masking to the photic zone and thresholding by the sum of maximum sensitivity and specificity (mSSS) and the threshold that excludes the lowest 5% of occurrence points (L95P). Horizontal lines demarcate the equator and tropics.

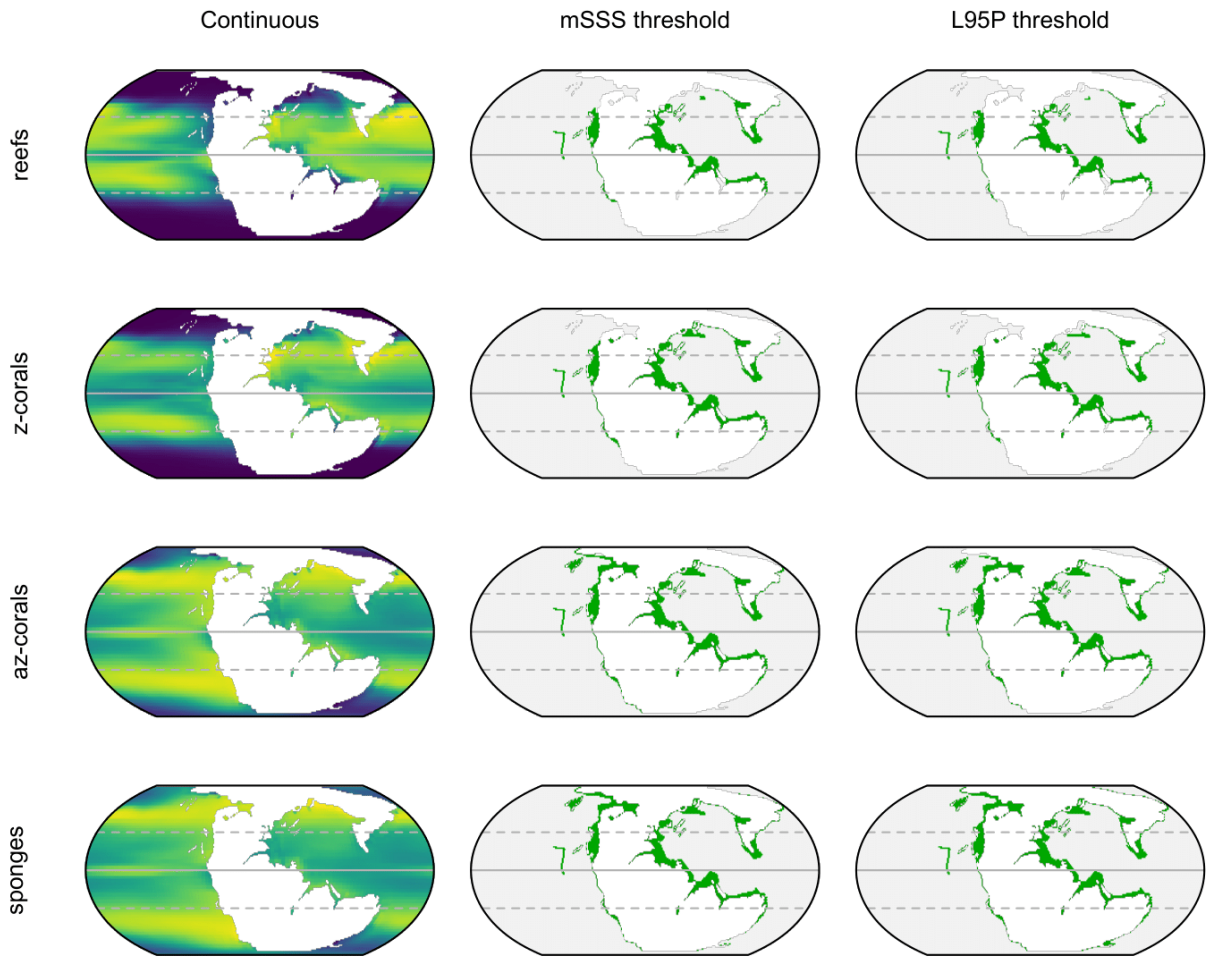


Fig. S79. Niche model projections for the middle Norian (Getech + HadCM3L). The continuous maps display the probability of habitability (brighter = more probable). The other maps display habitable area after masking to the photic zone and thresholding by the sum of maximum sensitivity and specificity (mSSS) and the threshold that excludes the lowest 5% of occurrence points (L95P). Horizontal lines demarcate the equator and tropics.

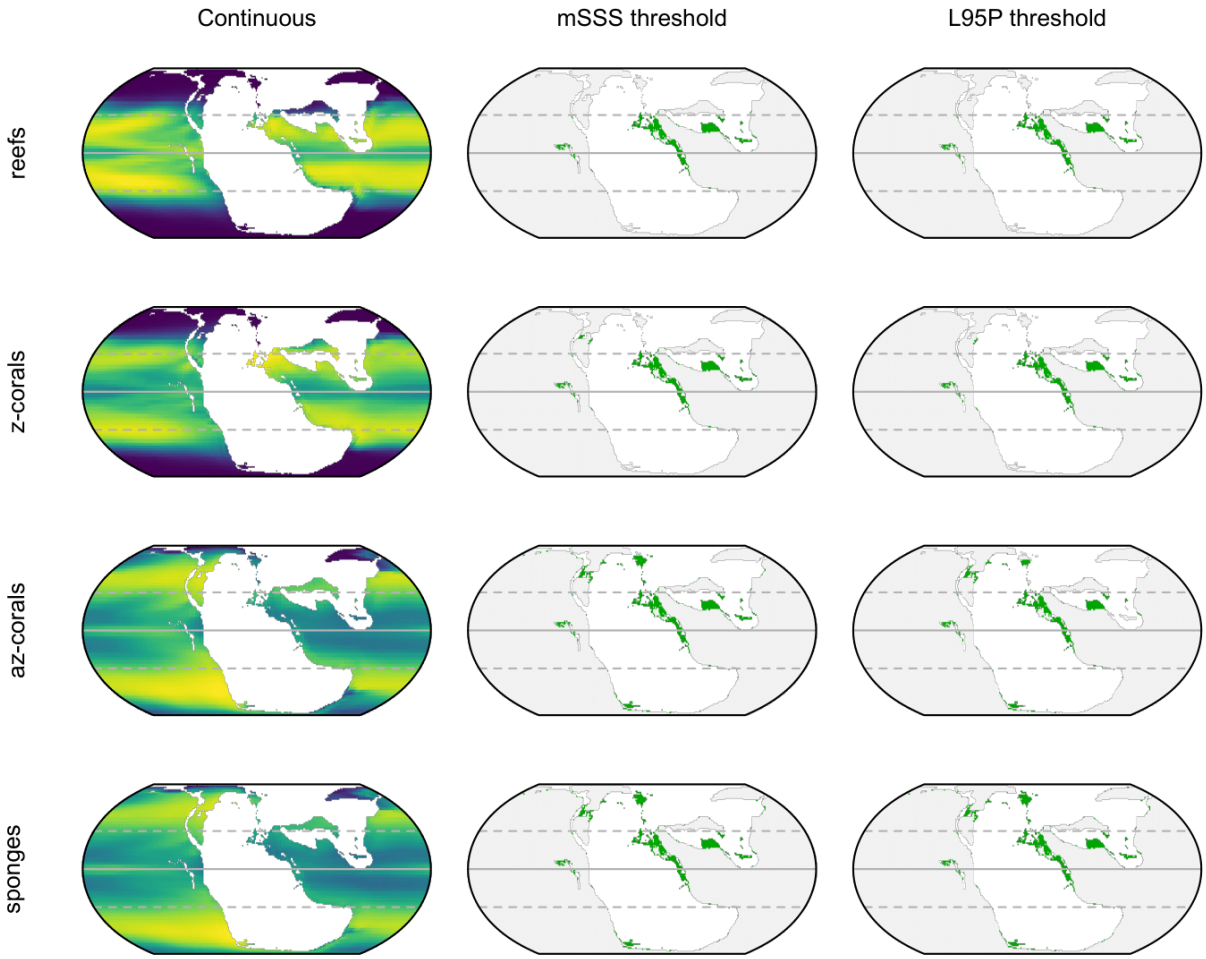


Fig. S80. Niche model projections for the middle Norian (PALEOMAP + HadCM3L). The continuous maps display the probability of habitability (brighter = more probable). The other maps display habitable area after masking to the photic zone and thresholding by the sum of maximum sensitivity and specificity (mSSS) and the threshold that excludes the lowest 5% of occurrence points (L95P). Horizontal lines demarcate the equator and tropics.

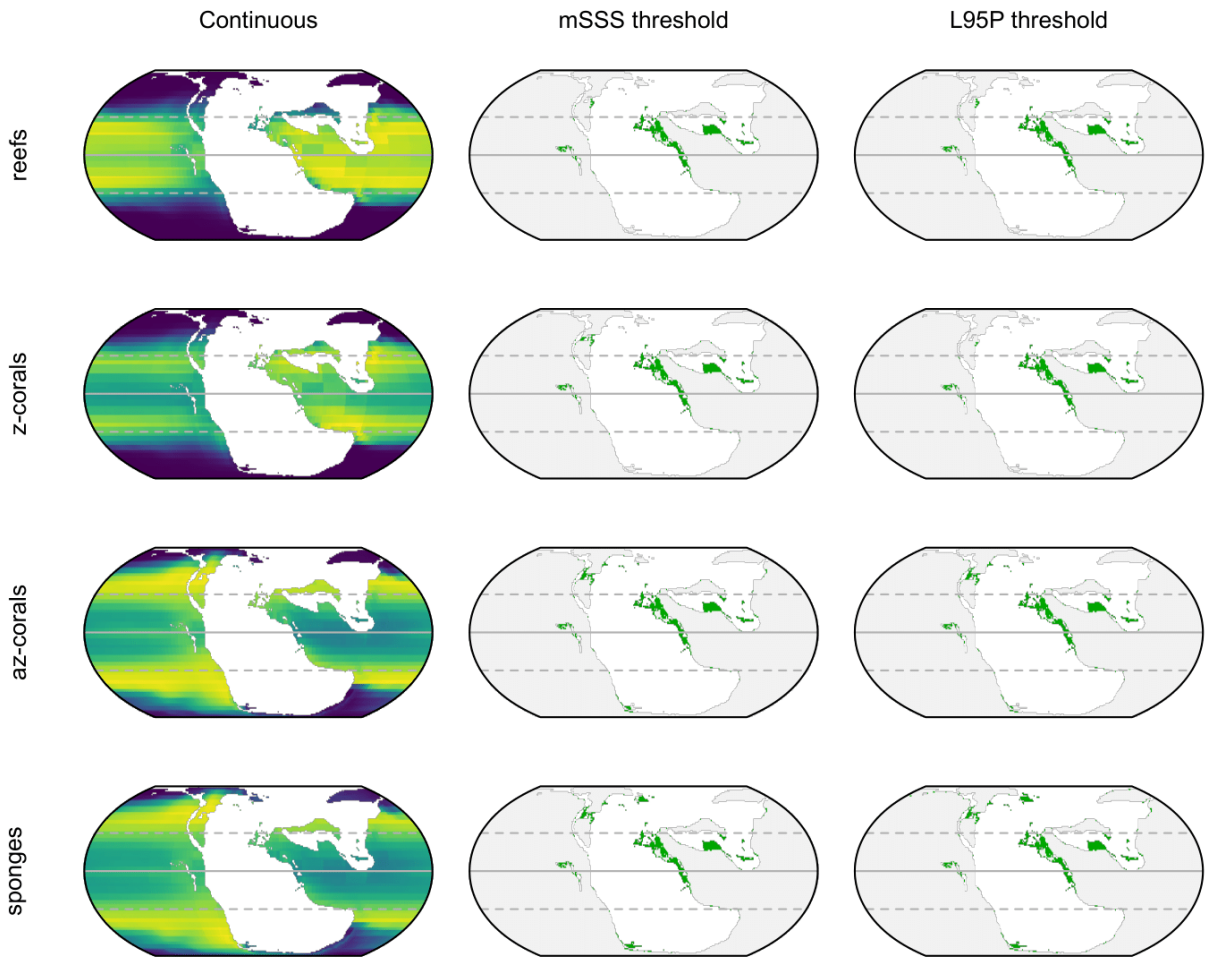


Fig. S81. Niche model projections for the middle Norian (PALEOMAP + CLIMBER-3*α*). The continuous maps display the probability of habitability (brighter = more probable). The other maps display habitable area after masking to the photic zone and thresholding by the sum of maximum sensitivity and specificity (mSSS) and the threshold that excludes the lowest 5% of occurrence points (L95P). Horizontal lines demarcate the equator and tropics.

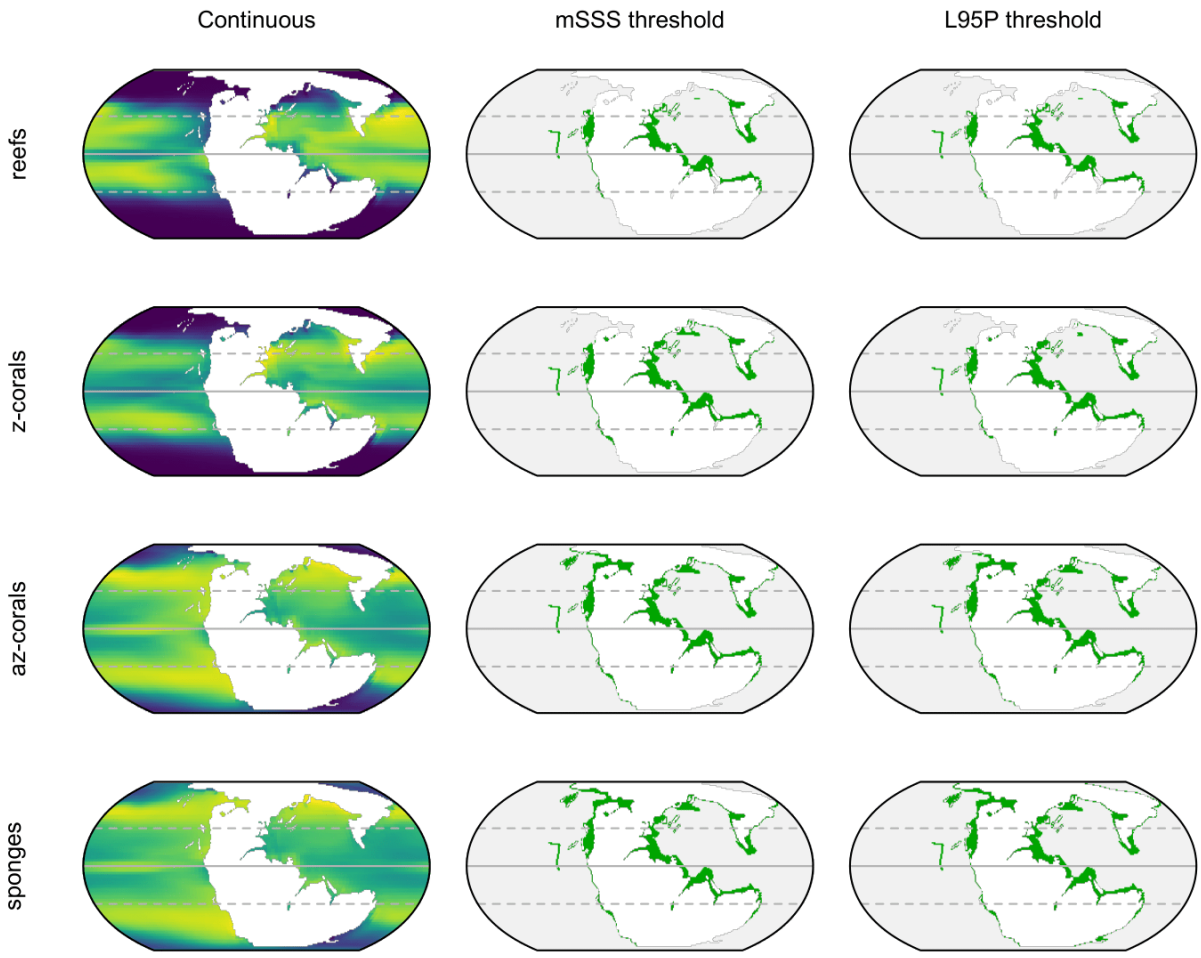


Fig. S82. Niche model projections for the late Norian (Getech + HadCM3L). The continuous maps display the probability of habitability (brighter = more probable). The other maps display habitable area after masking to the photic zone and thresholding by the sum of maximum sensitivity and specificity (mSSS) and the threshold that excludes the lowest 5% of occurrence points (L95P). Horizontal lines demarcate the equator and tropics.

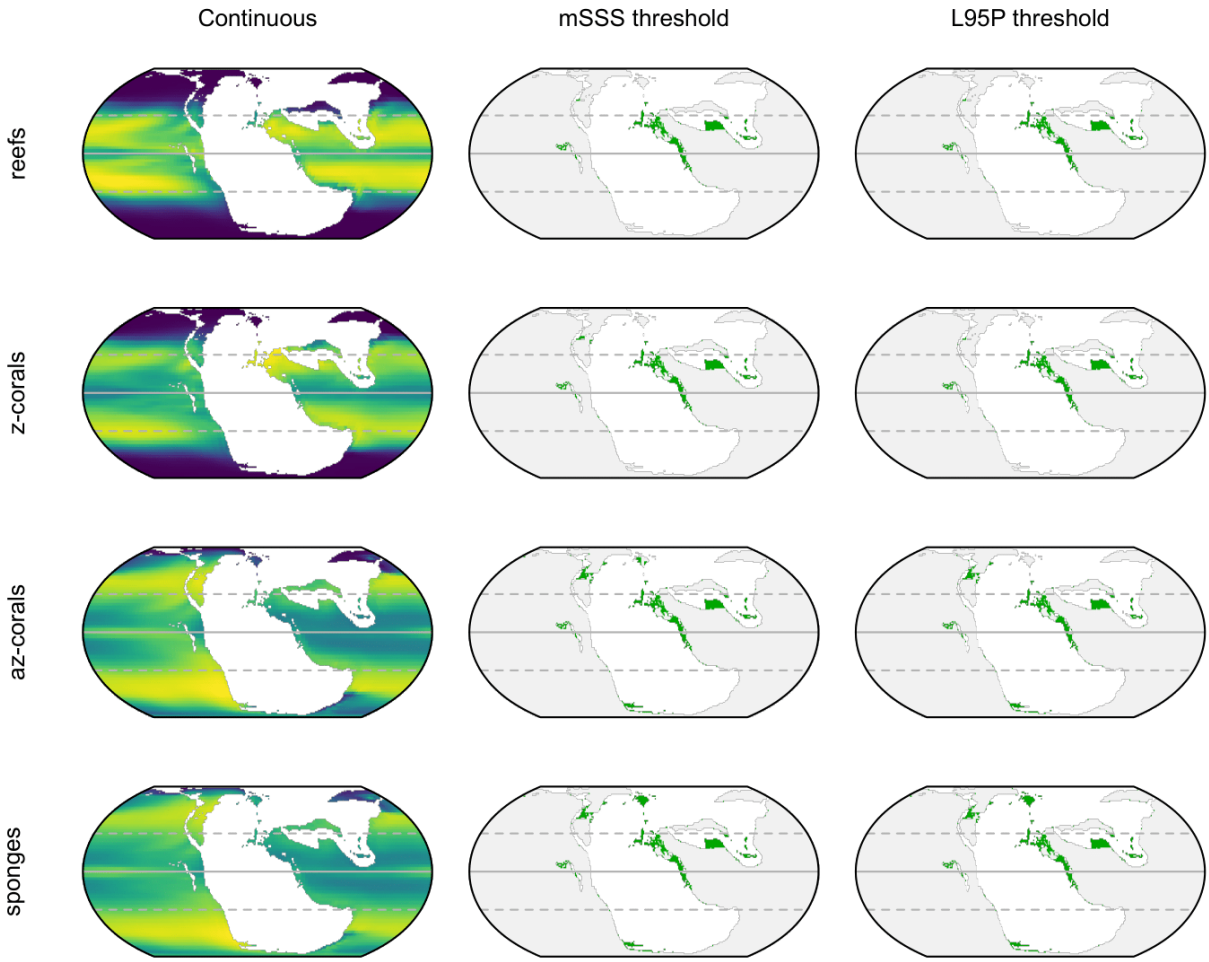


Fig. S83. Niche model projections for the late Norian (PALEOMAP + HadCM3L). The continuous maps display the probability of habitability (brighter = more probable). The other maps display habitable area after masking to the photic zone and thresholding by the sum of maximum sensitivity and specificity (mSSS) and the threshold that excludes the lowest 5% of occurrence points (L95P). Horizontal lines demarcate the equator and tropics.

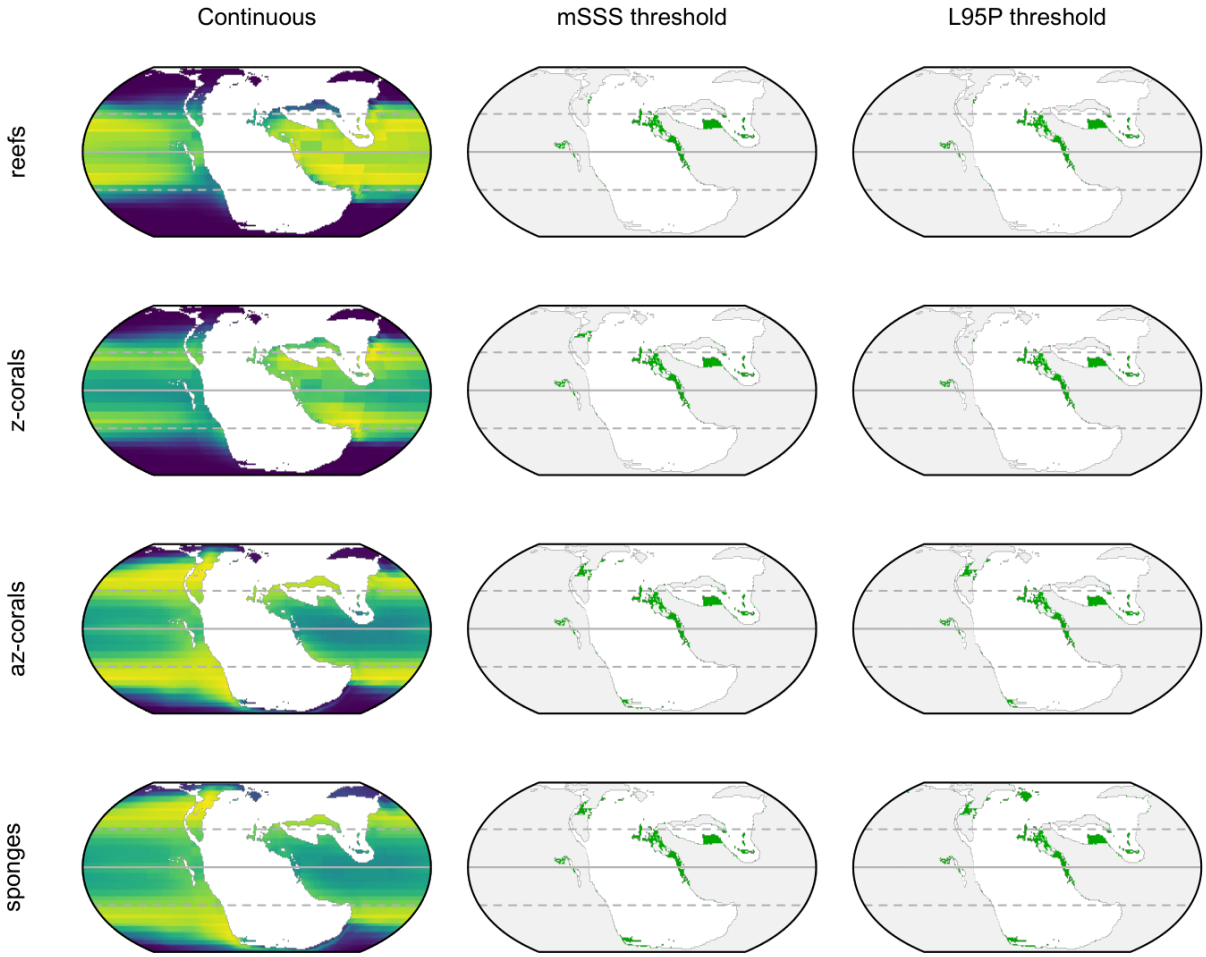


Fig. S84. Niche model projections for the late Norian (PALEOMAP + CLIMBER-3*α*). The continuous maps display the probability of habitability (brighter = more probable). The other maps display habitable area after masking to the photic zone and thresholding by the sum of maximum sensitivity and specificity (mSSS) and the threshold that excludes the lowest 5% of occurrence points (L95P). Horizontal lines demarcate the equator and tropics.

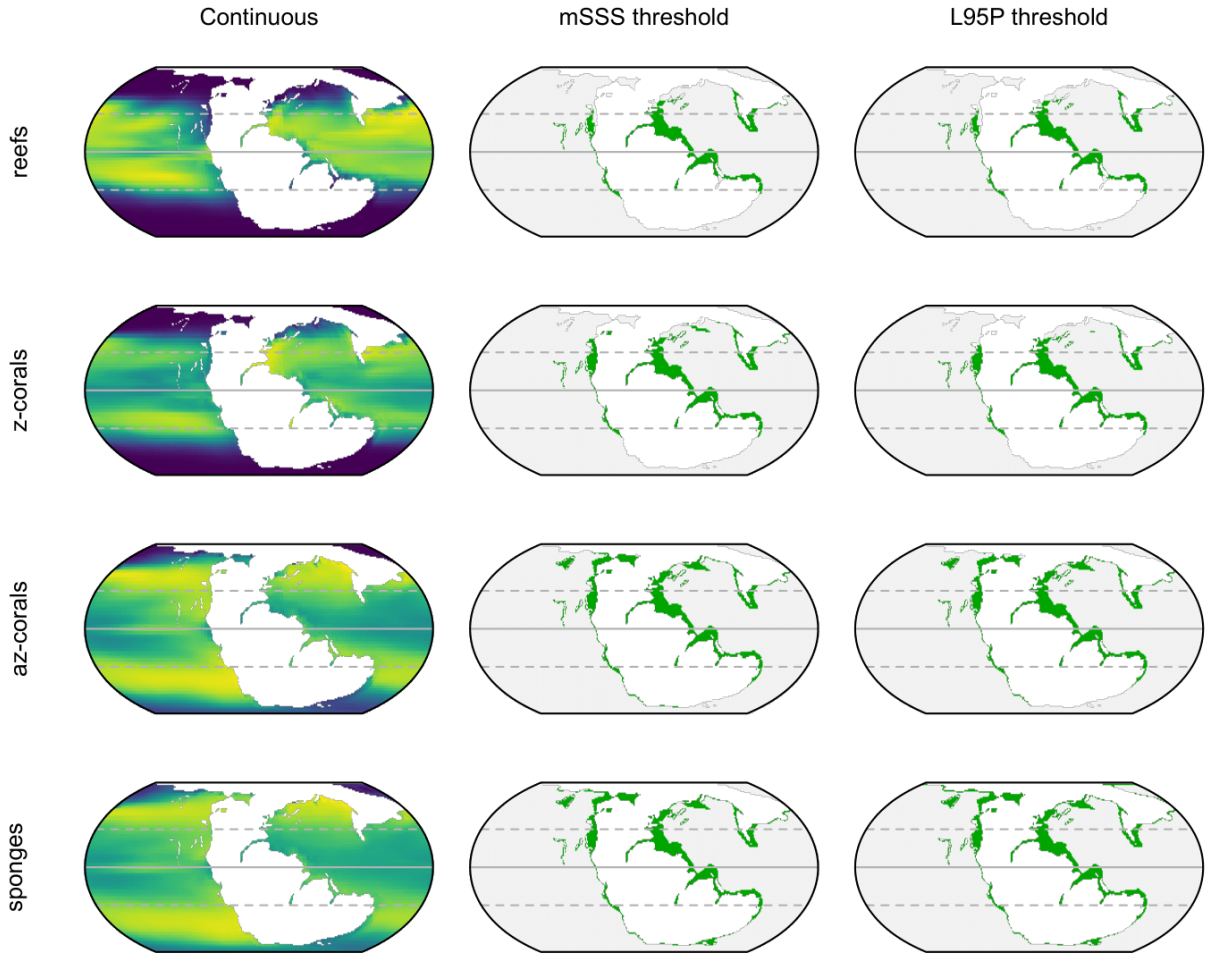


Fig. S85. Niche model projections for the Rhaetian (Getech + HadCM3L). The continuous maps display the probability of habitability (brighter = more probable). The other maps display habitable area after masking to the photic zone and thresholding by the sum of maximum sensitivity and specificity (mSSS) and the threshold that excludes the lowest 5% of occurrence points (L95P). Horizontal lines demarcate the equator and tropics.

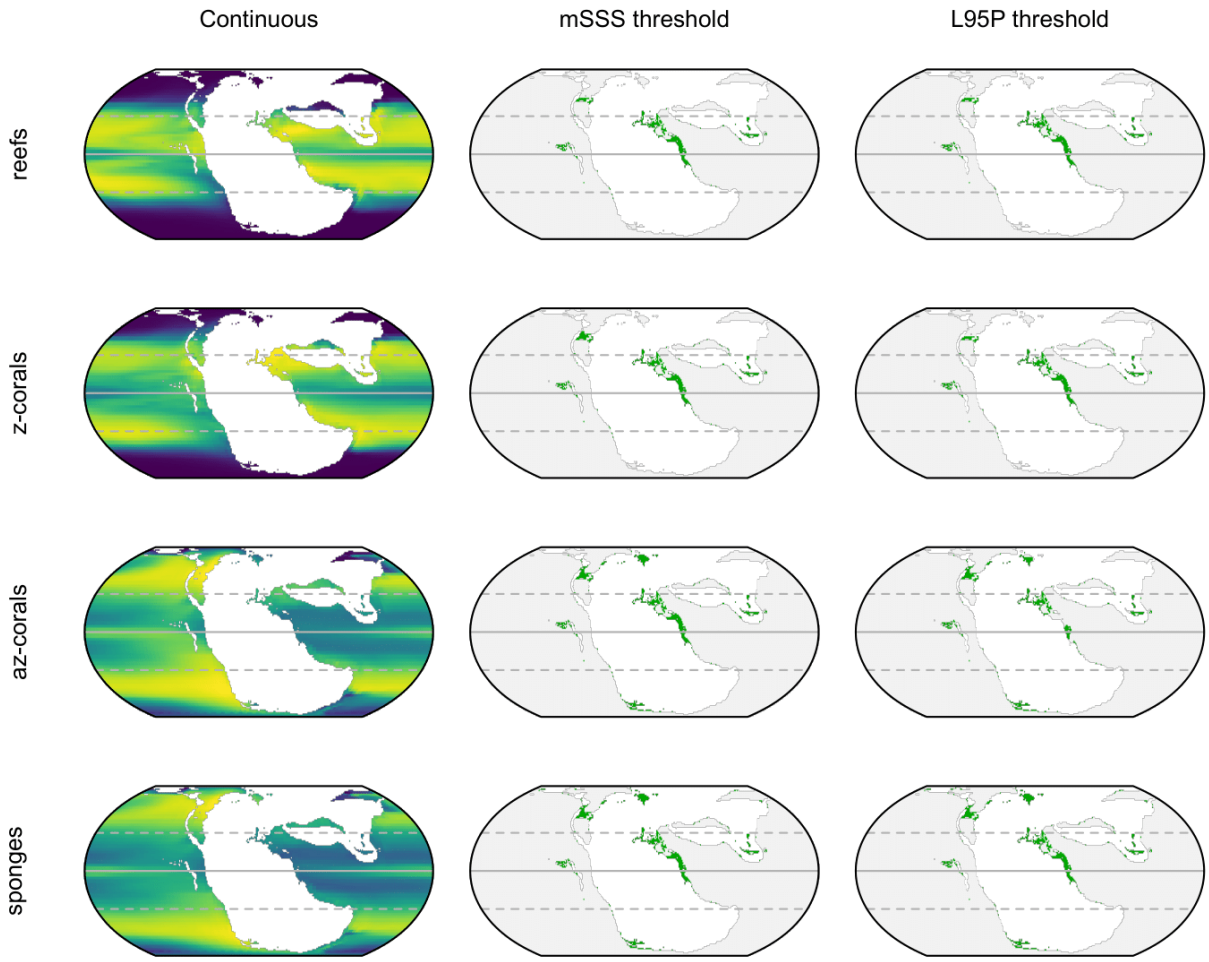


Fig. S86. Niche model projections for the Rhaetian (PALEOMAP + HadCM3L). The continuous maps display the probability of habitability (brighter = more probable). The other maps display habitable area after masking to the photic zone and thresholding by the sum of maximum sensitivity and specificity (mSSS) and the threshold that excludes the lowest 5% of occurrence points (L95P). Horizontal lines demarcate the equator and tropics.

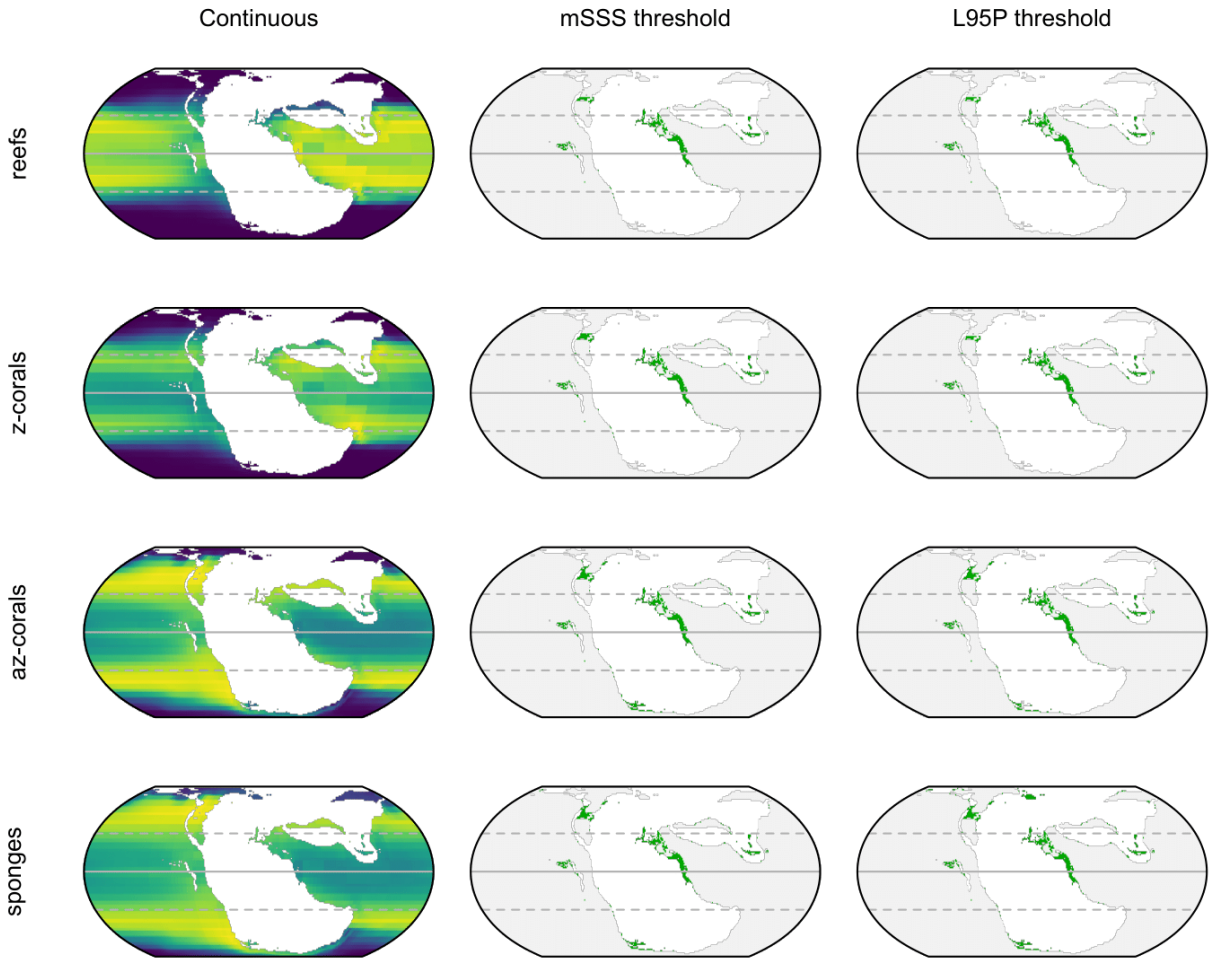


Fig. S87. Niche model projections for the Rhaetian (PALEOMAP + CLIMBER-3α). The continuous maps display the probability of habitability (brighter = more probable). The other maps display habitable area after masking to the photic zone and thresholding by the sum of maximum sensitivity and specificity (mSSS) and the threshold that excludes the lowest 5% of occurrence points (L95P). Horizontal lines demarcate the equator and tropics.

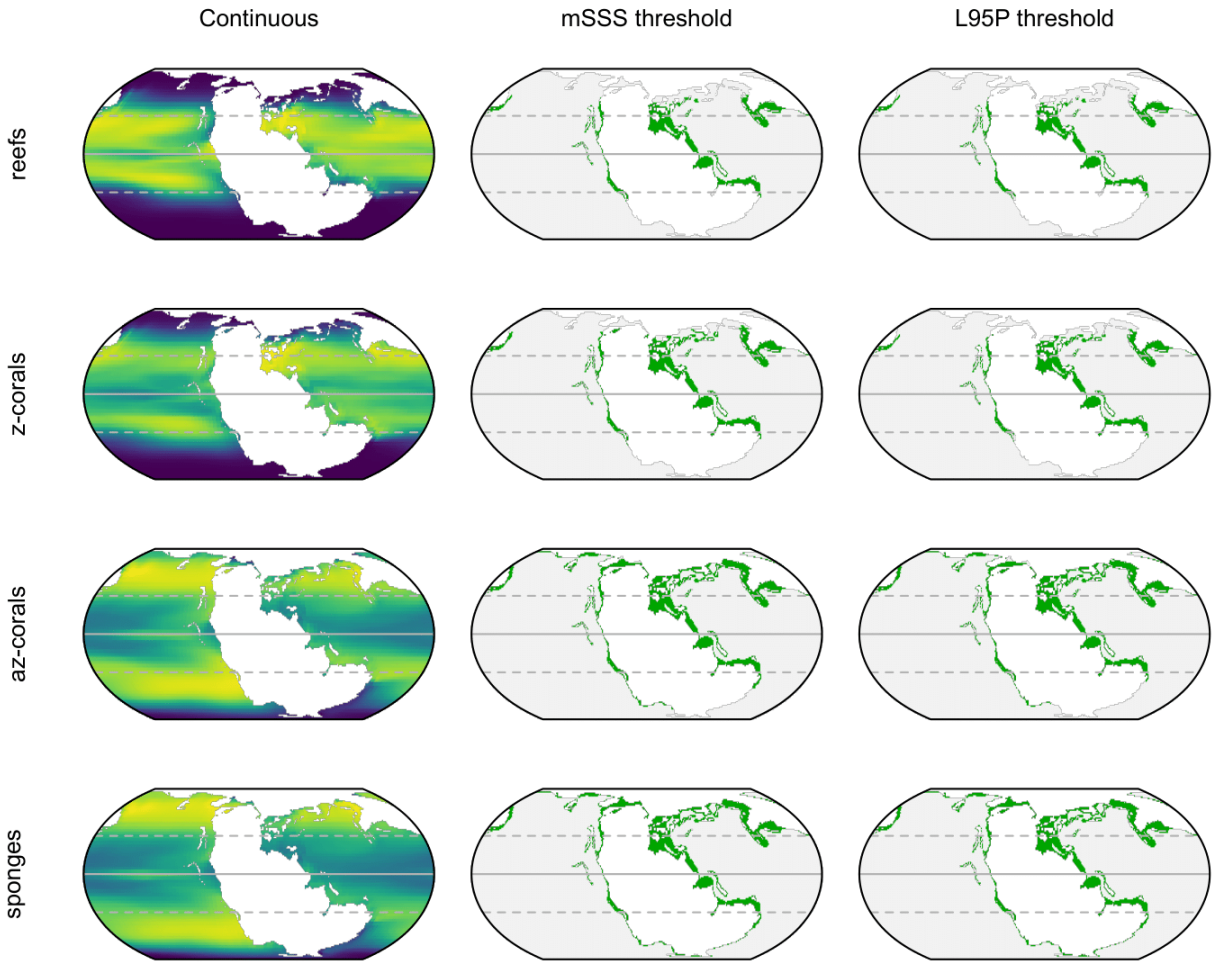


Fig. S88. Niche model projections for the Hettangian (Getech + HadCM3L). The continuous maps display the probability of habitability (brighter = more probable). The other maps display habitable area after masking to the photic zone and thresholding by the sum of maximum sensitivity and specificity (mSSS) and the threshold that excludes the lowest 5% of occurrence points (L95P). Horizontal lines demarcate the equator and tropics.

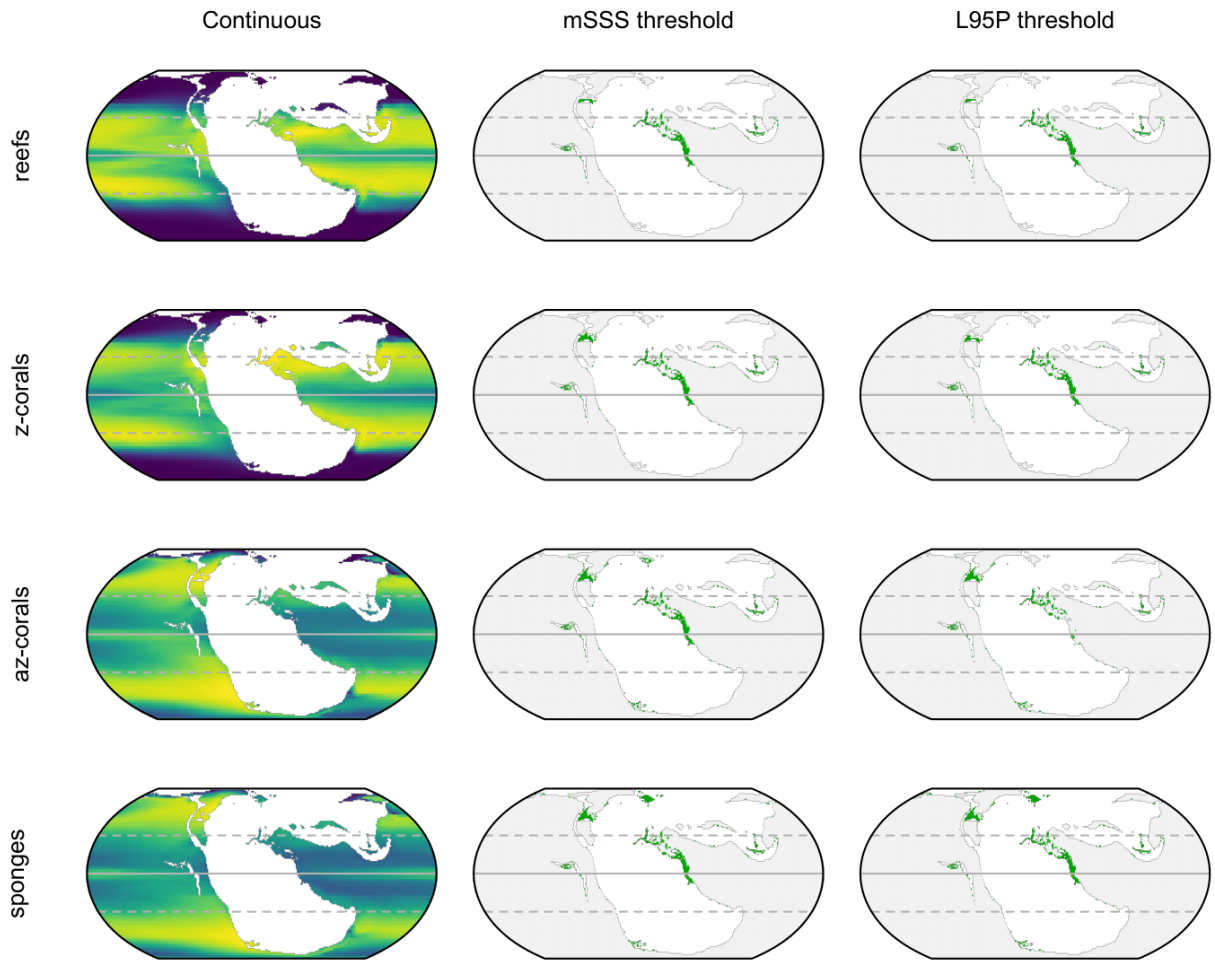


Fig. S89. Niche model projections for the Hettangian (PALEOMAP + HadCM3L). The continuous maps display the probability of habitability (brighter = more probable). The other maps display habitable area after masking to the photic zone and thresholding by the sum of maximum sensitivity and specificity (mSSS) and the threshold that excludes the lowest 5% of occurrence points (L95P). Horizontal lines demarcate the equator and tropics.

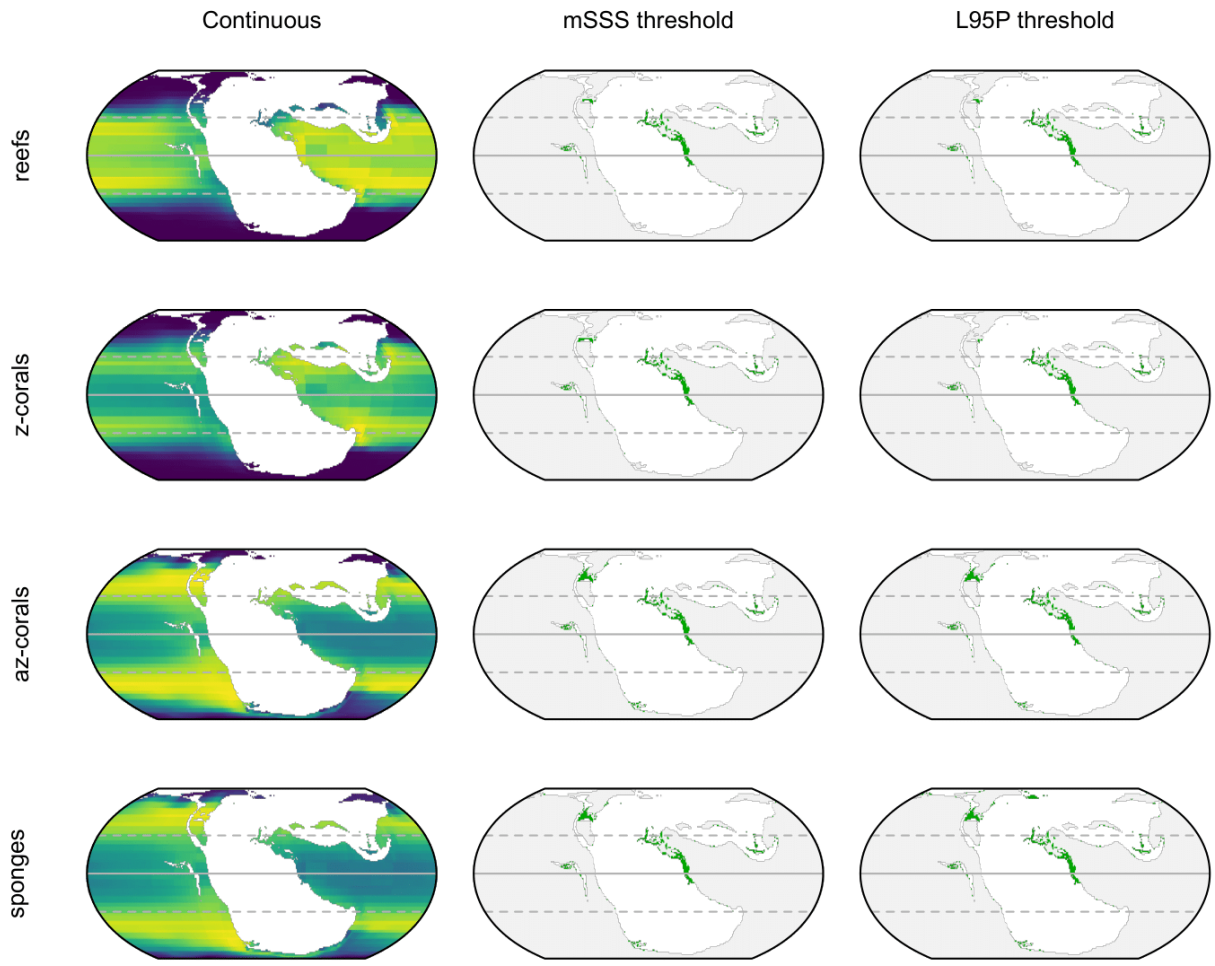


Fig. S90. Niche model projections for the Hettangian (PALEOMAP + CLIMBER-3α). The continuous maps display the probability of habitability (brighter = more probable). The other maps display habitable area after masking to the photic zone and thresholding by the sum of maximum sensitivity and specificity (mSSS) and the threshold that excludes the lowest 5% of occurrence points (L95P). Horizontal lines demarcate the equator and tropics.

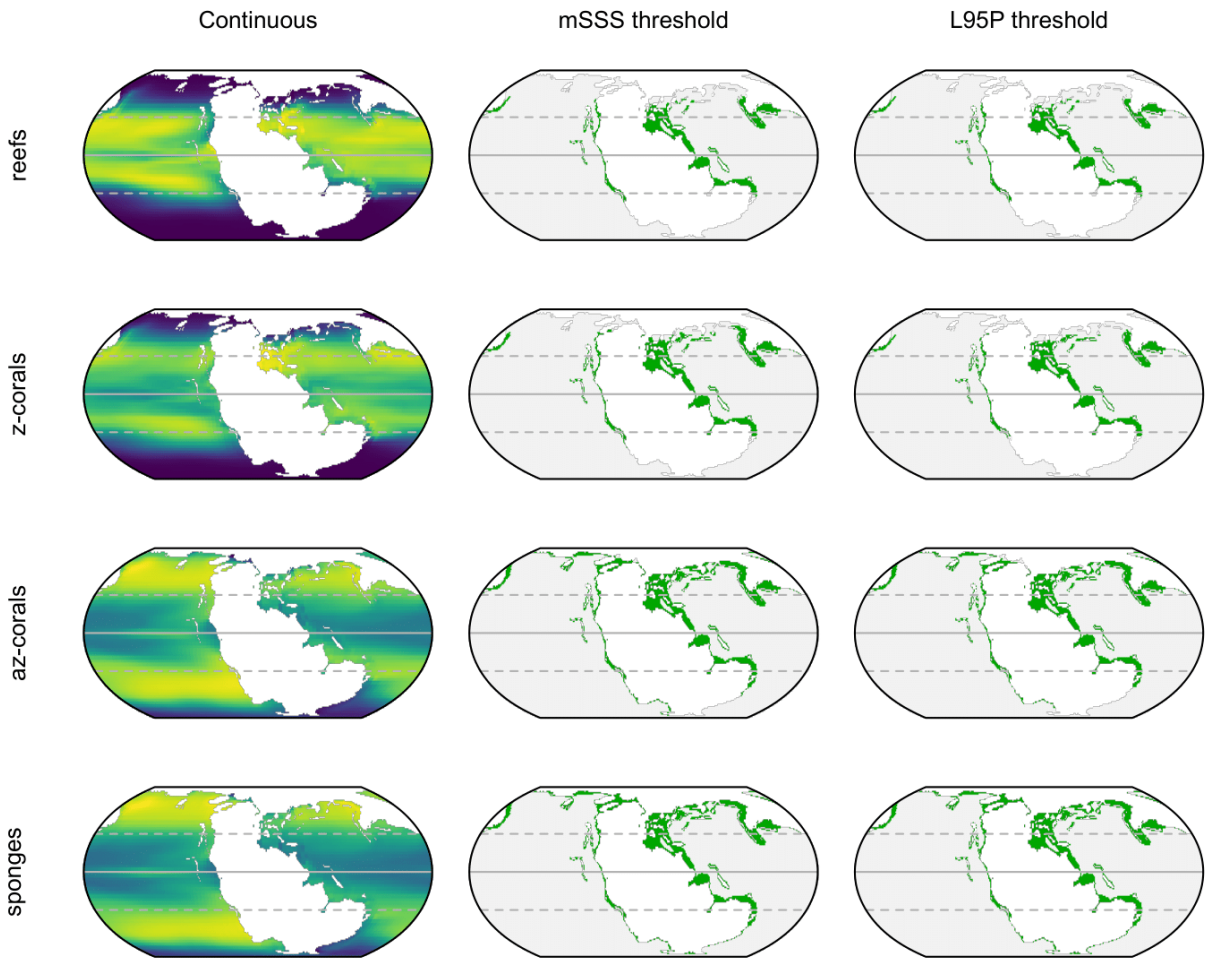


Fig. S91. Niche model projections for the Sinemurian (Getech + HadCM3L). The continuous maps display the probability of habitability (brighter = more probable). The other maps display habitable area after masking to the photic zone and thresholding by the sum of maximum sensitivity and specificity (mSSS) and the threshold that excludes the lowest 5% of occurrence points (L95P). Horizontal lines demarcate the equator and tropics.

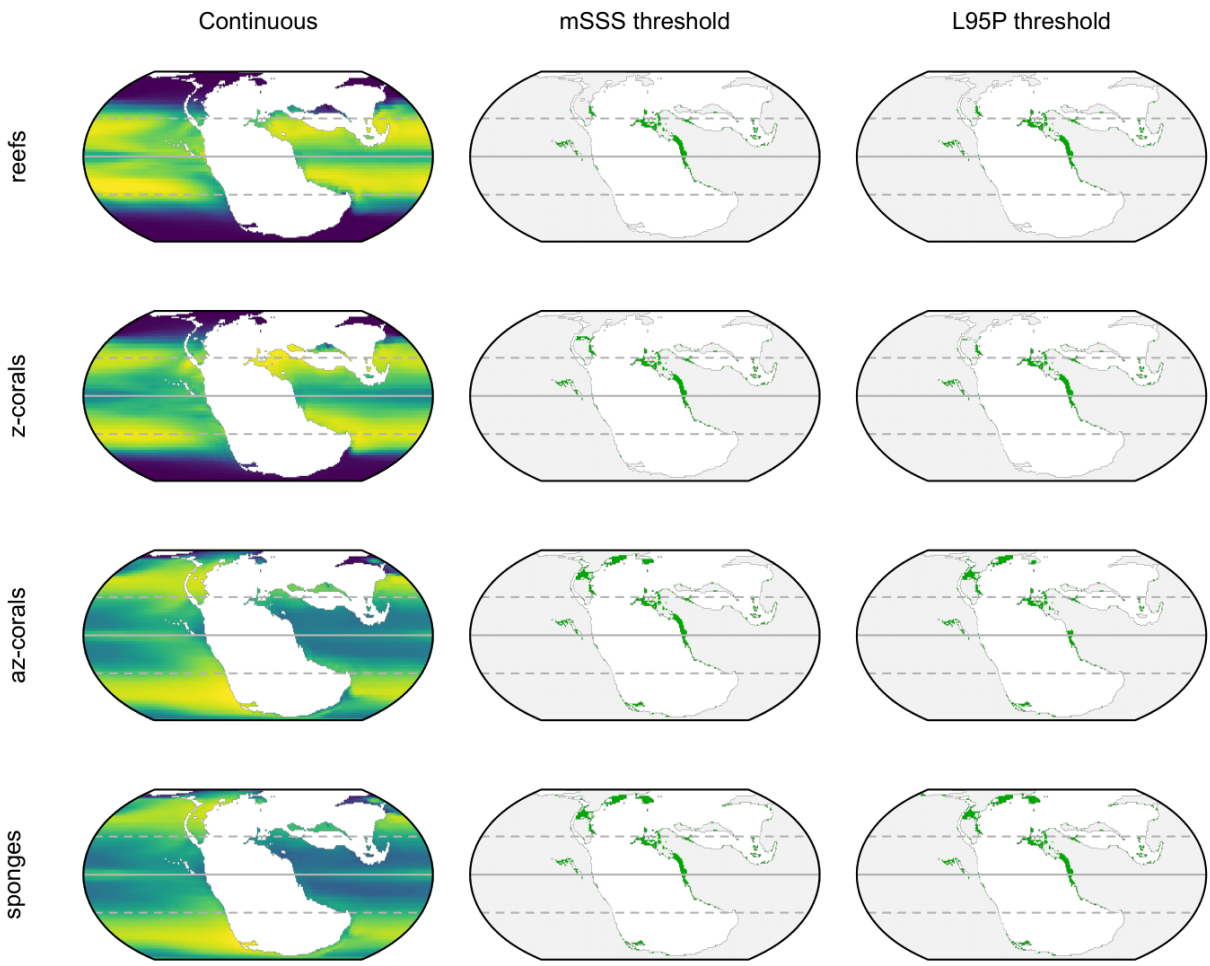


Fig. S92. Niche model projections for the Sinemurian (PALEOMAP + HadCM3L). The continuous maps display the probability of habitability (brighter = more probable). The other maps display habitable area after masking to the photic zone and thresholding by the sum of maximum sensitivity and specificity (mSSS) and the threshold that excludes the lowest 5% of occurrence points (L95P). Horizontal lines demarcate the equator and tropics.

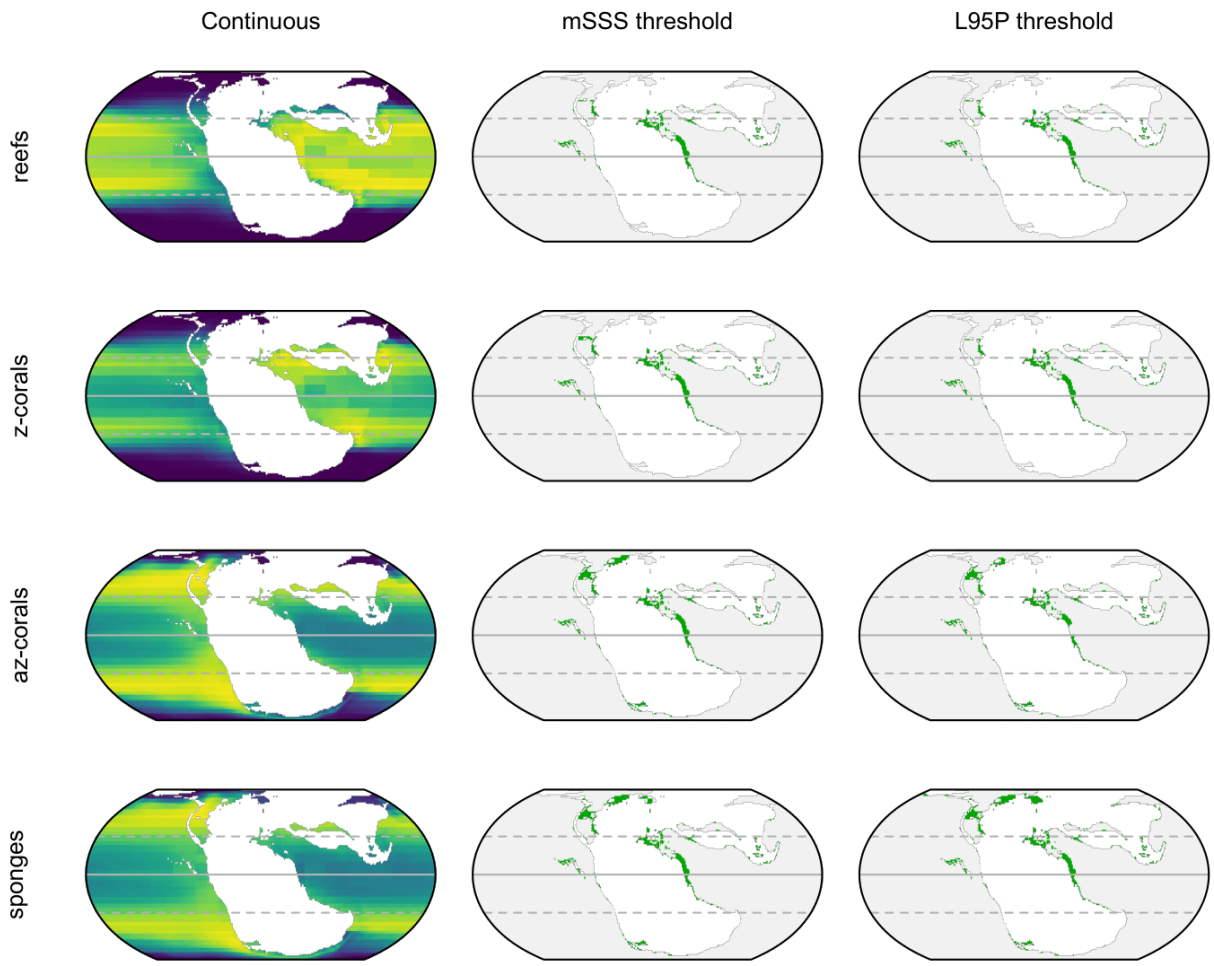


Fig. S93. Niche model projections for the Sinemurian (PALEOMAP + CLIMBER-3α). The continuous maps display the probability of habitability (brighter = more probable). The other maps display habitable area after masking to the photic zone and thresholding by the sum of maximum sensitivity and specificity (mSSS) and the threshold that excludes the lowest 5% of occurrence points (L95P). Horizontal lines demarcate the equator and tropics.

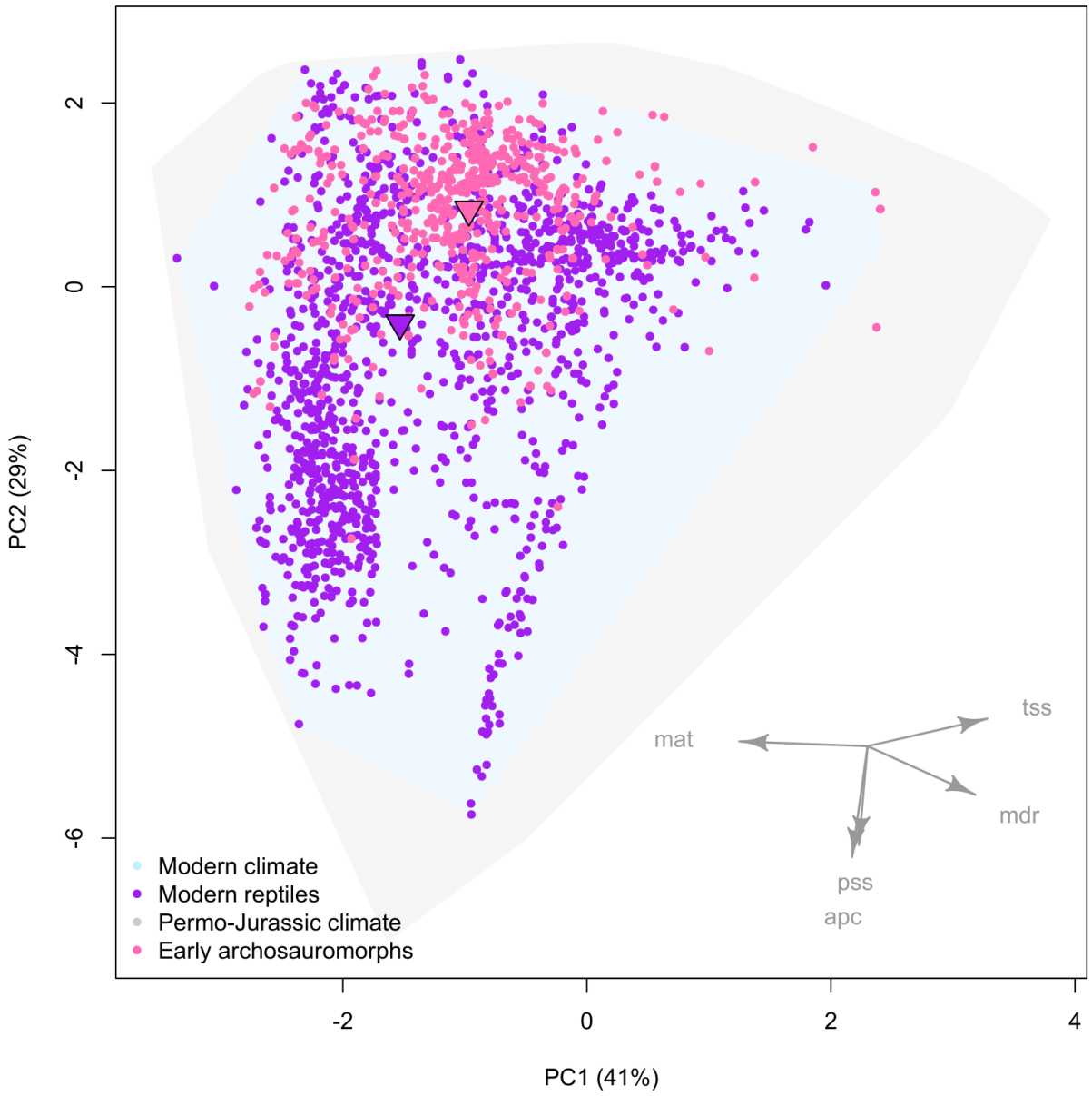


Fig. S94. Climate space of modern reptiles versus Permo-Triassic archosauromorphs. Triangles indicate the centroids of each group. mat = mean annual temperature, tss = temperature seasonality, mdr = mean diurnal range, pss = precipitation seasonality, apc = annual precipitation

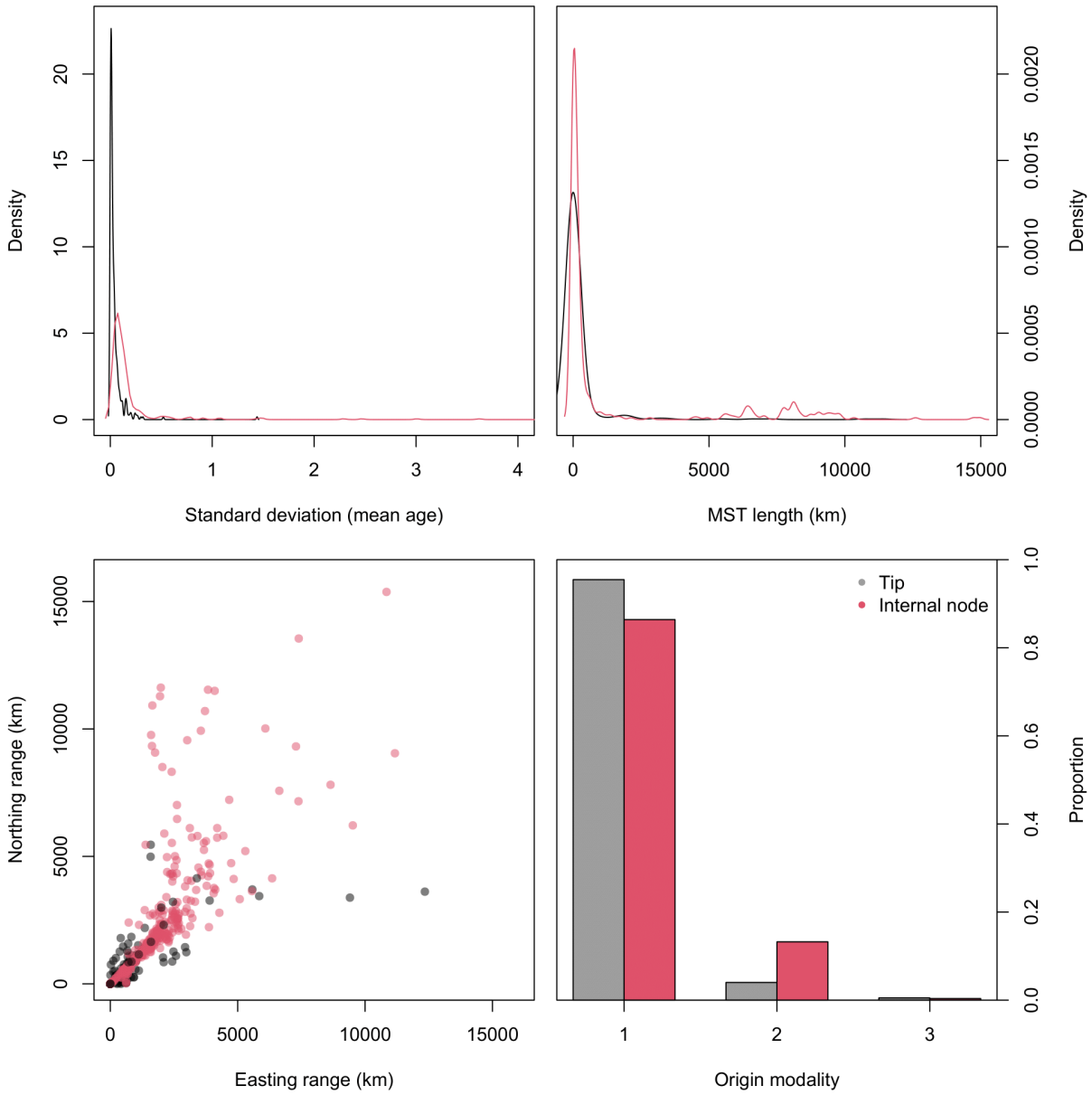


Fig. S95. Spatiotemporal uncertainty in archosauromorph origins. Generally, uncertainty is small, with low standard deviations between estimated node ages from the fossilised birth-death model, tight clustering of proposed longitude-latitude coordinates for each node, and unimodal positions estimated for most nodes.

Table S1. Class-wise range error proportions for Sepkoski Compendium genera

			Error proportion (%)	
Higher clade	Class	Genera (n)	Pre-revision	Post-revision
Porifera	Hexactinellids	336	37.2	22
	Calcarea	220	11.8	7.7
	Archaeocyatha	184	78.8	6
	Demospongea	114	77.2	43
	Regulares	75	1.3	0
	Irregulares	36	5.6	0
Mollusca	Gastropoda	3963	54.3	44.3
	Cephalopoda	3636	55.6	31.4
	Bivalvia	2801	52.6	40.7
	Polyplacophora	94	41.5	13.8
	Tergomya	57	35.1	28.1
	Rostroconchia	42	69	38.1
	Paragastropoda	34	58.8	29.4
	Scaphopoda	32	56.2	50
	Helcionelloida	4	100	75
Hemichordata	Pterobranchia	143	90.9	49.7
Echinodermata	Crinoidea	1122	41.9	18.3
	Echinoidea	751	34.5	23.3
	Asteroidea	145	44.8	24.1
	Holothuroidea	92	43.5	37
	Ophiuroidea	76	35.5	23.7
	Edrioasteroidea	65	50.8	13.8
	Stylophora	56	57.1	28.6
	Eocrinoidea	35	62.9	14.3
	Paracrinoidea	18	77.8	5.6
	Somasteroidea	8	37.5	0
	Ophiocistioidea	6	83.3	16.7
	Helicoplacoidea	3	66.7	0
	Ctenocystoidea	2	100	0
Cnidaria	Anthozoa	2151	53	34

	Hydrozoa	68	25	14.7
	Scyphozoa	51	31.4	15.7
	Medusae	27	25.9	0
	Hydroconozoa	8	12.5	0
Chordata	Actinopterygii	405	63.5	35.1
	Chondrichthyes	374	62	43.3
	Mammalia	302	55.6	33.4
	Reptilia	280	51.8	25.4
	Placodermi	163	8.6	2.5
	Aves	28	75	28.6
	Osteichthyes	5	80	20
	Petromyzontida	2	50	0
Bryozoa	Stenolaemata	897	35.3	21
	Gymnolaemata	641	29.6	27
Brachiopoda	Rhynchonellata	1180	76.7	48
	Lingulata	244	54.1	31.6
	Strophomenata	236	83.9	53
	Obolellata	7	85.7	28.6
	Paterinata	7	100	85.7
	Kutorginata	5	100	20
	Chileata	1	100	0
Arthropoda	Trilobita	2920	58.7	24.7
	Ostracoda	1877	32.8	22.4
	Thylacocephala	24	12.5	8.3
	Merostomoidea	11	9.1	9.1
	Xiphosura	10	70	30
	Arachnida	5	20	20
	Remipedia	2	50	0
	Megacheira	1	100	0
Annelida	Polychaeta	219	21	15.5

Table S2. Class-wise range error proportions for occurrence density methods

			Error proportion (%)	
Higher clade	Class	Genera (n)	Pacmacro	Interpeak
Protista	Rhizopodea	30	3.3	36.7
	Florideophyceae	63	9.5	47.6
Rhodophyta	Rhodophyceae	38	2.6	39.5
Porifera	Demospongea	697	2.6	21.7
	Hexactinellida	345	0.9	15.4
	Archaeocyatha	273	0.7	0
	Stromatoporoidea	120	7.5	29.2
	Calcarea	56	1.8	25
	Heteractinida	28	3.6	17.9
Plantae	Magnoliopsida	874	4	25.7
	Pteridopsida	201	13.9	47.3
	Pinopsida	83	7.2	42.2
	Cycadopsida	51	13.7	33.3
	Ginkgoopsida	47	14.9	42.6
	Peltaspermopsida	46	13	23.9
	Lycopsida	42	14.3	50
	Sphenopsida	31	25.8	32.3
	Cycadeoideopsida	29	13.8	48.3
	Psilophytopsida	20	10	10
	Pteropsida	20	5	20
	Gymnospermopsida	17	11.8	17.6
	Dictyopteridiopsida	11	27.3	27.3
	Equisetopsida	11	9.1	9.1
	Cladoxylopsida	9	0	11.1
	Arberiopsida	8	0	12.5
	Bryopsida	7	0	42.9
	Calamitopsida	6	0	16.7
	Marchantiopsida	4	25	25
	Archaeopteridopsida	3	33.3	0
	Hepaticae	3	33.3	66.7

	Progymnospermopsida	33.3	0	3
	Rhyniopsida	3	0	33.3
	Voltziopsida	3	0	33.3
	Gnetopsida	2	0	50
	Liliopsida	2	0	50
Ochrophyta	Bacillariophyceae	91	11	31.9
	Phaeophyceae	6	0	16.7
	Chrysophyceae	12	0	16.7
Nematoda	Enoplea	46	0	19.6
Mollusca	Gastropoda	4649	5	31.2
	Cephalopoda	3946	6.2	8.8
	Bivalvia	3076	7.4	26
	Polyplacophora	104	1	13.5
	Helcionelloida	91	4.4	12.1
	Tergomya	80	1.2	23.8
	Rostroconchia	63	3.2	34.9
	Scaphopoda	36	11.1	52.8
	Paragastropoda	31	0	9.7
Lophotrochozoa	Tentaculita	33	15.2	30.3
Hyolitha	Hyolithomorpha	14	0	14.3
	Orthothecimorpha	14	0	14.3
Hemichordata	Pterobranchia	221	12.7	11.3
Haptophyta	Coccolithophyceae	126	7.1	32.5
Foraminifera	Tubothalamea	143	4.2	31.5
	Nodosariata	57	7	59.6
	Globothalamea	38	5.3	42.1
Echinodermata	Crinoidea	916	1.7	19.4
	Echinoidea	554	3.2	25.1
	Asteroidea	206	0.5	21.8
	Holothuroidea	125	4.8	26.4
	Ophiuroidea	112	1.8	22.3
	Rhombifera	66	3	18.2
	Stylophora	55	0	20

	Edrioasteroidea	53	0	17
	Diploporita	46	0	10.9
	Eocrinoidea	46	0	15.2
	Stenuroidea	21	0	4.8
	Cyclocystoidea	20	0	20
	Cincta	17	5.9	0
	Paracrinoidea	17	0	23.5
	Ophiocistioidea	16	0	25
	Soluta	13	0	7.7
	Coronoidea	4	0	25
Cyanobacteria	Cyanophyceae	6	16.7	50
	Porostromata	4	25	75
Cnidaria	Anthozoa	1952	6.2	25.9
	Hydrozoa	46	4.3	39.1
	Scyphozoa	31	0	22.6
Chordata	Mammalia	5804	2.3	4
	Reptilia	2654	1.4	7.4
	Actinopterygii	1947	1.3	19.9
	Aves	1471	1	5.4
	Osteichthyes	1460	1.2	6.4
	Saurischia	874	1.7	5
	Chondrichthyes	710	4.9	27.7
	Ornithischia	407	2.5	4.9
	Conodonta	353	16.7	18.1
	Eufolivora	110	2.7	1.8
	Pteraspidomorpha	68	0	8.8
	Coelacanthimorpha	56	0	12.5
	Placodermi	47	0	6.4
	Galeaspida	25	0	8
	Craniata	35	5.7	40
	Agnatha	11	0	9.1
	Ascidiacea	4	0	50
	Cladistia	4	0	25

Chlorophyta	Dasycladophyceae	129	2.3	42.6
	Bryopsidophyceae	36	2.8	30.6
	Chlorophyceae	18	5.6	38.9
	Ulvophyceae	1	0	100
Charophyta	Charophyceae	31	0	32.3
Bryozoa	Stenolaemata	644	5.1	23
	Gymnolaemata	458	1.5	22.9
Brachiopoda	Rhynchonellata	2354	8.8	21.9
	Strophomenata	1053	10.1	23
	Lingulata	212	6.1	22.2
	Chileata	11	9.1	27.3
	Kutorginata	10	10	0
	Paterinata	9	11.1	33.3
Arthropoda	Insecta	13121	0.8	10.6
	Trilobita	2751	5.3	7.8
	Malacostraca	1206	2.2	18.3
	Ostracoda	1166	3.3	27
	Arachnida	1027	0.5	7
	Hexanauplia	97	3.1	27.8
	Myriapoda	85	1.2	12.9
	Collembola	59	0	10.2
	Xiphosura	54	1.9	14.8
	Branchiopoda	32	0	37.5
	Chilopoda	21	0	9.5
	Thylacocephala	19	0	15.8
	Symphyla	4	0	25
Annelida	Polychaeta	127	1.6	34.6
Dinoflagellata	Dinophyceae	146	5.5	39

Table S3. Pearson correlation for spatial extent and SQS diversity (Circumtethys)

	q = 0.4		q = 0.5		q = 0.6		q = 0.7	
Treatment + metric	Corr.	p	Corr.	p	Corr.	p	Corr.	p
Unstandardised								
MST length (km)	0.344	0.165	0.348	0.162	0.351	0.16	0.356	0.156
Lat. range (deg)	-0.241	0.749	-0.247	0.754	-0.258	0.764	-0.275	0.779
Lon. range (deg)	0605	0.032	0.601	0.033	0.592	0.036	0.575	0.041
MST-standardised								
MST length (km)	-0.605	0.968	-0.613	0.97	-0.619	0.972	-0.626	0.974
Lat. range (deg)	-0.272	0.776	-0.278	0.782	-0.292	0.793	-0.314	0.812
Lon. range (deg)	0.623	0.027	0.619	0.028	0.614	0.030	0.600	0.03
Lon-lat standardised								
MST length (km)	0.351	0.160	0.359	0.154	0.365	0.150	0.375	0.143
Lat. range (deg)	0.145	0.344	0.146	0.344	0.138	0.352	0.124	0.367
Lon. range (deg)	0.307	0.194	0.292	0.203	0.283	0.214	0.266	0.228
Both standardised								
MST length (km)	0.257	0.237	0.258	0.236	0.261	0.233	0.268	0.227
Lat. range (deg)	0.074	0.419	0.073	0.420	0.066	0.428	0.049	0.446
Lon. range (deg)	0.356	0.156	0.341	0.167	0.328	0.177	0.310	0.192

Table S4. Pearson correlation for spatial extent and SQS diversity (West Circumtethys)

	q = 0.4		q = 0.5		q = 0.6		q = 0.7	
Treatment + metric	Corr.	p	Corr.	p	Corr.	p	Corr.	p
Unstandardised								
MST length (km)	0.843	0.001	0.848	0.001	0.851	0.001	0.852	0.001
Lat. range (deg)	0.265	0.230	0.267	0.228	0.268	0.227	0.271	0.224
Lon. range (deg)	0.519	0.062	0.511	0.066	0.506	0.068	0.505	0.068
MST-standardised								
MST length (km)	0.237	0.255	0.250	0.243	0.253	0.241	0.243	0.249
Lat. range (deg)	0.260	0.234	0.262	0.232	0.262	0.232	0.264	0.230
Lon. range (deg)	0.517	0.063	0.511	0.066	0.503	0.069	0.500	0.070
Lon-lat standardised								
MST length (km)	0.690	0.014	0.700	0.012	0.708	0.011	0.708	0.011
Lat. range (deg)	0.226	0.265	0.228	0.264	0.227	0.264	0.228	0.264
Lon. range (deg)	0.263	0.231	0.261	0.233	0.259	0.235	0.263	0.232
Both standardised								
MST length (km)	0.302	0.198	0.295	0.204	0.291	0.207	0.288	0.210
Lat. range (deg)	0.226	0.266	0.226	0.265	0.225	0.266	0.225	0.266
Lon. range (deg)	0.262	0.232	0.259	0.235	0.257	0.237	0.260	0.234

Table S5. Pearson correlation for spatial extent and SQS diversity (East Circumtethys)

	q = 0.4		q = 0.5		q = 0.6		q = 0.7	
Treatment + metric	Corr.	p	Corr.	p	Corr.	p	Corr.	p
Unstandardised								
MST length (km)	0.089	0.404	0.085	0.408	0.077	0.416	0.077	0.417
Lat. range (deg)	-0.108	0.617	-0.126	0.636	-0.157	0.667	-0.192	0.702
Lon. range (deg)	-0.35	0.539	-0.033	0.536	-0.047	0.551	-0.069	0.576
MST-standardised								
MST length (km)	0.129	0.361	0.178	0.311	-0.196	0.293	0.199	0.291
Lat. range (deg)	-0.206	0.716	-0.236	0.745	-0.275	0.779	-0.318	0.815
Lon. range (deg)	-0.187	0.697	-0.184	0.695	-0.203	0.713	-0.231	0.739
Lon-lat standardised								
MST length (km)	0.137	0.353	0.130	0.360	0.121	0.369	-0.116	0.374
Lat. range (deg)	-0.047	0.552	-0.064	0.570	-0.097	0.605	-0.138	0.648
Lon. range (deg)	-0.013	0.515	-0.018	0.520	-0.033	0.537	-0.055	0.560
Both standardised								
MST length (km)	0.535	0.056	0.549	0.050	0.556	0.047	0.566	0.044
Lat. range (deg)	-0.125	0.636	-0.149	0.659	-0.189	0.699	-0.235	0.743
Lon. range (deg)	-0.259	0.765	-0.267	0.772	-0.286	0.788	-0.308	0.807

Table S6. Pearson correlation for spatial extent and SQS diversity (Boreal)

	q = 0.4		q = 0.5		q = 0.6		q = 0.7	
Treatment + metric	Corr.	p	Corr.	p	Corr.	p	Corr.	p
Unstandardised								
MST length (km)	0.688	0.014	0.692	0.013	0.700	0.012	0.716	0.010
Lat. range (deg)	-0.213	0.722	-0.212	0.712	-0.215	0.724	-0.214	0.724
Lon. range (deg)	-0.264	0.769	-0.266	0.771	-0.271	0.776	-0.275	0.779
MST-standardised								
MST length (km)	0.391	0.132	0.413	0.118	0.447	0.098	-0.507	0.068
Lat. range (deg)	-0.001	0.501	0.015	0.484	0.000	0.500	-0.043	0.547
Lon. range (deg)	-0.389	0.867	-0.387	0.865	-0.395	0.871	-0.382	0.862
Lon-lat standardised								
MST length (km)	0.742	0.007	0.743	0.007	0.752	0.006	0.765	0.005
Lat. range (deg)	0.036	0.461	0.033	0.464	0.036	0.460	0.038	0.458
Lon. range (deg)	0.334	0.173	0.308	0.193	0.282	0.215	0.257	0.237
Both standardised								
MST length (km)	0.604	0.032	0.578	0.040	0.581	0.584	0.584	0.038
Lat. range (deg)	0.003	0.496	0.000	0.500	0.008	0.008	0.008	0.491
Lon. range (deg)	0.253	0.241	0.238	0.254	0.213	0.181	0.181	0.309

Table S7. Pearson correlation for spatial extent and SQS diversity (North Panthalassic – P)

	q = 0.4		q = 0.5		q = 0.6		q = 0.7	
Treatment + metric	Corr.	p	Corr.	p	Corr.	p	Corr.	p
Unstandardised								
MST length (km)	0.565	0.056	0.580	0.051	0.600	0.044	0.616	0.039
Lat. range (deg)	0.481	0.095	0.490	0.090	0.504	0.083	0.510	0.080
Lon. range (deg)	-0.035	0.536	-0.022	0.522	-0.003	0.503	0.020	0.480
MST-standardised								
MST length (km)	0.196	0.307	0.203	0.300	0.219	0.285	0.236	0.271
Lat. range (deg)	0.444	0.115	0.457	0.108	0.473	0.100	0.486	0.092
Lon. range (deg)	-0.066	0.567	-0.050	0.550	-0.027	0.528	0.002	0.498
Lon-lat standardised								
MST length (km)	0.756	0.009	0.769	0.008	0.781	0.006	0.793	0.005
Lat. range (deg)	0.546	0.064	0.561	0.058	0.577	0.052	0.595	0.046
Lon. range (deg)	0.114	0.385	0.133	0.366	0.160	0.340	0.184	0.317
Both standardised								
MST length (km)	0.612	0.040	0.631	0.034	0.652	0.029	0.673	0.023
Lat. range (deg)	0.541	0.066	0.558	0.059	0.576	0.052	0.595	0.045
Lon. range (deg)	0.114	0.385	0.132	0.367	0.159	0.341	0.183	0.319

Table S8. Pearson correlation for spatial extent and SQS diversity (North Panthalassic)

	q = 0.4		q = 0.5		q = 0.6		q = 0.7	
Treatment + metric	Corr.	p	Corr.	p	Corr.	p	Corr.	p
Unstandardised								
MST length (km)	0.599	0.034	0.608	0.031	0.620	0.028	0.628	0.026
Lat. range (deg)	0.510	0.066	0.512	0.065	0.516	0.063	0.515	0.064
Lon. range (deg)	0.155	0.334	0.162	0.327	0.173	0.317	0.184	0.305
MST-standardised								
MST length (km)	0.342	0.167	0.343	0.166	0.348	0.162	0.352	0.159
Lat. range (deg)	0.490	0.075	0.494	0.073	0.498	0.071	0.500	0.070
Lon. range (deg)	0.137	0.353	0.146	0.344	0.158	0.332	0.174	0.315
Lon-lat standardised								
MST length (km)	0.756	0.006	0.767	0.005	0.777	0.004	0.786	0.004
Lat. range (deg)	0.506	0.068	0.515	0.064	0.523	0.060	0.530	0.058
Lon. range (deg)	0.242	0.251	0.251	0.242	0.264	0.231	0.272	0.224
Both standardised								
MST length (km)	0.618	0.028	0.634	0.025	0.650	0.021	0.665	0.018
Lat. range (deg)	0.502	0.07	0.512	0.065	0.522	0.061	0.530	0.057
Lon. range (deg)	0.240	0.252	0.250	0.243	0.263	0.232	0.271	0.224

Table S9. Pearson correlation for spatial extent and SQS diversity (Tangaroan)

	q = 0.4		q = 0.5		q = 0.6		q = 0.7	
Treatment + metric	Corr.	p	Corr.	p	Corr.	p	Corr.	p
Unstandardised								
MST length (km)	-0.247	0.754	-0.122	0.631	0.076	0.417	0.317	0.186
Lat. range (deg)	-0.079	0.585	-0.070	0.576	-0.024	0.527	0.035	0.462
Lon. range (deg)	0.514	0.064	0.504	0.069	0.427	0.109	0.274	0.222
MST-standardised								
MST length (km)	-0.321	0.817	-0.246	0.753	-0.123	0.632	0.039	0.458
Lat. range (deg)	-0.125	0.634	-0.123	0.632	-0.116	0.625	-0.082	0.589
Lon. range (deg)	0.512	0.065	0.493	0.074	0.414	0.117	0.265	0.230
Lon-lat standardised								
MST length (km)	-0.187	0.697	-0.053	0.558	0.177	0.313	0.410	0.120
Lat. range (deg)	0.472	0.084	0.492	0.074	0.484	0.078	0.416	0.116
Lon. range (deg)	0.585	0.038	0.622	0.028	0.570	0.043	0.442	0.100
Both standardised								
MST length (km)	-0.250	0.757	-0.147	0.658	0.042	0.455	0.285	0.212
Lat. range (deg)	0.479	0.081	0.490	0.075	0.474	0.083	0.373	0.145
Lon. range (deg)	0.580	0.039	0.626	0.026	0.596	0.035	0.464	0.089

Table S10. Pearson correlation for spatial extent and SQS diversity (South Panthalassic)

	q = 0.4		q = 0.5		q = 0.6		q = 0.7	
Treatment + metric	Corr.	p	Corr.	p	Corr.	p	Corr.	p
Unstandardised								
MST length (km)	-0.806	0.950	-0.792	0.945	-0.746	0.926	-0.651	0.883
Lat. range (deg)	-0.949	0.998	-0.947	0.998	-0.926	0.996	0.874	0.989
Lon. range (deg)	-0.924	0.996	-0.927	0.996	-0.905	0.993	0.849	0.984
MST-standardised								
MST length (km)	-0.944	0.992	-0.940	0.991	-0.920	0.986	-0.878	0.975
Lat. range (deg)	-0.950	0.998	-0.948	0.998	-0.936	0.997	-0.909	0.994
Lon. range (deg)	-0.938	0.997	-0.940	0.997	-0.929	0.996	-0.900	0.993
Lon-lat standardised								
MST length (km)	-0.633	0.874	-0.792	0.945	-0.861	0.970	-0.853	0.967
Lat. range (deg)	-0.576	0.884	-0.692	0.936	-0.756	0.959	-0.785	0.968
Lon. range (deg)	-0.583	0.888	-0.704	0.941	-0.769	0.963	-0.795	0.971
Both standardised								
MST length (km)	-0.467	0.786	-0.608	0.862	-0.681	0.897	-0.720	0.915
Lat. range (deg)	-0.568	0.880	-0.654	0.921	-0.699	0.939	-0.723	0.948
Lon. range (deg)	-0.583	0.888	-0.673	0.928	-0.719	0.946	-0.743	0.955

Table S11. Spearman correlation for spatial extent and SQS diversity (Circumtethys)

	q = 0.4		q = 0.5		q = 0.6		q = 0.7	
Treatment + metric	Corr.	p	Corr.	p	Corr.	p	Corr.	p
Unstandardised								
MST length (km)	0.139	0.354	0.139	0.354	0.139	0.354	0.224	0.268
Lat. range (deg)	-0.406	0.884	-0.406	0.884	-0.406	0.884	-0.370	0.861
Lon. range (deg)	0.600	0.037	0.600	0.037	0.600	0.037	0.552	0.052
MST-standardised								
MST length (km)	-0.576	0.960	-0.661	0.980	-0.661	0.989	0.685	0.984
Lat. range (deg)	-0.467	0.917	-0.539	0.948	-0.539	0.948	-0.552	0.952
Lon. range (deg)	0.709	0.014	0.721	0.012	0.721	0.012	0.697	0.016
Lon-lat standardised								
MST length (km)	0.200	0.292	0.127	0.366	0.248	0.246	0.309	0.194
Lat. range (deg)	0.091	0.406	0.164	0.328	0.152	0.341	0.079	0.419
Lon. range (deg)	0.188	0.304	0.285	0.214	0.176	0.316	0.115	0.379
Both standardised								
MST length (km)	-0.103	0.621	-0.079	0.594	-0.079	0.594	-0.079	0.594
Lat. range (deg)	-0.212	0.732	-0.248	0.765	-0.248	0.765	-0.248	0.765
Lon. range (deg)	0.382	0.139	0.358	0.156	0.358	0.156	0.358	0.156

Table S12. Spearman correlation for spatial extent and SQS diversity (West Circumtethys)

	q = 0.4		q = 0.5		q = 0.6		q = 0.7	
Treatment + metric	Corr.	p	Corr.	p	Corr.	p	Corr.	p
Unstandardised								
MST length (km)	0.770	0.007	0.758	0.008	0.758	0.008	0.758	0.008
Lat. range (deg)	0.152	0.341	0.188	0.304	0.188	0.304	0.188	0.304
Lon. range (deg)	0.442	0.102	0.358	0.156	0.358	0.156	0.358	0.156
MST-standardised								
MST length (km)	0.042	0.459	0.042	0.459	0.042	0.459	0.042	0.459
Lat. range (deg)	0.188	0.304	0.188	0.304	0.188	0.304	0.188	0.304
Lon. range (deg)	0.358	0.156	0.358	0.358	0.358	0.358	0.358	0.358
Lon-lat standardised								
MST length (km)	0.661	0.022	0.661	0.022	0.661	0.022	0.661	0.030
Lat. range (deg)	0.103	0.393	0.103	0.393	0.103	0.393	0.103	0.459
Lon. range (deg)	0.280	0.217	0.280	0.217	0.280	0.217	0.280	0.207
Both standardised								
MST length (km)	0.297	0.203	0.297	0.203	0.297	0.203	0.297	0.203
Lat. range (deg)	0.055	0.446	0.055	0.446	0.055	0.446	0.055	0.446
Lon. range (deg)	0.328	0.117	0.328	0.177	0.328	0.177	0.328	0.177

Table S13. Spearman correlation for spatial extent and SQS diversity (East Circumtethys)

	q = 0.4		q = 0.5		q = 0.6		q = 0.7	
Treatment + metric	Corr.	p	Corr.	p	Corr.	p	Corr.	p
Unstandardised								
MST length (km)	0.248	0.246	0.261	0.235	0.261	0.235	0.248	0.246
Lat. range (deg)	-0.176	0.696	-0.164	0.684	-0.164	0.684	-0.176	0.696
Lon. range (deg)	0.139	0.354	0.176	0.316	0.176	0.316	0.139	0.354
MST-standardised								
MST length (km)	0.018	0.486	0.042	0.459	0.030	0.473	0.018	0.486
Lat. range (deg)	-0.188	0.708	-0.176	0.696	-0.164	0.684	-0.224	0.743
Lon. range (deg)	0.030	0.473	0.055	0.446	0.030	0.473	-0.079	0.594
Lon-lat standardised								
MST length (km)	0.164	0.328	0.164	0.328	0.164	0.328	0.164	0.328
Lat. range (deg)	0.030	0.473	0.030	0.473	0.030	0.473	0.030	0.473
Lon. range (deg)	-0.067	0.581	-0.067	0.581	-0.067	0.581	-0.067	0.581
Both standardised								
MST length (km)	0.418	0.116	0.345	0.165	0.382	0.139	0.370	0.148
Lat. range (deg)	0.103	0.393	0.079	0.419	0.067	0.432	-0.030	0.541
Lon. range (deg)	-0.333	0.835	-0.479	0.923	-0.406	0.884	-0.382	0.869

Table S14. Spearman correlation for spatial extent and SQS diversity (Boreal)

	q = 0.4		q = 0.5		q = 0.6		q = 0.7	
Treatment + metric	Corr.	p	Corr.	p	Corr.	p	Corr.	p
Unstandardised								
MST length (km)	0.782	0.006	0.782	0.006	0.782	0.006	0.806	0.004
Lat. range (deg)	-0.152	0.672	-0.224	0.743	-0.236	0.754	-0.273	0.786
Lon. range (deg)	-0.467	0.917	0.418	0.891	-0.479	0.923	-0.394	0.876
MST-standardised								
MST length (km)	0.176	0.316	0.127	0.366	0.164	0.328	0.285	0.214
Lat. range (deg)	-0.152	0.672	-0.200	0.720	-0.212	0.732	-0.309	0.816
Lon. range (deg)	-0.394	0.876	-0.358	0.852	-0.479	0.923	-0.418	0.891
Lon-lat standardised								
MST length (km)	0.770	0.007	0.733	0.011	0.733	0.011	0.745	0.009
Lat. range (deg)	-0.079	0.594	-0.176	0.6996	-0.176	0.696	-0.212	0.732
Lon. range (deg)	0.467	0.089	0.418	0.116	0.418	0.116	0.406	0.124
Both standardised								
MST length (km)	0.321	0.184	0.455	0.095	-0.442	0.102	0.442	0.102
Lat. range (deg)	-0.079	0.594	-0.127	0.646	-0.139	0.659	-0.200	0.720
Lon. range (deg)	0.115	0.379	-0.006	0.514	0.018	0.486	0.006	0.500

Table S15. Spearman correlation for spatial extent and SQS diversity (North Panthalassic – P)

	q = 0.4		q = 0.5		q = 0.6		q = 0.7	
Treatment + metric	Corr.	p	Corr.	p	Corr.	p	Corr.	p
Unstandardised								
MST length (km)	0.250	0.260	0.250	0.260	0.417	0.135	0.333	0.193
Lat. range (deg)	0.383	0.156	0.383	0.156	0.533	0.074	0.433	0.125
Lon. range (deg)	-0.050	0.560	-0.050	0.560	-0.083	0.595	0.000	0.509
MST-standardised								
MST length (km)	-0.033	0.544	-0.033	0.544	-0.033	0.544	-0.033	0.544
Lat. range (deg)	0.383	0.156	0.383	0.156	0.383	0.156	0.383	0.156
Lon. range (deg)	-0.050	0.560	-0.050	0.560	-0.050	0.560	-0.050	0.560
Lon-lat standardised								
MST length (km)	0.600	0.048	0.600	0.048	0.600	0.048	0.600	0.048
Lat. range (deg)	0.550	0.066	0.550	0.066	0.550	0.066	0.550	0.066
Lon. range (deg)	-0.100	0.612	-0.100	0.612	-0.100	0.612	-0.100	0.612
Both standardised								
MST length (km)	0.600	0.048	0.600	0.048	0.600	0.048	0.600	0.048
Lat. range (deg)	0.550	0.066	0.550	0.066	0.550	0.066	0.550	0.066
Lon. range (deg)	-0.100	0.612	-0.100	0.612	-0.100	0.612	-0.100	0.612

Table S16. Spearman correlation for spatial extent and SQS diversity (North Panthalassic)

	q = 0.4		q = 0.5		q = 0.6		q = 0.7	
Treatment + metric	Corr.	p	Corr.	p	Corr.	p	Corr.	p
Unstandardised								
MST length (km)	0.455	0.095	0.455	0.095	0.576	0.044	0.515	0.066
Lat. range (deg)	0.552	0.052	0.552	0.052	0.661	0.022	0.588	0.040
Lon. range (deg)	0.236	0.257	0.236	0.257	0.212	0.280	0.273	0.224
MST-standardised								
MST length (km)	0.164	0.328	0.164	0.328	0.164	0.328	0.164	0.328
Lat. range (deg)	0.455	0.095	0.455	0.095	0.455	0.095	0.455	0.095
Lon. range (deg)	0.200	0.292	0.200	0.292	0.200	0.292	0.200	0.292
Lon-lat standardised								
MST length (km)	0.600	0.037	0.600	0.037	0.600	0.037	0.600	0.037
Lat. range (deg)	0.564	0.048	0.564	0.048	0.564	0.048	0.564	0.061
Lon. range (deg)	0.018	0.486	0.018	0.486	0.018	0.486	0.018	0.527
Both standardised								
MST length (km)	0.600	0.370	0.600	0.370	0.600	0.370	0.600	0.370
Lat. range (deg)	0.564	0.048	0.564	0.048	0.564	0.048	0.564	0.061
Lon. range (deg)	0.018	0.486	0.018	0.486	0.018	0.486	0.018	0.427

Table S17. Spearman correlation for spatial extent and SQS diversity (Tangaroan)

	q = 0.4		q = 0.5		q = 0.6		q = 0.7	
Treatment + metric	Corr.	p	Corr.	p	Corr.	p	Corr.	p
Unstandardised								
MST length (km)	-0.224	0.743	-0.200	0.720	0.067	0.432	0.370	0.148
Lat. range (deg)	-0.103	0.621	-0.139	0.659	-0.164	0.684	0.042	0.459
Lon. range (deg)	0.382	0.139	0.345	0.165	0.285	0.214	0.273	0.224
MST-standardised								
MST length (km)	-0.127	0.646	-0.127	0.646	0.127	0.366	0.152	0.341
Lat. range (deg)	0.200	0.720	-0.200	0.720	-0.236	0.754	-0.224	0.743
Lon. range (deg)	0.345	0.165	0.345	0.165	0.321	0.184	0.309	0.194
Lon-lat standardised								
MST length (km)	-0.165	0.684	-0.091	0.607	0.212	0.280	0.345	0.165
Lat. range (deg)	0.358	0.156	0.261	0.235	0.261	0.235	0.261	0.235
Lon. range (deg)	0.612	0.033	0.539	0.057	0.491	0.077	0.394	0.131
Both standardised								
MST length (km)	-0.164	0.684	-0.164	0.684	0.018	0.486	0.285	0.214
Lat. range (deg)	0.406	0.124	0.370	0.148	0.333	0.174	0.261	0.235
Lon. range (deg)	0.564	0.048	0.612	0.033	0.503	0.072	0.394	0.131

Table S18. Spearman correlation for spatial extent and SQS diversity (South Panthalassic)

	q = 0.4		q = 0.5		q = 0.6		q = 0.7	
Treatment + metric	Corr.	p	Corr.	p	Corr.	p	Corr.	p
Unstandardised								
MST length (km)	-0.900	0.992	-0.900	0.992	-0.900	0.992	-0.900	0.992
Lat. range (deg)	-0.886	0.992	-0.886	0.992	-0.886	0.992	-0.886	0.992
Lon. range (deg)	-1.000	1.000	-1.000	1.000	-1.000	1.000	-1.000	1.000
MST-standardised								
MST length (km)	-1.000	1.000	-1.000	1.000	-1.000	1.000	-1.000	1.000
Lat. range (deg)	-0.886	0.992	-0.886	0.992	-0.886	0.992	-0.886	0.992
Lon. range (deg)	-1.000	1.000	-1.000	1.000	-1.000	1.000	-1.000	1.000
Lon-lat standardised								
MST length (km)	-0.400	0.775	-0.700	0.933	-0.900	0.992	-0.900	0.992
Lat. range (deg)	-0.314	0.751	-0.486	0.851	-0.714	0.949	-0.714	0.949
Lon. range (deg)	-0.429	0.822	-0.657	0.932	-0.829	0.983	-0.829	0.983
Both standardised								
MST length (km)	0.000	0.525	-0.400	0.775	-0.700	0.933	-0.700	0.933
Lat. range (deg)	-0.314	0.751	-0.486	0.851	-0.714	0.949	-0.714	0.949
Lon. range (deg)	-0.429	0.822	-0.657	0.932	-0.829	0.983	-0.829	0.983

Table S19. Predictive performance of z-coral ENM hindcasts (mSSS binarized)

			Getech, HadCM3L		PALEOMAP, HadCM3L		PALEOMAP, CLIMBER-3α	
Interval	Hits	Misses	v	p	v	p	v	p
Wuchiapingian	-	-	-	-	-	-	-	-
Changhsingian	-	-	-	-	-	-	-	-
Induan	-	-	-	-	-	-	-	-
Olenekian	-	-	-	-	-	-	-	-
Anisian	9	0	0	> 0.001	0	> 0.001	0	> 0.001
Ladinian	7	5	0	> 0.001	0	> 0.001	0	> 0.001
Early Carnian	3	4	0	> 0.001	0	> 0.001	0	> 0.001
Late Carnian	10	4	0	> 0.001	0	> 0.001	0	> 0.001
Early Norian	9	7	0	> 0.001	0	> 0.001	0	> 0.001
Middle Norian	4	1	0	> 0.001	0	> 0.001	0	> 0.001
Late Norian	13	8	0	> 0.001	0	> 0.001	0	> 0.001
Rhaetian	9	17	0	> 0.001	0	> 0.001	0	> 0.001
Hettangian	11	1	0	> 0.001	0	> 0.001	0	> 0.001
Sinemurian	13	6	0	> 0.001	0	> 0.001	0	> 0.001

Test statistic and p values reported for the Wilcoxon signed ranks test

Table S20. Predictive performance of z-coral ENM hindcasts (L95P binarized)

			Getech, HadCM3L		PALEOMAP, HadCM3L		PALEOMAP, CLIMBER-3α	
Interval	Hits	Misses	v	p	v	p	v	p
Wuchiapingian	-	-	-	-	-	-	-	-
Changhsingian	-	-	-	-	-	-	-	-
Induan	-	-	-	-	-	-	-	-
Olenekian	-	-	-	-	-	-	-	-
Anisian	9	0	0	> 0.001	0	> 0.001	0	> 0.001
Ladinian	7	5	0	> 0.001	0	> 0.001	0	> 0.001
Early Carnian	3	4	0	> 0.001	0	> 0.001	0	> 0.001
Late Carnian	10	4	0	> 0.001	0	> 0.001	0	> 0.001
Early Norian	9	7	0	> 0.001	0	> 0.001	0	> 0.001
Middle Norian	4	1	0	> 0.001	0	> 0.001	0	> 0.001
Late Norian	13	8	0	> 0.001	0	> 0.001	0	> 0.001
Rhaetian	9	17	0	> 0.001	0	> 0.001	0	> 0.001
Hettangian	11	1	0	> 0.001	0	> 0.001	0	> 0.001
Sinemurian	13	6	0	> 0.001	0	> 0.001	0	> 0.001

Test statistic and p values reported for the Wilcoxon signed ranks test

Table S21. Predictive performance of az-coral ENM hindcasts (mSSS binarized)

			Getech, HadCM3L		PALEOMAP, HadCM3L		PALEOMAP, CLIMBER-3α	
Interval	Hits	Misses	v	p	v	p	v	p
Wuchiapingian	-	-	-	-	-	-	-	-
Changhsingian	-	-	-	-	-	-	-	-
Induan	-	-	-	-	-	-	-	-
Olenekian	-	-	-	-	-	-	-	-
Anisian	6	0	0	> 0.001	0	> 0.001	0	> 0.001
Ladinian	3	3	0	> 0.001	0	> 0.001	0	> 0.001
Early Carnian	4	0	0	> 0.001	0	> 0.001	0	> 0.001
Late Carnian	10	0	0	> 0.001	0	> 0.001	0	> 0.001
Early Norian	5	4	0	> 0.001	0	> 0.001	0	> 0.001
Middle Norian	2	1	0	> 0.001	471.5	> 0.001	0	> 0.001
Late Norian	2	2	0	> 0.001	0	> 0.001	461.5	> 0.001
Rhaetian	6	3	0	> 0.001	0	> 0.001	61.5	> 0.001
Hettangian	8	0	0	> 0.001	0	> 0.001	0	> 0.001
Sinemurian	5	5	0	> 0.001	0	> 0.001	0	> 0.001

Test statistic and p values reported for the Wilcoxon signed ranks test

Table S22. Predictive performance of az-coral ENM hindcasts (L95P binarized)

			Getech, HadCM3L		PALEOMAP, HadCM3L		PALEOMAP, CLIMBER-3α	
Interval	Hits	Misses	v	p	v	p	v	p
Wuchiapingian	-	-	-	-	-	-	-	-
Changhsingian	-	-	-	-	-	-	-	-
Induan	-	-	-	-	-	-	-	-
Olenekian	-	-	-	-	-	-	-	-
Anisian	6	0	0	> 0.001	0	> 0.001	0	> 0.001
Ladinian	3	3	1	> 0.001	0	> 0.001	0	> 0.001
Early Carnian	4	0	0	> 0.001	0	> 0.001	0	> 0.001
Late Carnian	10	0	0	> 0.001	0	> 0.001	0	> 0.001
Early Norian	5	4	0	> 0.001	0	> 0.001	0	> 0.001
Middle Norian	2	1	0	> 0.001	0	> 0.001	0	> 0.001
Late Norian	2	2	0	> 0.001	932	> 0.001	2327.5	> 0.001
Rhaetian	6	3	0	> 0.001	0	> 0.001	0	> 0.001
Hettangian	8	0	0	> 0.001	0	> 0.001	0	> 0.001
Sinemurian	5	5	0	> 0.001	0	> 0.001	0	> 0.001

Test statistic and p values reported for the Wilcoxon signed ranks test

Table S23. Predictive performance of coral reef ENM hindcasts (mSSS binarized)

			Getech, HadCM3L		PALEOMAP, HadCM3L		PALEOMAP, CLIMBER-3α	
Interval	Hits	Misses	v	p	v	p	v	p
Wuchiapingian	7	3	0	> 0.001	0	> 0.001	0	> 0.001
Changhsingian	9	3	0	> 0.001	0	> 0.001	0	> 0.001
Induan	-	-	-	-	-	-	-	-
Olenekian	-	-	-	-	-	-	-	-
Anisian	11	0	0	> 0.001	0	> 0.001	0	> 0.001
Ladinian	8	4	0	> 0.001	0	> 0.001	0	> 0.001
Early Carnian	4	8	0	> 0.001	0	> 0.001	0	> 0.001
Late Carnian	6	6	0	> 0.001	0	> 0.001	0	> 0.001
Early Norian	10	15	0	> 0.001	0	> 0.001	0	> 0.001
Middle Norian	2	2	0	> 0.001	969	> 0.001	0	> 0.001
Late Norian	8	18	0	> 0.001	0	> 0.001	0	> 0.001
Rhaetian	11	14	0	> 0.001	0	> 0.001	0	> 0.001
Hettangian	3	1	0	> 0.001	0	> 0.001	0	> 0.001
Sinemurian	5	2	0	> 0.001	0	> 0.001	0	> 0.001

Test statistic and p values reported for the Wilcoxon signed ranks test

Table S24. Predictive performance of coral reef ENM hindcasts (L95P binarized)

			Getech, HadCM3L		PALEOMAP, HadCM3L		PALEOMAP, CLIMBER-3α	
Interval	Hits	Misses	v	p	v	p	v	p
Wuchiapingian	7	3	0	> 0.001	0	> 0.001	0	> 0.001
Changhsingian	9	3	0	> 0.001	0	> 0.001	0	> 0.001
Induan	-	-	-	-	-	-	-	-
Olenekian	-	-	-	-	-	-	-	-
Anisian	11	0	0	> 0.001	0	> 0.001	0	> 0.001
Ladinian	8	4	0	> 0.001	0	> 0.001	0	> 0.001
Early Carnian	4	8	0	> 0.001	0	> 0.001	0	> 0.001
Late Carnian	6	6	0	> 0.001	0	> 0.001	0	> 0.001
Early Norian	10	15	0	> 0.001	0	> 0.001	0	> 0.001
Middle Norian	2	2	0	> 0.001	481.5	> 0.001	0	> 0.001
Late Norian	8	18	0	> 0.001	0	> 0.001	0	> 0.001
Rhaetian	11	14	0	> 0.001	0	> 0.001	0	> 0.001
Hettangian	3	1	0	> 0.001	0	> 0.001	0	> 0.001
Sinemurian	5	2	0	> 0.001	0	> 0.001	0	> 0.001

Test statistic and p values reported for the Wilcoxon signed ranks test

Table S25. Predictive performance of sponge ENM hindcasts (mSSS binarized)

			Getech, HadCM3L		PALEOMAP, HadCM3L		PALEOMAP, CLIMBER-3α	
Interval	Hits	Misses	v	p	v	p	v	p
Wuchiapingian	21	6	0	> 0.001	0	> 0.001	0	> 0.001
Changhsingian	23	1	0	> 0.001	0	> 0.001	0	> 0.001
Induan	1	0	0	> 0.001	0	> 0.001	0	> 0.001
Olenekian	11	0	0	> 0.001	0	> 0.001	0	> 0.001
Anisian	14	1	0	> 0.001	0	> 0.001	0	> 0.001
Ladinian	20	3	0	> 0.001	0	> 0.001	0	> 0.001
Early Carnian	14	3	0	> 0.001	0	> 0.001	0	> 0.001
Late Carnian	11	4	0	> 0.001	0	> 0.001	0	> 0.001
Early Norian	15	14	0	> 0.001	0	> 0.001	0	> 0.001
Middle Norian	4	1	0	> 0.001	0	> 0.001	0	> 0.001
Late Norian	18	14	0	> 0.001	0	> 0.001	0	> 0.001
Rhaetian	14	12	0	> 0.001	0	> 0.001	0	> 0.001
Hettangian	2	0	0	> 0.001	0	> 0.001	0	> 0.001
Sinemurian	9	0	0	> 0.001	0	> 0.001	0	> 0.001

Test statistic and p values reported for the Wilcoxon signed ranks test

Table S26. Predictive performance of coral reef ENM hindcasts (L95P binarized)

			Getech, HadCM3L		PALEOMAP, HadCM3L		PALEOMAP, CLIMBER-3α	
Interval	Hits	Misses	v	p	v	p	v	p
Wuchiapingian	21	6	0	> 0.001	0	> 0.001	0	> 0.001
Changhsingian	22	1	0	> 0.001	0	> 0.001	0	> 0.001
Induan	1	0	0	> 0.001	0	> 0.001	0	> 0.001
Olenekian	11	0	0	> 0.001	0	> 0.001	0	> 0.001
Anisian	14	1	0	> 0.001	0	> 0.001	0	> 0.001
Ladinian	20	3	0	> 0.001	0	> 0.001	0	> 0.001
Early Carnian	14	3	0	> 0.001	0	> 0.001	0	> 0.001
Late Carnian	11	4	0	> 0.001	0	> 0.001	0	> 0.001
Early Norian	15	14	0	> 0.001	0	> 0.001	0	> 0.001
Middle Norian	4	1	0	> 0.001	0	> 0.001	0	> 0.001
Late Norian	18	14	0	> 0.001	0	> 0.001	0	> 0.001
Rhaetian	14	12	0	> 0.001	0	> 0.001	0	> 0.001
Hettangian	2	0	0	> 0.001	0	> 0.001	0	> 0.001
Sinemurian	9	0	0	> 0.001	0	> 0.001	0	> 0.001

Test statistic and p values reported for the Wilcoxon signed ranks test