

This file is part of the following work:

**Siqueira Correa, Alexandre (2020) *The evolution, macroecology and biogeography of coral reef fishes: a trophic perspective*. PhD Thesis, James Cook University.**

Access to this file is available from:

<https://doi.org/10.25903/6fqw%2Dqd46>

Copyright © 2020 Alexandre Siqueira Correa.

The author has certified to JCU that they have made a reasonable effort to gain permission and acknowledge the owners of any third party copyright material included in this document. If you believe that this is not the case, please email

[researchonline@jcu.edu.au](mailto:researchonline@jcu.edu.au)

# The evolution, macroecology and biogeography of coral reef fishes: a trophic perspective

Thesis submitted by  
Alexandre Siqueira Correa, MSc

for the degree of Doctor of Philosophy (PhD)  
November 2020



ARC Centre of Excellence for Coral Reef Studies,  
James Cook University  
Townsville, Australia

## Acknowledgements

I am immensely grateful for the guidance provided by my two supervisors, Peter Cowman and David Bellwood. It is nice to think back and remember that this thesis sprouted from ideas discussed at Guarujá beach during the Evolution Meeting of 2015, which have now flourished into a body of work. From that moment on, I have been fortunate enough to count on two exceptional scientists that provided much-needed support throughout my PhD. Needless to say, it was not only about intellectual guidance (which was obviously fundamental), but also about showing me how to navigate the intricate path of science in a fun way. Thank you, Pete and Dave, for your enthusiasm, generosity, expertise, and for the opportunities and insights you provided along the way. I look forward to continuing this fruitful collaboration in the future!

Many of the ideas developed here represent the culmination of a long learning trajectory that started during my Masters back in Brazil. I am indebted to my mentor Sergio Floeter who first introduced me into the world of phylogenetics and biogeography. Without your enthusiasm and encouragement, I would probably not be where I am today, thank you for opening this door. I also thank Cadu Ferreira for encouraging me as an undergrad to work with reef fishes throughout the many editions of the 'Curso de Peixes Recifais'. In one of those editions, I had the pleasure to meet Howard Choat, whose vast knowledge about reef fish ecology and evolution has always been a source of inspiration. Thank you, Howard, for being so generous with your time and for sharing important insights about my work during productive coffee break discussions.

Indeed, the PhD was quite a journey, and I was extremely lucky to have had the opportunity to be joined on this trip by an amazing group of people that made everything a lot easier. I am grateful to the Bellwood Lab family: Orpha Bellwood, Chris Hemingson, Mike Mihalitsis, Renato Morais, Rob Streit, Victor Huertas, Sterling Tebbett, Arun Oakley-Cogan, François Latrille, Casey Bowden, Will Collins, Helen Yan, Jodie Schlaefer, Jess Valenzuela and Juliano Morais. I hope we can keep sharing adventures and the taste for good coffee for a long time together. To my dear 'bicho' friends, Pauline

Narvaez and Renato Morais, I am thankful for your support since the moment I arrived in Townsville. I really appreciate the opportunities of joining you on Lizard Island trips, which were essential to remind me why I love what I do. I also thank Ruby Grantham for the companionship, care and dedication, which were truly important on this ride along the bumpy PhD road. If it wasn't for your relentless effort in having a good work-life balance (and taking me with you, despite my silly resistance), I would have probably drowned in work and not maintained my sanity.

The JCU and Centre of Excellence for Coral Reef Studies staff were also essential in providing the backstage assistance during my candidature, particularly Alana Grech, David Bourne, Liz Tynan, Murray Logan, Janet Swanson, Vivian Doherty, Jennifer Lappin, Olga Bazaka, Greg Suosaari, Glenn Ewels, Claire Meade and Andrew Norton. I really appreciate your help.

Last but not least, I must thank those who form the foundation of my life and without whom there would be no 'me' nor PhD: family and friends. From BH to Floripa and now Townsville, I am grateful for the friendship of many who help make my world a better place to live with their smiles and helping hands. And finally, I am forever grateful for the support of my family who showed me that no matter how far I go, there will always be a safe port back home. Muito obrigado, pai, mãe e irmãos por todo o apoio durante essa longa caminhada! The geographical distance that separates us only brings us closer in thoughts, thank you for making this dream possible.

## Statement of the Contribution of Others

This thesis was conducted under the supervision of Peter Cowman and David Bellwood. I was responsible for conceiving and designing the research projects, collecting and curating data, developing figures and tables, performing statistical analyses, and writing and editing manuscripts. My supervisors provided intellectual guidance in the conception and throughout the implementation of research projects, technical and editorial advice, as well as financial support. In **chapters 1 and 2**, I also benefited from the contribution of Renato Morais in terms of data collection and curation, statistical guidance and editorial assistance.

Financial support for this thesis was provided by funding from the ARC Centre of Excellence for Coral Reef Studies and the College of Science and Engineering at James Cook University. Stipend and tuition support was provided by a Postgraduate Research Scholarship from James Cook University. Additional research funding was acquired through a David Yellowlees Excellence in Research Award from the ARC Centre of Excellence for Coral Reef Studies. Travel to conferences was supported by grants from my supervisors Peter Cowman and David Bellwood, and additional support was provided by a John Glover Travel Fund Award by the Australian Society for Fish Biology.

Every reasonable effort has been made to gain permission and acknowledge the owners of copyright material. I would be pleased to hear from any copyright owner who has been omitted or incorrectly acknowledged.

## Abstract

Biotic interactions are known to shape local species assemblages, but can potentially scale-up to determine larger spatial and temporal patterns in the distribution of species. Although we recognize a myriad of interactions between species (e.g. predation, competition and mutualisms), they are all united by a common link: the energetic demands of interacting elements. The nature of biotic interactions depends on the dynamics of energy acquisition and transfer, which is ultimately determined by the trophic identity of interacting species. This identity provides a proxy for a wide range of morphological, physiological and behavioural traits, which represent the raw material for natural selection and evolution. The effects of biotic interactions, and more specifically the trophic identity of species, therefore, is likely to be crucial in understanding the origins and maintenance of high-diversity systems.

Coral reefs stand out amongst high-diversity systems for harbouring the vast majority of marine species, despite occupying only a small fraction of the global oceans. Within these biodiversity cradles, fishes represent one of the most important energetic conduits, taking part in a large proportion of the recognized biotic interactions. Although reef-associated fishes span a broad range of dietary strategies, the evolutionary processes that determined their diversity among trophic guilds have thus far remained unclear. In this thesis, therefore, I aimed to understand the macroevolutionary, macroecological and biogeographical processes underpinning the present-day distribution of coral reef fishes across trophic guilds. By combining information from fossils, molecular phylogenies, geographical distributions and ecology, I explored the trophic evolution of coral reef fishes from multiple viewpoints.

This thesis is structured in four data chapters (2 to 5) that follow a hierarchical order of taxonomic comprehensiveness. In **chapter 2**, I gathered ecological and geographical data on over six thousand species of reef-associated fishes and asked: what factors dictated the pace of their evolution? To answer this question, I built near-complete reef fish phylogenies and estimated lineage-

specific diversification rates. I then applied gradient boosting techniques to assess the most important variables for predicting diversification patterns. Surprisingly, species trophic identity and body size were the only variables found to have higher importance in predicting diversification rates than expected by chance, with trophic guild having almost double the importance of body size. From the six classified guilds, herbivores/detritivores stood out when compared to other groups. Not only were they found to have diversified faster, but this effect was shown to be more pronounced in large-bodied species. After a more in-depth analysis of trophic evolution, I also found that transitions to planktivory are common throughout reef fish evolution, while omnivory seems to represent a transient state between high and low trophic levels. These results indicate that trophic evolution was one of the main determinants of lineage diversification in reef-associated fishes.

After finding this strong trophic signal in reef fish evolution, in **chapter 3**, I investigated whether present-day diversity distribution patterns display a similar imprint. To do this, I used data on species geographical ranges, along with an underwater fish survey dataset, to quantify species richness patterns per trophic guild across the globe. By focusing on thirteen consensus coral reef fish families (approximately 3600 species), I found that planktivores are far more diverse within the Indo-Australian Archipelago marine biodiversity hotspot – disproportionately so when compared to other trophic groups. These results were extremely consistent across spatial scales and highlight a unique trophic link in the global distribution of coral reef fish species. Interestingly, the evolutionary signal behind this accumulation of planktivorous fish species within the global biodiversity hotspot could not be detected from molecular phylogenies. This suggests that ecological and environmental factors, mediated through differential extinctions, are the likely drivers of the pattern. This chapter therefore highlights the important role of trophic ecology in shaping global species richness gradients.

In **chapters 4 and 5**, I narrowed down the taxonomic scope to focus specifically on herbivorous coral reef fishes and describe their associated large-scale biogeographical patterns. In **chapter 4**, I applied a comparative framework to trace the evolution of ecological traits and ecosystem functions

in the three common herbivorous groups on coral reefs: surgeonfishes, rabbitfishes and parrotfishes. By comparing evolutionary patterns between two major oceanic basins, I found that the present-day herbivorous fish trait composition in the Indo-Pacific results from a temporal expansion of the Tethyan fossil assemblage. The Atlantic, on the other hand, was found to carry the imprint of past extinction events that likely shaped the modern trait composition of herbivorous fishes therein. This history of extinction is also reflected in the ecosystem functions displayed by the herbivorous fish assemblages. The Atlantic harbours far fewer lineages performing each function when compared to the Indo-Pacific. These results highlight the role of history as a critical driver of the global functional composition of herbivorous coral reef fishes.

Finally, in **chapter 5**, I used modern phylogenetic comparative methods to describe the historical biogeography of the same herbivorous groups analysed in the previous chapter. This time, however, I built phylogenetic trees that included both living and fossil species to assess scenarios of range evolution from the early origins of the groups. This framework revealed that the Palaeocene–Eocene was an important period for lineage origination in the central Tethys region. However, it was only during the Miocene that extant herbivorous fish genera originated, deriving from surviving Tethyan lineages. I found that these lineages expanded within the Indo-Pacific, while the Atlantic remained isolated and went through extinctions. As a result, four out of five Atlantic herbivorous lineages were found to have Indo-Pacific origins, while only one endemic clade retains a Tethyan origin. This chapter points to the prominence of extinctions in structuring the herbivorous coral reef fish assemblages, particularly in the Atlantic Ocean.

Overall, this thesis demonstrates that the global composition of extant coral reef fishes results largely from the interplay between historical contingencies and the evolution of novel trophic strategies. It thus bridges two fundamental fields in biology: ecology and evolution. This evolutionary perspective focused on species roles in ecosystems, rather than a purely taxonomic view, offers an exciting future research avenue, particularly in high-diversity systems such as coral reefs.

# Table of Contents

<b>Acknowledgements</b>	<b>i</b>
<b>Statement of the Contribution of Others</b>	<b>iii</b>
<b>Abstract</b>	<b>iv</b>
<b>Table of Contents</b>	<b>vii</b>
<b>List of Figures</b>	<b>x</b>
<b>Chapter 1. General Introduction</b>	<b>1</b>
1.1 Macroevolution and macroecology in a trophic context	1
1.2 The evolution of fishes on coral reefs	3
1.3 Thesis aims and outline	6
<b>Chapter 2. Evolution of trophic guilds in coral reef fishes</b>	<b>8</b>
2.1 Abstract	8
2.2 Introduction	9
2.3 Methods	11
2.3.1 Reef fish phylogeny	11
2.3.2 Diversification rates	12
2.3.3 Explanatory variables	13
2.3.4 Predicting diversification rates	15
2.3.5 Trait-dependent diversification	17
2.3.6 Trophic transitions	19
2.4 Results	19
2.4.1 Diversification rate heterogeneity	19
2.4.2 Predictors of reef fish diversification	20
2.4.3 Historical patterns of trophic evolution	23
2.5 Discussion	26
2.5.1 Drivers of reef fish diversification	26
2.5.2 Trophic transitions	31
2.5.3 Model Considerations	32
2.5.4 Conclusions	33
<b>Chapter 3. Macroecology of trophic guilds in coral reef fishes</b>	<b>34</b>
3.1 Abstract	34
3.2 Introduction	35
3.3 Methods	38
3.3.1 Species distribution and survey data	38
3.3.2 Species trait data	39
3.3.3 Statistical analyses	40
3.3.4 Phylogenetic comparative methods	41
3.4 Results	42
	vii

3.5 Discussion	49
3.5.1 Geological and oceanographic drivers	50
3.5.2 Ecological drivers	51
<b>Chapter 4. Evolution of traits and functions in herbivorous coral reef fishes</b>	<b>54</b>
4.1 Abstract	54
4.2 Introduction	55
4.3 Methods	57
4.3.1 Chronograms	57
4.3.2 Herbivorous reef fish data	58
4.3.3 Ancestral states & biogeography	59
4.3.4 Multidimensional trait space	61
4.4 Results	62
4.5 Discussion	66
4.5.1 Trait combinations in space and time	67
4.5.2 Ecosystem functions in space and time	70
4.5.3 Caveats	72
4.5.4 Conclusions	72
<b>Chapter 5. Historical biogeography of herbivorous coral reef fishes</b>	<b>73</b>
5.1 Abstract	73
5.2 Introduction	74
5.3 Methods	76
5.3.1 Phylogenetic inferences	76
5.3.2 Historical biogeography	79
5.3.3 Biogeographical stochastic mappings	81
5.4 Results	81
5.5 Discussion	88
5.5.1 Origins of biogeographical patterns	89
5.5.2 Colonization of the Atlantic Ocean	91
5.5.3 Model considerations	94
5.5.4 Conclusions	95
<b>Chapter 6. Concluding Discussion</b>	<b>96</b>
6.1 The trophic component of coral reef fish evolution and macroecology	96
6.2 Historical contingencies in the evolution of coral reef fishes	98
6.3 Implications and future avenues	101
<b>References</b>	<b>103</b>
<b>Appendix A. Supplementary Material to Chapter 2</b>	<b>120</b>
Supplementary Figures	120
Supplementary Tables	126
<b>Appendix B. Supplementary Material to Chapter 3</b>	<b>127</b>
Supplementary Figures	127

<b>Appendix C. Supplementary Material to Chapter 4</b>	<b>132</b>
Supplementary Methods	132
<i>Phylogenetic inferences</i>	132
<i>Herbivorous reef fish data</i>	133
<i>Ancestral range estimation</i>	135
<i>Uncertainties in ancestral state reconstructions</i>	136
Supplementary References	137
Supplementary Figures	143
Supplementary Tables	163
<b>Appendix D. Supplementary Material to Chapter 5</b>	<b>177</b>
Supplementary Figures	177
Supplementary Tables	180
Supplementary References	183
<b>Appendix E. Publications during candidature</b>	<b>185</b>
Publications arising from thesis chapters	185
Other peer-reviewed papers published during candidature	186

## List of Figures

	<b>Page</b>
<b>Figure 2.1  </b> Near-complete reef fish phylogeny mapped with net diversification rates.	<b>20</b>
<b>Figure 2.2  </b> Ecological and geographical factors driving reef fish tip diversification patterns.	<b>21</b>
<b>Figure 2.3  </b> Tip diversification rates predicted for reef fish trophic groups while varying body size.	<b>22</b>
<b>Figure 2.4  </b> Historical net diversification rate estimates for six reef fish trophic groups.	<b>24</b>
<b>Figure 2.5  </b> Directionality of transitions out of each reef fish trophic group.	<b>25</b>
<b>Figure 3.1  </b> Global coral reef fish richness per trophic group.	<b>43</b>
<b>Figure 3.2  </b> Global coral reef fish species richness (A) and proportion (B) per guild with distance from the Indo-Australian Archipelago (IAA).	<b>45</b>
<b>Figure 3.3  </b> Coral reef fish species richness at the regional scale per trophic group.	<b>47</b>
<b>Figure 3.4  </b> Potential evolutionary mechanisms underpinning the richness gradient in coral reef fish planktivores.	<b>48</b>
<b>Figure 4.1  </b> Multidimensional trait space occupied by surgeonfish (circles), rabbitfish (squares) and parrotfish (triangles) lineages in two biogeographical regions through time.	<b>63</b>
<b>Figure 4.2  </b> Multidimensional trait space occupied by extant surgeonfish (circles), rabbitfish (squares) and parrotfish (triangles) species in two biogeographical regions.	<b>64</b>
<b>Figure 4.3  </b> Surgeonfish, rabbitfish and parrotfish lineages performing each ecosystem function through time in the Indo-Pacific (a) and Atlantic (b) realms.	<b>65</b>
<b>Figure 5.1  </b> Maximum clade credibility phylogeny of Acanthuridae (surgeonfishes) derived from the fossilized birth-death model.	<b>83</b>
<b>Figure 5.2  </b> Maximum clade credibility phylogeny of Siganidae (rabbitfishes) derived from the fossilized birth-death model.	<b>84</b>
<b>Figure 5.3  </b> Maximum clade credibility phylogeny of Scarini (parrotfishes) derived from the fossilized birth-death model.	<b>86</b>
<b>Figure 5.4  </b> Routes of colonization of the Atlantic by surgeonfishes and parrotfishes.	<b>87</b>

## Chapter 1. General Introduction

### 1.1 Macroevolution and macroecology in a trophic context

Uncovering how biodiversity emerges at different spatial and temporal scales has long been a major aim in evolutionary research. Since the recognition of global patterns in the distribution of diversity in the nineteenth century, naturalists have highlighted macroevolutionary processes as the fundamental drivers of large-scale biotic patterns (e.g. Darwin, 1859; Wallace, 1878). The higher species diversity in tropical habitats is possibly the most pervasive of these biotic patterns (Willig *et al.*, 2003; Hillebrand, 2004). However, species diversity varies unevenly, with many other geographical factors, such as area and elevation, exerting an influence (Gaston, 2000). According to long-standing theories, these global patterns in biological diversity represent a snapshot of past species origination, extinction, dispersal and adaptation throughout evolution (Brooks & McLennan, 1991; Jablonski *et al.*, 2006). In turn, these evolutionary processes are influenced by biological traits that determine the performance of species in the varied environments in which they occur (Wainwright, 2007). Thus, to be able to describe how present-day macroecological patterns were built up, we need to understand the processes through which new species form, go extinct, disperse and get modified in both space and time.

This macroevolutionary approach to species diversity patterns has gained momentum in the last twenty years with a rapid increase in information on species distributions, traits and molecular phylogenies. Across multiple taxa, large-scale patterns of species origination, extinction, dispersal and trait evolution have been revealed (e.g. Bininda-Emonds *et al.*, 2007; Alfaro *et al.*, 2009a; Jetz *et al.*, 2012; Louca *et al.*, 2018; Smith & Brown, 2018; Upham *et al.*, 2019; Varga *et al.*, 2019). In vertebrates, recent studies have highlighted the association between rates of evolution and key geographical factors (e.g. Quintero & Jetz, 2018; Rabosky *et al.*, 2018). However, despite the importance of geography in determining broad patterns of origination and extinction, it is increasingly recognized

that species ecologies have also played a significant role in macroevolutionary trajectories (Coyne & Orr, 2004; Ricklefs, 2007). Species body size is arguably the ecological trait that has received most attention in the recent macroevolutionary literature (e.g. Harmon *et al.*, 2010; Venditti *et al.*, 2011; Rabosky *et al.*, 2013; Benson *et al.*, 2014; Womack & Bell, 2020), given its association with important life-history characteristics such as metabolic rates and generation times (Martin & Palumbi, 1993; Brown *et al.*, 2004). Yet, other ecological traits have also been shown to correlate with large-scale patterns of lineage diversification.

Among these traits, the trophic identity of species seems to be pivotal in the macroevolutionary dynamics of animal taxa. For instance, it has long been suggested that the evolution of novel trophic modes, particularly herbivory, might be associated with bursts of diversification in insects (Mitter *et al.*, 1988; Farrell, 1998). Although these studies were performed over twenty years ago, more recent applications of phylogenetic comparative methods provided support for those predictions in some arthropod groups (Wiens *et al.*, 2015; Poore *et al.*, 2017; McKenna *et al.*, 2019). The trophic component of macroevolution has also been shown to drive patterns of lineage origination in some important vertebrate radiations. In mammals, for example, carnivory has been reconstructed as the likely ancestral state of the group, but interestingly, herbivory was associated with a faster pace of diversification (Price *et al.*, 2012). Similarly, herbivorous bird lineages were shown to diversify faster than other trophic guilds, whereas omnivory has been revealed as an avian macroevolutionary sink (Burin *et al.*, 2016). Although these studies emphasise the importance of understanding patterns of trophic evolution in animal radiations, they were mostly focused on terrestrial groups. In marine systems, previous works have identified links between species diets and rates of evolution in coral reef fishes (Lobato *et al.*, 2014; Gajdzik *et al.*, 2019). However, these studies had a limited phylogenetic scope and were performed with a few fish families. To-date, there has been no study focused on marine vertebrates that includes comprehensive phylogenetic sampling to investigate macroevolutionary and macroecological patterns across major trophic guilds. This approach would be particularly promising in biodiversity cradles such as coral reefs.

## 1.2 The evolution of fishes on coral reefs

Amongst marine systems, coral reefs stand out because of their incredibly high number of species, despite occupying a relatively small area of the global oceans (Spalding & Grenfell, 1997). Although these shallow, tropical environments formed by carbonate substrates have historically acted as cradles for marine biodiversity (Kiesling *et al.*, 2010), their present status as coral-dominated systems is largely a product of the last 60 million years (Myr) of evolution (Kiesling, 2009; Bellwood *et al.*, 2017). Along with the expansion of corals as major reef builders in the Cenozoic, came the intensification of the relationships between fishes and reefs (Bellwood & Wainwright, 2002; Bellwood *et al.*, 2015). Today, the fishes that live on reefs comprise one of the most diverse vertebrate assemblages in the world, with over six thousand species globally (Parravicini *et al.*, 2013). Paralleling the number of described mammals (Burgin *et al.*, 2018), reef fishes present a remarkable ecological diversity with species fulfilling widely distinct niches. The association with reef environments seems to have promoted high rates of diversification in extant fish lineages when compared to other habitats (Alfaro *et al.*, 2007; Cowman & Bellwood, 2011). However, the relationship between fishes and reefs has a long and complex history of origination and extinction that spans over 400 million years.

This complex history could only be revealed through combined evidence from geology, palaeontology and molecular phylogenetics. In a recent review, Bellwood *et al.* (2015) used elements from all of those disciplines to demonstrate that the evolution of the relationship between fishes and reefs can be divided into six major phases. Phase 1 started in the Devonian (420 – 359 Million years ago [Ma]) and is related to the presence of ancient fish forms in the vicinity of reefs (Long & Trinajstić, 2010). At this point in time, there was no evidence for a close relationship between fishes and reefs, since fishes were habitat generalists. It is only in Phase 2 (230 – 90 Ma) that we see the origins of morphologies that indicate the capacity of predating on benthic invertebrates in a manner comparable to those seen in some modern fishes (Bellwood *et al.*, 2015). While this represents an important ecological breakthrough, these extinct fish lineages still had very little resemblance to modern

acanthomorphs, which form the bulk of the extant piscine fauna on coral reefs. Then, by the end of the Mesozoic, Phase 3 (90 – 66 Ma) marks the appearance of stem acanthomorph lineages in the fossil record (Friedman & Sallan, 2012). These origins are also supported by molecular phylogenies (Near *et al.*, 2012; Alfaro *et al.*, 2018). The generalized morphologies of these initial acanthomorphs, however, provide little evidence for a change in the nature of fish-reef interactions. It is only during the Cenozoic (last 66 Myr) that modern coral reef fish assemblages arose and it is when we see the diversification of both taxonomic and functional groups that are typical of modern reefs. Although this geological era is famously known as the ‘Age of Mammals and Birds’, in the marine realm it is, in some respects, the ‘Second Age of Fishes’ after the Devonian Period (Friedman & Sallan, 2012).

The Cenozoic was, therefore, the most important period for the evolution of fishes on coral reefs. This geological era has been divided in Phases 4, 5 and 6 (*sensu* Bellwood *et al.* 2015), and it marks fundamental changes in the way fishes interact with reefs with escalating complexity (Bellwood *et al.*, 2017). Straight after the Cretaceous–Palaeogene (K–Pg) mass extinction event, Phase 4 (66 – 34 Ma) is characterized by the origins of most extant reef fish families in the fossil record (Bellwood, 1996). More importantly perhaps, it was when some key reef fish functional groups first appeared. For instance, Phase 4 marks the origins of herbivory by fishes in marine systems (Bellwood, 2003; Bellwood *et al.*, 2014a), which profoundly altered the nature of reefs (Steneck, 1983). Geographically, most of these lineages were present in the central region of the extinct Tethys Sea, which was the global hotspot for marine biodiversity at the time (Renema *et al.*, 2008). This is an important element that differentiates phases 4 and 5. With major tectonic rearrangements, Phase 5 (34 – 5 Ma) was when the marine biodiversity hotspot shifted from its former position in the central Tethys region to its current location within the Indo-Australian Archipelago (IAA) (Renema *et al.*, 2008). Arguably, this was the most important phase for lineage origination and expansion in reef fishes, given that most extant generic diversity arose in the last 30 Myr (Cowman & Bellwood, 2011, 2013a; Bellwood *et al.*, 2017). In addition, Phase 5 underpins the dawn of critical functions that sustain reefs to this day. It marks the origins of bioeroding parrotfishes (Cowman *et al.*, 2009), detritus-feeding surgeonfishes (Bellwood *et*

*al.*, 2014a) and blennies (Bellwood *et al.*, 2014b), and coral feeding butterflyfishes (Bellwood *et al.*, 2010) and wrasses (Cowman *et al.*, 2009). Given that some of these fishes are amongst the most abundant faunal components on present-day reefs (Bellwood *et al.*, 2014a), this phase represents a major and crucial shift in the ecology of coral reefs. Finally, in Phase 6 (5 – 0 Ma) reef fishes experienced extensive speciation events with marked changes in coloration (Bellwood *et al.*, 2015, 2017), despite presenting little functional differentiation.

What becomes evident from this historical process and formation of fish assemblages on coral reefs is that it was largely intertwined with the evolution of trophic characteristics in fish species. From its early origins, the relationship between fishes and reefs have been essentially characterized based on how fishes feed. Thus, describing aspects of trophic evolution in fishes on coral reefs is paramount if we want to understand the genesis of these high-diversity systems. Previous efforts have been made to explore the evolution of trophic modes in some key coral reef fish families. For instance, Cowman *et al.* (2009) used inferred dates estimated from molecular sequences to trace the origins of feeding modes in the Labridae, one of the most speciose coral reef fish families. By relying on morphological comparisons between extant and fossil species, Bellwood *et al.* (2014 a,b) reconstructed the temporal appearance of some important morphological attributes in herbivorous and detritivorous reef fish lineages. More recently, Borstein *et al.* (2018) applied a comprehensive phylogenetic framework to investigate how rates of morphological evolution vary between fishes in different trophic categories. These and other studies provided important insights into particular aspects of reef fish trophic evolution, however, a more integrative approach combining information from fossils, molecular phylogenies, biogeography and ecology is still lacking. Considering the recent advances in all these different fields, the timing could not be more appropriate to begin addressing important knowledge gaps in our understanding of how trophic characteristics evolved in reef fishes.

### 1.3 Thesis aims and outline

Fishes perform a key role within the energetic network of coral reefs, occupying diversified guilds ranging from consumers of primary productivity to top predators. However, we still lack an in-depth analysis of the macroevolutionary mechanisms shaping their diversity among trophic guilds. In this thesis, therefore, I aimed to understand the processes underpinning the extant distribution of coral reef fishes across trophic guilds in space and time. In **chapter 2**, I used a comprehensive approach to investigate the mechanisms through which reef fish lineage diversification has been enhanced in the recent geological past. To do so, I employed gradient boosting techniques in association with phylogenetic comparative methods to assess the relative importance of geographical and ecological factors in predicting species-specific net diversification rates. In addition, I explored reef fish trophic evolution from the perspective of historical evolutionary rates and transitions among guilds. Following up, in **chapter 3**, I explored the macroecological signal of trophic guilds in coral reef fishes. By relying on two independent datasets on species distributions, I dissected species richness patterns into major coral reef fish trophic groups to then model the relationship between diversity and distance to the global hotspot for marine biodiversity (IAA). To have a better grasp of the evolutionary mechanisms behind the species richness patterns, I also applied phylogenetic comparative methods using the comprehensive reef fish tree produced in the previous chapter. Then, in **chapters 4 and 5**, I focused on herbivorous coral reef fish groups to describe in more detail their global biogeographical patterns from both functional and taxonomic perspectives. In **chapter 4**, I analysed the evolution of important morphological and behavioural traits, and their resulting ecosystem functions between two major oceanic basins: Indo-Pacific and Atlantic. This was achieved by comparing the herbivorous fossil assemblage with extant lineages, and tracing their evolution through the use of ancestral state reconstructions in molecular phylogenies. Finally, in **chapter 5**, I further explored the evolutionary differences between the Atlantic and the Indo-Pacific Oceans, in terms of piscine herbivory. However, instead of only using extant species phylogenies, as in chapter 4, I used a robust evolutionary model that permits the inclusion of fossil species within the trees. With this framework, I was then able to

## Chapter 1. General introduction

build realistic biogeographical reconstructions of herbivorous reef fish lineages from their presumed origins in the fossil record. Hence, this thesis integrates fundamental aspects of evolutionary biology, historical biogeography, palaeontology and ecology with the overarching aim of explaining present-day patterns in coral reef fish biodiversity distribution and the role of trophic identity in this evolutionary history.

## Chapter 2.

### Evolution of trophic guilds in coral reef fishes

*This chapter is published as:*

Siqueira, A. C., Morais, R. A., Bellwood, D. R., Cowman, P. F. (2020). Trophic innovations fuel reef fish diversification. *Nature Communications*, 11, 2669. doi: 10.1038/s41467-020-16498-w

#### 2.1 Abstract

Reef fishes are an exceptionally speciose vertebrate assemblage, yet the main drivers of their diversification remain unclear. It has been suggested that Miocene reef rearrangements promoted opportunities for lineage diversification, however, the specific mechanisms are not well understood. Here, we assemble near-complete reef fish phylogenies to assess the importance of ecological and geographical factors in explaining lineage origination patterns. We reveal that reef fish diversification is strongly associated with species' trophic identity and body size. Large-bodied herbivorous fishes outpace all other trophic groups in recent diversification rates, a pattern that is consistent through time. Additionally, we show that omnivory acts as an intermediate evolutionary step between higher and lower trophic levels, while planktivory represents a common transition destination. Overall, these results suggest that Miocene changes in reef configurations were likely driven by, and subsequently promoted, trophic innovations. This highlights trophic evolution as a key element in enhancing reef fish diversification.

## 2.2 Introduction

The heterogeneity in rates of species formation across the tree of life is a widely recognized macroevolutionary pattern (Alfaro *et al.*, 2009a; Louca *et al.*, 2018; Maliet *et al.*, 2019). As a product of speciation and extinction rates, diversification varies through time (Stadler, 2011) and among lineages (Alfaro *et al.*, 2009a), being ultimately influenced by both biotic and abiotic factors (Benton, 2009). Consequently, understanding how biotic and abiotic factors interact in space and time is paramount in explaining underlying patterns of species diversification. For instance, recent studies have shown that distinct diversification trajectories within vertebrate groups can be explained by geographical (Jetz *et al.*, 2012; Rabosky *et al.*, 2018) and species-specific biological traits (Price *et al.*, 2012; Burin *et al.*, 2016). Although these studies provided important insights into individual drivers of vertebrate radiations, disentangling the simultaneous influence of multiple factors on rate heterogeneities is still challenging. This will require a comprehensive approach with methods that can account for multi order interactions among the underlying drivers of species diversification.

Coral reefs constitute an excellent system for applying such broad macroevolutionary approaches, given their status as cradles for biodiversity (Kieffer *et al.*, 2010). Particularly important within these high-diversity systems, fishes represent key energetic conduits, taking part in a large proportion of the recognized biotic interactions. This ecological diversity is also reflected in taxonomic terms, with reef-associated fishes being one of the most speciose vertebrate assemblages in the world (Eschmeyer *et al.*, 2010). Although it has been shown that the association with reefs was an important promoter of fish cladogenesis (Cowman & Bellwood, 2011), the specific mechanisms driving this diversification are not yet fully understood. Historical and geological processes have clearly influenced global distribution patterns of reef fishes at large temporal scales (Bellwood & Wainwright, 2002; Renema *et al.*, 2008; Cowman & Bellwood, 2013; Pellissier *et al.*, 2014; Leprieur *et al.*, 2016; Siqueira *et al.*, 2019a [chapter 5 in this thesis]), with Miocene (23–5.3 Million years ago [Ma]) changes in reef configuration being posited as one of the most important drivers of lineage expansion (Bellwood *et*

*al.*, 2017). However, it is still unknown whether this change in the pace of lineage formation in the Miocene occurred under the influence of biotic or predominantly abiotic factors.

As recently suggested (Bellwood *et al.*, 2017), the drivers of reef fish diversification in the Miocene seem to have involved a complex mix of history and ecology. On the historical side, this epoch was marked by major geomorphological changes that reshaped marine biogeography with the formation of the Indo-Australian-Archipelago (IAA) marine biodiversity hotspot (Renema *et al.*, 2008). This process was likely associated with rapid diversification of reef fish lineages (Cowman & Bellwood, 2011), given the extensive opportunities for vicariance and range expansion provided by the geographical complexity of the IAA. On the ecological side, key trophic innovations (i.e. evolutionary novelties that granted access to previously unexplored resources; Wainwright, 2007) in reef-associated fishes have fundamentally altered the nature of Miocene reefs (Bellwood *et al.*, 2017). While major reef fish trophic groups were already represented in the Eocene (56–33.9 Ma), specialised morphologies associated with the exploitation of detrital and corallivore trophic pathways, for example, only arose in the Miocene (Bellwood *et al.*, 2010, 2014b,a, 2017; Siqueira *et al.*, 2019b [chapter 4 in this thesis]). These morphological and trophic innovations have also been linked to increased lineage origination in selected reef fish groups (Bellwood *et al.*, 2010; Price *et al.*, 2010; Frédérich *et al.*, 2013; Lobato *et al.*, 2014; Clements *et al.*, 2017), suggesting that trophic evolution might have had a prominent role in driving patterns of reef fish diversification. Although recent studies have independently explored these potential ecological mechanisms or geographical factors underlying reef fish diversification patterns (e.g. Lobato *et al.*, 2014; Pellissier *et al.*, 2014; Siqueira *et al.*, 2016; Borstein *et al.*, 2019; Donati *et al.*, 2019; Gajdzik *et al.*, 2019), they are yet to be examined in a comparative analytical framework capable of quantifying relative support.

To fill this knowledge gap, we applied phylogenetic comparative methods in near-complete phylogenetic trees to examine the relative importance of ecological and geographical factors in explaining recent lineage origination patterns in reef fishes. More specifically, we first estimate the

rates of diversification for all lineages of reef-associated fishes. Then we apply extreme gradient boosting techniques to assess the most important variables in explaining these lineage-specific rates. Finally, we investigate the historical patterns of reef fish trophic evolution in terms of evolutionary rates and guild transitions, after having identified trophic evolution as a major driver of recent reef fish diversification. These approaches provide a complementary picture of both recent and historical rates of evolution in an important vertebrate radiation, reef fishes.

## 2.3 Methods

### 2.3.1 Reef fish phylogeny

We built a comprehensive phylogeny of reef fish species, based on a recently published chronogram of ray-finned fishes (Rabosky *et al.*, 2018). This chronogram was constructed using a 27-gene alignment for 11,638 actinopterygian species and was time-calibrated using a comprehensive dataset of fossil occurrences. We downloaded the Rabosky *et al.* (2018) chronogram from *fishtreeoflife.org*. Then we used the ‘ape’ (Paradis *et al.*, 2004) R package to prune down the tree, restricting it to reef-associated taxa. Since the definition of what constitutes a reef fish is a contentious subject (Bellwood & Wainwright, 2002), we used a systematic approach in selecting the species to be kept in the tree. Starting from the full list of fish families with reef-associated species from Bellwood & Wainwright (2002), we used the ‘*rfishbase*’ (Boettiger *et al.*, 2012) R package to access the list of all valid species within each of those families and then calculate the proportion that were classified as reef-associated. Finally, we selected families with more than 20% of reef-associated species and kept them in the tree. The final pruned chronogram contained 2,585 species in 65 families.

This time-calibrated pruned tree was subsequently used as a backbone for the imputation of all missing species within each of the selected families. To do this, we generated a list of all valid reef-associated species belonging to the selected families based on FishBase (Froese & Pauly, 2019) and the Eschemeyer’s Catalog of Fishes (Fricke *et al.*, 2019). We then assigned taxonomic ranks to all

species present in the list using the same online datasets, but also using information from the backbone tree to better define monophyletic groups. With this taxonomic dataset, we applied the Taxonomic Addition for Complete Trees [TACT] stochastic polytomy resolution algorithm (Chang *et al.*, 2020), which uses birth-death models to calculate diversification rates for taxonomic ranks and inputs missing species within the most restrictive ranks according to the respective calculated rate. This method has the advantage of estimating local diversification rates, as opposed to global rates, being more suitable for large phylogenies with heterogeneous rate regimes (Rabosky *et al.*, 2018). Although our approach is very similar to the one implemented by Rabosky *et al.* (2018) to build a near-complete tree, we used more restrictive taxonomic ranks in an attempt to narrow down the placement of missing species. In most cases, missing species were placed within their respective genera or, at least, within their respective subfamilies where available. Finally, to account for stochastic variability in the placement of missing species within genera/subfamily, we generated a distribution of 100 near-complete reef fish trees, each containing 6,257 tips.

### 2.3.2 Diversification rates

To estimate diversification rates within our phylogenies, we used the program BAMM 2.5.0 (Rabosky, 2014). This program uses a Bayesian framework and a reversible-jump Markov Chain Monte Carlo (rjMCMC) process to find distinct diversification regimes within a phylogeny and estimate lineage-specific speciation and extinction rates. For each of our trees, we ran time-variable models for 30 million generations using default operators and priors generated through the ‘*BAMMtools*’ (Rabosky *et al.*, 2014) R package. To facilitate convergence, we set a prior expectation of 100 diversification regime shifts. Since we were using near-complete trees, we set the `globalSamplingFraction` parameter to one. At the end of each run, we removed the initial 10% of the samples as burn-in and assessed convergence through the effective sample sizes using the ‘*coda*’

(Plummer *et al.*, 2006) R package. After running BAMM independently in each of our 100 trees, we combined their results by assessing the median estimated tip diversification rates.

Although concerns related to BAMM have been raised (e.g. Moore *et al.*, 2016; Meyer *et al.*, 2018), they have been largely addressed in subsequent studies and program refinements (e.g. Rabosky *et al.*, 2017; Rabosky, 2019). The current program, therefore, remains a robust framework for estimating diversification rates in large phylogenetic trees. Recently, another framework has been proposed (ClDS ; Maliet *et al.*, 2019), providing model improvements in terms of lineage-specific rate estimates. Although this model represents a very strong alternative to BAMM, its implementation is still computationally very intensive, making analyses in large phylogenies such as ours impractical. Therefore, to be able to use other method as a cross-validation for our main BAMM analysis, we applied the 'DR statistic' (Jetz *et al.*, 2012) in our near-complete trees. Although this method is mainly focused on speciation, rather than diversification rates (Title & Rabosky, 2019), it is a very useful metric to study speciation rate dynamics alongside BAMM (Title & Rabosky, 2019). We applied this method in our 100 trees and assessed the median lineage-specific speciation rates. The median BAMM and the DR tip estimates were then used independently as the response variables in our predictive model (Methods section 2.3.4). Finally, since we focused on patterns of recent (tip) diversification rates, our estimates are unlikely to be influenced by the recently described issues of parameter non-identifiability (Louca & Pennell, 2020) in extant species phylogenies.

### 2.3.3 Explanatory variables

To assess the main drivers of diversification in reef fishes, we generated a dataset with potential explanatory variables. These variables consisted of a set of species' ecological traits and geographical factors hypothesised to influence the pace of reef fish lineage formation. We used information from the literature, online datasets and expert assessments (Mouillot *et al.*, 2014; Floeter *et al.*, 2018; Morais & Bellwood, 2018; Fricke *et al.*, 2019; Froese & Pauly, 2019) to classify species

according to a continuous trait reflecting body size (maximum body length), and three categorical traits related to species' ecologies (trophic identity, activity pattern and position in the water column). All the body length data available for our studied species was downloaded from FishBase through 'rfishbase' (Boettiger *et al.*, 2012). For the trophic identity, we grouped species into six major categories related to their diets in the adult life stages: generalized carnivores (GC), mobile invertivores (MI), omnivores (OM), planktivores (PK), sessile invertivores (SI) and herbivores/detritivores (HD). These categories are related to previously defined dietary groups for reef fishes (Mouillot *et al.*, 2014), however, we merged the herbivores/macroalgivores category with the general herbivores/detritivores group. This was done to avoid biases in the predictive and the trait-dependent diversification models, given the very small sample size of macroalgivores in our dataset. Additionally, we used a broader categorization for carnivores than Mouillot *et al.* (2014). Species that feed on larger prey (i.e. fish and cephalopods) were classified as piscivores by Mouillot *et al.* (2014), however, we adopted a more encompassing category that includes species that feed more generally on larger elusive prey (including larger crustaceans). Our classification considered the most common diets described for each species regardless of potential geographical variation. We also split species between diurnal, nocturnal or both (Mouillot *et al.*, 2014), according to their circadian pattern of activity. Lastly, we used the vertical position where fishes are commonly found in the water column as a proxy for their degree of association with the reef matrix, so we classified species as benthic, benthopelagic and pelagic (Mouillot *et al.*, 2014).

To classify species according to geographical variables, we downloaded the presence-absence dataset from Rabosky *et al.* (2018, 2019). This dataset consists of estimated geographic ranges of marine fishes using a set of environmental predictors through the AquaMaps algorithm (Ready *et al.*, 2010). From this presence-absence data, we filtered those species that were present in our trees and we calculated their geographical range by summing the number of occupied cells. Additionally, we classified species according to their presence in each major oceanic basin (Atlantic, Indo-Pacific or both), and we calculated the absolute latitude of the centroid of their geographical distribution. By

combining the absolute latitude value with the longitudinal centroid of each species, we calculated the distance between that centroid and a central point in the Indo-Australian Archipelago (IAA; Lat 0°; Long 121°W). These variables were added to the model (see Methods section 2.3.4) to assess predictions related to the influence of biogeography into reef fish diversification rates (Cowman & Bellwood, 2011; Bellwood *et al.*, 2017).

In addition to the presence-absence dataset, we also downloaded the supplementary data from Rabosky *et al.* (2018, 2019) that contained environmental variables per grid cell. With this data, we accessed the mean sea surface temperature (SST) and the mean primary productivity (Pprod) at the centroid grid of each species. Since tropical reef fish lineages have been found to sustain higher net diversification rates (Siqueira *et al.*, 2016), we used these variables to assess if this might be associated with higher temperatures or energy availability. Our complete dataset containing species' ecological traits (body size, trophic identity, activity and position) and geographical variables (geographic range, oceanic basin, distance to IAA, mean SST and mean Pprod) had a total of 4,875 species.

#### 2.3.4 Predicting diversification rates

To evaluate the importance of each ecological and geographical variables in predicting reef fish diversification rates, we used the Gradient Boosted Regression Tree method XGBoost (Chen & Guestrin, 2016). This machine learning technique represents a state-of-the-art method for modelling complex nonlinear relationships (Elith *et al.*, 2008). It has advantages over other modelling techniques because it automatically handles multi order interactions among predictors, it does not require prior data transformation or outlier exclusion (Elith *et al.*, 2008), and it provides fast and accurate predictions (Chen & Guestrin, 2016). We used the 'xgboost' (Chen *et al.*, 2019) R package to build our predictive model. Before running the predictive model, we performed two tuning steps to obtain the combination of parameters (learning rate, maximum tree depth, gamma and subsampling rate) that

would result in the minimum root mean square error (rmse). In the first tuning step, we fit models with a range of predefined parameter combinations that were varied systematically to assess which would provide the minimum rmse. In the second tuning step, we refit 1,000 models by randomly sampling parameters from a uniform distribution with upper and lower bounds defined as values from the best parameter combination of the first step plus or minus 10%. The parameter combination with the minimum rmse from the second tuning step was then used in the final predictive model. Both tuning steps and the final predictive model were fitted using a Gamma distribution for the median tip diversification rates resulting from our BAMM analysis as the response variable.

We used a cross-validation procedure to assess the model's accuracy and precision in predicting diversification rates. To do that, we divided our dataset into training and testing parts by randomly subsetting 80% and 20% of the datapoints, respectively. We used the training dataset to refit the final model and assess the coefficients of prediction. These coefficients were then used to predict the tip diversification rates in the testing dataset. Accuracy was calculated as the average bias by subtracting each predicted tip diversification from its actual value in the training dataset. Precision was assessed using the  $R^2$  of a linear model fitted between the measured and predicted diversification rates. These cross-validation tests were performed 1,000 times to assess the mean accuracy and precision values.

We ran the predictions for all levels of the categorical variables, and for a range of values spanning the minimum and maximum measured continuous variables. These predictions were bootstrapped for 1,000 iterations to assess the relative importance of each explanatory variable. Finally, we did another 1,000 bootstrap iterations of the final predictive model varying only trophic group and maximum body length (the most important variables; see Results section 2.4.2), while keeping all the other continuous variables in their mean values and the categorical variables in their most common category.

All of these steps were replicated using the 'DR statistic' results as the response variable. Moreover, two model sensitivity analyses were performed. First, we ran the *xgboost* analysis selecting only the reef fish families considered 'consensus' families (Bellwood & Wainwright, 2002) (i.e. Acanthuridae, Apogonidae, Blenniidae, Carangidae, Chaetodontidae, Gobiidae, Holocentridae, Labridae, Lutjanidae, Mullidae, Pomacanthidae, Pomacentridae and Serranidae). Second, we ran the predictive model excluding the families defined by Brandl *et al.* (2018) as cryptobenthic reef fishes. With the first analysis we intended to eliminate potential issues of defining what constitutes reef fishes. In the second analysis we wanted to exclude the potential taxonomic bias associated with smaller body sized species, i.e. we expect more undescribed cryptobenthic species than larger bodied ones.

### 2.3.5 Trait-dependent diversification

After detecting trophic identity as the main explanatory variable for recent (tip) patterns of lineage diversification in reef fishes (see Results section 2.4.2), we explored the historical patterns of trophic evolution using the whole structure of the phylogenetic trees. This was achieved by building multistate speciation and extinction models (MuSSE; FitzJohn, 2012) for the classified trophic groups. Our two sets of diversification analyses differ in the sense that the first (BAMM) was used to estimate rates independently of trait evolution, whereas the second (MuSSE) was specifically used to investigate trait-dependent patterns of diversification. These trait-dependent diversification models allow the analysis of character state evolution coupled with changes in speciation and extinction rates.

For each of our reef fish trees, we estimated the parameters (speciation, extinction and transition rates) associated with each trophic group using an unconstrained MuSSE model with the maximum likelihood function of the '*diversitree*' (FitzJohn, 2012) R package. Subsequently, we used the resulting maximum likelihood coefficients to apply the Bayesian framework of '*diversitree*' and sample the posterior probability distribution of parameters. We ran the MCMC chain for 2,000

generations with exponential priors from a preliminary run of 100 generations. After each run, we excluded 10% of the samples as burn-in and assessed convergence using the effective sample sizes. Finally, we combined the post burn-in samples from all trees and calculated net diversification rates by subtracting extinction rates from speciation rates.

Issues related to the model selection procedure of trait-dependent diversification models have been previously identified (Rabosky & Goldberg, 2015), however, they are unlikely to affect our analysis. This is because we did not use MuSSE to perform model selection and thus imply that trophic group is the only trait affecting reef fish diversification. Based on our BAMM results (**Fig. 2.1**), we know that the diversification regime in the full reef fish tree is highly heterogeneous and it was unlikely influenced by only one trait. Because we detected trophic identity as an important variable for explaining tip diversification rate variability in reef fishes, we used this method exclusively to explore full-tree patterns. Thus, our trait-dependent analysis should be viewed as a complementary resource to the results found with the trait-independent one (BAMM). As a way to alleviate potential issues with the trait-dependent analysis, we applied the HiSSE method (Beaulieu & O'Meara, 2016) by splitting our trophic categories between ancestral (generalized carnivores, mobile invertivores and omnivores) and more recently derived (herbivores/detritivores, sessile invertivores and planktivores) groups. Using the HiSSE framework, we built an unconstrained model that considered rates to be different between analysed character states (trophic group) with one hidden diversification regime per state, and compared it to a model in which rates were constrained between states but different from the hidden diversification regime. Results from this HiSSE analysis supported our MuSSE results (see Results section 2.4.3) and estimated higher diversification rates for the recently derived trophic groups compared to the ancestral ones (Supplementary Table 2.1).

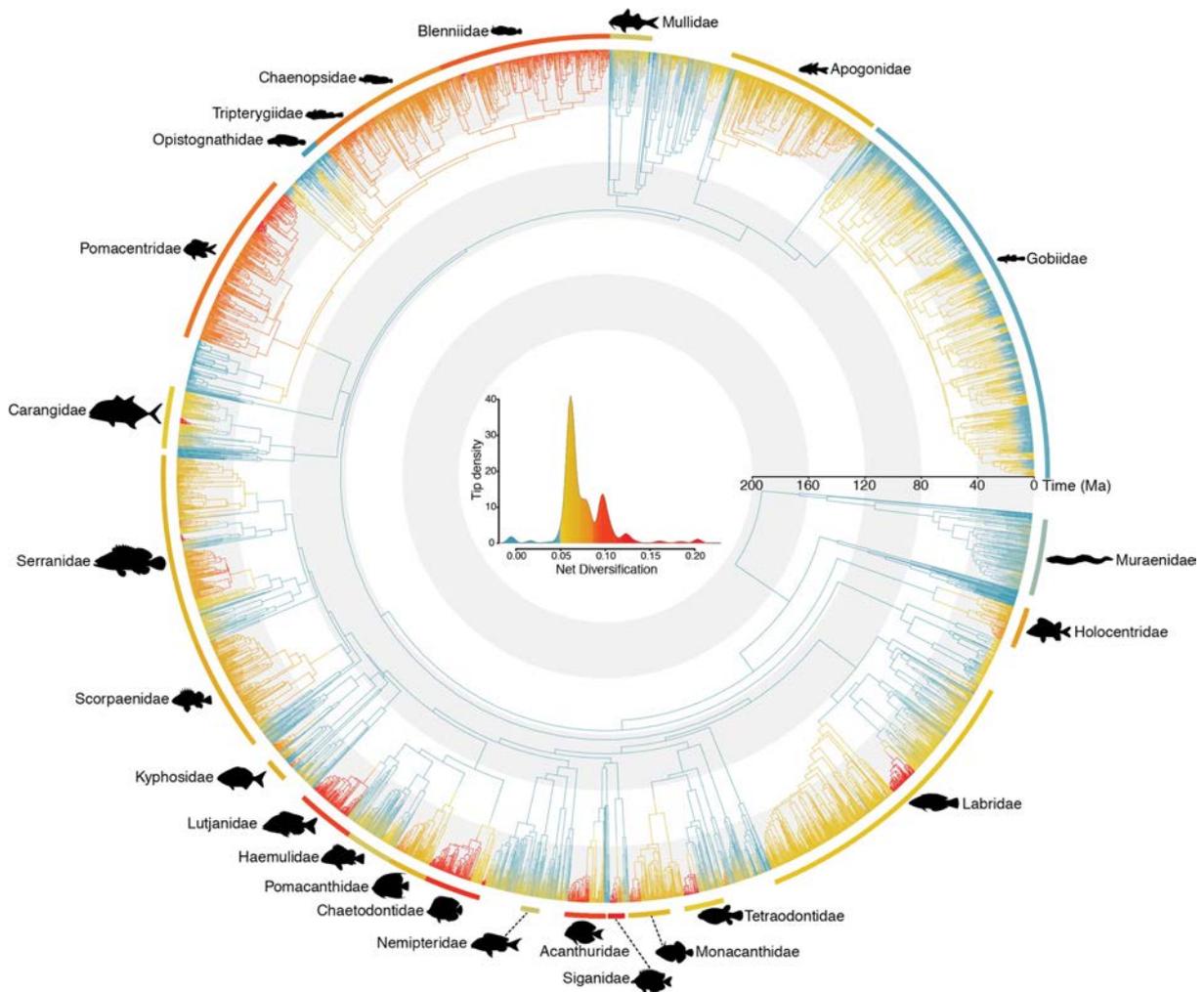
### 2.3.6 Trophic transitions

We quantified the transitions between classified trophic groups by using stochastic character mappings (Huelsenbeck *et al.*, 2003). Considering that rate heterogeneity can affect the results of ancestral state reconstructions (Maddison, 2006), we used the results of our trait-dependent diversification model (MuSSE) to perform this analysis. For each of our near-complete trees, we simulated 10 stochastic maps using a modified version of the *make.simmap* function from the 'phytools' (Revell, 2012) R package. We customized the aforementioned function to use the transition rates and the ancestral state reconstruction results derived from the original MuSSE model (*asr.marginal* function in 'diversitree') as inputs for the stochastic mappings. The combined results of all stochastic maps were summarized to assess the mean number of transitions per trophic group. These estimates were then used to plot chord diagrams representing the directionality of transitions using the 'circlize' (Gu *et al.*, 2014) R package.

## 2.4 Results

### 2.4.1 Diversification rate heterogeneity

We found extremely heterogeneous diversification rates throughout our comprehensive reef fish trees. Net diversification rates (speciation minus extinction) varied by more than two orders of magnitude, ranging from slightly negative ( $-0.007$  lineages  $\text{Myr}^{-1}$ ) in the genus *Megalops*, to extreme values ( $1.2$  lineages  $\text{Myr}^{-1}$ ) in *Hypoplectrus*. Most extant lineages (inset, **Fig. 2.1**) and reef fish families had intermediate rates of diversification, although families such as the Siganidae, Acanthuridae, Chaetodontidae and Lutjanidae presented a noticeably faster pace of species formation (**Fig. 2.1**). Remarkably, despite being the most speciose family on reefs, gobies presented generally low rates of diversification, particularly in the last 20 million years (Myr) (**Fig. 2.1**).

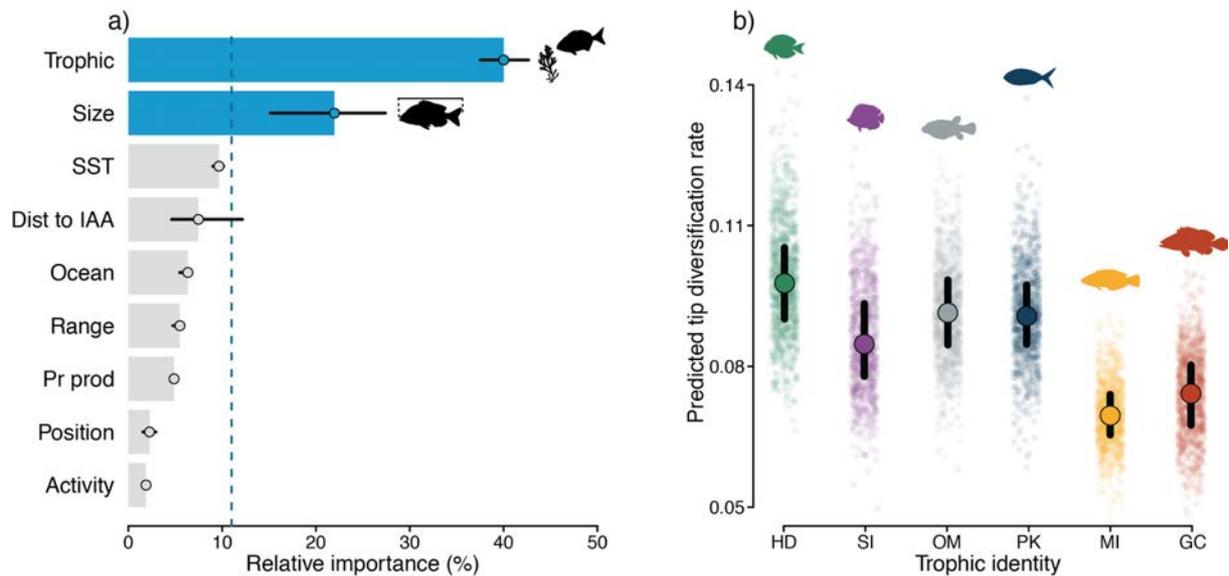


**Figure 2.1 | Near-complete reef fish phylogeny mapped with net diversification rates.** Inset shows the overall distribution of diversification rate values for all tree tips. Blue colours represent low diversification rates, yellow intermediate, and red colours depict high diversification values. Rates were estimated through BAMM (Rabosky, 2014). External arcs show median diversification rates estimated for some iconic reef fish families, represented by the silhouettes.

#### 2.4.2 Predictors of reef fish diversification

Our extreme gradient boosting analysis showed that species trophic identity is the most important variable in explaining patterns in tip diversification for reef fishes. This variable had a mean relative importance of 40% in our final models (**Fig. 2.2a**). Besides trophic identity, body size was the only other variable that had a higher importance in predicting recent diversification rates than expected by chance, with a mean of 22% (**Fig. 2.2a**). All other ecological and geographical variables remained at or below the relative importance expected by chance. Overall, our model performed well,

with very high prediction accuracy (mean average bias of 0.002 or 2.5%) and moderate precision (30% mean prediction variance explained).

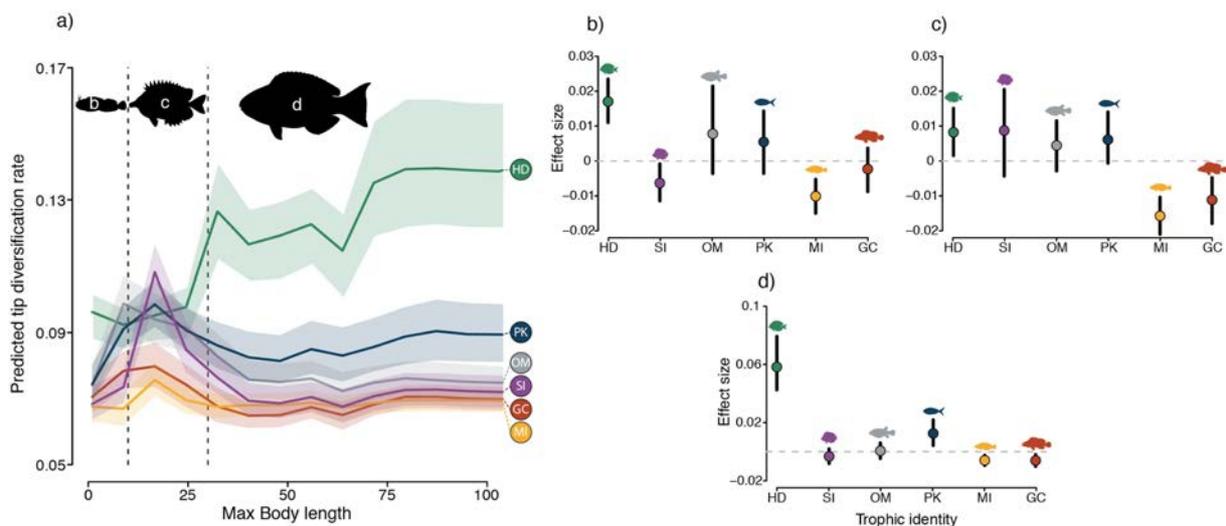


**Figure 2.2 | Ecological and geographical factors driving reef fish tip diversification patterns. (a)** Mean relative importance (%) of explanatory variables based on an extreme gradient boosting model. Blue bars show variables above chance expectation (dashed line). Black lines represent importance quantiles (25% and 75%) derived from 1000 model bootstraps. Trophic: trophic group; Size: maximum body length; SST: sea surface temperature; Dist to IAA: distance to the Indo-Australian Archipelago; Ocean: oceanic basin; Range: geographic range; Pr Prod: primary productivity; Position: position in the water column; Activity: circadian activity period (see Methods section 2.3.3). **(b)** Predicted tip diversification rates per trophic group. In this analysis, all other continuous variables are kept at their mean values and categorical variables in the most common category. HD: herbivores/detritivores (green); SI: sessile invertivores (purple); OM: omnivores (grey); PK: planktivores (blue); MI: mobile invertivores (yellow); GC: generalized carnivores (red). Semi-transparent dots are bootstrapped predictions ( $n = 1000$ ), with larger points representing median values with respective 25% and 75% prediction quantiles (black lines).

Tip diversification rates predicted per trophic group, while keeping body size at the mean value (25 cm), were found to be highest for herbivores/detritivores (mean 0.097 [0.090 - 0.105; 75% prediction quantiles]) (**Fig. 2.2b**). Omnivores, planktivores and sessile invertivores had intermediate tip diversification values (0.091 [0.084 - 0.098], 0.090 [0.084 - 0.097], 0.084 [0.077 - 0.093])

respectively), while generalized carnivores and mobile invertivores were found to be the slowest diversifying groups (0.074 [0.067 - 0.080], 0.069 [0.065 - 0.073] respectively; **Fig. 2.2b**).

We also found a clear interaction between body size and trophic group (**Fig. 2.3a**). Larger herbivores/detritivores were predicted to have significantly higher diversification rates than smaller bodied ones. Moreover, diversification in this group was higher than in other groups, where the rate ~ body size relationship flattened toward larger body-sized species (**Fig. 2.3a**). Interestingly, we found three different diversification rate regimes by dividing the results between body size classes containing a similar number of species between them. Smaller sized species (< 10 cm) were predicted to have lower tip diversification rates than larger sized ones in most trophic groups (**Fig. 2.3a**). Nevertheless, herbivores/detritivores were the fastest diversifying lineages in this body size class (**Fig. 2.3b**). In the intermediate size class (10 – 30 cm), predicted rates were higher for herbivores/detritivores, planktivores, sessile invertivores and omnivores, when compared to the other groups (**Fig. 2.3c**). Finally, in the large body size class (> 30 cm), herbivore/detritivore lineages were predicted to diversify considerably faster than any other trophic group (**Fig. 2.3d**).



**Figure 2.3 | Tip diversification rates predicted for reef fish trophic groups while varying body size. (a)** Predicted tip diversification rates for species of various maximum body lengths in different trophic groups, based on an extreme gradient boosting model ( $n = 1000$  model bootstraps). All other variables are kept at their mean values

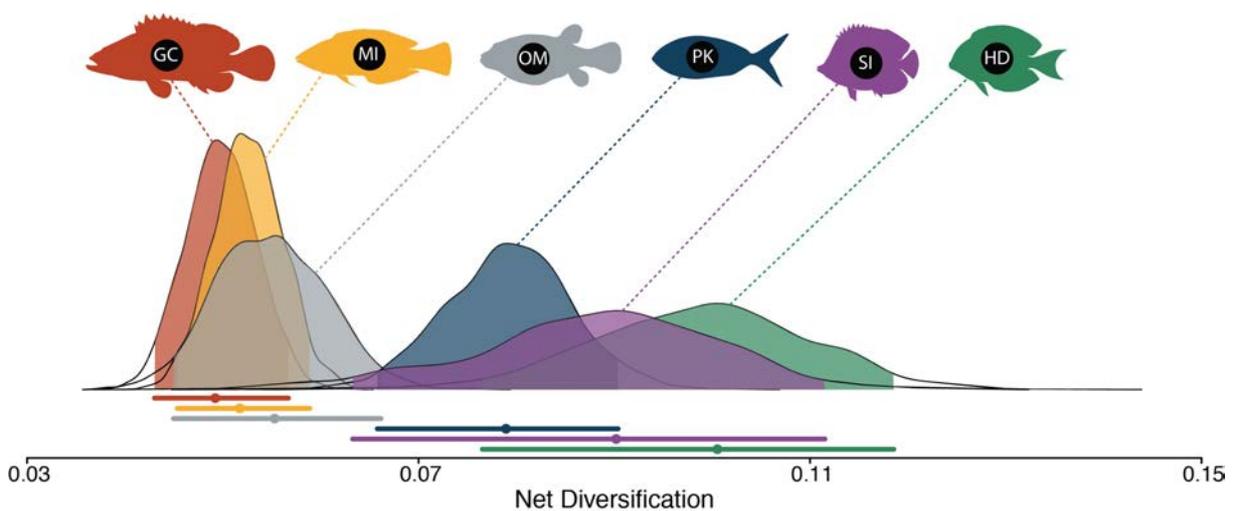
and categorical variables (except trophic identity) in the most common category. Solid lines show median predictions per trophic group with respective prediction quantile intervals (25% and 75%). Dashed line separates size classes for which we show effect sizes per trophic group: **(b)** below 10 cm; **(c)** between 10 and 30 cm; **(d)** above 30 cm. In b–d, circles show the median effects (trophic group median minus global median in each size class) and black lines show 25% and 75% effect quantiles. HD: herbivores/detritivores (green); SI: sessile invertivores (purple); OM: omnivores (grey); PK: planktivores (blue); MI: mobile invertivores (yellow); GC: generalized carnivores (red).

Model results and predictions were consistent when we used the estimates derived from the ‘DR statistic’ as an alternative to the BAMM estimates (see Methods section 2.3.2), with only geographic range and temperature slightly increasing in importance (Supplementary Figure 2.1). Furthermore, when we considered only the ‘consensus’ reef fish families (i.e. universally occurring families on coral reefs, rather than ‘reefs’ *sensu lato*; see Methods section 2.3.2), we found a higher model precision (36%) with trophic identity increasing in importance (55%) when compared to other variables (Supplementary Figure 2.2a). Predictions of diversification rates were similar to the main model, although rates were slightly higher in smaller to medium-sized omnivores and planktivores (Supplementary Figures 2.2b and c). Similarly, after removing cryptobenthic fish families from the model, we found higher precision (35%) and comparable predictions, with higher rates predicted for smaller to medium omnivores and planktivores (Supplementary Figure 2.3). This time, however, the importance of trophic group was reduced (33%) in comparison with body size (29%), suggesting that cryptobenthic fishes contribute to the trophic signal found in the main model.

#### 2.4.3 Historical patterns of trophic evolution

Complementing the tip diversification rate results, our trophic-dependent diversification models revealed that, historically, trophic groups with more recent evolutionary origin diversified faster when compared to ancestral trophic states. Herbivores/detritivores, sessile invertivores and planktivores had significantly higher rates of lineage formation than generalized carnivores, mobile

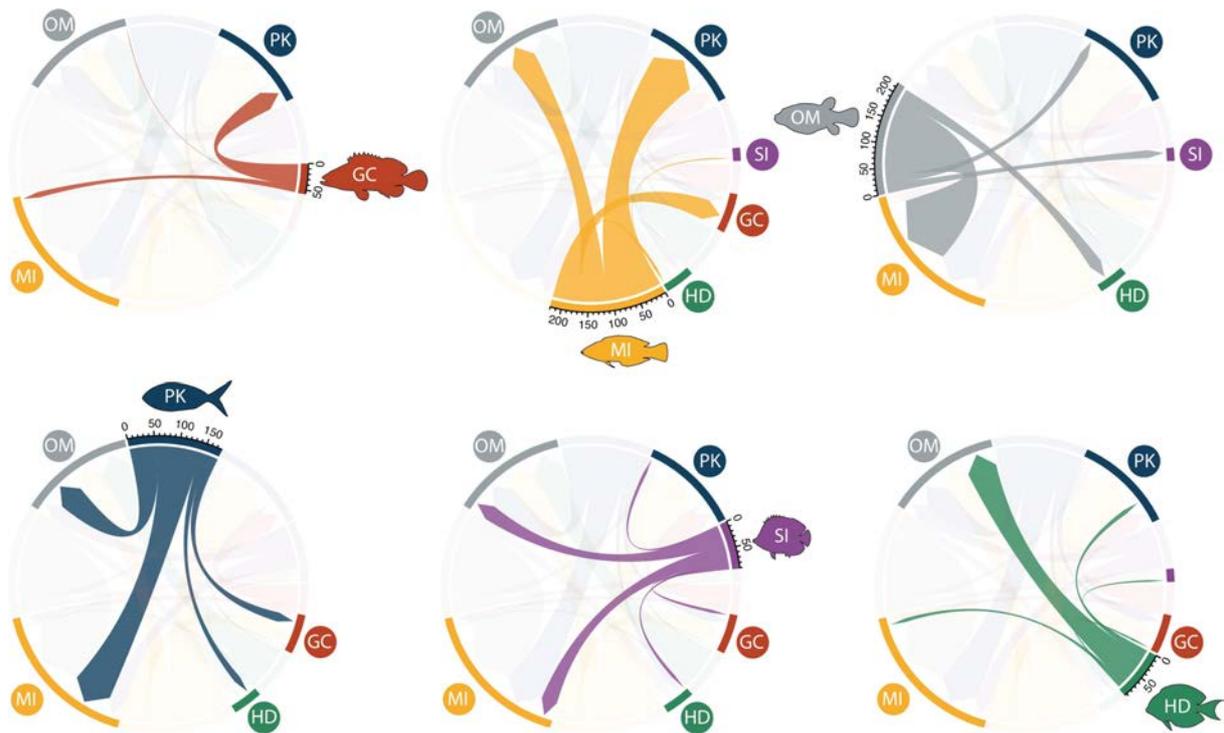
invertivores and omnivores (**Fig. 2.4**). Apart from a few exceptions (e.g. Lutjanidae), these results are similar to those found for tip diversification rates (**Fig. 2.2b**), indicating that patterns of lineage diversification among trophic groups have been historically consistent, with recent trophic groups diversifying rapidly in the last 20 Myr (Supplementary Figures 2.4 and 2.5). Additionally, we found that estimated speciation rates were higher than extinction rates in all groups (Supplementary Figure 2.6), which resulted in positive net diversification rates for all trophic groups (**Fig. 2.4**).



**Figure 2.4 | Historical net diversification rate estimates for six reef fish trophic groups.** Lines below the distributions show mode values (solid circles) with respective 95% credibility intervals. Rates represent the values estimated with *MuSSE*. HD: herbivores/detritivores (green); SI: sessile invertivores (purple); OM: omnivores (grey); PK: planktivores (blue); MI: mobile invertivores (yellow); GC: generalized carnivores (red).

Stochastic character mappings revealed a clear sequential pattern of transitions between trophic groups. Generalized carnivore and mobile invertivore lineages transitioned frequently between them and to planktivory (**Fig. 2.5**). However, these groups very rarely transitioned to other trophic groups such as herbivory/detritivory or sessile invertivory. The transitions to these groups happened, almost exclusively, from omnivorous lineages (**Fig. 2.5**). Herbivore/detritivore and sessile invertivore lineages occasionally transitioned back to omnivory, while planktivores frequently

transitioned back to mobile invertivory. Most groups exhibited frequent transitions to planktivory, making it a common trophic destination in reef fish evolution. Finally, the transitions to omnivory were predominantly made by mobile invertivore lineages (**Fig. 2.5**), suggesting omnivory as an intermediate evolutionary step between lower and higher trophic levels.



**Figure 2.5 | Directionality of transitions out of each reef fish trophic group.** Maximum chord width represents the number of lineages averaged between 100 trees, shown in the scale. HD: herbivores/detritivores (green); SI: sessile invertivores (purple); OM: omnivores (grey); PK: planktivores (blue); MI: mobile invertivores (yellow); GC: generalized carnivores (red).

## 2.5 Discussion

Using near-complete phylogenies, coupled with a comprehensive ecological and geographical dataset, we identified species trophic guild and body size as major drivers of diversification in reef-associated fishes. Although the role of different types of resource use has been previously suggested as a driver of evolutionary rates in some reef fish groups (Cowman *et al.*, 2009; Price *et al.*, 2010; Frédérick *et al.*, 2013; Lobato *et al.*, 2014; Clements *et al.*, 2017; Gajdzik *et al.*, 2019), we reveal its full potential across the complete reef fish tree of life. Through an intricate relationship with body size, the trophic identity of species was more important in predicting the pace of reef fish evolution than any other ecological or geographical factor examined. On average, herbivorous/detritivorous fish lineages diversified faster than other trophic groups. However, rate differences are amplified, rather than diminished, in large-bodied species. Alongside herbivores/detritivores, planktivores and sessile invertivores showed faster than average historical rates of evolution, particularly in the last 20 Myr. This highlights the potential importance of new reef configurations in the Miocene (Bellwood *et al.*, 2017) in promoting trophic innovations within these guilds. Complementing the patterns of reef fish trophic evolution, we also show that planktivory and omnivory constitute key evolutionary pathways. Planktivory is the main evolutionary destination in trophic transition episodes, while omnivory appears to represent a transient state between high and low trophic levels. The major drivers of reef fish diversification and the evolutionary pathways of trophic transitions will be discussed separately below.

### 2.5.1 Drivers of reef fish diversification

Trophic innovations have been previously identified as a key element in the radiation of one of the most speciose fish families on coral reefs, the Labridae (Cowman *et al.*, 2009). Expanding the taxonomic scope, Lobato *et al.* (2014) suggested that ecological opportunity might have underpinned higher diversification rates in some reef fish lineages that shifted towards lower-level trophic guilds.

Even though this latter study was based on a coarse trophic distinction between guilds feeding on low-quality and high-quality food, our results using finer trophic categories largely agree that lower-quality feeding guilds (herbivores/detritivores and sessile invertivores) have higher diversification rates than higher-quality feeding ones. However, in addition to these lower-quality feeding guilds, we found that planktivorous lineages also diversified disproportionately fast. This suggests that it is recently acquired trophic strategies, rather than low-quality feeding per se, that may have opened up opportunities for shifts in the pace of lineage origination. These trophic innovations in reef fishes predominantly occurred in the last 20 Myr (Cowman *et al.*, 2009; Lobato *et al.*, 2014; Bellwood *et al.*, 2017), a time that closely matches the highest diversification rates of key lineages in our study (**Fig. 2.1**) and generally across the tree of life (Henaó Diaz *et al.*, 2019). Thus, we suggest that this increased diversification in herbivores, sessile invertivores and planktivores may be explained by ecological opportunities unveiled by fundamental changes in reef configuration occurring during the Miocene.

This geological period was marked by the rise of high-turnover reef ecosystems in which both fast-growing corals and large bioeroding fishes first appeared (Bellwood *et al.*, 2017). These fundamental changes in the dynamics of reef structure likely promoted new opportunities for trophic innovation in fishes (Cowman *et al.*, 2009). Particularly important for the expansion of recently derived trophic groups in the Miocene appears to have been the colonization of reef flats. Evidence from present-day reefs suggest that this habitat is by far the most productive reef zone for benthic organisms, and this is reflected in their yield to grazing fishes (Russ, 2003; Bellwood *et al.*, 2018). However, these shallow areas of the reef are also exposed to high wave energy and fish populations may be shaped by the availability of flow and predatory refuges (Fulton & Bellwood, 2005; Bejarano *et al.*, 2017). Thus, these habitats appear to offer potential benefits but they also present substantial challenges for most fishes (Fulton & Bellwood, 2005; Bejarano *et al.*, 2017). In evolutionary terms, although some typical herbivorous reef fish families arose and expanded in the Paleocene-Eocene (66–33.9 Ma) (Siqueira *et al.*, 2019a [*chapter 5 in this thesis*]), it was not until the Miocene that they acquired necessary body and fin morphologies to move into this challenging reef zone (Bellwood *et*

*al.*, 2014a; Siqueira *et al.*, 2019b [chapter 4 in this thesis]). This apparently simple move may have driven profound trophodynamic changes in shallow reefs (Bellwood *et al.*, 2018), which might help explain our results.

There are three key components to this explanation. First, the colonization of the productive reef flats probably allowed herbivorous fishes to expand their population sizes (Bellwood *et al.*, 2018). Second, the intense grazing pressure promoted by large herbivorous populations may have facilitated the expansion of corals in shallow waters, by altering the coral–algal competitive balance (Bellwood *et al.*, 2017). This is supported by the paleontological evidence, which suggests that, despite some peripheral scleractinian reef formation in the Eocene (Wallace & Rosen, 2006), the rise of modern scleractinian-dominated reefs only took place in the late Miocene (Wallace & Rosen, 2006; Mihaljević *et al.*, 2014; Santodomingo *et al.*, 2015). Finally, once corals dominated shallow waters, they had the capacity to promote the expansion of sessile invertivorous and planktivorous lineages. In the case of sessile invertivores, this expansion was likely related to more opportunities for resource exploitation, given the general increase in the availability of both shelter and the abundance of organisms exploited as food sources (Bellwood *et al.*, 2017). In planktivores, the expansion was potentially linked to the shelter provided by topographical complexity against predators and water flow in highly-productive and highly-hydrodynamic shallow reef environments (Johansen *et al.*, 2008). Although intense hydrodynamics might offer a constant flow of planktonic resources, without the refuge provided by corals these shallow reef habitats would probably be uninhabitable for many planktivorous species (Morais & Bellwood, 2019).

This hypothesized scenario provides not only a logical explanation for the observed diversification rates among trophic groups, but it also helps elucidating the patterns found for large body sized herbivores. To meet metabolic demands, herbivorous fishes have to maintain higher feeding rates when compared to other trophic groups (Choat, 1991). However, by doing this, these fishes become more exposed to predation (Hay, 1981). As body size is a major determinant of

predation risk in reef fishes (Goatley & Bellwood, 2016), being large may provide herbivorous fishes with a size refuge from predation. Consequently, the colonization of the reef flats in the Miocene might have been particularly beneficial for large-bodied herbivorous fishes. This is because they were free to maintain high grazing rates on highly productive reef flats, while avoiding the typically high predation pressure in these habitats (Bellwood *et al.*, 2018). Thus, predation could have been a key component in driving differences in the diversification rate of small-medium and large herbivorous fishes. Although speculative, these ideas provide a fertile ground for future studies willing to compare different models of size evolution in herbivorous reef fishes.

In addition to the predation effect, body size has also been shown to correlate positively with geographical range in reef fishes (Luiz *et al.*, 2013), which highlights the importance of this trait for species' long-distance colonization capabilities. Although this might promote genetic connectivity in ecological time scales, in evolutionary time scales it might also increase the chances of vicariance, given the variability in effectiveness of marine biogeographical barriers through time (Bellwood & Wainwright, 2002; Cowman & Bellwood, 2013b; Hodge & Bellwood, 2016) and the likelihood of fragmentation of previously contiguous populations. While body size can be considered a 'universal trait' related to multiple biological processes (Bellwood *et al.*, 2019b), the key element here might be related to use of shallow reef flat habitats. It appears that the remarkably higher diversification rates found for large-bodied herbivores/detritivores (**Fig. 2.2**) was probably related to a combination of higher population sizes, driven by colonization of highly productive reef flats and low mortality, coupled with long-distance dispersal potential within these lineages.

For most trophic identities, our recent and historical approaches provided similar results. However, we found a decoupling between high tip and low historical rates estimated for omnivores (**Figs. 2.2b and 2.4**). This suggests that omnivorous lineages might have experienced only limited rates of origination in the past, counterbalancing recent expansions. Alternatively, this decoupling might be related to the transient nature of omnivory through time, which would result in short-lived

evolutionary lineages within that trophic group. Interestingly, omnivorous lineages have previously been shown to be the slowest evolving groups in both mammals (Price *et al.*, 2012) and birds (Burin *et al.*, 2016). In the latter group, extinction rates were estimated to be even higher than speciation, leading the authors to flag omnivory as a macroevolutionary sink. Although this was not the case for reef fishes in both small (Gajdzik *et al.*, 2019) and large taxonomic scales, low historical rates of diversification in omnivorous lineages seem to be a common pattern in vertebrate evolution.

Herbivores have also been found to be the fastest diversifying lineages in many disparate vertebrate and invertebrate taxa (e.g. Price *et al.*, 2012; Burin *et al.*, 2016; Poore *et al.*, 2017), suggesting that animal trophic evolution might follow common rules. However, to our knowledge, this is the first time that the synergistic effects of species body size and trophic identity have been considered simultaneously when exploring diversification patterns in vertebrates. Body size is regarded as an important component of organismal evolution, given its influence on metabolic rates (Brown *et al.*, 2004) and generation times (Martin & Palumbi, 1993). Thus, our results showing lower diversification rates for smaller bodied species seem counterintuitive considering evolutionary theories. For example, small cryptobenthic fishes contribute to a large proportion of the species richness found on coral reefs (Brandl *et al.*, 2018), and their high population turnover and low connectivity should promote faster rates of diversification (Brandl *et al.*, 2019). Yet, our results show that gobies, for instance, might be amongst the slowest evolving families on coral reefs. We propose two possible explanations for these seemingly counterintuitive results. First, although the fast life history of cryptobenthic fishes should be reflected in rapid diversification, some groups might be experiencing high rates of extinction. Unfortunately, estimating extinction rates from phylogenetic trees can be problematic (Rabosky, 2010), making the test of this hypothesis difficult without a good fossil record. Second, our rates of diversification for cryptobenthic fishes might be underestimated due to taxonomic sampling (judging by the rate of species descriptions for these groups and the expected number of undescribed species; Brandl *et al.*, 2018). While plausible, when we controlled for this effect by removing key cryptobenthic families (Brandl *et al.*, 2018), our trophic results

remained practically unchanged and we still found slightly lower diversification rates for smaller-bodied species (Supplementary Figure 2.3). Nevertheless, this might be an important topic for further investigation in attempts to clarify the relationship between diversification rates and body size in reef fishes.

### 2.5.2 Trophic transitions

In terms of evolutionary trophic pathways, transitions to planktivory have long been recognized as one of the most recurrent patterns in reef fish evolution (Hobson, 1991), with examples occurring consistently across a broad range of families (Floeter *et al.*, 2018). However, our study represents the first effort to quantify this pattern using a large-scale phylogenetic framework. It is also recognized that these shifts are associated with specific morphological and behavioural changes related to food acquisition (e.g. Cooper & Westneat, 2009; Friedman *et al.*, 2016). Despite being unusual in other trophic identities, these morphological modifications (e.g. slender fusiform bodies and deeply forked caudal fins; Floeter *et al.* 2018) associated with planktivory seem to arise frequently, no matter the trophic group of the originating lineage (**Fig. 2.5**). Not surprisingly, reef fish planktivores nested within groups with more generalised morphologies are often described as separate genera due to differences in body and caudal shape, despite only representing a shift to a feeding mode higher in the water column (Floeter *et al.*, 2018). One hypothesis that may explain this pattern is that recurrent transitions to planktivory in adult stages should be an easier evolutionary step compared to other trophic transitions simply because most reef fishes have already been planktivorous in early life stages (Hobson, 1991).

In other recently-derived trophic groups, however, transitions occur almost exclusively from omnivorous lineages (**Fig. 2.4**), a finding that matches previously described patterns in the Labridae and Pomacentridae (Gajdzik *et al.*, 2019). Herbivores/detritivores and sessile invertivores have numerous specific morphological, physiological and behavioural attributes (e.g. Choat & Clements,

1998; Konow & Ferry-Graham, 2013) that are unlikely to be simply acquired in evolutionary terms. Not coincidentally, these trophic identities represent the most taxonomically restricted groups of reef fishes. Thus, as suggested for selected reef fish families (Gajdzik *et al.*, 2019), the pathway to transition within these trophic groups appears to involve an intermediate generalist stage in which lineages have not yet fully developed the biological traits related to the exploration of specific resources. Interestingly, omnivorous reef fishes have been shown to have very slow rates of morphological evolution (Borstein *et al.*, 2019). Alongside our results, this suggests that omnivory might not be an evolutionary stable trophic strategy, rather, it may represent a transitional stage between reef fish trophic groups.

### 2.5.3 Model Considerations

While it has recently been demonstrated that deep temporal trends in speciation and extinction rates cannot be reliably identified from phylogenies containing extant species only (Louca & Pennell, 2020), our study is unlikely to suffer from this issue. This is because our model relies on estimates of very recent diversification rates (tip-rates), which have been shown to be relatively robust to the issues of parameter non-identifiability (Louca & Pennell, 2020). Furthermore, considering the extreme heterogeneity in diversification rates found in reef fishes (**Fig. 2.1**), and the multitude of other potential explanatory variables that were not included in our model, an average of 30% of explained variance can be regarded as a good performance for an intuitively simple model such as ours. Reef fishes have extraordinarily diverse life and evolutionary histories; therefore, it is remarkable that a coarse trophic distinction and maximum species body size alone can explain almost one third of the variability in diversification rates. It is hard to conceive another single factor that could have a higher explanatory power than the ones found herein. Additionally, when we considered only the 'consensus' reef fish families, our model explained an even higher proportion of the variability (36%), with trophic group increasing considerably in importance (55%). This suggests that our diversification rate results

were most strongly associated with the history of coral reefs and not with peripheral environments that also support 'reef-associated' fish species.

#### 2.5.4 Conclusions

Trophic innovations are closely tied to evolutionary rate shifts in reef-associated fishes. Relative to all other trophic groups, herbivorous fishes have sustained remarkably fast diversification rates, a pattern that is particularly pronounced in large body sized species. This combination is likely related to their ecological success after colonizing the productive reef flat during the Miocene. Acting through an evolutionary cascade, the colonization of this zone appears to have triggered profound changes in reef configuration, which in turn underpinned critical trophodynamic shifts and the diversification of other trophic groups. These cascading effects were likely mediated by recurrent transitions between guilds. While planktivory represents a common evolutionary route in reef fish evolution, omnivory might have provided the critical transitional link between higher and lower trophic levels. Overall, our results suggest the existence of a mechanistic basis underpinning the role of trophic evolution in determining macroevolutionary patterns in reef fishes.

## Chapter 3. Macroecology of trophic guilds in coral reef fishes

*This chapter is currently under review in a peer-reviewed journal:*

Siqueira, A. C., Morais, R. A., Bellwood, D. R., Cowman, P. F. Planktivores as trophic drivers of global coral reef fish diversity patterns.

### 3.1 Abstract

One of the most prominent features of life on Earth is the uneven number of species across large spatial scales. Despite being inherently linked to energetic constraints, these gradients in species richness distribution have rarely been examined from a trophic perspective. Here, we dissect the global diversity of over 3600 coral reef fishes to reveal patterns across major trophic groups. By analysing multiple nested spatial scales, we show that planktivores contribute disproportionately to the formation of the Indo-Australian Archipelago (IAA) marine biodiversity hotspot. Besides being 'hotter' at the hotspot, planktivorous fishes display the steepest decline in species numbers with distance from the IAA when compared to other trophic groups. Surprisingly, the evolutionary imprint of this remarkable gradient in planktivorous fish richness could not be detected in extant species phylogenies. Thus, we identify two potential complementary drivers for this pattern. First, exceptional levels of partitioning among planktivorous coral reef fishes were driven by temporally stable oceanographic conditions and abundant planktonic resources in the IAA. Second, extinctions of planktivores outside the IAA have been particularly pronounced during Quaternary climate fluctuations. Overall, our results highlight trophic ecology as an important component of global species richness gradients.

### 3.2 Introduction

The uneven distribution of species numbers across the globe was one of the earliest macroecological patterns to be described (Darwin, 1859). With congruent cases in multiple taxonomic groups, the accumulation of species in lower latitudes is perhaps the most striking example (Hillebrand, 2004). While many contemporary environmental drivers have been proposed to explain such discrepancies (Willig *et al.*, 2003; Fine, 2015), it is widely recognized that present-day richness patterns represent the conclusion of an extensive history of past species origination, extinction and dispersal events (Jablonski *et al.*, 2006). Thus, explaining global biodiversity patterns requires the integration of mechanisms acting over both ecological (survival and coexistence) and evolutionary timescales (lineage persistence and divergence) (Fine, 2015).

Global gradients in terrestrial biodiversity have received considerably more attention than their marine counterparts. Yet, the marine realm hosts one of the most remarkable diversity patterns in the world with a major global biodiversity hotspot (Bellwood & Hughes, 2001; Mora *et al.*, 2003; Tittensor *et al.*, 2010). While latitudinal trends in marine species richness parallel those observed in terrestrial taxa (Willig *et al.*, 2003; Tittensor *et al.*, 2010), the longitudinal accumulation of species in the Indo-Australian Archipelago (IAA) is a unique feature of marine systems (Hoeksema, 2007; Tittensor *et al.*, 2010; Bellwood *et al.*, 2012). This uniqueness stems from the high seascape connectivity and the dispersive nature of the pelagic larval stages of marine organisms (Lindsay, 2012). As a result, the superposition of marine latitudinal and longitudinal gradients forms a very distinct radial bullseye pattern of species distribution, with richness peaking in the IAA.

This bullseye pattern of marine biodiversity is particularly pronounced in coastal habitats (Tittensor *et al.*, 2010), and explanations for its origin have revolved around it being a centre of lineage origination, overlap (or accumulation), and survival (reviewed in Hoeksema, 2007; Bellwood *et al.*, 2012; and Gaither & Rocha, 2013). These explanations are not mutually exclusive and geological, palaeontological and molecular evidence point to a complex and temporally dynamic combination of these alternatives throughout the Cenozoic, leading to the formation of the modern IAA hotspot.

During the Paleocene-Eocene (66–33.9 million years ago [Ma]), the marine biodiversity hotspot was situated in the central region of the extinct Tethys Sea (Renema *et al.*, 2008; Leprieur *et al.*, 2016). At that time, the islands that form the present-day mosaic of the IAA were just starting to emerge (Hall, 2002; Lohman *et al.*, 2011), and it likely served as a peripheral area of species accumulation (Cowman & Bellwood, 2013a). With tectonic changes throughout the Oligocene-Miocene (33.9–5.3 Ma), the region around the IAA became progressively more complex (Hall, 2002; Lohman *et al.*, 2011), whilst biodiversity in the central Tethys started to wane (Renema *et al.*, 2008). As a consequence, new lineages began to originate in the IAA during the Miocene (23–5.3 Ma), giving rise to most of the extant diversity in multiple marine taxa (e.g. Renema *et al.*, 2008; Williams & Duda, 2008; Cowman & Bellwood, 2013; Bellwood *et al.*, 2017; Miller *et al.*, 2018). Finally, in the last three million years the IAA appears to have acted mainly as a centre of survival, protecting species against extinctions including those related to Quaternary climatic oscillations (Pellissier *et al.*, 2014). This historical sequence of events was, therefore, largely responsible for the genesis of the bullseye pattern of marine species richness distribution.

While history is the key element underpinning the origin of global marine biodiversity patterns, ecological factors may also play an important role. Particularly for coral reef fishes, which are the major contributors for the IAA hotspot (Tittensor *et al.*, 2010), the availability of shallow-water habitat area has been repeatedly shown to be an important predictor of species richness (Bellwood & Hughes, 2001; Bellwood *et al.*, 2005; Parravicini *et al.*, 2013). This area effect overrides the predicted mid-domain model of species ranges stacking in a bounded domain within the Indo-Pacific (Connolly *et al.*, 2003; Mora *et al.*, 2003). More recently, species traits have been considered for their role in structuring assemblages, and species maximum body size has also been revealed as a strong predictor of species richness across multiple spatial scales (Barneche *et al.*, 2019). Locations within the IAA tend to have more reef fish species with smaller body sizes (Barneche *et al.*, 2019). In turn, species body size has been correlated with dispersal potential (Luiz *et al.*, 2013), which reinforces the disparities in species numbers between the centre and the periphery of the marine biodiversity hotspot (Donati *et*

*al.*, 2019). Thus, the bullseye pattern of reef fish richness is essentially the culmination of historical processes that resulted in the accumulation and maintenance of small-bodied, low-dispersive species in the IAA.

Although many elements of this story have already been revealed, there is a fundamental component missing from the macroevolutionary narrative: trophic status. The trophic identity of species has recently been shown to be inherently linked to the pace of species formation in coral reef fishes (Siqueira *et al.*, 2020 [*chapter 2 in this thesis*]). Yet, it remains unclear how marine richness gradients are compartmentalized among species with different trophic ecologies. Therefore, to address this knowledge gap, we assess the trophic component of diversity distributions in coral reef fishes. More specifically, we first describe the global patterns of reef fish species richness across major trophic groups. Subsequently, we explore the relationship between guild richness and distance to the centre of marine diversity (IAA) at both global and regional scales, accounting for species body size. Finally, after finding a disproportional accumulation of planktivorous species within the IAA, we investigate the potential evolutionary mechanisms underpinning this pattern. Our results reveal a previously undescribed trophic link to the bullseye pattern of coral reef fish biodiversity distribution.

### 3.3 Methods

#### 3.3.1 Species distribution and survey data

We used two independent datasets of coral reef fish distributions: a global presence-absence record of species in 150 km<sup>2</sup> resolution grids (Rabosky *et al.*, 2018, 2019); and a fish community survey dataset (Edgar & Stuart-Smith, 2014). The presence-absence dataset was downloaded from a publicly available repository (Rabosky *et al.*, 2019), and was built using the AquaMaps algorithm (Ready *et al.*, 2010). The authors estimated geographic ranges of marine fishes based on species occurrence records and a set of environmental predictors (Rabosky *et al.*, 2018). From this presence-absence dataset, we filtered those species that belong to the consensus coral reef fish families (*sensu* Bellwood, 1996). These thirteen families are always found on coral reefs irrespective of their biogeographical location (i.e. Acanthuridae, Apogonidae, Blenniidae, Carangidae, Chaetodontidae, Gobiidae, Holocentridae, Labridae, Lutjanidae, Mullidae, Pomacanthidae, Pomacentridae and Serranidae). Therefore, our focus here was on fish families that universally occur on coral reefs, rather than ‘reefs’ *sensu lato*, to avoid potentially confounding effects of habitat type. Following previous molecular phylogenetic analyses, we considered Caesioninae as a subfamily of Lutjanidae (Miller & Cribb, 2007), and Microdesminae and Ptereleotrinae as part of Gobiidae (Tornabene *et al.*, 2013). Altogether, the families considered here comprise approximately 3600 described species. Based on the geographic ranges of these coral reef fishes within the dataset, we calculated the number of overlapping species per grid cell. Subsequently, we divided the species richness per cell according to the classified trophic groups (see Methods section 3.3.2). Finally, we kept only cells that had at least one species per trophic group in the dataset to avoid distribution extremes where very few species occur. Our final presence-absence dataset consisted of 2800 geographic cells containing the number of species per trophic group along with the respective latitudinal and longitudinal coordinates of the centroid of these cells.

The community survey dataset was downloaded from the publicly available Reef Life Survey website ([reeflifesurvey.com](http://reeflifesurvey.com)). This dataset consists of global fish surveys, systematically collected using standardized methods (Edgar & Stuart-Smith, 2014). Each individual survey (transect) involves an

underwater visual census of fish communities that covers two blocks of 250m<sup>2</sup> each, totalling 500m<sup>2</sup> per survey (Edgar & Stuart-Smith, 2014). We averaged the species counts between these two blocks to get the mean number of species in a final area of 250m<sup>2</sup> per transect. Our goal with this dataset was to assess the richness per transect across the Indo-Pacific, therefore, we downloaded the surveys ranging from the Western Indian Ocean and Red Sea to the Central Pacific islands (**Fig. 3.3A**). After filtering data available in these regions that contained a minimum of four transects per ecoregion, we used the data from 848 sites. To calculate the mean richness per site across trophic groups, we categorized all species recorded in the transects according to our defined guilds (see Methods section 3.3.2). Finally, to be able to explore cross-scale patterns of species distributions, we aggregated individual sites into ecoregions (*sensu* Spalding *et al.*, 2007) and calculated the mean species richness per trophic group in each region. This regional dataset comprised 31 ecoregions containing at least three sites each.

### 3.3.2 Species trait data

We used a previously assembled dataset on reef fish ecological traits (Siqueira *et al.*, 2020 [*chapter 2 in this thesis*]) to assess species-specific trophic identity and body size. The maximum body length (body size) data for each species within this dataset was originally sourced from FishBase (Froese & Pauly, 2019). For the trophic identity, species were grouped into six major guilds: generalized carnivores, mobile invertivores, omnivores, planktivores, sessile invertivores and herbivores/detritivores. These guilds are based on species diets in the adult life stages and have been previously defined in the literature (Mouillot *et al.*, 2014). The major differences, however, are that the original herbivores/macroalgivores group has been merged with the general herbivores/detritivores guild, and that we used a broader categorization for carnivores (Siqueira *et al.*, 2020 [*chapter 2 in this thesis*]). While Mouillot *et al.* (2014) classifies species that feed on larger prey (i.e. fish and cephalopods) as piscivores, we adopt a broader category of generalized carnivores

to include species that feed more generally on larger elusive prey (including larger crustaceans). Our trophic categorization was used in combination with the distribution and survey datasets to calculate the number of species in each guild per geographic cell (presence-absence data), ecoregion and site (survey data). In addition to the richness per trophic group, we also calculated the mean species size per guild in each geographic cell, ecoregion and site using the body size data (Siqueira *et al.*, 2020 [chapter 2 in this thesis]). This body size dataset was then used in our statistical modelling procedures.

### 3.3.3 Statistical analyses

To assess the relationship between guild richness and distance from the centre of marine biodiversity, we first calculated the geographical distance (in km) between each grid cell, region and site to a central point in the IAA (IAAc = Lat 0°, Long 121°). This was done using the function *distHaversine* from the 'geosphere' R package (Hijmans, 2019). Subsequently, we applied negative binomial models to correlate the species richness per grid cell (presence-absence data) with the distance from the IAAc in each trophic guild. These models were built using the *gam* function within the 'mgcv' R package (Wood, 2017), and all accounted for spatial autocorrelation between geographic cells. We also calculated the proportional guild richness per geographic cell, and modelled it against distance from the IAAc using beta regressions implemented in the 'betareg' R package (Cribari-Neto & Zeileis, 2010). For the community survey dataset, we fitted generalized linear models per trophic group against distance from the IAAc, since the richness in this case represents averaged values of multiple transects (site) and sites (ecoregion). Both the site and the ecoregion models were fitted using a gamma distribution for the response variable (mean species richness) with a logarithmic link.

Species body size has recently been demonstrated to be a key predictor of coral reef richness across spatial scales (Barneche *et al.*, 2019). Therefore, all of our models were fitted using the mean body size per sampled area as an interactive factor with distance from the IAAc. Since we were interested in isolating the effect of distance from the IAAc in species richness, we performed our main

model predictions using the mean body size fixed in the estimated value for the cells, regions and sites closer to the IAAC. To calculate this fixed value, we first fitted a LOESS polynomial regression with an  $\alpha$  parameter of 0.7 between mean body size and distance to the IAAC. Then we extracted the first estimated value derived from this relationship and used it in our main model predictions. Finally, to assess the effect of varying body size values in our global model results, we performed predictions using the median, and the 2.5%, 25%, 75% quantiles of the distribution of mean species body size per geographic cell in each trophic group.

### 3.3.4 Phylogenetic comparative methods

After finding a disproportional contribution of planktivores to the IAA biodiversity hotspot (see Results section 3.4), we explored the potential evolutionary mechanisms driving this pattern. First, planktivorous species might have accumulated in the IAA as a result of higher diversification within that area. To test this, we calculated the mean tip diversification rate of planktivorous lineages in each geographic cell. The species-specific diversification values were extracted from the Siqueira *et al.* (2020; [chapter 2 in this thesis]) dataset and were originally calculated using the software BAMM (Rabosky, 2014). Then we fitted a generalized linear model of planktivorous tip diversification rate against distance to the IAAC to assess whether origination was higher in cells within the IAA. This model was fitted using a gamma distribution for the response variable (net diversification rate) with a logarithmic link.

Alternatively, the disproportional accumulation of planktivores within the IAA might have been the result of higher transition rates towards that guild, since transitions to planktivity have been shown to be prevalent throughout reef fish evolution (Siqueira *et al.*, 2020 [chapter 2 in this thesis]). To test this hypothesis, we calculated the proportion of transitions to planktivity between marine biogeographical provinces using stochastic character mappings (Huelsenbeck *et al.*, 2003). Firstly, we categorized all consensus species present in the Siqueira *et al.* (2020; [chapter 2 in this thesis])

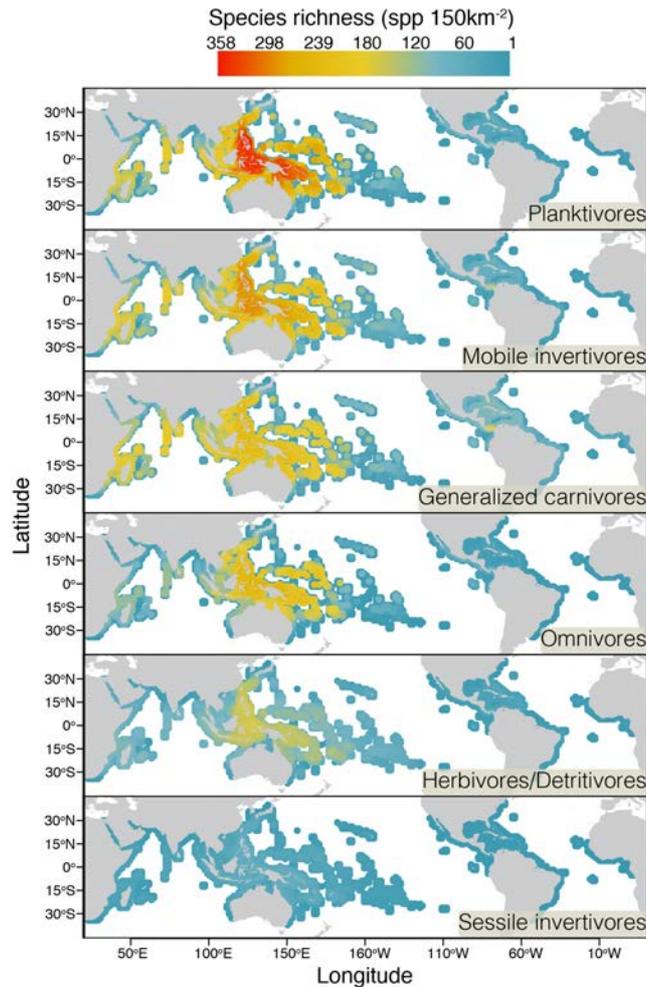
phylogenetic tree according to presence-absence in six biogeographical provinces: Indo-Australian Archipelago (IAA); Central Pacific (CP); Western Indian (WI); Tropical Eastern Pacific (TEP); Eastern Atlantic (EA); Western Atlantic (WA). Then we pruned the phylogeny to contain only species that were present in each province. Finally, for each pruned tree, we simulated 1000 stochastic maps of trophic guilds using a modified version of the *make.simmap* function (Revell, 2012) that considers rate heterogeneity across the tree (Siqueira *et al.*, 2020 [chapter 2 in this thesis]). From the *simmap* results, we calculated the proportion of transitions towards planktivory from the total trophic transitions.

Lastly, the planktivorous fish hotspot might have resulted from an accumulation of lineages via dispersal into the IAA. We assessed this hypothesis by applying the GeoSSE model (Goldberg *et al.*, 2011), within the '*diversitree*' R package (FitzJohn, 2012). This model allows the estimation of dispersal rates associated with geographical states along a phylogenetic tree. Therefore, we built an unconstrained GeoSSE model to calculate dispersal rates out and into the IAA, considering the presence or absence of species within that area. This model was applied to a phylogenetic tree that was pruned to only contain planktivorous species. We used the resulting model coefficients to implement the Bayesian framework of '*diversitree*' and sample the posterior distribution of dispersal parameters. We ran 4000 iterations of the MCMC chain with exponential priors from a preliminary run of 100 iterations. Finally, we eliminated initial 10% samples as burn-in and assessed convergence through the effective sample sizes.

### 3.4 Results

We found remarkable disparities in the distribution of coral reef fish species across trophic groups. In the thirteen consensus families examined (i.e. families that occur universally on coral reefs; see Methods section 3.3.1), our global presence-absence dataset revealed species richness to be highest in the IAA in all groups (**Fig. 3.1**). However, the absolute number of planktivorous reef fish species exceeds by far those in other trophic groups in grid cells (150 km<sup>2</sup>) around the IAA (**Fig. 3.1**). While over 350 species of planktivores can be found in most IAA grid cells, no other trophic group

exceeds 300 species per cell. Indeed, the group with the second highest number of species per cell (mobile invertivores) has approximately 20% less species in the richest cells when compared to planktivores. This accumulation of planktivorous fish species is particularly pronounced around the Philippines, Indonesia, Papua New Guinea and Solomon Islands (**Fig. 3.1**).



**Figure 3.1 | Global coral reef fish richness per trophic group.** Maps show the absolute number of species per grid cell ( $n = 2800$ ) in each classified trophic group. Cell colours correspond to the scale bar and range from low (blue), to intermediate (yellow), and high (red) richness. Grid cell resolution is  $150 \text{ km}^2$  (see Rabosky et al. 2018).

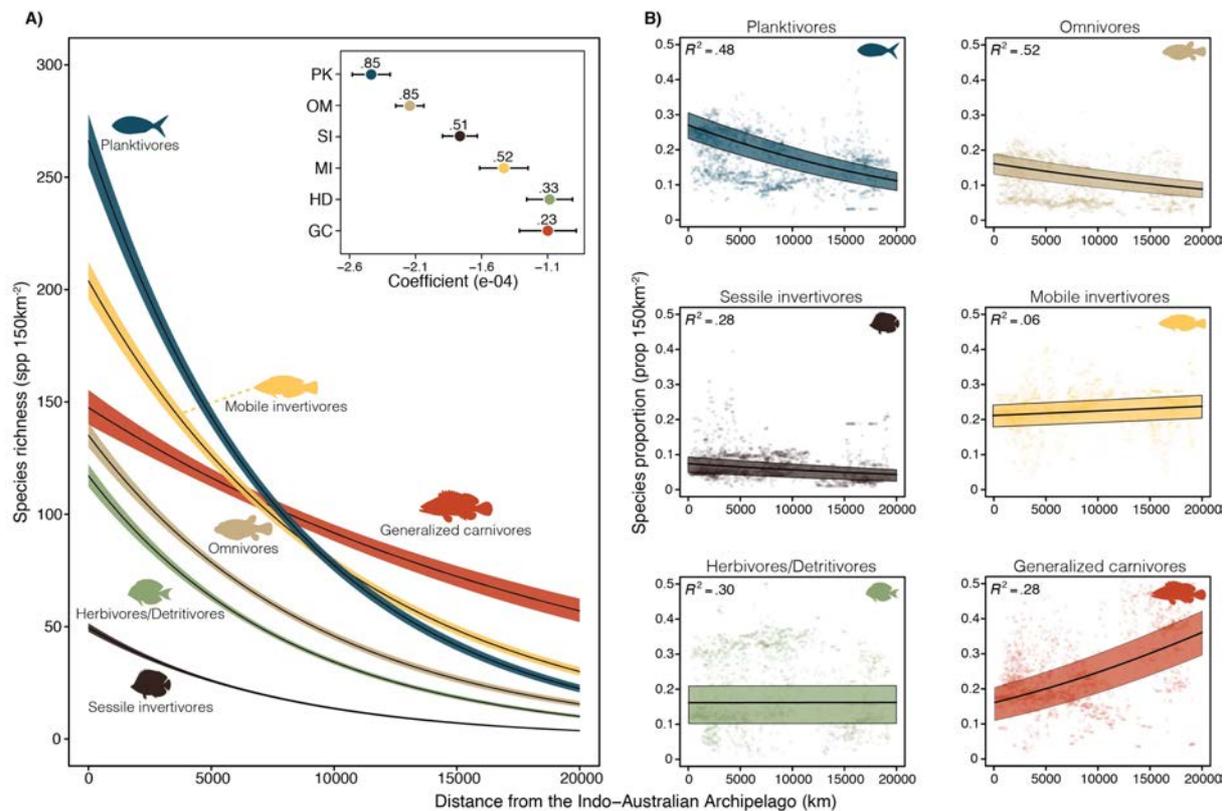
When we modelled species richness per trophic group against the distance to a central point in the IAA (IAAc = Lat  $0^\circ$ , Long  $121^\circ$ ; see Methods section 3.3.3), the disproportional contribution of planktivores to the bullseye pattern became even more evident. As the distance from the IAAc

increases, all trophic groups decrease in species richness (**Fig. 3.2A**; Supplementary Figure 3.1). However, planktivores display the steepest decline (**Fig. 3.2A**), with a significantly more negative slope than any other trophic group (**Fig. 3.2A inset**). Remarkably, this intuitively simple model including only the distance from the IAA and the mean species body size per grid cell was capable of explaining 85% of the global variance in planktivore richness. Besides planktivores, only omnivores had such a high model fit (with a less negative slope), while mobile invertivores, generalized carnivores, herbivores/detritivores and sessile invertivores had 50% or less explained variance (**Fig. 3.2A**).

Model results were consistent when we used the proportion of species per trophic group as the response variable. Planktivores comprised around 27% (23% – 30%; interquartile range) of species in IAA cells, a ratio that decreases steeply as one moves away from the IAAc (**Fig. 3.2B**). No other trophic group had such a high predicted proportion in cells close to the IAA and, once again, planktivores and omnivores had the highest model fits (48% and 52% of explained variance, respectively). Interestingly, despite the low amount of explained variance, generalized carnivores presented an inverse trend with an increasing proportion towards the most peripheral cells, while herbivores/detritivores and mobile invertivores seem to maintain even species proportions globally (**Fig. 3.2B**).

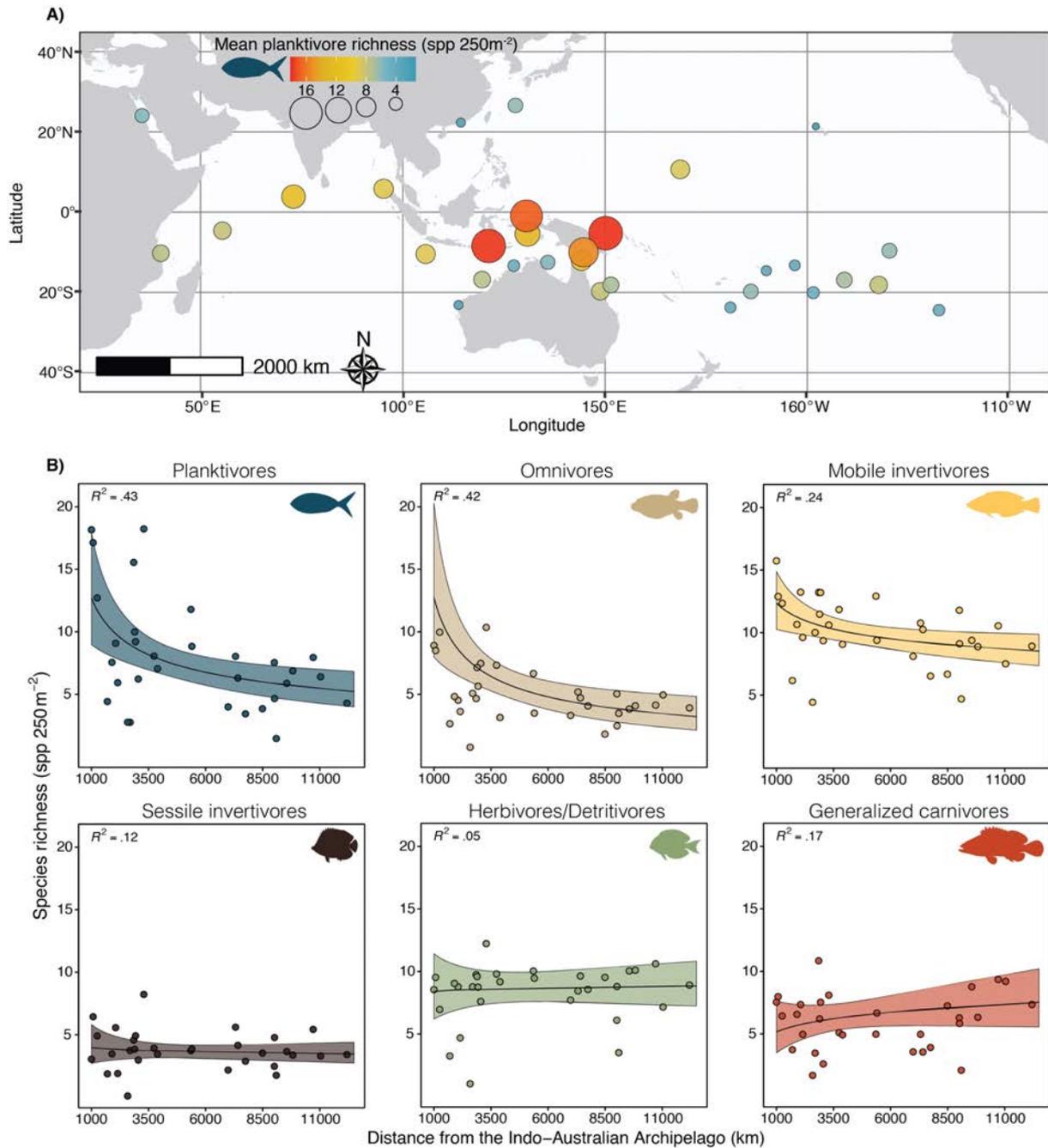
It is important to note that the mean species body size per grid cell contributed substantially to model performance. For instance, if we exclude body size as an interactive factor with distance from the IAAc in planktivores, the proportion of explained variance drops from 85% to 34% in our main model. However, this effect is not limited to planktivores. Across all trophic levels, mean species body size tends to be lower in cells closer to the IAAc (Supplementary Figure 3.2). Therefore, with the exception of herbivores/detritivores, species richness is predicted to be higher in cells that have lower mean species body sizes and are closer to the IAAc (Supplementary Figure 3.3). Since we were interested in isolating the effect of the distance from the IAA, we predicted richness values in our main model (**Fig. 3.2A**) using the body size fixed at the mean value for the cells closer to the centre of the

IAA (see Methods section 3.3.3). If anything, these model predictions are conservative given that the fixed mean cell body size value was still above the mean species body size value per trophic group.



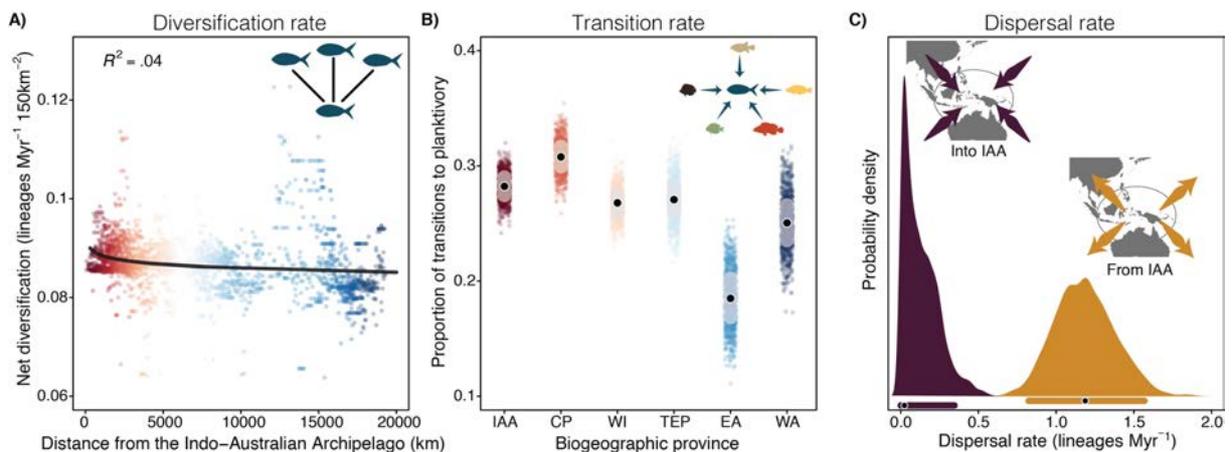
**Figure 3.2 | Global coral reef fish species richness (A) and proportion (B) per guild with distance from the Indo-Australian Archipelago (IAA).** (A) Number of species per grid cell (mean [black line]  $\pm$  95% confidence interval [polygons]) predicted from a negative binomial model per trophic group. The inset displays the model coefficient ( $\pm$  95% confidence interval) per trophic group along with the  $R^2$  value of each model. Lines in the main figure represent the interaction between mean body size and trophic group and, thus, their perceived inclination may not match the size-independent coefficient represented on the inset. (B) Proportion of species per grid cell in each trophic group (mean [black line] and interquartile range [polygons]) predicted from beta regression models. Respective pseudo- $R^2$  values are shown at the top-left corners. Semi-transparent dots represent sampled grid cells ( $n = 2800$ ). Predictions from all models were performed with body size fixed in the estimated value for the cells closer to the centre of the IAA. Trophic groups – GC = generalized carnivores (red); HD = herbivores/detritivores (green); MI = mobile invertivores (yellow); OM = omnivores (beige); PK = planktivores (blue); SI = sessile invertivores (brown).

When we analysed an Indo-Pacific fish survey dataset (848 sites in 31 ecoregions; see Methods section 3.3.1), we detected very similar trends. The mean planktivorous fish richness detected per visual census tended to be substantially higher within the IAA when compared to peripheral sampling locations (**Fig. 3.3A**). Furthermore, only visual censuses performed within the IAA (i.e. less than 3,000 km from the IAAc) contained means of 16 to 18 planktivorous species per 250 m<sup>2</sup>, which is almost double that of most other regions (**Fig. 3.3A**). By modelling species richness against distance from the IAAc, we found results that were highly consistent with our global analysis. Besides presenting better model fits, planktivores and omnivores were predicted to have steeper negative slopes with distance from the IAAc than other trophic groups (**Fig. 3.3B**). Finally, the results were similar when we analysed the dataset at the scale of individual sites (Supplementary Figure 3.4), as opposed to using the sites combined into regions. Alongside our global analysis, these results reveal a robust cross-scale pattern of accumulation of planktivorous species in the IAA.



**Figure 3.3 | Coral reef fish species richness at the regional scale per trophic group. (A)** Mean planktivorous fish species richness per transect in 31 ecoregions (sensu Spalding et al. 2007) across the Indo-Pacific. Ecoregions comprise of multiple aggregated values from sites (see Methods section 3.3.1), with each site sampled using standardized fish counts covering 250m<sup>2</sup> in area (Reef Life Survey; Edgar & Stuart-Smith, 2014). **(B)** Regional-level mean species richness per transect for each trophic group (from visual surveys, points) at increasing distances from the Indo-Australian Archipelago (IAA). Curves show predictions from a generalized linear model (mean [black line]  $\pm$  95% confidence interval [polygons]) with respective  $R^2$  values (top-left corner). Model predictions were performed with body size fixed in the estimated value for the regions closer to the centre of the IAA. Trophic groups – Generalized carnivores (red); Herbivores/Detritivores (green); Mobile invertivores (yellow); Omnivores (beige); Planktivores (PK; blue); Sessile invertivores (brown).

Finally, we assessed the potential evolutionary mechanisms driving the global species richness pattern in planktivores. First, we found that net diversification rates (speciation minus extinction, as calculated from a near-complete reef fish phylogeny; see Methods section 3.3.4) did not present any geographic signal with distance from the IAAc (**Fig. 3.4A**; Supplementary Figure 3.5). Although there seems to be a slight increase in diversification in cells close to the centre of the IAA (**Fig. 3.4A**), the model fit was very low, suggesting that diversification differences alone would not be sufficient to explain observed richness patterns. Second, the proportion of transitions towards planktivity throughout reef fish evolution was not higher in the IAA when compared to other biogeographical provinces (**Fig. 3.4B**). Lastly, rates of dispersal in planktivorous lineages were found to be substantially higher from the IAA towards other regions (**Fig. 3.4C**), which is the exact opposite of what would be expected under a scenario of species accumulation via dispersal. Altogether, these results suggest that the evolutionary mechanisms underpinning the accumulation of planktivorous species in the IAA cannot be revealed by the signals detectable through the use of extant species molecular phylogenies alone. Hence, environmental factors, mediated through differential extinctions are likely to be involved.



**Figure 3.4 | Potential evolutionary mechanisms underpinning the richness gradient in coral reef fish planktivores. (A)** Mean net diversification rate per geographic cell ( $n = 2800$ ) in planktivorous lineages against

distance from the Indo-Australian Archipelago (IAA). Black line shows prediction from a generalized linear model with respective  $R^2$  value (top-left). **(B)** Proportion of transitions towards planktivory per biogeographical province. Black points show mean proportions and grey shades the interquartile range. Semi-transparent points represent individual simmap simulations per province ( $n = 1000$ ). Provinces: Central Pacific (CP); Western Indian (WI); Tropical Eastern Pacific (TEP); Eastern Atlantic (EA); Western Atlantic (WA). Colours in (A) and (B) depict a gradient of distance from the IAA, with red shades representing locations closest and blue shades farthest from the IAA. **(C)** Dispersal rates from (orange) and into (purple) the IAA in planktivorous reef fish lineages. Black points underneath the posterior distributions represent modal values with respective 95% credibility intervals ( $n = 3600$  iterations).

### 3.5 Discussion

Multiple, independent, lines of evidence suggest that one of the most remarkable gradients in species richness distribution in the world has a strong trophic signal. Although the IAA bullseye pattern of marine species richness has been described for over fifty years (reviewed in Bellwood *et al.*, 2012), to our knowledge, this is the first time that trophic characteristics have been analysed concomitantly with global distribution and survey data to uncover patterns in a speciose vertebrate assemblage. Our analyses revealed that, while all coral reef fish trophic guilds contain more species in the IAA, planktivores contribute disproportionately to the foundation of this hotspot. This disproportional contribution is consistently recovered across sampling scales and when species proportion per trophic group is considered. Interestingly, although evolutionary processes (speciation, extinction and dispersal) form the fundamental basis for disparities in species richness globally (Jablonski *et al.*, 2006), our phylogenetic analyses did not provide support for any single evolutionary mechanism underlying the planktivore-based hotspot. This suggests that either: (i) the drivers of the disproportional richness of planktivorous species in the IAA are not detectable through extant species phylogenies alone (cf. Quental & Marshall, 2010; Louca & Pennell, 2020); or (ii) the pattern has an ecological basis. Below we argue that the explanation may lie in a combination of the two alternatives.

### 3.5.1 Geological and oceanographic drivers

The Indo-Australian Archipelago is known to be the most complex and dynamic geological region in the tropics (Hall, 2002; Lohman *et al.*, 2011). Hence, it is widely hypothesized that the exceptional accumulation of marine species within the IAA (Tittensor *et al.*, 2010) is ultimately linked to this geological complexity (Bellwood *et al.*, 2012). For instance, the large shallow water habitat area in the IAA has been consistently demonstrated to be a strong predictor of species richness in coral reef fishes (Bellwood & Hughes, 2001; Bellwood *et al.*, 2005; Parravicini *et al.*, 2013; Barneche *et al.*, 2019). However, we show that the effect of the IAA is disproportionately pronounced in planktivorous fishes, which suggests that simple species  $\sim$  area relationships provide only a partial explanation. While the geological complexity of the IAA may indeed provide larger coral reef habitat area (Parravicini *et al.*, 2013), it also promotes highly dynamic oceanographic conditions (Gordon & Fine, 1996; Gordon, 2005) which might help explaining the high planktivorous fish richness found therein.

The first important element of this oceanographic explanation is associated with the constancy of planktonic resource availability within the IAA. Planktivorous coral reef fishes are heavily reliant on allochthonous food sources that are brought to the reef by complex water movements related to oceanographic currents and tidal regimes (Hamner *et al.*, 1988; Hobson, 1991). Although these particle transportation processes tend to operate mostly on reef-scale topographic features (Hamner *et al.*, 1988), it is likely that the wider-scale oceanographic dynamics in the IAA promotes a constant input of resources for planktivorous fishes. Besides supporting an intense flow of water driven by the strong exchange between Pacific and Indian Ocean waters (Gordon & Fine, 1996; Gordon, 2005), the IAA is also under the influence of strong upwelling systems and tidal regimes (Robertson & Field, 2005). The IAA may thus provide a constant flow of abundant planktonic resources. However, these energetic inputs could only explain the disproportional planktivorous species richness in the IAA if they had been maintained through time (Fischer, 1960; Sanders, 1968). Geological evidence provides support for this hypothesis.

Although the initial tectonic history of the IAA dates back to the Eocene (56–33.9 Ma), its highest geomorphological complexity was only achieved around five million years ago (Hall, 2002; Lohman *et al.*, 2011). Thus, it is reasonable to infer that the major oceanographic processes directly related to the geological features of the archipelago were already in place at that time. This suggests that, despite obvious variations related to sea-level changes (Hoeksema, 2007), the large-scale oceanography of the IAA has remained relatively constant over the last five million years. Not coincidentally, this date matches the mean divergence time of extant species in coral reef fishes (Hodge *et al.*, 2014). More importantly, however, for planktivorous reef fish species, this means a five-million-year period with an almost uninterrupted flow of food particles.

#### 3.5.2 Ecological drivers

On the ecological side, there is clear evidence that planktivorous fishes partition the abundant planktonic food resources in multiple ways. First, reef fish planktivores display clear within-reef spatial distribution patterns (Hamner *et al.*, 1988; Hobson, 1991), with remarkable composition heterogeneity. This heterogeneity is often associated with morphological features that allow some species to benefit from high availability of larger zooplankton in forereef habitats, while dealing with increased hydrodynamics or predation pressure (Hobson & Chess, 1978; Hobson, 1991). Second, planktonic resources are partitioned in time (Hobson, 1991). Whilst most planktivorous fish groups are diurnal (e.g. Pomacentridae, Serranidae, Labridae and Caesioninae), two speciose families are predominantly nocturnal (Apogonidae and Holocentridae). Finally, planktivorous reef fishes may exhibit strong partitioning depending on the resources being targeted. For instance, fairy wrasses (genus *Cirrhilabrus*) appear to target predominantly gelatinous material, in contrast to crustacean zooplankton that is targeted by other planktivorous species within the Labridae (Huertas & Bellwood, 2020). Taken together, these lines of evidence suggest that the disproportional amount of

planktivorous fish species in the IAA might be the result of successful partitioning of constant and abundant resources.

Evidence from productivity patterns on present-day coral reefs provide further support for the ecological drivers of the IAA planktivorous fish hotspot. In a recent analysis across coral reef fish trophic pathways, Morais & Bellwood (2019) revealed that water column-derived productivity may surpass the productivity of any other trophic pathways explored by reef fishes. Although this work was performed on a single coral reef, these pelagic subsidies sustaining high planktivorous fish productivity appear to be widespread, as evidenced by the importance of planktivores to the fish biomass reported by other large-scale studies, particularly in the IAA (Campbell *et al.*, 2020; Heenan *et al.*, 2020). Alongside the parallel evolution of morphological features permitting water column usage in multiple independent reef fish lineages (Floeter *et al.*, 2018), our results indicate that planktivory is a successful evolutionary and ecological strategy, provided that a constant supply of planktonic resources can be maintained. Hence, the patterns described herein agree with longstanding ecological hypotheses that correlate species diversity and coexistence with resource availability, temporal stability, productivity and niche partitioning (Connell & Orias, 1964; MacArthur, 1970).

These resource-related factors appear to provide a compelling case for the differences in species richness between planktivores and other reef fish trophic groups within the IAA. However, explaining the disproportional drop in planktivorous species with distance from the IAA hinges on the understanding that the mechanisms driving such declines are unlikely to be detectable in extant species phylogenies. Past research has shown that the distance from stable coral reef habitats during Quaternary climate fluctuations (last 2.6 million years) outweighs present-day environmental factors in explaining global reef fish richness patterns (Pellissier *et al.*, 2014). This highlights the potential role of areas that maintained suitable coral reef habitat over geological time as extinction refugia for fishes (Pellissier *et al.*, 2014). Our results provide an analogous productivity-based scenario. Given that most

historical coral reef refugia lie within the IAA, it seems likely that planktivorous fishes have been disproportionately affected by extinction in areas away from the IAA. In other words, planktivorous species distributions point strongly to differential survival within the IAA vs peripheral locations during the last five million years. Thus, reef refuges and historically stable oceanographic conditions in the IAA provided constant habitat and food resources for planktivorous fishes through time. This scenario helps to explain why we did not find a geographic signal in the diversification rates of planktivores (**Fig. 3.4A**). The pattern was probably driven by extinction and extinction rates are virtually impossible to estimate without a detailed fossil record (Quental & Marshall, 2010). Finally, this theory also supports our results for transition rates (**Fig. 3.4C**): the IAA served as a source of surviving planktivorous lineages that recolonized depauperate peripheries after extinction events.

In conclusion, our study highlights a unique association between large-scale marine diversity gradients and species trophic identities. Splitting the global coral reef fish species distribution data between trophic groups revealed that planktivores are major contributors to the disparate richness within the IAA biodiversity hotspot. This is likely related to the persistent oceanographic conditions promoted by the geological complexity of the IAA over the last five million years. By providing an abundant and constant flow of food through time, this oceanographic setting fostered high levels of resource partitioning among planktivorous reef fishes within the IAA. Peripheral regions, on the other hand, almost certainly experienced periods of intense extinction of planktivorous fish lineages associated with habitat loss and oceanographic changes. Despite having been recolonized by some surviving lineages from the IAA, these peripheral regions appear to carry the imprint of past extinctions within planktivorous coral reef fishes.

## Chapter 4.

### Evolution of traits and functions in herbivorous coral reef fishes

*This chapter is published as:*

Siqueira, A. C., Bellwood, D. R., Cowman, P. F. (2019). The evolution of traits and functions in herbivorous coral reef fishes through space and time. *Proceedings of the Royal Society B: Biological Sciences*, 286, 20182672. doi: 10.1098/rspb.2018.2672

#### 4.1 Abstract

Herbivory by fishes has been identified as a key ecological process shaping coral reefs through time. Although taxonomically limited, herbivorous reef fishes display a wide range of traits, which results in varied ecosystem functions on reefs around the world. Yet, we understand little about how these trait combinations and functions in ecosystems changed through time and across biogeographical realms. Here, we used fossils and phylogenies in a functional ecological framework to reveal temporal changes in nominally herbivorous fish assemblages among oceanic basins in both trait space and lineage richness among functions. We show that the trait space occupied by extant herbivorous fishes in the Indo-Pacific resulted from an expansion of traits from the ancestral Tethyan assemblages. By contrast, trait space in the Atlantic is the result of lineage turnover, with relatively recent colonization by lineages that arose in the east Tethys/Indo-Pacific. From an ecosystem function perspective, the Atlantic supports a depauperate fauna, with few extant herbivorous reef fish lineages performing each function. Indo-Pacific fishes support both more functions and more lineages within each function, with a marked Miocene to Pleistocene expansion. These disparities highlight the importance of history in explaining global variation in fish functional composition on coral reefs.

## 4.2 Introduction

The diversity of extant ecological systems results from complex interactions between biotic and abiotic factors acting in space and time. While abiotic processes tend to influence biodiversity at larger spatial and temporal scales (Benton, 2009), biotic interactions are increasingly being recognized as important drivers of evolutionary change at smaller scales and in high-diversity systems (Jablonski, 2008). As a consequence, our ability to describe how diversity has been built up through time depends on integrating distinct, but not independent, sources of information. This realization has stimulated recent calls to unite paleontological and neontological data into a single comprehensive framework that facilitates a more integrative approach to biodiversity research (Fritz *et al.*, 2013; Price & Schmitz, 2016). At the core of this integration is the focus on a species' function rather than its taxonomic identity (Price & Schmitz, 2016), reflecting its role in ecosystem dynamics (Mouillot *et al.*, 2011) and evolution (Vermeij, 1977). Species functional diversity is not only an important metric that can illuminate changes in ecosystem functioning through time (Mouillot *et al.*, 2013), it can also reveal insights into larger scale biogeographical processes (Violle *et al.*, 2014).

Although it is clear that integrating fossil and extant species information in a trait-based framework can promote a better understanding of biodiversity assembly through time (e.g. Villéger *et al.*, 2011; Bellwood *et al.*, 2014a), there have been few attempts, to date, to frame it in a historical biogeography context. Functional biogeography (*sensu* Violle *et al.*, 2014) provides an important framework for analysing large scale patterns in trait combinations (Whittaker *et al.*, 2014; Toussaint *et al.*, 2016). However, the role of history remains largely unexplored, despite promising insights from high-diversity systems, such as coral reefs (Hemingson & Bellwood, 2018). These high diversity environments have been shaped through time by major geological events that have left detectable traces in extant reef assemblages (Pellissier *et al.*, 2014), fossil deposits (Renema *et al.*, 2008) and molecular phylogenies (Cowman & Bellwood, 2013a; Cowman *et al.*, 2017).

While biodiversity in reef environments has a long history (Wood, 1999), the modern incarnation of coral-dominated reef systems is largely restricted to the last 60 million years (Myr)

(Bellwood & Wainwright, 2002; Bellwood *et al.*, 2017). Throughout this time, fish and coral assemblages have displayed marked functional changes (Bellwood *et al.*, 2017). One of the most notable early changes was the expansion of herbivorous fishes (Bellwood, 2003). Today, this group plays a critical role in reef ecosystems, mediating the competitive balance between corals and algae, and contributing to the resilience of coral reefs (Bellwood *et al.*, 2004). Without this top-down control, coral reef environments can shift to algal-dominated states (Hughes *et al.*, 2007; Goatley *et al.*, 2016) with detrimental effects on associated biodiversity (Wilson *et al.*, 2006). Herbivorous fishes also play a pivotal role in carbonate dynamics by reworking and transporting calcareous sediments (Bellwood & Choat, 1990). Thus, documenting how herbivorous fish traits and their associated ecosystem functions evolved through time and geographical space is essential if we wish to understand variation in ecosystem functions in present-day ecosystems.

The history of herbivorous reef fishes is tightly associated with the formation of modern coral reef systems (Bellwood *et al.*, 2017). Before the early Cenozoic, there is no clear evidence of herbivorous marine vertebrates (Steneck *et al.*, 2017). Herbivory was performed primarily by invertebrates (Steneck, 1983). The Eocene fossil deposits of Monte Bolca in northern Italy, therefore, provide the first evidence of a fundamental change in the evolution of coral reef assemblages (Bellwood, 1996); the first unequivocal shift in fish-benthos interactions (Bellwood, 2003). Ancient surgeonfish and rabbitfish lineages from Monte Bolca laid the foundations for piscine herbivory on coral reefs and, along with later arising parrotfishes (Choat *et al.*, 2012), form the main herbivorous components of present-day coral reefs (Choat, 1991). Although Bolca fossils are extremely valuable for understanding the evolution of herbivorous fishes (Bellwood, 2003; Bellwood *et al.*, 2014b,a), it remains a unique assemblage. As a consequence, the only way to trace the functional history of herbivorous fishes on coral reefs in space and time is by combining fossils, molecular phylogenies and extant species ecology in a holistic framework.

Recent efforts have been made to understand reef fish functional traits (Floeter *et al.*, 2018) and biogeography (Cowman *et al.*, 2017) through ancestral reconstructions, however, an integrative approach is still lacking. Although herbivorous fishes present a vast combination of traits and functions on coral reefs (Green & Bellwood, 2009), their assemblages are unevenly distributed among biogeographical realms, with broad implications for ecosystem functioning (Bellwood *et al.*, 2004). Unfolding the historical factors that drive this disparity would shed light on the processes that shaped the assembly of essential functions. By combining information from fossils and extant species in a comparative phylogenetic framework, we provide an integrated functional and biogeographical overview of the global evolution of key herbivorous groups on tropical reefs. Furthermore, by specifically separating traits from ecosystem functions, we track: (a) how the herbivorous fish trait space has been occupied through time among major marine biogeographical realms since the Eocene; and (b) the origin of lineages performing each of the key ecosystem functions of herbivorous fishes in space and time as indicated by extant taxa. Our framework provides a sequential view of the formation of one of the most important ecological groups on coral reefs, herbivores.

## 4.3 Methods

### 4.3.1 Chronograms

We built the most comprehensive time-calibrated phylogenies to-date for the three most important nominally herbivorous fish groups on modern coral reefs (Choat, 1991): Acanthuridae (surgeonfishes), Siganidae (rabbitfishes) and Scarini (parrotfishes). The surgeonfishes and rabbitfishes are recognized as separate families, while the parrotfishes are a tribe within the Labridae (Westneat & Alfaro, 2005), containing the subtribes Scarina and Sparisomatina. For each group, we downloaded sequences from GenBank and used Bayesian inferences in BEAST2 (Bouckaert *et al.*, 2014) to construct chronograms. The surgeonfish phylogeny incorporated two mitochondrial (*Cox1* and *Cytb*) and seven nuclear genes (*ENC1*, *myh6*, *plagl2*, *Rag1*, *Rh*, *zic1* and *ETS2*) and comprised 72 species (~90% total diversity) from all extant genera. The rabbitfish phylogeny was based on two mitochondrial genes

(*Cytb* and *16s*) and one nuclear gene (*ITS1*). It contained 24 species (~80% total diversity) in its single genus *Siganus*. Finally, the parrotfish phylogeny was based on five mitochondrial (*Cox1*, *Cytb*, *12s*, *16s* and control region) and six nuclear markers (*Bmp4*, *Dlx2*, *Otx1*, *Rag2*, *S7I1* and *Tmo-4C4*), for 87 species (~87% total diversity) belonging to all extant genera.

The chronograms were built using partitioned analysis, birth-death models with a relaxed lognormal clock prior and fossil calibrations. Details on sequence alignment, model construction, and tree calibration can be found in the Supplementary Methods.

#### 4.3.2 Herbivorous reef fish data

We categorized all species in the phylogenies according to seven traits related to feeding: two categorical morphological traits related to food processing (tooth and alimentary tract morphology), a continuous trait reflecting body size (maximum length), and four categorical behavioural traits (feeding mode, diet, feeding habitat and schooling behaviour), summarizing how, what and where species feed. Trait assignments were based on the literature, online datasets and expert assessments. These seven traits, combined, provide a broad indication of each species' ecological role. However, the relationship between traits and ecosystem functions is complex. While a given function may depend on a single trait state, this is not always the case and, in many instances, functions may be correlated with numerous traits within and among states (Bellwood *et al.*, 2019b). We therefore undertook a separate analysis, assigning ecosystem functions for all species, based on published literature and expert assessments. Details of trait and ecosystem function coding can be found in the Supplementary Methods.

We also assembled a database for the putative herbivorous fossil fish species from the Eocene *Lagerstätten* of Monte Bolca. These fossils from a single locality (2 deposits) in northern Italy, represent the richest and most well-preserved fossil record of present-day reef-associated fish families (Bellwood, 1996). Because no parrotfish fossil has been recorded from the Eocene (Bellwood

& Schultz, 1991), we only included surgeonfishes and rabbitfishes in our fossil database. Moreover, we excluded fossil species described from incomplete or larval-stage specimens. We assigned the same traits to fossils as the extant species by correlating their morphology with modern analogues (*cf.* Brandl & Bellwood, 2013; Bellwood *et al.*, 2014a). We also classified the fossils according to ecosystem functions based on their combination of traits. Most of our fossils show indications of being turf-algae croppers with some potential pair-forming species (Bellwood *et al.*, 2014a), however, others could potentially be macroalgae browsers or planktivores. To account for this uncertainty, we considered the ambiguous fossils as the most distinct states (browsers and planktivores); as a result, our fossil multidimensional space (see Methods section 4.3.4) is most likely overestimated based on available evidence.

All extant species were also classified according to their geographical ranges based on data from the literature and IUCN's Red List (IUCN, 2017). We built a presence/absence matrix of species considering the six recognized biogeographical regions for reef-associated fishes (Kulbicki *et al.*, 2013): Western-Indian (WI), Central-Indo-Pacific (CIP), Central-Pacific (CP), Tropical-Eastern-Pacific (TEP), Western-Atlantic (WA) and Eastern-Atlantic (EA). We also classified the fossil species biogeographically as being present in the Tethys sea, since all were located in the ancient marine biodiversity hotspot in the Eocene (Renema *et al.*, 2008).

#### 4.3.3 Ancestral states & biogeography

We retrieved the ancestral states in lineages present in four time-slices - 20, 15, 10 and 5 Million years ago (Ma) - chosen to encompass the changes that took place in the most important phases for coral reef fish diversification (Bellwood *et al.*, 2017). This was achieved by using the Bayesian framework of BayesTraits (Meade & Pagel, 2018). Within this software, we performed ancestral state reconstructions for each trait in our maximum clade credibility (MCC) trees using the VarRates model (Venditti *et al.*, 2011). This model accounts for heterogeneity in the rates of trait

evolution within the trees. For each discrete trait, we first set multistate models with uniform priors for the transition rates based on results from a maximum likelihood analysis. We then ran three independent MCMC chains of five million iterations each, sampling node state probabilities every 4000 iterations. For traits that had more than four states, we used the rjMCMC option that handles a higher number of parameters (Meade & Pagel, 2018). After discarding 20% burnin, we obtained 1000 samples of node state posterior probabilities (PP) for each run. We assessed convergence within chains using the effective sample size (ESS) scores, and between chains using the marginal likelihoods, calculated using the stepping stone sampler (Xie *et al.*, 2011) in BayesTraits. After ensuring convergence, we used the results of one chain per trait in downstream analysis. The same procedure was used for reconstructing the ecosystem functions. Finally, for the continuous trait (body size), we also used VarRates in BayesTraits with the same MCMC steps as the discrete analysis, but we applied the independent contrasts model (Freckleton, 2012) to assess ancestral state values.

The results from BayesTraits were used to retrieve the ancestral states in the four time-slices (20, 15, 10 and 5 Ma). We first assessed, for each internal node in the phylogenies, the states with the highest modal PP for the discrete traits and the reconstructed modal value for the continuous trait. With these states and values, we built a trait database for the branch points cut by the time-slices in our trees. Each of these intersecting points was classified according to the states of the closest node for the discrete traits and the proportional value of change between the two adjacent nodes for the continuous trait. These analyses were performed using R (R Core Team, 2019). For the ecosystem functions, we applied the same approach as the discrete traits, however, we used more time-slices to assess the presumed origin of each function within ocean basins and to have a minimum estimate of lineages performing them through time based on extant taxa. We can only provide a minimum estimate of lineages in each function because, without the fossil record, molecular phylogenies are prone to a perception bias and the number of species can only increase or saturate (Quental & Marshall, 2010; Marshall, 2017). Specifically, our reconstructions indicate the timing of origin of traits

and functions that have led to the construction of modern faunas. They do not preclude earlier originations in extinct taxa.

To assess the ancestral ranges in our groups, we built biogeographical models using the 'BioGeoBEARS' R package (Matzke, 2013). We used this framework to build models according to time-constraints from the past geological history of marine environments that are well known to influence coral reef fish biogeography (Cowman & Bellwood, 2013a,b). Details of model constraints and selection can be found in the Supplementary Methods. Based on results from range reconstructions, we split the lineages from each time-slice between two major oceanic realms (Atlantic and Indo-Pacific; Supplementary Figures 4.4 – 4.6). Lineages present in the TEP before 3.1 Myr were considered part of the Atlantic realm, given its connection with the WA prior to the closure of the Isthmus of Panama (Cowman *et al.*, 2017). The extant species in our dataset were also biogeographically split between Atlantic and Indo-Pacific realms based on their current ranges.

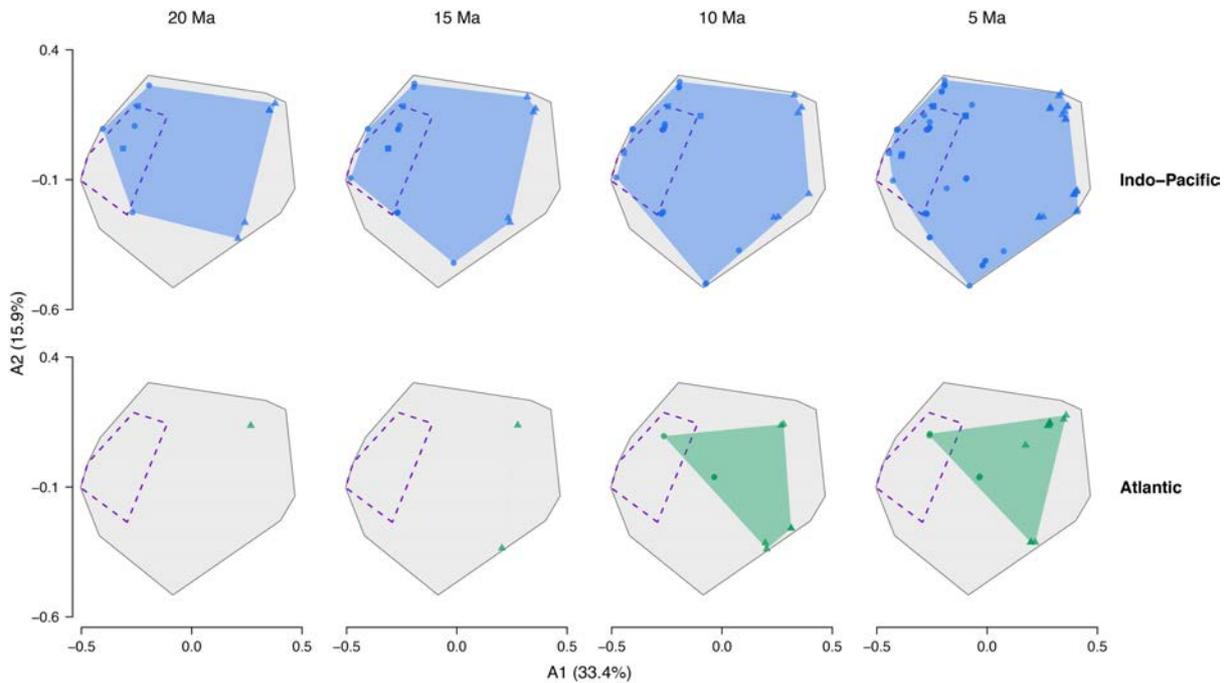
#### 4.3.4 Multidimensional trait space

With the congregated trait dataset (including fossil and extant species, and lineages from time-slices) we were able to plot multidimensional convex hulls for each time period and biogeographical realm. Using the R package 'FD' (Laliberté *et al.*, 2014), we calculated the Gower dissimilarity matrix between all lineages and extracted the axes of the principal coordinate (PCoA) analysis. We then assessed the number of axes that would adequately reflect our trait space using the 'quality\_funct\_space' R function (Maire *et al.*, 2015). As was expected from a mainly categorical space (Maire *et al.*, 2015), the quality of representation increased with the number of axes (Supplementary Figure 4.7), however, we kept only the first four axes for convenience of graphical representation and because they represent more than 70% of the explained variance in the data. Each principal coordinate axis is correlated differently with the classified traits (Supplementary Figures 4.11 – 4.13), therefore, changes in the multidimensional space are better explained by the combination of traits rather than

individual traits. Finally, we assessed the robustness of our results by performing ancestral reconstructions with a recently developed hidden Markov model (Herrera-Alsina *et al.*, 2019) for discrete trait evolution. Moreover, we examined the effects of two issues that could potentially affect our results: phylogenetic uncertainties (topology and node dates); and uncertainty about reconstructed node states. Details of these sensitivity analyses can be found in the Supplementary Methods.

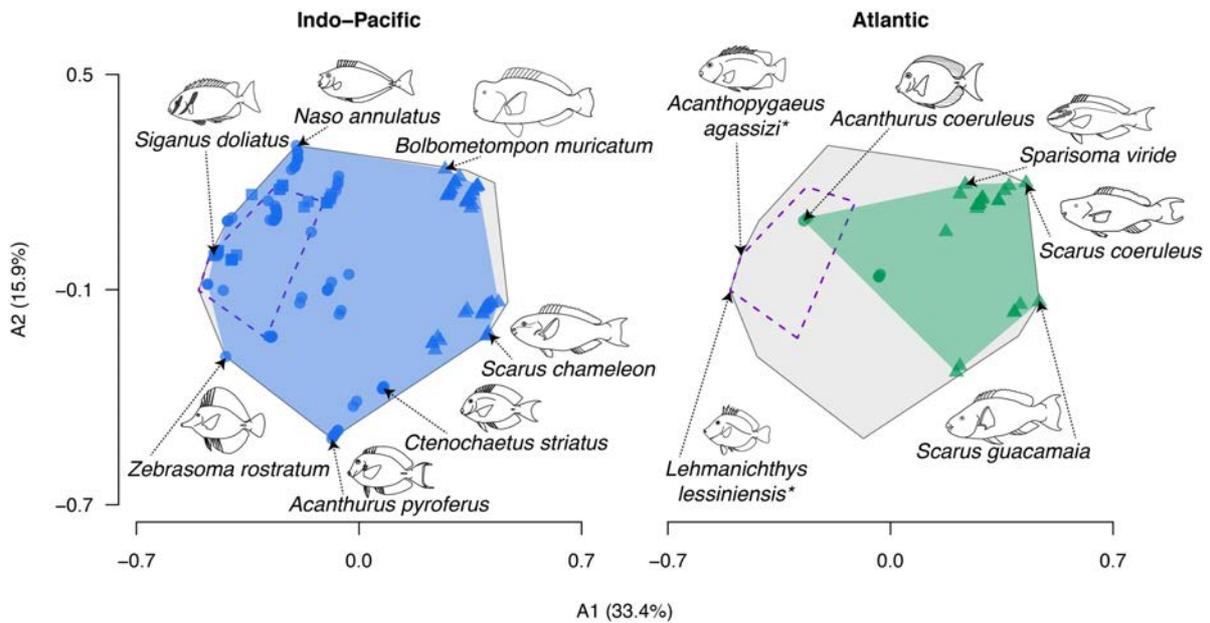
#### 4.4 Results

Our Bayesian phylogenetic inferences yielded well-supported trees for all taxa (Supplementary Figures 4.1 – 4.3). When we performed ancestral state and range reconstructions, we found remarkable differences in the multidimensional trait space occupied through time between biogeographical realms (**Fig. 4.1**). The trait space for fossil species overlaps with lineages retrieved from the Indo-Pacific in the 20 Ma time-slice (**Fig. 4.1**). From this time-slice onwards, the Indo-Pacific realm showed a marked expansion from the ancestral Tethyan trait space. This expansion is associated with the rise of parrotfishes between the Eocene and the Oligocene and the diversification of trait combinations of both surgeonfishes and parrotfishes in the Indo-Pacific in the last 15 Ma (Supplementary Figures 4.5 – 4.6). By comparison, the 20 - 15 Ma Atlantic trait space exhibited a complete turnover when compared to the ancestral Tethyan assemblage (**Fig. 4.1**). The trait combinations retrieved from the Atlantic 20 - 15 Ma were not recorded from the Tethys sea 50 Ma. After 15 Ma, the Atlantic trait space expanded with the initial diversification of Sparisomatina parrotfishes and the dispersal of a surgeonfish lineage from the Indo-Pacific (Supplementary Figures 4.5 – 4.6). This colonization event at 10 Ma represented the return of a trait combination to the Atlantic that is within the boundaries of the Tethyan trait space (**Fig. 4.1**). As a consequence of these historical differences, the total herbivorous reef fish trait space is occupied almost entirely by extant Indo-Pacific lineages, with space occupied by remnant Atlantic and Tethys lineages being largely nested within it (**Fig. 4.2**).



**Figure 4.1 | Multidimensional trait space occupied by surgeonfish (circles), rabbitfish (squares) and parrotfish (triangles) lineages in two biogeographical regions through time.** Plots show the first two axes (A1-A2) derived from a principal coordinate analysis (PCoA) performed on seven traits related to feeding. Each column represents a time-slice (20-5 Ma) in which we assessed the traits through ancestral reconstructions. Background grey area shows the total space occupied combining fossils, time-slices and extant species. Convex hulls represent space occupied by Indo-Pacific (blue), Atlantic (green) and Tethys (50 Ma; purple dashed line) lineages in each time-slice. Symbols represent lineages present in each time-slice.

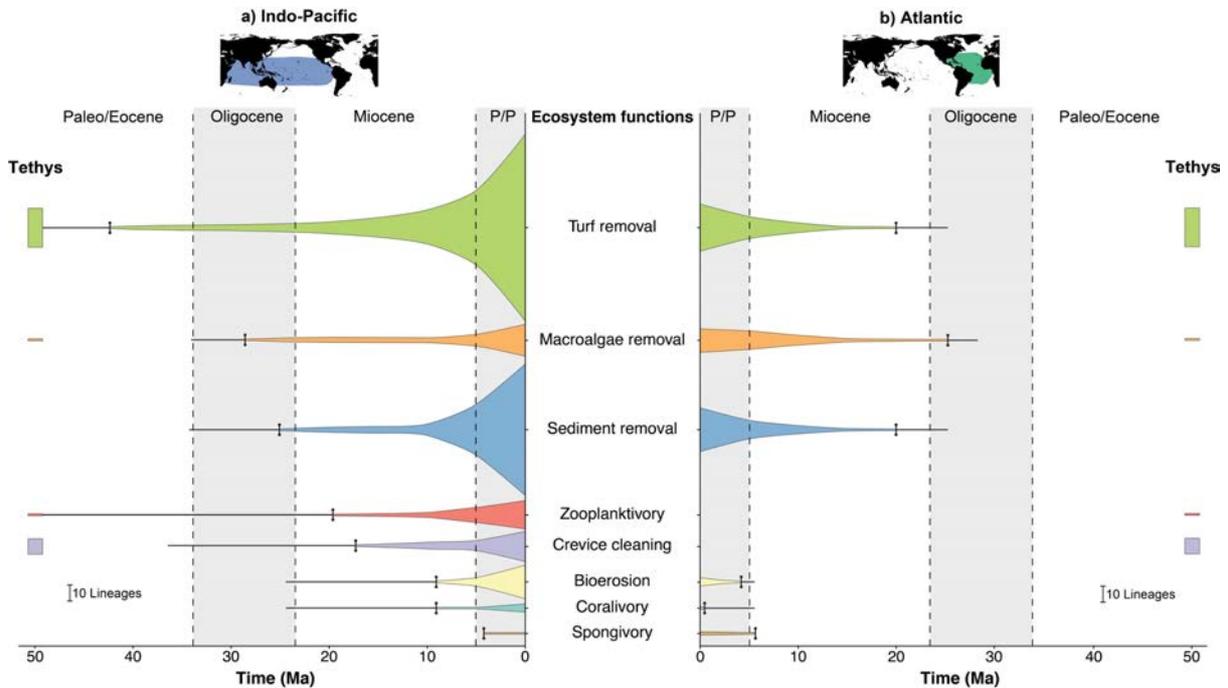
Our sample of posterior trees derived from the phylogenetic analyses showed that uncertainties in node dates and tree topology had very little effect on the patterns of trait space occupation through time (Supplementary Figure 4.14). Moreover, the observed differences between biogeographical realms were consistently recovered even when we accounted for trait-dependent diversification and uncertainties in node states derived from the ancestral reconstructions (Supplementary Figures 4.15 – 4.17).



**Figure 4.2 | Multidimensional trait space occupied by extant surgeonfish (circles), rabbitfish (squares) and parrotfish (triangles) species in two biogeographical regions.** Plots show the first two axes (A1-A2) derived from a principal coordinate analysis (PCoA) performed on seven traits related to feeding. Background grey area shows the total space occupied combining fossils, time-slices and extant species. Convex hulls represent space occupied by Indo-Pacific (blue), Atlantic (green) and Tethys (50 Ma; purple dashed line) species. Illustrations show representatives from each biogeographical realm and two fossil species (\*).

For the ecosystem functions, we identified marked differences between biogeographical realms in the origins of lineages performing each function through time (Fig. 4.3). The dominant functions present in the fossil assemblage were the removal of algal turfs and crevice feeding, which were performed by both surgeonfishes and rabbitfishes. Lineages retrieved from our reconstructions showed older origins for all extant ecosystem functions in the Indo-Pacific when compared to the Atlantic. In the Indo-Pacific, the removal of algal turfs and macroalgae by present-day lineages can be traced back to the Eocene/Oligocene, while the lineages performing other functions probably originated in the Oligocene/Miocene (Fig. 4.3a). By contrast, in the Atlantic, reconstructions suggested that removal of turf, macroalgae and sediment by extant fish lineages started in the early Miocene, while bioerosion and corallivory are even more recent, with origins in the Pliocene/Pleistocene (Fig.

**4.3b).** Besides the absence of rabbitfishes in the Atlantic, this region also lacks herbivorous lineages that transitioned to zooplanktivory or are capable of feeding in crevices.



**Figure 4.3 | Surgeonfish, rabbitfish and parrotfish lineages performing each ecosystem function through time in the Indo-Pacific (a) and Atlantic (b) realms.** Funnel width through time was calculated based on the ancestral state reconstructions for ecosystem functions, therefore, numbers are limited to inferences from extant taxa. The split between regions was based on ancestral range reconstructions. The time of origin of each function (vertical dash) was based on the earliest reconstruction of an ancestral node performing that function, and the line traced from the funnels represent the length of the branch that lead to that node, indicating the possible origin of the function at any point along the branch. Rectangles at 50 Ma represent the number of fossil lineages from Monte Bolca performing each ecosystem function in the Tethys Sea. Dashed lines separate geological epochs (P/P: Pliocene/Pleistocene).

Although the origin of lineages performing each ecosystem function varied among biogeographical realms, we found an expansion of lineages in the mid-Miocene continuing up to the Pliocene/Pleistocene in both the Atlantic and the Indo-Pacific (**Fig. 4.3**). Nevertheless, this expansion was much more pronounced in the Indo-Pacific, particularly for the removal of algal turfs and sediment (**Fig. 4.3a**), which reflects the diversification of the most speciose clades in these functional

groups (e.g. *Scarus* and *Acanthurus*; Supplementary Figures 4.2 – 4.3). In the Atlantic, the Pliocene/Pleistocene expansion in turf and sediment removing lineages (**Fig. 4.3b**) is related to the diversification of *Sparisoma* and *Scarus* parrotfishes, however, these lineages represent less than one third of the number of parrotfish lineages present in the Indo-Pacific. Comparatively, the only ecosystem functions that have a similar number of extant lineages between biogeographical realms are the removal of macroalgae and sponges (**Fig. 4.3**). This reflects the fact that most *Sparisoma* species can feed on both turfs and macroalgae (with some also feeding on sponges), and that macroalgal browsers and spongivores are not very diverse in our focal taxa on Indo-Pacific coral reefs.

By mapping our classified traits with their respective ecosystem functions (Supplementary Figures 4.18 – 4.19), we observed an intricate link between morphology, behaviour and ecology. The removal of turfs, macroalgae and sediment can be performed by multiple trait combinations (Supplementary Figure 4.18) and species occupying similar areas of trait space tend to perform similar ecosystem functions irrespective of biogeography (Supplementary Figure 4.19). However, there are specificities in each biogeographical realm. For example, most scraping parrotfishes in the Atlantic are also capable of removing macroalgae, which does not happen in the Indo-Pacific. There, the removal of macroalgae is exclusively performed by browsing species. Other ecosystem functions such as zooplanktivory and crevice cleaning are performed by lineages with more restricted combinations of traits. Since these functions are absent in lineages from the Atlantic (**Fig. 4.3**), this ocean basin also lacks associated traits (**Fig. 4.1**; Supplementary Figures 4.18 – 4.19). Interestingly, spongivory is the only ecosystem function that is performed by lineages with very distinct trait combinations in different realms (i.e. *Sparisoma* spp. in the Atlantic and *Siganus* in the Indo-Pacific).

## 4.5 Discussion

Through a comprehensive framework combining fossils, phylogenies and the ecology of extant species, we have revealed two very distinct scenarios for the evolution of herbivorous coral reef fishes between major marine biogeographical realms. The Indo-Pacific showed a clear history of

continuity and expansion from the ancestral Tethyan herbivore assemblage. This result uncovered a trait-based component to the historical biogeography of coral reef fishes (Cowman & Bellwood, 2013a), with the fossil assemblage of the Tethys hotspot forming the core foundations of both taxa (Renema *et al.*, 2008) and herbivorous trait combinations in the Indo-Pacific. By contrast, the Atlantic herbivore composition was marked by the isolation from the marine diversity hotspots and the loss of ancient trait combinations. As a consequence, the extant Atlantic reef fish herbivore assemblage is a combination of a clade that arose there and a wider range of more recently derived lineages that have invaded from the east Tethys/Indo-Pacific (Supplementary Figures 4.5 – 4.6). These different histories were also reflected in the lineages performing ecosystem functions within each realm. The Atlantic fauna contains only a small subset of the functions seen in the Indo-Pacific, with few lineages within each represented function. These two components of piscine herbivory in coral reefs, trait combinations and ecosystem functions, will be discussed separately below.

#### 4.5.1 Trait combinations in space and time

The major differences in herbivorous trait combinations between the two oceanic basins can be traced back to the earliest known herbivorous assemblage in marine environments. The absence of morphological features related to herbivory in Mesozoic marine fish fossils indicates that the early Cenozoic was likely a starting point for marine piscine herbivory (Bellwood, 2003). However, it is hard to accurately place this origin in space and time. The first unequivocal assemblage of marine herbivorous fishes is in the Eocene fossils of Monte Bolca (Bellwood, 2003), with many groups already showing striking morphological similarities with their modern counterparts (Bellwood *et al.*, 2014a). Considering the central location of this fossil assemblage and its connections with both the east and the west Tethys (presently Indo-Pacific and Atlantic), it is likely to represent a reference point from which global reef fish assemblages evolved (Bellwood *et al.*, 2017). From this reference point, the subsequent history of tropical extant herbivorous fishes mirrored that of other reef fish families

(Cowman & Bellwood, 2013a) with two geographically independent components: the east and the west Tethyan provinces (*sensu* Bellwood & Wainwright, 2002) following different trajectories.

The eastern component was subject to biogeographical shifts in marine biodiversity. In the Eocene the global marine biodiversity hotspot was located around the Monte Bolca deposits. Following tectonic events, it subsequently ‘hopped’ to the east during the Oligocene and Miocene, ultimately forming the current-day hotspot in the Indo-Australian-Archipelago (IAA) (Renema *et al.*, 2008). This biogeographical shift was reflected in our results, since surgeonfish and rabbitfish trait combinations that were present in the Eocene are maintained across time in the Indo-Pacific. However, as marine hotspots were shifting to the east, herbivorous fish lineages also diversified, resulting in the rise of new trait combinations. The Miocene (23-5.3 Ma) encompasses a period in which coral reefs were experiencing a marked restructuring with expansion of habitats occupied by reef organisms, diversification of lineages, trophic rearrangements and new fish-coral interactions (Bellwood *et al.*, 2017; Siqueira *et al.*, 2020 [*chapter 2 in this thesis*]). Herbivorous fishes were major players in this restructuring of reef ecosystems, particularly parrotfishes and surgeonfishes. These groups underwent evolutionary shifts that resulted in new traits related to the exploitation of detritus and epi/endolithic components of the marine benthos (Bellwood *et al.*, 2014b; Clements *et al.*, 2017), which led to the trait space expansion that we detected in the Indo-Pacific. As an example of ecological opportunity driving faster diversification rates (Lobato *et al.*, 2014; Siqueira *et al.*, 2020 [*chapter 2 in this thesis*]), these shifts may also have underpinned the high abundance of herbivorous fishes on modern Indo-Pacific coral reefs.

In marked contrast, in the Atlantic, the trait combinations found in herbivorous reef fishes reflect the isolation of the west Tethyan province. With the eastern migration of hotspots, the western side of the Tethys became increasingly isolated from the centre of marine biodiversity (Cowman & Bellwood, 2013a). This isolation promoted a separate biogeographical dynamic in the Atlantic (Joyeux *et al.*, 2001; Floeter *et al.*, 2008) that resulted in a high proportion of internally originated lineages

(Cowman & Bellwood, 2013a). Within the Atlantic, the Caribbean reefs are a hotspot for marine biodiversity (Briggs & Bowen, 2013), however, instead of showing a continuous history of connectivity with ancestral Tethys assemblages, like the IAA, the Caribbean was marked by profound faunal turnover as a result of extinction events (Budd, 2000; O’Dea *et al.*, 2007). Most of these extinction events are reported from the Pliocene (O’Dea *et al.*, 2007). However, for herbivorous reef fishes, the key extinction events were probably much earlier (Eocene and Oligocene) (Tyler & Sorbini, 1998; Bellwood *et al.*, 2017) and resulted in the loss of trait combinations that were present in the Tethys sea. As a consequence, the present-day herbivore fish composition of the Atlantic represents a combination of traits that were derived from lineages that arose in the east Tethys/Indo-Pacific (*Acanthurus* and *Scarus* species) with only one parrotfish clade (*Cryptotomus* + *Nicholsina* + *Sparisoma*) that originated in the Atlantic. Although this clade presents some particular features when compared to other parrotfishes (Streelman *et al.*, 2002), the trait space occupied by these Atlantic parrotfishes is nested entirely within the Indo-Pacific parrotfish space. The recent invasion of the Mediterranean by rabbitfishes and their potential future expansion for Caribbean reefs (Bellwood & Goatley, 2017) might represent yet another return of an ancient Tethys set of traits to the Atlantic.

Trait combinations provide an interesting picture of the evolution and global biogeography of herbivorous reef fishes, yet, they represent a partial view of the functional history of herbivory on tropical reefs. The morphological and behavioural traits used herein may be good proxies of their ecological role, which may, ultimately, drive ecosystem processes on coral reefs. However, to better understand how these organismal traits scale-up to larger-scale ecological processes, we also have to consider the ecosystem consequences of such combinations (Bellwood *et al.*, 2019b). The effect of a species on an ecosystem, as an ecosystem function, should be considered as a distinct measure when compared to the traditional views of functions based on organismal traits (Violle *et al.*, 2007). Thus, our analysis of herbivorous ecosystem functions provided a separate but complementary view of the global evolution of herbivory in the marine tropics.

#### 4.5.2 Ecosystem functions in space and time

The diversification of ecosystem functions among herbivorous fishes coincides with a period of major realignment of the marine hotspots that took place in the east Tethys/Indo-Pacific during the Miocene. This period not only marks the rise of most modern reef fish genera with increased subsequent lineage diversification (Cowman & Bellwood, 2011; Bellwood *et al.*, 2017), but coincides with the shift of the centre of tropical marine diversity to the IAA (Renema *et al.*, 2008). In parallel with the taxonomic diversification, this was also an important time for reef fish functional and trophic innovation (Alfaro *et al.*, 2009b; Cowman *et al.*, 2009; Lobato *et al.*, 2014; Siqueira *et al.*, 2020 [*chapter 2 in this thesis*]), which was clearly reflected in our ecosystem functions results. Bellwood *et al.* (2017) describe this as a distinct phase (phase 5) in the evolution of reef fishes, that was characterized by the formation of a high turnover, dynamic reef ecosystem. However, this important phase for functional expansion seems to be a phenomenon that is largely restricted to the Indo-Pacific. The Atlantic Ocean lags behind the Indo-Pacific both in the time of origin of most modern reef processes executed by extant herbivorous fishes and in the number of extant lineages performing each function.

Environmental changes have been a consistent feature in the Atlantic throughout the last 10 Myr, particularly in the Caribbean (O’Dea *et al.*, 2007). This has resulted in unstable conditions for corals and reefs through time (Johnson *et al.*, 1995; Budd, 2000), prompting the suggestion that lineages that could thrive in peripheral non-coral environments might have resisted periods of high extinction for coral reef-associated faunas (Bellwood & Wainwright, 2002). Interestingly, most of the extant Atlantic herbivorous fish species still have macroalgae and seagrass as important components of their diet (Bonaldo *et al.*, 2014) and some of them are still associated with peripheral reef habitats in at least one of their life stages. These food habits and habitat associations might thus have been one of the main reasons for the persistence through time of the only Atlantic herbivorous fish clade restricted to the western Tethys (Sparisomatina). However, this may be a partial explanation given

that some rabbitfish species are also not closely associated with coral reefs, yet, the family is no longer present in the Atlantic (its presence in Bolca presumably indicates it as an Atlantic component in the Eocene). Selective or not, extinction events shaped the Atlantic herbivore composition leaving the Sparisomatina parrotfishes as major contributors to ecosystem functions through time in this ocean basin.

Our results from the trait approach were largely complementary with those from the ecosystem functions. The first has provided a clear picture of how ecological innovations have been a prevalent feature in the last 20 Ma for herbivorous fishes on coral reefs. The second has yielded insights about how these novel ecological features relate to ecosystem functions that enabled these lineages to play pivotal roles in modern coral reef ecosystems. For example, the initial diversification of parrotfishes and surgeonfishes in the Early Miocene resulted in the rise of new trait combinations in the Indo-Pacific (**Fig. 4.1**). However, it was not until the Late Miocene/Pliocene that these lineages started to expand in terms of functional roles, particularly in turf and sediment removal (**Fig. 4.3**). This showcases the complementarity of trait and function approaches. Parrotfishes and surgeonfishes have very distinct trait combinations reflecting different feeding strategies. However, when translated to ecosystem functions most lineages within these groups fit into only two processes: the removal of algal turfs and sediment (Supplementary Figure 4.18). Interestingly, we see a similar pattern of lineage expansion within these functions in both biogeographical realms, although much more pronounced in the Indo-Pacific. Not coincidentally, turf and sediment removal are critical functions that facilitate coral-dominated states on present-day reefs (Hughes *et al.*, 2007; Goatley *et al.*, 2016). This suggests that trait innovations in herbivorous fishes from the Miocene might have been key to the evolution of coral reefs as we know today, and that recent human impacts might be shifting them back to pre-Miocene conditions before those traits evolved.

#### 4.5.3 Caveats

Although we found striking dissimilarities between biogeographical realms through time, our results are limited to what can be inferred from present-day faunas. It is important to note that the lack of trait combinations and ecosystem functions shown here at certain times or locations does not represent evidence of absence, merely absence of evidence. Our phylogenies indicate the presumed origins in terms of the time and location of traits and functions within extant lineages and are, therefore, blind to the attributes of extinct lineages unknown from the fossil record (Quental & Marshall, 2010; Marshall, 2017). Indeed, the Eocene fossil evidence indicates that some traits and functions can predate these initial occurrences based on phylogenies. Moreover, extinctions may have erased the signal of trait and function trajectories that are not found in extant species (Marshall, 2017). Unfortunately, the incompleteness of the fossil record for herbivorous reef fishes after the Eocene leaves no alternatives other than using the only tool available to trace their history after Monte Bolca, molecular phylogenies. Finally, our inferences about the Tethyan traits and functions are limited to what is known from a single assemblage (Monte Bolca), which might be an underrepresentation of the global Eocene reef fish fauna.

#### 4.5.4 Conclusions

Our study represents the first effort to collectively analyse the evolutionary processes behind the formation of major extant herbivorous coral reef fish groups among marine biogeographical realms. Character optimizations and fossil information showed that both trait space and the origin of lineages performing ecosystem functions differed between the Indo-Pacific and the Atlantic through time. While modern reef processes related to herbivory developed and expanded in the Indo-Pacific, the Atlantic was very likely shaped by extinction events. Present-day differences in the functional composition of herbivorous fish assemblages between the two major oceanic basins are, therefore, largely a result of their disparate evolutionary histories.

## Chapter 5.

# Historical biogeography of herbivorous coral reef fishes

*This chapter is published as:*

Siqueira, A. C., Bellwood, D. R., Cowman, P. F. (2019). Historical biogeography of herbivorous coral reef fishes: The formation of an Atlantic fauna. *Journal of Biogeography*, 46, 1611–1624. doi: 10.25903/5cd265eb0a405

### 5.1 Abstract

**Aim:** To describe the global biogeography of key herbivorous coral reef fish groups since their presumed origins, using data from both fossil and extant species.

**Location:** Global Cenozoic reefs.

**Taxon:** Acanthuridae (surgeonfishes), Siganidae (rabbitfishes) and Scarini (parrotfishes).

**Methods:** We applied the fossilized birth-death model to build chronograms including a comprehensive sampling of extant species and all the fossil occurrences described for each group. With the resulting chronograms, we built biogeographical models considering the geological changes in reef habitat availability since the ancient Tethys Sea. Finally, we used biogeographical stochastic mappings to trace the routes of colonization of the Atlantic Ocean by lineages in our focal taxa.

**Results:** We found that the Paleocene–Eocene was a period of intense lineage origination for surgeonfishes and rabbitfishes in the central Tethys Sea with the appearance of ancient genera. Most of these genera were probably extinct by the Eocene–Oligocene boundary as they do not correspond with modern taxa. Parrotfishes, however, originated in the early Oligocene, an epoch that corresponds with the geographical transition of the marine biodiversity hotspot. In all groups, extant genera had similar origin times and all expanded in the Miocene, mainly in the Indo-Pacific. In the Atlantic, only one parrotfish lineage with Tethyan ancestry appears to have survived. It subsequently gave rise to extant endemic genera (*Sparisoma* and *Cryptotomus*). The other extant lineages in the Atlantic all have Indo-Pacific origins and colonized more recently using different dispersal pathways.

**Main conclusions:** The Indo-Pacific herbivorous fish fauna is the result of ongoing lineage expansion that started in the central Tethys. The Atlantic is a composite fauna with just one endemic lineage and at least four colonization events from the Indo-Pacific.

## 5.2 Introduction

Herbivorous fishes have long been recognized as one of the most important ecological groups in coral reef ecosystems (Choat, 1991). By constantly removing algal turfs, macroalgae and sediment, these fishes provide a competitive release for corals, facilitating their settlement and survival (Burkepile & Hay, 2006; Hughes *et al.*, 2007; Goatley *et al.*, 2016). Without herbivorous fishes, coral reefs can have compromised resilience, leading to a shift to alternate algal-dominated states (Bellwood *et al.*, 2004; Goatley *et al.*, 2016). However, these shifts have clear biogeographical idiosyncrasies, and coral reefs in the Atlantic seem to be less resilient when compared to those in the Indo-Pacific (Adam *et al.*, 2015; Pawlik *et al.*, 2016). Among the underlying causes of this pattern, the uneven biogeographical composition of herbivorous fishes has been hypothesized to be a key factor in explaining differences in coral reef dynamics and resilience (Bellwood *et al.*, 2004; Adam *et al.*, 2015; Bellwood & Goatley, 2017). Therefore, tracing the historical biogeography of herbivorous fishes will shed light on the evolutionary processes that have shaped these assemblages within and between biogeographical regions.

The early history of some extant herbivorous reef fish groups is well documented in the fossil record (Bannikov *et al.*, 2010; Tyler & Micklich, 2011). The richness of surgeonfish (Acanthuridae) and rabbitfish (Siganidae) fossils from the Eocene suggests that they have been important components of reef-associated faunas for at least 50 million years (Myr) (Bellwood, 1996). Besides being extremely informative about the timing of origin of the two extant fish families, these fossils also provide essential information about the geographical location of their early evolution in the central Tethys Sea (Renema *et al.*, 2008). Furthermore, they have broadened our understanding of the formation of crucial ecosystem processes on present-day coral reefs (Bellwood, 2003; Bellwood *et al.*, 2014a,b; Siqueira *et al.*, 2019b [*chapter 4 in this thesis*]). However, despite their importance, herbivorous fish fossils have not yet been put into a phylogenetic context for the exploration of biogeographical patterns. The biogeography of surgeonfishes and rabbitfishes was first investigated with the use of

morphological cladograms (Winterbottom & McLennan, 1993), however, a more comprehensive description of their historical biogeography remains to be undertaken.

In addition to surgeonfishes and rabbitfishes, parrotfishes (tribe Scarini, family Labridae) are the third major component of roving herbivorous fish faunas on coral reefs (Choat, 1991). These fishes are functionally unique when compared to the other roving herbivores, possessing particular features that allowed them to occupy distinct trophic niches (Bonaldo *et al.*, 2014; Clements *et al.*, 2017). With these features (e.g. fused premaxillary teeth and a pharyngeal mill), parrotfishes perform key ecosystem roles on global reefs such as the production and transport of calcareous sediment (Bonaldo *et al.*, 2014). The phylogenetic relationships and biogeography of extant parrotfishes have been examined multiple times (e.g. Bellwood, 1994; Streelman *et al.*, 2002; Robertson *et al.*, 2006; Choat *et al.*, 2012), yet, there has only been one attempt to reconstruct their ancestral ranges (Cowman & Bellwood, 2013a). This latter study, however, described major patterns for the entire Labridae, without focusing on Scarini. Although scarce in the fossil record (Bellwood *et al.*, 2019a), parrotfishes might be the most informative group in terms of the different historical processes between the Atlantic and the Indo-Pacific regions (Bellwood & Wainwright, 2002). Therefore, a more detailed analysis of the biogeography of parrotfishes, alongside other prominent herbivorous clades, would aid the understanding of the historical processes underpinning the development of reef herbivory.

The historical biogeography of four reef fish families has been examined previously (Cowman & Bellwood, 2013a), providing important insights into global patterns of origination and dispersal. However, this study was based on molecular phylogenies built with extant species only, which makes deep-time patterns of lineage distribution hard to infer. So far, only one study has attempted to unite neontological and paleontological data to explore biogeographical patterns in reef fishes (Dornburg *et al.*, 2015), yet, it was limited to one family (Holocentridae). A comparative analysis of multiple reef fish groups combining fossil and extant species is lacking. Considering their evolutionary and ecological importance for coral reefs, herbivorous fishes make the perfect models for this goal. Our overall aim,

therefore, was to analyse the global biogeography of herbivorous coral reef fishes since their earliest records incorporating both phylogenetic evidence and the fossil record. Specifically, our objectives were: (1) to build phylogenies of surgeonfishes, rabbitfishes and parrotfishes including extant and fossil species; (2) to trace their historical biogeography; and (3) to explore the pathways through which surgeonfish and parrotfish lineages colonized the Atlantic Ocean.

## 5.3 Methods

### 5.3.1 Phylogenetic inferences

We used previously assembled genetic databases (Siqueira *et al.*, 2019b [*chapter 4 in this thesis*]), combined with fossil information to build chronograms for the Acanthuridae (surgeonfishes), Siganidae (rabbitfishes) and Scarini (parrotfishes). The genetic sequences for the Acanthuridae included two mitochondrial (*Cox1* and *Cytb*) and seven nuclear markers (*ENC1*, *myh6*, *plagl2*, *Rag1*, *Rh*, *zic1* and *ETS2*) for 72 species in all extant genera, representing ~90% of the extant diversity. The Siganidae phylogeny included two mitochondrial markers (*Cytb* and *16s*) and one nuclear region (*ITS1*) for 24 species (~80% extant diversity). For the Scarini, we assembled sequences for 87 species including five mitochondrial (*Cox1*, *Cytb*, *12s*, *16s* and *control region*) and six nuclear genes (*Bmp4*, *Dlx2*, *Otx1*, *Rag2*, *S711* and *Tmo-4C4*). This dataset encompassed 87% of parrotfish diversity in all extant genera. The genetic sequences were all available in Genbank and were downloaded using Geneious Pro version 11.1 (Kearse *et al.*, 2012). Additionally, we downloaded available sequences for the species *Zanclus cornutus* and *Luvarus imperialis*, and for nine species from the family Labridae (two Hypsigenyines and seven Cheilines) to be used as outgroups for the Acanthuridae and Scarini, respectively. For the Siganidae, we included data for *Zanclus cornutus* and *Prionurus scalprum* as outgroups. We aligned the gene datasets using the Muscle algorithm (Edgar, 2004) in Geneious, checking for inconsistencies by eye. Finally, we assessed the best partitioning scheme for each gene

with PartitionFinder2 (Lanfear *et al.*, 2016). Accession numbers for genetic sequences and the results from partitioning model selection can be found in Siqueira *et al.* (2019; [chapter 4 in this thesis]).

In addition to the genes, we also assembled a database that included all the fossils described for each group with its respective ages and locations. The fossil record of the Acanthuridae includes 25 species belonging to 19 extinct genera, most described from the Eocene of Monte Bolca (Tyler & Bannikov, 2000; Tyler, 2005b,a; Tyler & Micklich, 2011; Bannikov & Tyler, 2012). Similarly, the Siganidae has 10 described fossil species in eight extinct genera (Tyler & Bannikov, 1997; Bannikov & Tyler, 2002; Bannikov *et al.*, 2010). Parrotfishes are not represented in the fossil record of Monte Bolca and have only three described fossils: the disarticulated head of *Calotomus preisli*; a dental plate belonging to a putative *Bolbometopon* species (Bellwood & Schultz, 1991); and *Pacuarescarus kussmauli* described from fragmentary pharyngeal bones (Laurito *et al.*, 2014). The information from all these fossils is summarized in Supplementary Table 5.1. With the genetic database and the fossil information we were able to implement fossilized birth-death (FBD) models (Gavryushkina *et al.*, 2014; Heath *et al.*, 2014) in the Bayesian framework of BEAST2 (Bouckaert *et al.*, 2014) to simultaneously estimate tree topology, branch lengths and node ages for each of our focal taxa. To do that, we first included the fossils with tip dates corresponding to their approximate occurrences in geological time. Then, we set gene partition models according to the results from PartitionFinder and specified relaxed lognormal priors for the clock model.

Since we did not have morphological data for the fossils to estimate their relationships with extant species, we constrained their positions in the phylogenies. For the Scarini, we constrained the fossils to be monophyletic with their corresponding extant genus. This approach was not possible with *P. kussmauli* because, although the authors suggest that the material might be similar to the extant *Cetoscarus* (Laurito *et al.*, 2014), there are no synapomorphies that allow its placement among extant taxa or lineages. It was therefore excluded from our phylogenetic analysis. In the Acanthuridae and Siganidae all fossils belong to extinct genera. We therefore used published morphological cladograms

including fossils and extant species in these two groups to constrain the fossil positions. Our constraints for the surgeonfish phylogeny were firstly based on subfamily groupings, with two extinct subfamilies (Padovathurinae and Gazolachthyinae, Tyler 2005a, 2005b), and the Nasinae and Acanthurinae including extant and extinct taxa (Winterbottom, 1993; Tyler & Micklich, 2011). The later was also split between higher and lower Acanthurinae according to the groupings in Tyler & Micklich (2011). The fossils of unknown status within the Acanthurinae (Tyler & Micklich, 2011) were treated as a separate clade in the subfamily. In addition, the genera *Frigosorbinia* and *Pesciarichthys* were treated as sister taxa within the Acanthurinae (Bannikov & Tyler, 2012), and the Nasinae fossils were grouped together as a sister group to the clade *Eonaso* + *Naso* according to Tyler (2000). Finally, the fossils belonging to the same genus were constrained to be monophyletic. For the rabbitfishes, we also considered fossils of the same genus to be monophyletic and we used the most parsimonious cladogram produced by the unordered analysis of Bannikov *et al.* (2010) to constrain the fossil positions. Although we did not directly estimate the position of the fossils in our phylogenies, we took advantage of the best knowledge available for their morphology and placed them accordingly to be informative enough for our biogeographical models (see Methods section 5.3.2).

Among the fossils used, *Eonaso deani* (Acanthuridae) is of unknown age, however, we decided to keep it in our analysis given its biogeographical importance: it is the only Nasinae species described from the Atlantic Ocean (Tyler & Sorbini, 1998). We placed it in the Oligocene following the suggestion of Tyler (2000) that it is far more recent than the Eocene. Moreover, we included Monte Bolca fossils related to the outgroups to have a more comprehensive fossil sampling in our models. The fossil *Eozanclus brevirostris* (Blot & Voruz, 1970) was included as sister taxa to the extant *Zanclus cornutus* in the Zanclidae outgroup for the surgeonfish and rabbitfish phylogenies, and *Phyllopharyngodon longipinnis* (Bellwood, 1990) was included as monophyletic with the outgroup Hypsigenyines for the parrotfish phylogeny.

Our phylogenetic inferences were then conducted by first performing one MCMC run per taxon for 100 million generations each, saving trees every 10 thousand generations (10,000 trees per run). From the results of this first run, we removed 20% initial generations as burnin and combined the trees into a maximum clade credibility (MCC) tree for each taxon using TreeAnnotator 2.5.0 (Bouckaert *et al.*, 2014). This MCC tree was then used as a starting tree in five independent MCMC runs per taxon for the same number of generations as the first run. These runs were all conducted within the CIPRES Science Gateway (Miller *et al.*, 2010) computing environment. To assess convergence and stationarity in our runs, we used the effective sample size (ESS) scores analysed in the software Tracer v1.7.0 (Rambaut *et al.*, 2018). Finally, we removed 20% burnin from each run, and combined all trees per taxon in LogCombiner v2.5.0 (Bouckaert *et al.*, 2014). Each set of trees was then compiled into a MCC tree in TreeAnnotator.

### 5.3.2 Historical biogeography

We built a dataset for the distribution of extant species in our focal taxa considering six marine biogeographical regions (Kulbicki *et al.*, 2013): Western Indian (WI), Central Indo-Pacific (CIP), Central Pacific (CP), Tropical Eastern Pacific (TEP), Western Atlantic (WA) and Eastern Atlantic (EA). The classification was based on the literature and distribution maps from IUCN's Red List (IUCN, 2017). The fossils were also included in this dataset, however, we created a seventh biogeographical region to include most of them, representing the central region of the ancient Tethys Sea. This region encompasses the area where the Mediterranean Sea lies today, and part of continental Eurasia that formed a major oceanic basin between the Eocene and the Miocene (Hou & Li, 2018). Thus, the fossils described from Italy, Austria, Switzerland, Iran, Turkmenistan and Russia (Caucasus) were all classified as being present in the central Tethys. Importantly, the presence of these fossils in the central Tethys region does not preclude their presence in other biogeographical regions. Fossil species could have had a larger distribution; however, the incompleteness of the fossil record limits our inferences to

known occurrences. Therefore, the limited geographical range assigned to fossils in our dataset and models only represent absence of evidence and not evidence of absence.

By combining the distribution dataset with the phylogenies, we were able to estimate ancestral ranges in our groups. Within the maximum likelihood framework of 'BioGeoBEARS' (Matzke, 2013), we built biogeographical models constrained according to well-known geological events that shaped distribution patterns of coral reef fishes (Bellwood & Wainwright, 2002; Cowman & Bellwood, 2013a). To do that, we first constrained the root node of the trees to be present in the extinct Tethys sea. Both the presence of surgeonfish and rabbitfish fossils in the Eocene (50 Ma) and the estimated origin of parrotfishes in our phylogeny (~32 Ma; see Results section 5.4), suggest that these groups originated in the central Tethys region. Second, we constructed matrices of areas allowed and dispersal multipliers (Supplementary Tables 5.2 – 5.3) to reflect the dynamic nature of marine biogeographical barriers through time (Bellwood & Wainwright, 2002; Cowman & Bellwood, 2013b). Our analysis was stratified in three time-slices, the first being between 90 and 12 Ma, in which we allowed lineages to disperse between the EA and WI regions passing through the open Tethys seaway. From 12 Ma onwards, we excluded the central Tethys from the analysis to reflect the Terminal Tethyan Event (TTE) (Steininger & Rögl, 1979). However, we kept a low dispersal multiplier value between the WI and EA to allow the possibility of dispersal around southern Africa (Bowen *et al.*, 2006). Finally, at 3.1 Ma we prohibited the dispersal between the TEP and WA regions to reflect the final closure of the Isthmus of Panama (IOP) (Lessios, 2008). In all time-slices, we kept a low dispersal value between the CP and TEP regions to represent the soft nature of the East Pacific Barrier (EPB) (Bellwood & Wainwright, 2002; Lessios & Robertson, 2006; Cowman & Bellwood, 2013b).

These time-stratified analyses were performed following the notation of three widely recognized models in historical biogeography: DEC (Ree & Smith, 2008); DIVA (Ronquist, 1997); and BayAREA (Landis *et al.*, 2013). From these basic models, we built combinations of models including, or not, two parameters: the founder-speciation event ( $j$ ); and the dispersal matrix power exponential

(w). The first one adds the possibility of inheritance of a new area by a daughter lineage while the sister-splitting lineage inherits the original ancestral range (Matzke, 2014). The second reduces the subjectivity in user-defined dispersal multiplier matrices by estimating an exponential for its values through maximum likelihood (Dupin *et al.*, 2017). Altogether, 12 biogeographical models were fitted to each taxon and compared through AIC scores.

### 5.3.3 Biogeographical stochastic mappings

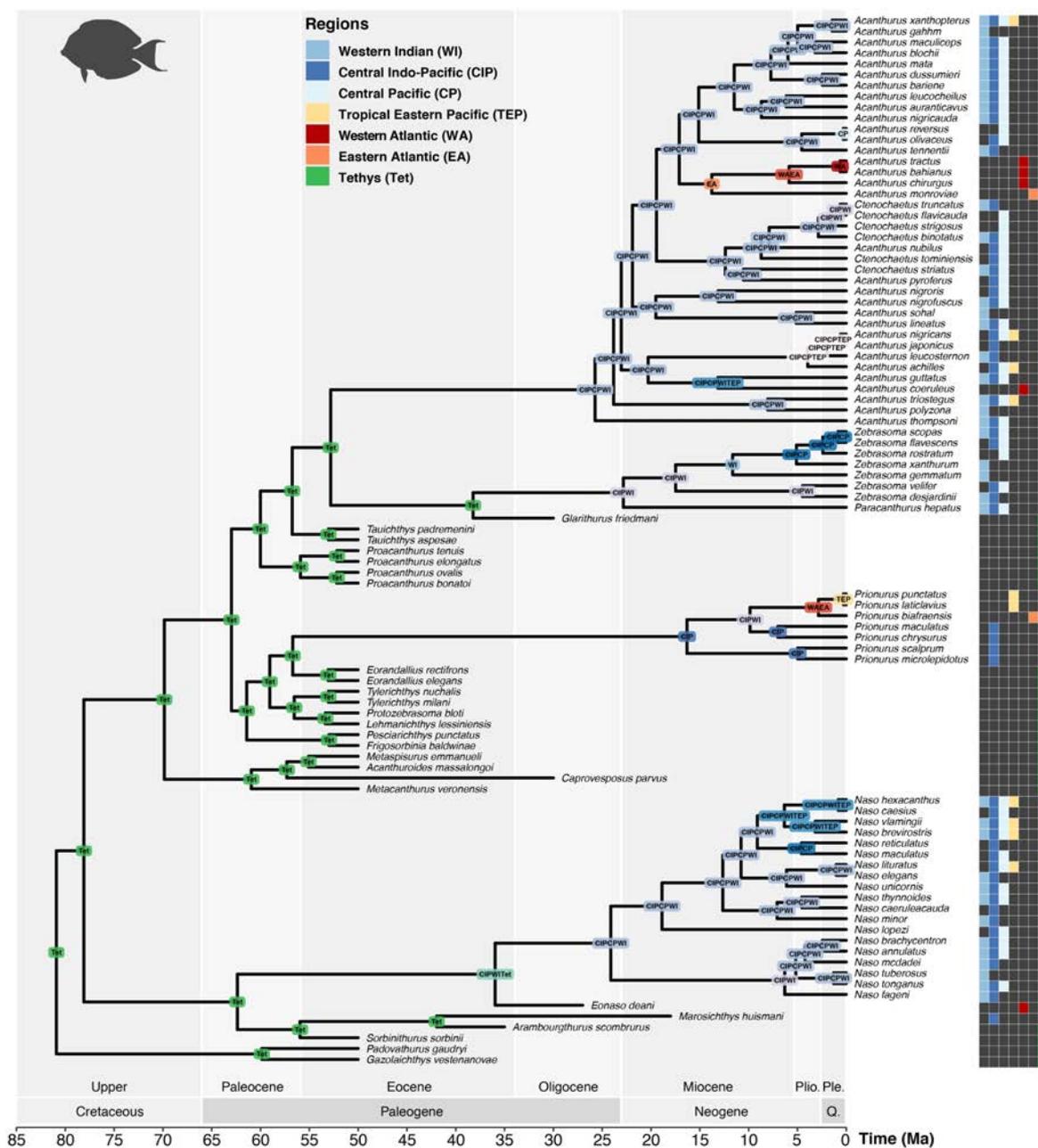
To investigate the pathways through which surgeonfish and parrotfish lineages colonized the Atlantic Ocean, we simulated biogeographical histories using stochastic mappings implemented in 'BioGeoBEARS' (Dupin *et al.*, 2017). Through time, there were only four possible pathways for the colonization of the Atlantic by reef fishes (Floeter *et al.*, 2008): 1) relict lineages from the ancestral Tethys sea; 2) lineages that originated in the Indo-Pacific and colonized the WA through the open Tethys seaway before the TTE; 3) Indo-Pacific origin and crossing through TEP before the closure of IOP; and 4) colonization of the WA via southern Africa. We quantified the support for each of these scenarios by creating 1000 stochastic mappings for each group based on their best biogeographical model for ancestral range estimation. From each of these mappings, we then extracted the node and corner states for the lineages that colonized the Atlantic on the original model to calculate the percentage of times the original scenario was repeated.

## 5.4 Results

Our fossilized birth-death models resulted in well-supported phylogenetic trees for all groups (Supplementary Figures 5.1 – 5.3). While the origin of crown Scarini was estimated to have happened around the early Oligocene (32.2 Ma [23.3 - 41.3 95% highest posterior density, HPD]; Supplementary Figure 5.3), the Acanthuridae and Siganidae crown ages were estimated to be a lot older, spanning the Upper Cretaceous (80.9 Ma [67.5 - 95.6 HPD] and 69.2 Ma [58.5 - 84.7 HPD], respectively;

Supplementary Figures 5.1 – 5.2). Surgeonfish and rabbitfish fossil taxa originated throughout the Paleocene–Eocene. However, in all groups, the origins of extant generic diversity only occurred in the Oligocene–Miocene (Supplementary Figures 5.1 – 5.3). In the parrotfishes, the two most speciose genera represent particularly young divergences when compared to other genera, with the origin of *Scarus* estimated at 9.5 Ma [6.4 - 13.1 HPD] and *Chlorurus* at 5.7 Ma [3.7 - 8.5 HPD] (Supplementary Figure 5.3).

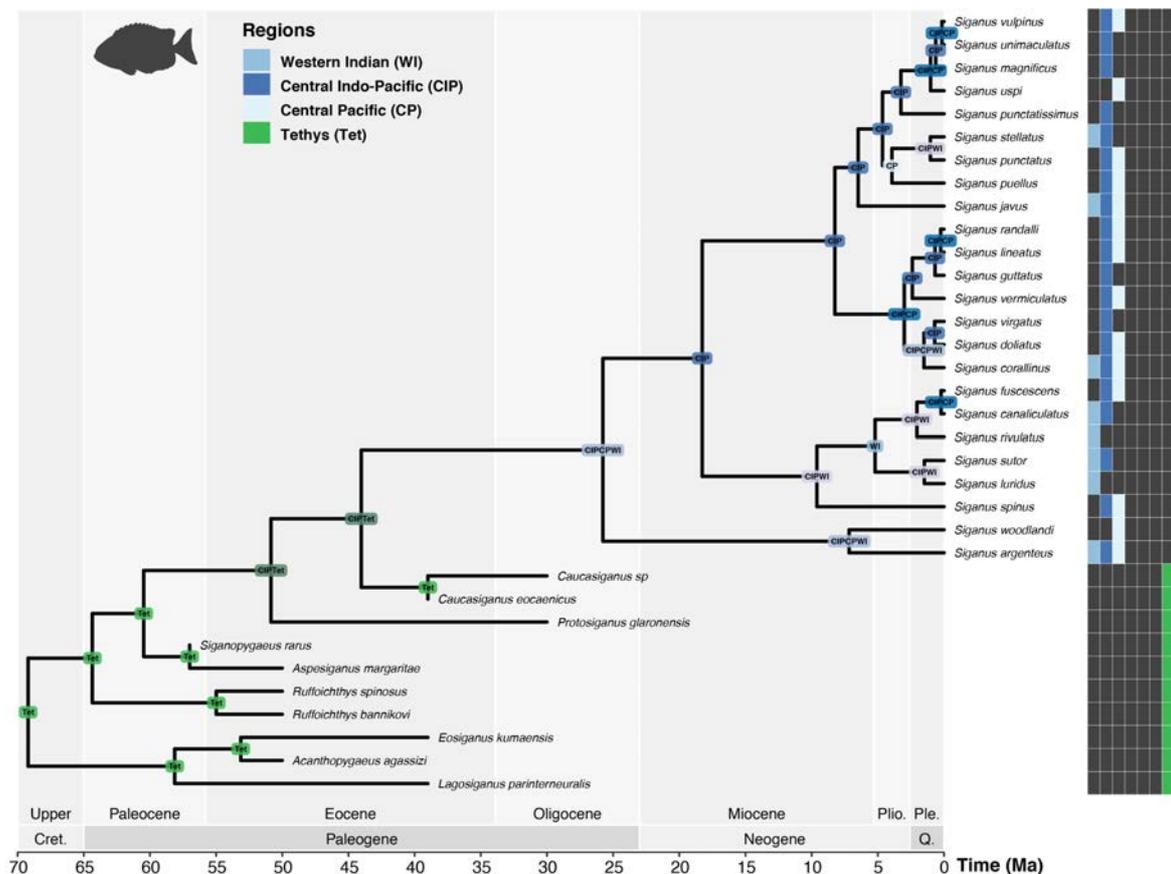
The best-supported biogeographical model for the surgeonfishes was the BayAREA with free *j* and *w* (Supplementary Table 5.4). Within this model, all extant acanthurid genera originated and diversified in the eastern Tethys/Indo-Pacific regions (**Fig. 5.1**). Although the split between the Nasinae and Acanthurinae was estimated to have occurred in the Tethys Sea, early in the history of the family, the extant genera within each subfamily all had their origins in the Miocene. Both the *Acanthurus* + *Ctenochaetus* and the *Naso* clades originated from a widespread ancestor (WI + CIP + CP) at 25.7 Ma [19.3 - 31.9 HPD] and 24.1 Ma [16.2 - 32.5 HPD], respectively (**Fig. 5.1**). Moreover, the *Zebrasoma* + *Paracanthurus* clade originated from an ancestor present in the WI + CIP at 22.8 Ma [15.4 - 30.6 HPD]. Finally, the younger *Prionurus* lineage (16.3 Ma [9.6 - 23.8 HPD]) originated from a CIP ancestor (**Fig. 5.1**). A few extant surgeonfish lineages are shown to have crossed the EPB and colonized the TEP, however, most TEP colonization events happened within the last five Myr (e.g. *Acanthurus xanthopterus* and *Naso hexacanthus*; **Fig. 5.1**).



**Figure 5.1 | Maximum clade credibility phylogeny of Acanthuridae (surgeonfishes) derived from the fossilized birth-death model, depicting node states according to the results from the best-supported biogeographical model (BayAREA+j+w). Boxes represent the geographical distribution of both extant and fossil species. Please, see Discussion section 5.5.3 with regards to fossil age estimates.**

In the rabbitfishes, the best supported model (DEC, Supplementary Table 5.5) suggests that the ancestor of the extant *Siganus* was a widespread lineage (WI + CIP + CP) in the late Oligocene (25.7

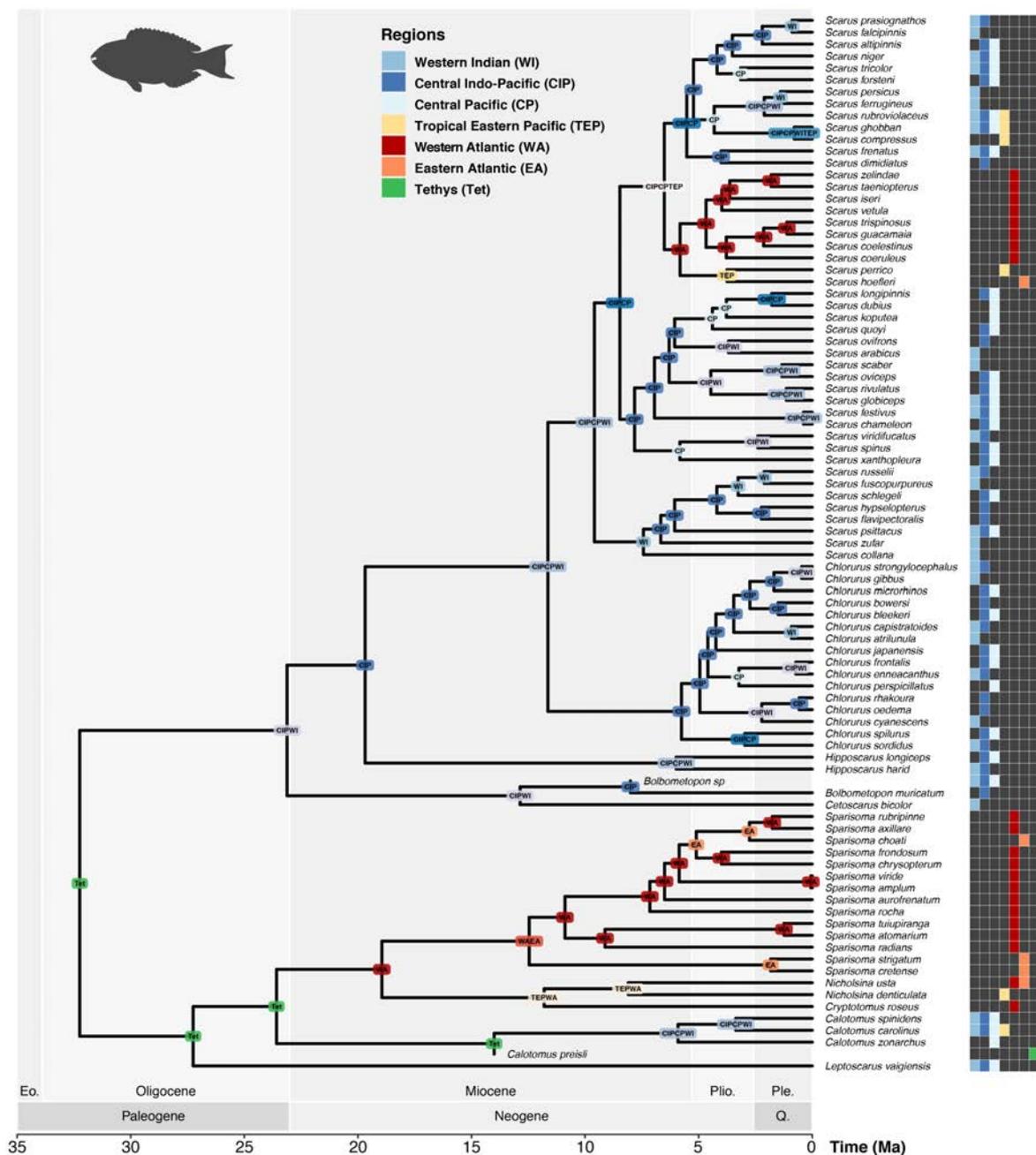
Ma [16.2 - 37.6 HPD]; **Fig. 5.2**). This ancestor derived from a lineage that was still present in the Tethys Sea, where the fossil lineages diversified throughout the Paleocene–Eocene. The main diversification of extant siganid lineages occurred within the CIP throughout the Pliocene–Pleistocene (**Fig. 5.2**). However, none of these extant lineages have crossed the EPB or have naturally invaded the Atlantic realm.



**Figure 5.2 | Maximum clade credibility phylogeny of Siganidae (rabbitfishes) derived from the fossilized birth-death model, depicting node states according to the results from the best-supported biogeographical model (DEC). Boxes represent the geographical distribution of both extant and fossil species. Please, see Discussion section 5.5.3 with regards to fossil age estimates.**

For the parrotfishes, the DEC with free *j* and *w* parameters was recovered as the best-supported biogeographical model (Supplementary Table 5.6). This model supported an eastern Tethyan (CIP + WI) origin of the Scarina clade (*Scarus* + *Chlorurus* + *Hipposcarus* + *Cetoscarus* +

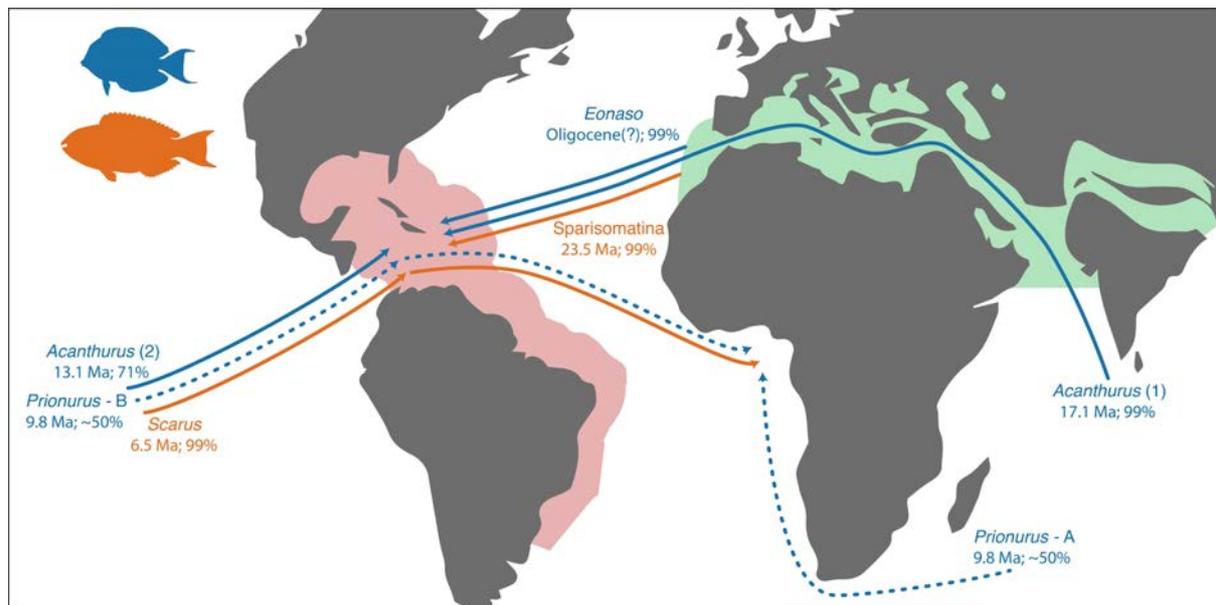
*Bolbometopon*) at 23.1 Ma [15.6 - 32.6 HPD] (**Fig. 5.3**). The subsequent diversification within this clade took place predominantly in the CIP throughout the Miocene, with many vicariant and dispersal events between this region and both the WI and CP. There have been several dispersal events across the EPB by scarine parrotfishes, most notably by *Scarus rubroviolaceus*, the ancestor of *Scarus ghobban* + *Scarus compressus*, and the common ancestor between these lineages and the Atlantic *Scarus* clade (**Fig. 5.3**). By contrast, our models suggest that the Sparisomatina (*Sparisoma* + *Nicholsina* + *Cryptotomus* + *Calotomus* + *Leptoscarus*) had a central Tethyan origin at 27.2 Ma [19.4 - 35.3 HPD]. After splitting from the *Leptoscarus* lineage, the remaining Sparisomatina lineages were subsequently split between an Atlantic clade and the Indo-Pacific *Calotomus* clade (**Fig. 5.3**). Within the Atlantic Sparisomatina, most of the diversification occurred in the WA, mirroring the Atlantic *Scarus* clade.



**Figure 5.3 | Maximum clade credibility phylogeny of Scarini (parrotfishes) derived from the fossilized birth-death model, depicting node states according the results from the best-supported biogeographical model (DEC+j+w). Boxes represent the geographical distribution of both extant and fossil species.**

Combining the biogeographical models and the stochastic mappings, we found distinct scenarios for the colonization of the Atlantic Ocean among taxa. The models suggest that parrotfishes

initially colonized the Atlantic from the central Tethys region at 23.5 Ma [16.5 - 30.9 HPD] (**Fig. 5.4**). This lineage gave rise to two endemic Atlantic genera (*Sparisoma* and *Cryptotomus*), and *Nicholsina* that subsequently colonized the TEP before the closure of the IOP (*N. denticulata*) (**Fig. 5.3**). More recently, a *Scarus* lineage colonized from the opposite direction, crossing the EPB and entering the WA through the open IOP at 6.5 Ma [4.4 - 9.1 HPD] (**Fig. 5.4**). Although the core diversification within the Atlantic *Scarus* occurred in the WA, the EA has also been colonized by one lineage (*S. hoefleri*) (**Fig. 5.3**). Thus, the *Scarus* and the Sparisomatina lineages represent the only two colonization events of the Atlantic by parrotfishes.



**Figure 5.4 | Routes of colonization of the Atlantic by surgeonfishes and parrotfishes.** Each arrow represents a lineage with its respective age of colonization in Million years ago (Ma) and the percentage (%) of support in the Biogeographical Stochastic Mappings (BSM). Dashed lines represent alternative scenarios (A and B) for the colonization of the Atlantic by the Prionurus lineage that received similar BSM support. Acanthurus (1) represents the clade of *A. monroviae*, *A. chirurgus*, *A. bahianus* and *A. tractus*; and Acanthurus (2) represents the lineage of *A. coeruleus*. The map represents the configuration of continents during the Middle Miocene that allowed the colonization through the Tethys seaway (green) and the open Isthmus of Panama. The red area represents the Western Atlantic region.

For surgeonfishes, four Atlantic colonization events were retrieved in our range reconstructions. Firstly, the fossil *Eonaso deani* marked the dispersal of a Nasinae lineage from the central Tethys to the WA (**Fig. 5.4**). Secondly, an *Acanthurus* lineage colonized the Atlantic at 17.1 Ma [14.5 - 24.8 HPD] through the open Tethys seaway (**Fig. 5.4**). This lineage gave rise to the most speciose extant surgeonfish clade in the Atlantic Ocean (**Fig. 5.1**). More recently, another *Acanthurus* lineage (*A. coeruleus*) colonized the WA at 13.1 Ma [7.6 - 18.8 HPD], however, this likely happened through the open IOP after initial colonization of the TEP (**Fig. 5.4**). Finally, the Atlantic was colonized by a *Prionurus* lineage at 9.8 Ma [6.0 - 14.1 HPD]. In this case, though, our stochastic mapping offered similar support for west to east (i.e. from TEP through open IOP) and east to west (i.e. from WI around southern Africa) colonization pathways (**Fig. 5.4**).

## 5.5 Discussion

Combining data from extant and fossil species, we explored the historical biogeography of the three most important nominally herbivorous fish groups on coral reefs. Our reconstructions suggest that the Paleocene–Eocene was a period of high lineage origination in the central Tethys region (**Figs. 5.1 – 5.2**). They also highlight the Oligocene as a transition period for marine biodiversity with the hotspot moving from the central Tethys region to the Indo-Australian-Archipelago (IAA) (cf. Renema et al., 2008). In addition to these deep biogeographical events, we show a strong signature of isolation in the evolutionary dynamics of the Tropical Eastern Pacific and Atlantic. The colonization of these biogeographical regions by Indo-Pacific herbivorous lineages was underpinned by at least four distinct events spread throughout the Miocene. Below, we discuss the origins of the biogeographical patterns in herbivorous groups and the colonization of the Atlantic in two separate sections.

### 5.5.1 Origins of biogeographical patterns

The richness of fossil deposits from the central region of Tethys highlights its role as a hotspot for marine biodiversity in the Eocene (Renema *et al.*, 2008). For fishes, these deposits also represent the first unequivocal evidence of morphologies associated with piscine herbivory in marine systems (Bellwood, 2003). Surgeonfishes and rabbitfishes present in the Eocene closely resemble their extant counterparts (Bellwood *et al.*, 2014a). However, it is hard to tell whether these fossil species were geographically widespread. So far, no herbivorous fish fossils have been recorded to the east (Indo-Pacific) or to the west (Atlantic) of the central Tethys region in the Eocene. Thus, the available evidence points to this region as the potential place of origin of surgeonfishes and rabbitfishes, a pattern that is strongly supported in the Labridae for which we have a far larger fossil record (Bellwood *et al.*, 2019a).

In surgeonfishes, the central Tethyan fossil occurrences extend from the Eocene (the Monte Bolca *lagerstätten*) to the Oligocene (*Caprovesposus* [Russia] and *Glarithurus* [Switzerland]). In rabbitfishes the records extend from the late Paleocene (*Siganopygaeus* [Turkmenistan]) to the Oligocene (*Caucasiganus* [Russia] and *Protosiganus* [Switzerland]). By placing these occurrences in a phylogenetic context, we highlight the central Tethys as the biogeographical region where the early history of these families most likely occurred. Instead of just being associated with the expansion into niches left by the ancient Mesozoic fishes that went extinct in the K–Pg event (Friedman, 2010), this diversification appears to be associated with the exploitation a trophic niche previously unexplored by fishes (Bellwood *et al.*, 2017). Surgeonfishes and rabbitfishes in the central Tethys during the Paleocene–Eocene had morphological features associated with the exploitation of benthic resources that were not present in Mesozoic fishes (Bellwood, 2003). Although Paleogene reefs from the central Tethys differed markedly from present-day coral reefs (Zamagni *et al.*, 2012; Bellwood *et al.*, 2017), they provided the ecological stage for the initial diversification of herbivorous fishes.

From the late Oligocene onwards, however, the evolutionary history of herbivorous reef fishes presents a clear geographical shift (**Figs. 5.1 – 5.2**). The Eocene–Oligocene boundary was marked by extensive tectonic, eustatic, climatic, oceanographic and geomorphological ('TECOG') changes (Bellwood *et al.*, 2012) that promoted mass extinction in marine faunas (Ivany *et al.*, 2000). These changes also affected reef fish assemblages (Cowman & Bellwood, 2011), possibly leading to the extinction of some surgeonfish and rabbitfish genera in the Tethys. However, from the Oligocene to the Miocene, the availability of shallow water marine environments was increasing to the east of Tethys (Williams & Duda, 2008; Leprieur *et al.*, 2016), which might have promoted the initial diversification of Indo-Pacific fish lineages. The nasin *Marosichthys huismani* from Indonesia (Tyler, 1997) shows that surgeonfish assemblages were already developing in the east Tethys region by the Early Miocene (**Fig. 5.1**). This is reinforced by our biogeographical reconstructions which point to the origins of most extant herbivorous fish genera in the Indo-Pacific during the Miocene. Therefore, while the Oligocene was a transition period with the marine biodiversity hotspot shifting from its central Tethyan location to the IAA (Renema *et al.*, 2008), the Miocene marks the rise and expansion of extant genera of herbivorous reef fishes in the Indo-Pacific.

Interestingly, the second oldest parrotfish fossil also represents one of the youngest records of herbivorous lineages in the central Tethys. The presence of *C. preisli* from the Middle Miocene of Austria (Bellwood & Schultz, 1991) provides evidence that herbivorous lineages survived in the region even after the IAA hotspot began to form. This is further demonstrated by the presence of surgeonfish teeth fossils from the same region (Schultz, 2003; Tripalo *et al.*, 2016) and the presence of an unidentified acanthurid larva in the Miocene of Greece (Gaudant *et al.*, 2005). Although extant species within *Calotomus* are restricted to the Indo-Pacific, the presence of a congeneric fossil species in the Paratethys region suggests that, like the surgeonfishes, the early history of parrotfishes also occurred within, or included, that region (Bellwood, 1994). If so, habitat partitioning might have been a key factor in the early divergence between clades (Streelman *et al.*, 2002). The phylogeny of parrotfishes indicates an early split between a reef (*Scarina*) and a seagrass (*Sparisomatina*) clade in the Oligocene

(**Fig. 5.3**), which might be related to the complex geomorphological history of the Paratethys region (Rogl, 1999) and the consequent habitat heterogeneity during that time. After this initial divergence, the reef clade flourished in the east Tethys/Indo-Pacific diversifying into five genera, whereas the seagrass clade remained in the Tethys throughout the Oligocene, until the split between the west Tethys/Atlantic lineage (*Cryptotomus*, *Nicholsina* and *Sparisoma*) and the east Tethys/Indo-Pacific *Calotomus* clade in the Miocene.

### 5.5.2 Colonization of the Atlantic Ocean

The soft nature of the central Atlantic barrier, the presence of a tropical east-west current, and oceanic islands as stepping stones, might have maintained some level of connectivity between the WA and Tethys through the late Mesozoic and early-mid Cenozoic (Bellwood & Wainwright, 2002; Hou & Li, 2018). However, with marked 'TECOG' changes in the Oligocene, the differences in composition between the Atlantic and the east Tethys/Indo-Pacific appears to have increased (Cowman *et al.*, 2017), and parrotfishes represent the most iconic example of this historical process (Bellwood & Wainwright, 2002). The presence of a lineage within the Sparisomatina that is endemic to the west Tethys/Atlantic in the early Miocene strongly suggests that marine provinciality started to develop before the TTE (12-18 Ma). Since the late Oligocene, this lineage has been evolving independently from its east Tethys/Indo-Pacific sparisomatine counterparts after splitting from its central Tethyan ancestor at about 24 Ma (**Fig. 5.3**). The parrotfish fossil *P. kussmauli* from the Lower Miocene of Costa Rica (Laurito *et al.*, 2014) also indicates that a relatively basal parrotfish (*cf.* Bellwood *et al.*, 1994) was present at this time in the WA. However, the fragmentary nature of this fossil precludes its precise phylogenetic placement and, unfortunately, we cannot tell whether it represents an ancient sparisomatine or a previously widespread scarine in the Atlantic during the Miocene. If future evidence supports the latter it would be another case of colonization of the Atlantic by scarines, along with the extant *Scarus*.

The Atlantic *Scarus* represent a recent colonization event by parrotfishes, close to the end of the Miocene. Our best biogeographical model reconstructed a member of the *Scarus* lineage invading the Atlantic through the TEP before the closure of the IOP (**Fig. 5.4**). This reconstructed colonization route supports a longstanding idea (Bellwood, 1994; Choat *et al.*, 2012) that was hitherto lacking empirical support, and represents an evolutionary step with important ecological consequences. These parrotfishes are the largest-bodied herbivorous fishes in the Atlantic, with two species (*S. guacamaia* and *S. coeruleus*) growing to much larger sizes than any other Indo-Pacific congeneric (Siqueira *et al.*, 2019b [*chapter 4 in this thesis*]). Since size is related to numerous functions (Bellwood *et al.*, 2019b), these species were probably important for Atlantic ecosystem functions prior to human-induced declines.

Although provinciality among marine realms was in place since the Oligocene, no other extant endemic herbivorous fish genus evolved in the Atlantic like the sparisomatine parrotfishes. All the surgeonfishes present in the Atlantic derive from Indo-Pacific lineages that colonized during the Miocene (**Fig. 5.4**). In this family, however, evidence for provinciality comes from fossils. The presence of *Eonaso deani* in the Caribbean (Tyler & Sorbini, 1998) shows that the Atlantic once had a nasin surgeonfish fauna. Moreover, since the entire Nasinae sub-family is now absent from the Atlantic, it also suggests that local extinctions have played a fundamental role in shaping Atlantic herbivorous assemblages. Interestingly, even though we could not distinguish between alternative scenarios for the colonization of *Prionurus* in the Atlantic, both point to the possibility of their presence in the WA just before the final closure of the IOP (**Fig. 5.1**). If that is the case, it would be another example of an extinct surgeonfish genus from the WA, along with *Eonaso*.

Further evidence for the role of extinction in the Atlantic comes from other reef taxa that have better representation in the fossil record. Corals (Budd, 2000) and bryozoans (Di Martino *et al.*, 2018), for instance, both show that extinction events have been prevalent in the Atlantic throughout the Cenozoic. Particularly after the closure of the IOP, environmental changes caused a dramatic faunal

turnover on Caribbean reefs (O'Dea *et al.*, 2007). This recent turnover potentially affected reef fishes (Wainwright *et al.*, 2018), however, for herbivorous fishes, it seems to have had less effect on the Atlantic fauna when compared to the extinction of ancient Tethyan lineages. Considering that most Atlantic lineages were shared with the TEP prior to the IOP closure, recent extinctions of herbivorous fishes might have been more marked in the TEP. This has been suggested for another reef fish group (Tavera *et al.*, 2018), and would explain why neither *Sparisoma* or *Cryptotomus* are present in TEP and why just a few surgeonfishes (*P. punctatus* and *P. laticlavus*) and parrotfishes (*N. denticulata* and *S. perrico*) in the TEP are derived from lineages that share a recent common ancestor with Atlantic taxa. As a result, the TEP herbivorous fish fauna is mainly composed of very recent invasions from Indo-Pacific lineages that have breached the East Pacific Barrier within the last 3 Myr (Lessios & Robertson, 2006). This pattern is also corroborated by the fossil record of coral species from the TEP (reviewed in López-Pérez, 2017), that shows a recent shift from a Caribbean-associated fauna to an Indo-Pacific derived one.

Given that they are well represented in the fossil record of the Tethys, rabbitfishes are perhaps the most intriguing case of absence from the Atlantic. The probable presence of rabbitfishes in the Tethys for over 30 Myr (Fig. 5.2) and the potential connectivity with the WA (Hou & Li, 2018) offer support for suggestions that rabbitfishes were present in the Atlantic in the early Cenozoic (Bellwood & Wainwright, 2002). Moreover, the similar timing of origin of the extant rabbitfish genus (*Siganus*), when compared to the other extant herbivorous fish genera (Figs. 5.1-5.3), suggests that there has been enough time for a rabbitfish lineage to colonize the Atlantic. However, the only colonization event out of the Indo-Pacific boundaries was the recent human-induced invasion of the Mediterranean Sea by two species (*S. luridus* and *S. rivulatus*) (Sala *et al.*, 2011). The explanation may lie in a combination of ancient extinctions and limited colonization abilities. Siganid species that were likely to be present in the Atlantic throughout the Paleogene are now extinct. Moreover, the extant *Siganus* appears to have evolved in the Indo-Pacific when connectivity with the Atlantic was already very limited and they were thus never able to naturally colonize that realm. The limited colonization

capacity of modern rabbitfishes is supported by the fact that even the species with the largest geographical range (*S. argenteus*) has not yet crossed the EPB and, therefore, is absent from the TEP. However, it does span two thirds of the tropics from Africa to Pitcairn (Woodland, 1990), which weakens the hypothesis of limited colonization capability to some extent. Only new fossil discoveries and a better understanding of the colonization capacities of rabbitfishes will provide answers to this intriguing case.

### 5.5.3 Model considerations

Although our estimated ages for the parrotfish clades were very similar to previous studies using a simple birth-death model (Choat *et al.*, 2012; Siqueira *et al.*, 2019b [*chapter 4 in this thesis*]), the estimated age of origin for surgeonfishes and rabbitfishes has been pushed back in our new model (**Figs. 5.1-5.2**). We believe that the accumulation of fossils in one assemblage at 50 Ma might have driven the older age estimates in these groups. The FBD model regards fossils as part of the same macroevolutionary process as the extant species in the phylogenies (Heath *et al.*, 2014). Consequently, the diversification of surgeonfish and rabbitfish fossils in the Paleocene–Eocene is required to conform to a similar pattern to that found in the genetic divergence between extant species in the model; this inevitably pushes the estimated ages of lineages with fossils back in time. This push to the past has already been reported in other studies that applied the FBD model (Saladin *et al.*, 2017; Silvestro *et al.*, 2019), although others did not find this effect (Heath *et al.*, 2014; Arcila *et al.*, 2015). Whether or not this represents an artefact of the model is a subject that requires further investigation. However, at least for the extant genera in our groups, the FBD provided similar date estimates when compared to previous studies using other models. Moreover, it allowed the inclusion of fossils as terminal tips in our phylogenies, providing increased precision over standard dating methods, and potentially more accurate biogeographical scenarios in historical reconstructions. Yet, the fact that it identified within-family divergence events before the K–Pg boundary, contrary to currently available fossil evidence

(reviewed in Friedman & Sallan, 2012), suggest that these Cretaceous divergences should be viewed with considerable caution.

#### 5.5.4 Conclusions

Our results show that herbivorous coral reef fishes have had two main phases of lineage origination, each with a distinct biogeographical scenario. Firstly, lineages originated in the central Tethys, underpinning a diverse Eocene fauna. Secondly, those Tethyan lineages that survived the Eocene/Oligocene boundary expanded in the Indo-Pacific during the Miocene to produce the modern fauna. On the other hand, the Atlantic and the TEP were shaped by extinction and isolation from the marine biodiversity hotspots. This history of isolation is reflected in the fact that four of the five Atlantic herbivorous lineages have Indo-Pacific ancestors with only one endemic clade retaining a Tethyan origin. The Atlantic herbivorous reef fish fauna represents a conspicuous example of the importance of history in explaining the structure of extant marine fish assemblages.

## Chapter 6. Concluding Discussion

With the overarching aim of explaining coral reef fish biodiversity distribution patterns from a trophic perspective, this thesis sheds light on important macroevolutionary, macroecological and biogeographical processes related to this prominent vertebrate radiation. Increasingly, evolutionary biologists and ecologists are stressing the importance of applying integrative approaches to biodiversity research (Fritz *et al.*, 2013; Price & Schmitz, 2016; Rapacchiolo & Blois, 2019). My research represents a contribution in this direction by integrating concepts from core disciplines in biology. As suggested by Price & Smith (2016), an improved synthesis and relevance of biodiversity studies requires: (i) a focus on form and function; (ii) a better comprehension of scale-dependency; and (iii) interdisciplinary integration. By combining these three elements in various ways, the chapters presented throughout this thesis provide a more holistic understanding of the processes generating and maintaining the diversity of fish species on coral reefs across temporal and spatial scales. Below, I discuss some key concluding remarks that can be drawn from the present work.

### 6.1 The trophic component of coral reef fish evolution and macroecology

The findings described in **chapters 2 and 3** provide support for the growing body of literature highlighting species trophic characteristics as a major driver of macroevolutionary and macroecological patterns (e.g. Price *et al.* 2012; Burin *et al.* 2016). In the marine realm, similar ideas have long been proposed based on the analyses of the fossil record. Vermeij (1977) introduced the term ‘Mesozoic Marine Revolution’ to refer to the escalation in sturdiness of marine gastropod shells in response to changes in predation pressure during the Cretaceous (145.5 – 66 Ma). This phenomenon describes the important effect that evolutionary innovations, related to species trophic status (in this case durophagy), have had on the morphology and composition of marine benthic

organisms. This evolutionary arms race, driven by predator-prey interactions, was posited as the main factor underpinning the origin and diversification of marine molluscs in the Mesozoic (Vermeij, 1977). More recently, the Cenozoic was the stage for another marine biotic revolution arising from species trophic interactions (Steneck, 1983; Bellwood, 2003). The appearance of specific morphological features in fossils from the Eocene (56 – 34 Ma) point to the origins of piscine herbivory in the marine realm (Bellwood 2003). This fundamentally altered the nature of fish-benthos interactions, which also had profound consequences for the benthic composition of reef systems (Steneck, 1983; Wood, 1999). As coined by Bellwood (2003), this event thus represents a ‘Marine Cenozoic Revolution’, that probably started with the rise and expansion of early surgeonfishes and rabbitfishes as shown in **chapter 5** herein.

These proposed ‘Marine revolutions’ emphasise how changes in the functional architecture of species through evolution might have cascading effects that alter large scale system properties such as species diversity and trophic relationships. The results described in **chapters 2 and 3** not only reinforce that idea, but also provide evidence that the Cenozoic revolution was a long-lasting process that culminated in the formation of present-day coral reef biodiversity patterns. Although marine piscine herbivory was in place since at least the Eocene, it was only during the Miocene (23 – 5.3 Ma) that the major extant groups (**chapter 5**) and their associated ecological processes (**chapter 4**) started to expand. This expansion might have been linked with the ecological opportunities provided by a previously unexploited resource by fishes (Harmelin-Vivien, 2002; Lobato *et al.*, 2014). However, there is more to the story than just the use of novel adaptive zones. Herbivory by marine fishes not only represented a Cenozoic innovation (Bellwood 2003), but it also opened up opportunities for trophic novelties and diversification in other coral reef fish guilds. This intricate story could only be revealed through the comprehensive framework applied in **chapter 2**.

Indeed, the rise of fish herbivores in marine systems appears to have been the main trigger of the so-called ‘Marine Cenozoic Revolution’ (Bellwood 2003). However, one of the most intriguing

elements of this revolution that has been unveiled by the present thesis was the cascading effect that those fishes produced when they expanded in the Miocene. Herbivorous coral reef fishes were pioneers in exploiting an important ecological resource (i.e. particulates on the reef flats), and more innovative ways of utilising those resources arose in the Oligocene-Miocene (**chapter 4**). As a consequence of the intense herbivorous feeding pressure, more complex coral communities started to develop in the Miocene, which then provided diversification opportunities for corallivore and planktivore fish lineages (Bellwood *et al.*, 2017; **chapter 2**; Siqueira *et al.*, 2020). This evolutionary cascade in trophic groups not only influenced the pace at which new fish lineages originate, but it also helped shaping global species richness patterns. However, it did this in an unexpected way. The trophic guild that has the highest rate of origination amongst reef fish lineages (herbivores/detritivores) is not the one that contributes the most to the present-day diversity disparities between coral reef areas. Although planktivore lineages were not outstanding in terms of diversification rates, they are the ones that seem to be best partitioning non-reef resources on coral reefs (**chapter 3**). Hence, planktivorous fishes might have been amongst the main beneficiaries of the shelter provided by the novel benthic communities that started forming in the Miocene. In combination, these results highlight two important complementary dimensions of trophic success that operated at different scales within coral reef systems. On an evolutionary time-scale (last 20 Myr), the rise of innovations related to herbivory allowed fish lineages to diversify at a faster pace when compared to other guilds. On an ecological time-scale (present), planktivores successfully partition oceanic resources, particularly in the coral-rich areas of the Indo-Australian Archipelago.

## 6.2 Historical contingencies in the evolution of coral reef fishes

The term historical contingency, when applied to biology, refers to non-deterministic historical events that fundamentally altered evolutionary trajectories through time (Gould, 1989). As proposed by the evolutionary biologist Stephen Jay Gould in his book *Wonderful Life*, rewinding the tape of evolution and replaying it would probably never lead to same outcomes. Therefore,

contingency can be better understood from the perspective of mass extinction events that have distorted the pathways through which lineages evolved (Gould, 1989; Jablonski, 2001). Determinism and contingency represent opposite ends of an evolutionary spectrum; however, they are not mutually exclusive and the goal is often to understand the strength of each mechanism at the appropriate scale (Blount *et al.*, 2018). For instance, in this thesis, I highlight a potential non-contingent feature of evolution, with the repeated instances of increased diversification rates associated with transitions towards herbivory in many different animal taxa (**chapter 2**; Siqueira *et al.* 2020; Price *et al.* 2012; Wiens *et al.* 2015; Burin *et al.* 2016; Poore *et al.* 2017; McKenna *et al.* 2019). However, results from **chapters 4 and 5** make it evident that historical contingency was a key determining factor in the deep-time evolutionary patterns of herbivorous coral reef fishes. In the groups analysed, both functional (**chapter 4**; Siqueira *et al.*, 2019b) and taxonomic (**chapter 5**; Siqueira *et al.* 2019a) perspectives unveiled contingent forces (i.e. extinction events) driving the differences between major marine biogeographical realms.

Although they were not part of the ‘big five’ extinction events that shaped life on earth (Raup & Sepkoski, 1982), the extinctions of herbivorous coral reef fishes in the Atlantic epitomize historical contingencies as important drivers of extant marine biodiversity patterns. The evolutionary history of fish lineages in the Atlantic Ocean is marked by increasing isolation from the marine biodiversity hotspots through time and lineage turnover (Bellwood & Wainwright, 2002; Cowman & Bellwood, 2013a; Cowman *et al.*, 2017). During the Paleocene-Eocene (66 – 34 Ma), the Atlantic was probably well connected to the central Tethyan marine biodiversity hotspot (Hou & Li, 2018) and had a diverse reef fish fauna. This is supported by recent fossil evidence from deposits in Mexico (Cantalice & Alvarado-Ortega, 2016; Cantalice *et al.*, 2018, 2020). However, with the eastern migration of the marine biodiversity hotspot (Renema *et al.*, 2008), related to tectonic changes throughout the Oligocene-Miocene (34 – 5 Ma), the Atlantic became gradually more isolated in evolutionary terms. This process of isolation was clearly demonstrated in **chapters 4 and 5** with herbivorous fish lineages. Nevertheless, isolation from the hotspot in itself is not enough to explain the disparities in the extant

reef fish faunal composition between the Atlantic and the Indo-Pacific. These disparities might be more profound than one would expect by area differences alone. Although the Indo-Pacific Ocean has more reef area when compared to the Atlantic, area differences may not be sufficient to account for the nearly four-fold disparity in reef fish richness between the two oceanic basins. Therefore, contingent extinctions had a prominent role in shaping the reef fish fauna in the Atlantic. There is sufficient evidence that the Caribbean (the marine biodiversity hotspot in the Atlantic; Briggs & Bowen, 2013) went through major environmental changes before and after the final closure of the Isthmus of Panama (3.1 Ma) that led to the extinction of many reef-associated organisms (Budd, 2000; O’Dea *et al.*, 2007; Di Martino *et al.*, 2018). Moreover, I show in this thesis that earlier extinction events (Eocene and Oligocene) may also have been important for shaping the Atlantic herbivorous reef fish composition. It appears that contingency has been a key feature throughout the evolution of marine organisms in general, but also for herbivorous fish ecologies (**chapter 4**) and lineages (**chapter 5**) in particular.

It remains important to provide cautionary notes about the inferred ecosystem consequences of the patterns described in **chapter 4**. Herbivorous fishes have long been recognized as essential components of coral reef ecosystems (Choat, 1991), and their uneven biogeographical composition has been hypothesized to be a key factor in explaining differences in coral reef resilience (Bellwood *et al.*, 2004). In this thesis I show how historical contingencies likely shaped this uneven biogeographical composition. However, by highlighting this disparate history, I do not intend to imply it as the main causal mechanism underlying differences in ecosystem functioning between Atlantic and Indo-Pacific coral reefs. Although herbivore functional diversity may influence community structure in coral reefs (Burkepile & Hay, 2008), there are three main reasons to be cautious about directly correlating biogeographical differences in fish composition with ecosystem functioning. First, other major biological and geographical dissimilarities are found between the coral reefs in both realms (Roff & Mumby, 2012; Pawlik *et al.*, 2016). This might moderate the influence of variation in herbivore assemblages. Second, the different evolutionary histories presented here do not carry any information

about species abundance or biomass through time, which will be important when considering the impact of a species in its ecosystem (Hughes *et al.*, 2007). And third, I focused on the main (most diverse and abundant) groups but there are other herbivorous lineages that might be important elements in coral reef ecosystems (e.g. Kyphosidae, Pomacentridae and Pomacanthidae). Nevertheless, patterns to-date in these other herbivorous groups closely reflect those shown here (Frédérich *et al.*, 2013; Knudsen & Clements, 2016; Baraf *et al.*, 2019), suggesting that the observed historical patterns may be relevant across entire herbivore assemblages. Only further investigations will allow us to better understand the ecosystem consequences of the disparate evolutionary history between the Atlantic and Indo-Pacific oceans.

### 6.3 Implications and future avenues

This thesis provides evidence for the potential role of trophic evolution in shaping macroevolutionary and macroecological patterns in coral reef fishes. The findings presented herein identify species ecology as a potential major driver of present-day biodiversity distribution patterns. While the research has been focused on a specific group of marine vertebrates, the patterns and processes described in each of my chapters might transcend taxonomic boundaries. Therefore, future research may benefit from an increased focus on the consequences of trophic evolution, examining the extent to which it affects extant species distribution patterns in other groups of organisms. This would be particularly important in the marine realm, where our ecological knowledge generally lags behind most terrestrial systems (Webb, 2012). Additionally, it would be interesting to investigate the evolutionary forces that drove the rise of trophic innovations in the first place. The coarse dietary classification used throughout my chapters was essential to provide enough resolution to answer questions at large temporal and spatial scales. However, describing species diets using a more detailed resolution will surely be important for understanding how finely niches are partitioned between species and how that translates into finer-scale evolutionary processes. For instance, recent studies using gut content metabarcoding techniques and stable isotope analyses are providing unprecedented

levels of resolution into species trophic ecologies in coral reef fishes (e.g. Casey *et al.*, 2019; Eurich *et al.*, 2019). Despite the potential challenges, combining these recent advances in trophic ecology with modern phylogenetic approaches is likely to be a fruitful avenue for future research by ecologists and evolutionary biologists alike.

In summary, my research bridges two fundamental fields in biology: ecology and evolution. This evolutionary perspective focused on species roles in ecosystems offered new insights into the processes generating and maintaining the extant diversity of fishes on coral reefs. In **chapter 2**, I showed that species trophic identity is the most important factor for predicting the pace at which new reef fish lineages originate. While herbivory was shown to drive rates of diversification, planktivores were shown to be a key transition destination throughout reef fish evolution. Furthermore, in **chapter 3**, planktivores were also shown to be major drivers of global species richness patterns in coral reef fishes. There are substantially more planktivores in the Indo-Australian Archipelago than any other trophic guild. These patterns probably arose in the recent geological past, but the deep time patterns of herbivorous lineage evolution explored in **chapters 4 and 5** have revealed that the Earth's geological history has had a prominent role in shaping the faunas between the two major marine realms: the Atlantic and Indo-Pacific. The Atlantic went through extinctions, while the Indo-Pacific has a history of connectivity with the ancient biodiversity hotspots. Overall, this thesis shows that the global composition of present-day coral reef fishes results largely from the interplay between historical contingencies and the evolution of novel trophic strategies. It is my hope that the research presented herein contributes to a better understanding of trophic evolution in coral reef fishes and that it will stimulate new and creative approaches that will unravel the history underpinning present-day biodiversity patterns.

## References

- Adam, T.C., Burkepile, D.E., Ruttenberg, B.I. & Paddock, M.J. (2015) Herbivory and the resilience of Caribbean coral reefs: knowledge gaps and implications for management. *Marine Ecology Progress Series*, **520**, 1–20.
- Alfaro, M.E., Alfaro, M.E., Santini, F., Santini, F., Brock, C., Brock, C., Alamillo, H., Alamillo, H., Dornburg, A., Dornburg, A., Rabosky, D.L., Rabosky, D.L., Carnevale, G., Carnevale, G., Harmon, L.J. & Harmon, L.J. (2009a) Nine exceptional radiations plus high turnover explain species diversity in jawed vertebrates. *Proceedings of the National Academy of Sciences of the United States of America*, **106**, 13410–13414.
- Alfaro, M.E., Brock, C.D., Banbury, B.L. & Wainwright, P.C. (2009b) Does evolutionary innovation in pharyngeal jaws lead to rapid lineage diversification in labrid fishes? *BMC evolutionary biology*, **9**, 255.
- Alfaro, M.E., Faircloth, B.C., Harrington, R.C., Sorenson, L., Friedman, M., Thacker, C.E., Oliveros, C.H., Černý, D. & Near, T.J. (2018) Explosive diversification of marine fishes at the Cretaceous-Palaeogene boundary. *Nature Ecology and Evolution*, **2**, 688–696.
- Alfaro, M.E., Santini, F. & Brock, C.D. (2007) Do Reefs Drive Diversification in Marine Teleosts? Evidence From the Pufferfish and Their Allies (Order Tetraodontiformes). *Evolution*, **61**, 2104–2126.
- Arcila, D., Pyron, A.R., Tyler, J.C., Ortí, G. & Betancur-R, R. (2015) An evaluation of fossil tip-dating versus node-age calibrations in tetraodontiform fishes (Teleostei: Percomorphaeae). *Molecular Phylogenetics and Evolution*, **82**, 131–145.
- Bannikov, A.F. & Tyler, J.C. (2002) A new genus and species of rabbitfish (Acanthuroidei: Siganidae) from the Eocene of Monte Bolca, Italy. *Studi e Ricerche sui Giacimenti Terziari di Bolca, Museo Civico di Storia Naturale di Verona, Miscellanea Paleontologica*, **9**, 37–45.
- Bannikov, A.F. & Tyler, J.C. (2012) Redescription of the Eocene of Monte Bolca, Italy, surgeon fish *Pesciarichthys punctatus* (Perciformes, Acanthuridae), and a new genus, *Frigosorbinia*, for *P. baldwinae*. *Studi e Ricerche sui Giacimenti Terziari di Bolca, Museo Civico di Storia Naturale di Verona, Miscellanea Paleontologica*, **15**, 19–30.
- Bannikov, A.F., Tyler, J.C. & Sorbini, C. (2010) Two new taxa of Eocene rabbitfishes (Perciformes, Siganidae) from the North Caucasus (Russia), with redescription of *Acanthopygaeus agassizi* (Eastman) from Monte Bolca (Italy) and a phylogenetic analysis of the family. *Bollettino del Museo Civico di Storia Naturale di Verona, Geologia, Paleontologia, Preistoria*, **34**, 3–21.
- Baraf, L.M., Pratchett, M.S. & Cowman, P.F. (2019) Ancestral biogeography and ecology of marine angelfishes (F: Pomacanthidae). *Molecular Phylogenetics and Evolution*, **140**, 106596.
- Barneche, D.R., Rezende, E.L., Parravicini, V., Maire, E., Edgar, G.J., Stuart-Smith, R.D., Arias-González, J.E., Ferreira, C.E.L., Friedlander, A.M., Green, A.L., Luiz, O.J., Rodríguez-Zaragoza, F.A., Vigliola, L., Kulbicki, M. & Floeter, S.R. (2019) Body size, reef area and temperature predict global reef-fish species richness across spatial scales. *Global Ecology and Biogeography*, **28**, 315–327.
- Beaulieu, J.M. & O'Meara, B.C. (2016) Detecting hidden diversification shifts in models of trait-dependent speciation and extinction. *Systematic Biology*, **65**, 583–601.

## References

- Bejarano, S., Jouffray, J.B., Chollett, I., Allen, R., Roff, G., Marshall, A., Steneck, R., Ferse, S.C.A. & Mumby, P.J. (2017) The shape of success in a turbulent world: wave exposure filtering of coral reef herbivory. *Functional Ecology*, **31**, 1312–1324.
- Bellwood, D.R. (1990) A new fossil fish *Phyllopharyngodon longipinnis* gen. et sp. nov. (family Labridae) from the Eocene, Monte Bolca, Italy. *Studi e Ricerche sui Giacimenti Terziari di Bolca, Museo Civico di Storia Naturale, Verona*, **6**, 149–160.
- Bellwood, D.R. (1994) A Phylogenetic study of the parrotfishes family Scaridae (Pisces: Labroidae), with a revision of genera. *Records of the Australian Museum, Supplement*, 1–86.
- Bellwood, D.R. (2003) Origins and escalation of herbivory in fishes: a functional perspective. *Paleobiology*, **29**, 71–83.
- Bellwood, D.R. (1996) The Eocene fishes of Monte Bolca: the earliest coral reef fish assemblage. *Coral Reefs*, **15**, 11–19.
- Bellwood, D.R. & Choat, J.H. (1990) A functional analysis of grazing in parrotfishes (family Scaridae): the ecological implications. *Environmental Biology of Fishes*, **28**, 189–214.
- Bellwood, D.R. & Goatley, C.H.R. (2017) Can biological invasions save Caribbean coral reefs? *Current Biology*, **27**, R13–R14.
- Bellwood, D.R., Goatley, C.H.R. & Bellwood, O. (2017) The evolution of fishes and corals on reefs: form, function and interdependence. *Biological Reviews*, **92**, 878–901.
- Bellwood, D.R., Goatley, C.H.R., Brandl, S.J. & Bellwood, O. (2014a) Fifty million years of herbivory on coral reefs: fossils, fish and functional innovations. *Proceedings of the Royal Society B: Biological Sciences*, **281**, 20133046.
- Bellwood, D.R., Goatley, C.H.R., Cowman, P.F., Bellwood, O. & Mora, C. (2015) *The evolution of fishes on coral reefs: fossils, phylogenies and functions*. *Ecology of Fishes on Coral Reefs* (ed. by C. Mora), p. 388. Cambridge University Press, Cambridge.
- Bellwood, D.R., Hoey, A.S., Bellwood, O. & Goatley, C.H.R. (2014b) Evolution of long-toothed fishes and the changing nature of fish-benthos interactions on coral reefs. *Nature communications*, **5**, 3144.
- Bellwood, D.R. & Hughes, T.P. (2001) Regional-scale assembly rules and biodiversity of coral reefs. *Science*, **292**, 1532–1535.
- Bellwood, D.R., Hughes, T.P., Connolly, S.R. & Tanner, J. (2005) Environmental and geometric constraints on Indo-Pacific coral reef biodiversity. *Ecology Letters*, **8**, 643–651.
- Bellwood, D.R., Hughes, T.P., Folke, C. & Nyström, M. (2004) Confronting the coral reef crisis. *Nature*, **429**, 827–833.
- Bellwood, D.R., Klanten, S., Cowman, P.F., Pratchett, M.S., Konow, N. & van Herwerden, L. (2010) Evolutionary history of the butterflyfishes (f: Chaetodontidae) and the rise of coral feeding fishes. *Journal of evolutionary biology*, **23**, 335–349.
- Bellwood, D.R., Renema, W. & Rosen, B.R. (2012) *Biodiversity hotspots, evolution and coral reef biogeography*. *Biotic Evolution and Environmental Change in Southeast Asia* (ed. by D.J. Gower, K. Johnson, J. Richardson, B. Rosen, L. Rüber, and S. Williams), pp. 216–245. Cambridge University Press, Cambridge.

## References

- Bellwood, D.R. & Schultz, O. (1991) A Review of the fossil record of the Parrotfishes (Labroidae:Scaridae) with a description of a new *Calotomus* species from the Middle Miocene (Badenian) of Austria. *Annalen Des Naturhistorischen Museums in Wien. Serie A, Fur Mineralogie und Petrographie, Geologie Und Palaontologie, Anthropologie und Prahistorie*, **92**, 55–71.
- Bellwood, D.R., Schultz, O., Siqueira, A.C. & Cowman, P.F. (2019a) A review of the fossil record of the Labridae. *Annalen des Naturhistorischen Museums in Wien, Serie A*, **121**, 125–193.
- Bellwood, D.R., Streit, R.P., Brandl, S.J. & Tebbett, S.B. (2019b) The meaning of the term “function” in ecology: a coral reef perspective. *Functional Ecology*, **33**, 948–961.
- Bellwood, D.R., Tebbett, S.B., Bellwood, O., Mihalitsis, M., Morais, R.A., Streit, R.P. & Fulton, C.J. (2018) The role of the reef flat in coral reef trophodynamics: Past, present, and future. *Ecology and Evolution*, **8**, 4108–4119.
- Bellwood, D.R. & Wainwright, P.C. (2002) *The history and biogeography of fishes on coral reefs. Coral reef fishes: dynamics and diversity on a complex ecosystem* (ed. by P.F. Sale), pp. 5–32. Academic Press, San Diego.
- Benson, R.B.J., Campione, N.E., Carrano, M.T., Mannion, P.D., Sullivan, C., Upchurch, P. & Evans, D.C. (2014) Rates of dinosaur body mass evolution indicate 170 million years of sustained ecological innovation on the avian stem lineage. *PLoS Biology*, **12**, e1001853.
- Benton, M.J. (2009) The Red Queen and the Court Jester: species diversity and the role of biotic and abiotic factors through time. *Science*, **323**, 728–732.
- Bininda-Emonds, O.R.P., Cardillo, M., Jones, K.E., MacPhee, R.D.E., Beck, R.M.D., Grenyer, R., Price, S.A., Vos, R.A., Gittleman, J.L. & Purvis, A. (2007) The delayed rise of present-day mammals. *Nature*, **446**, 507–512.
- Blot, J. & Voruz, C. (1970) Les poissons fossils du Monte Bolca. *Memorie del Museo Civico di Storia Naturale di Verona*, **18**, 31–42.
- Blount, Z.D., Lenski, R.E. & Losos, J.B. (2018) Contingency and determinism in evolution: Replaying life’s tape. *Science*, **362**, eaam5979.
- Boettiger, C., Lang, D.T. & Wainwright, P.C. (2012) Rfishbase: Exploring, manipulating and visualizing FishBase data from R. *Journal of Fish Biology*, **81**, 2030–2039.
- Bonaldo, R.M., Hoey, A.S. & Bellwood, D.R. (2014) The ecosystem roles of parrotfishes on tropical reefs. *Oceanography and Marine Biology: An Annual Review*, **52**, 81–132.
- Borstein, S.R., Fordyce, J.A., O’Meara, B.C., Wainwright, P.C. & McGee, M.D. (2019) Reef fish functional traits evolve fastest at trophic extremes. *Nature Ecology and Evolution*, **3**, 191–199.
- Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C.H., Xie, D., Suchard, M.A., Rambaut, A. & Drummond, A.J. (2014) BEAST 2: a software platform for Bayesian evolutionary analysis. *PLoS Computational Biology*, **10**, e1003537.
- Bowen, B.W., Muss, A., Rocha, L.A. & Grant, W.S. (2006) Shallow mtDNA coalescence in Atlantic pygmy angelfishes (genus *Centropyge*) indicates a recent invasion from the Indian Ocean. *Journal of Heredity*, **97**, 1–12.
- Brandl, S.J. & Bellwood, D.R. (2013) Morphology, sociality, and ecology: can morphology predict pairing

## References

- behavior in coral reef fishes? *Coral Reefs*, **32**, 835–846.
- Brandl, S.J., Goatley, C.H.R., Bellwood, D.R. & Tornabene, L. (2018) The hidden half: ecology and evolution of cryptobenthic fishes on coral reefs. *Biological Reviews*, **93**, 1846–1873.
- Brandl, S.J., Tornabene, L., Goatley, C.H.R., Casey, J.M., Morais, R.A., Côté, I.M., Baldwin, C.C., Parravicini, V., Schiattkatte, N.M.D. & Bellwood, D.R. (2019) Demographic dynamics of the smallest marine vertebrates fuel coral reef ecosystem functioning. *Science*, **364**, 1189–1192.
- Briggs, J.C. & Bowen, B.W. (2013) Marine shelf habitat: biogeography and evolution. *Journal of Biogeography*, **40**, 1023–1035.
- Brooks, D.R. & McLennan, D.A. (1991) *Phylogeny, Ecology, and Behavior: A Research Program in Comparative Biology*, University of Chicago Press, Chicago.
- Brown, J.H., Gillooly, J.F., Allen, A.P., Savage, V.M. & West, G.B. (2004) Toward a metabolic theory of ecology. *Ecology*, **85**, 1771–1789.
- Budd, A.F. (2000) Diversity and extinction in the Cenozoic history of Caribbean reefs. *Coral Reefs*, **19**, 25–35.
- Burgin, C.J., Colella, J.P., Kahn, P.L. & Upham, N.S. (2018) How many species of mammals are there? *Journal of Mammalogy*, **99**, 1–14.
- Burin, G., Kissling, W.D., Guimaraes, P.R., Sekercioglu, çagan H. & Quental, T.B. (2016) Omnivory in birds is a macroevolutionary sink. *Nature Communications*, **7**, 11250.
- Burkepile, D.E. & Hay, M.E. (2008) Herbivore species richness and feeding complementarity affect community structure and function on a coral reef. *Proceedings of the National Academy of Sciences*, **105**, 16201–16206.
- Burkepile, D.E. & Hay, M.E. (2006) Herbivore vs. nutrient control of marine primary producers: context-dependent effects. *Ecology*, **87**, 3128–3139.
- Campbell, S.J., Darling, E.S., Pardede, S., Ahmadi, G., Mangubhai, S., Amkieltiela, Estradivari & Maire, E. (2020) Fishing restrictions and remoteness deliver conservation outcomes for Indonesia’s coral reef fisheries. *Conservation Letters*, **13**, e12698.
- Cantalice, K.M. & Alvarado-Ortega, J. (2016) *Eekaulostomus cuevasae* gen. and sp. nov., an ancient armored trumpetfish (Aulostomoidea) from Danian (Paleocene) marine deposits of Belisario Domínguez, Chiapas, southeastern Mexico. *Palaeontologia Electronica*, **19.3.53A**, 1–24.
- Cantalice, K.M., Alvarado-Ortega, J. & Alaniz-Galvan, A. (2018) *Paleoserranus lakamhae* gen. et sp. nov., a Paleocene seabass (Perciformes: Serranidae) from Palenque, Chiapas, southeastern Mexico. *Journal of South American Earth Sciences*, **83**, 137–146.
- Cantalice, K.M., Alvarado-Ortega, J. & Bellwood, D.R. (2020) †*Chaychanus gonzalezorum* gen. et sp. nov.: A damselfish fossil (Percomorphaceae; Pomacentridae), from the Early Paleocene outcrop of Chiapas, Southeastern Mexico. *Journal of South American Earth Sciences*, **98**, 102322.
- Casey, J.M., Meyer, C.P., Morat, F., Brandl, S.J., Planes, S. & Parravicini, V. (2019) Reconstructing hyperdiverse food webs: Gut content metabarcoding as a tool to disentangle trophic interactions on coral reefs. *Methods in Ecology and Evolution*, **10**, 1157–1170.
- Chang, J., Rabosky, D.L. & Alfaro, M.E. (2020) Estimating diversification rates on incompletely sampled

## References

- phylogenies: theoretical concerns and practical solutions. *Systematic Biology*, **69**, 602–611.
- Chen, T. & Guestrin, C. (2016) *XGBoost: A scalable tree boosting system*. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ACM Press, New York, NY.
- Chen, T., He, T., Benesty, M., Khotilovich, V., Tang, Y., Cho, H., Chen, K., Mitchell, R., Cano, I., Zhou, T., Li, M., Xie, J., Lin, M., Geng, Y. & Li, Y. (2019) xgboost: Extreme Gradient Boosting. R package version 0.82.1.
- Choat, J.H. (1991) *The biology of herbivorous fishes on coral reefs*. *The Ecology of Fishes on Coral Reefs* (ed. by P.F. Sale), pp. 120–155. Academic Press, San Diego.
- Choat, J.H. & Clements, K.D. (1998) Vertebrate herbivores in marine and terrestrial environments: A nutritional ecology perspective. *Annual Review of Ecology and Systematics*, **29**, 375–403.
- Choat, J.H., Klanten, O.S., Van Herwerden, L., Robertson, R.D. & Clements, K.D. (2012) Patterns and processes in the evolutionary history of parrotfishes (Family Labridae). *Biological Journal of the Linnean Society*, **107**, 529–557.
- Clements, K.D., German, D.P., Piché, J., Tribollet, A. & Choat, J.H. (2017) Integrating ecological roles and trophic diversification on coral reefs: multiple lines of evidence identify parrotfishes as microphages. *Biological Journal of the Linnean Society*, **120**, 729–751.
- Connell, J.H. & Orias, E. (1964) The ecological regulation of species diversity. *The American Naturalist*, **98**, 399–414.
- Connolly, S.R., Bellwood, D.R. & Hughes, T.P. (2003) Indo-pacific biodiversity of coral reefs: deviations from a mid-domain model. *Ecology*, **84**, 2178–2190.
- Cooper, W.J. & Westneat, M.W. (2009) Form and function of damselfish skulls: Rapid and repeated evolution into a limited number of trophic niches. *BMC Evolutionary Biology*, **9**, 24.
- Cowman, P.F. & Bellwood, D.R. (2011) Coral reefs as drivers of cladogenesis: expanding coral reefs, cryptic extinction events, and the development of biodiversity hotspots. *Journal of Evolutionary Biology*, **24**, 2543–2562.
- Cowman, P.F. & Bellwood, D.R. (2013a) The historical biogeography of coral reef fishes: global patterns of origination and dispersal. *Journal of Biogeography*, **40**, 209–224.
- Cowman, P.F. & Bellwood, D.R. (2013b) Vicariance across major marine biogeographic barriers: temporal concordance and the relative intensity of hard versus soft barriers. *Proceedings of the Royal Society B: Biological Sciences*, **280**, 20131541.
- Cowman, P.F., Bellwood, D.R. & van Herwerden, L. (2009) Dating the evolutionary origins of wrasse lineages (Labridae) and the rise of trophic novelty on coral reefs. *Molecular phylogenetics and evolution*, **52**, 621–631.
- Cowman, P.F., Parravicini, V., Kulbicki, M. & Floeter, S.R. (2017) The biogeography of tropical reef fishes: endemism and provinciality through time. *Biological Reviews*, **92**, 2112–2130.
- Coyne, J.A. & Orr, H.A. (2004) *Speciation*, Oxford University Press USA, Sunderland (MA).
- Cribari-Neto, F. & Zeileis, A. (2010) Beta regression in R. *Journal of Statistical Software*, **34**, 1–24.
- Darwin, C. (1859) *On the origin of species by means of natural selection, or, the preservation of favoured races in the struggle for life*, John Murray, London.

## References

- Di Martino, E., Jackson, J.B.C., Taylor, P.D. & Johnson, K.G. (2018) Differences in extinction rates drove modern biogeographic patterns of tropical marine biodiversity. *Science Advances*, **4**, eaaq1508.
- Donati, G.F.A., Parravicini, V., Leprieur, F., Hagen, O., Gaboriau, T., Heine, C., Kulbicki, M., Rolland, J., Salamin, N., Albouy, C. & Pellissier, L. (2019) A process-based model supports an association between dispersal and the prevalence of species traits in tropical reef fish assemblages. *Ecography*, **42**, 2095–2106.
- Dornburg, A., Moore, J., Beaulieu, J.M., Eytan, R.I. & Near, T.J. (2015) The impact of shifts in marine biodiversity hotspots on patterns of range evolution: evidence from the Holocentridae (squirrelfishes and soldierfishes). *Evolution*, **69**, 146–161.
- Dupin, J., Matzke, N.J., Särkinen, T., Knapp, S., Olmstead, R.G., Bohs, L. & Smith, S.D. (2017) Bayesian estimation of the global biogeographical history of the Solanaceae. *Journal of Biogeography*, **44**, 887–899.
- Edgar, G.J. & Stuart-Smith, R.D. (2014) Systematic global assessment of reef fish communities by the Reef Life Survey program. *Scientific Data*, **1**, 1–140007.
- Edgar, R.C. (2004) MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, **32**, 1792–1797.
- Elith, J., Leathwick, J.R. & Hastie, T. (2008) A working guide to boosted regression trees. *Journal of Animal Ecology*, **77**, 802–813.
- Eschmeyer, W.N., Fricke, R., Fong, J.D. & Polack, D.A. (2010) Marine fish diversity: History of knowledge and discovery (Pisces). *Zootaxa*, **50**, 19–50.
- Eurich, J.G., Matley, J.K., Baker, R., McCormick, M.I. & Jones, G.P. (2019) Stable isotope analysis reveals trophic diversity and partitioning in territorial damselfishes on a low-latitude coral reef. *Marine Biology*, **166**, 17.
- Farrell, B.D. (1998) “Inordinate fondness” explained: Why are there so many beetles? *Science*, **281**, 555–559.
- Fine, P.V.A. (2015) Ecological and evolutionary drivers of geographic variation in species diversity. *Annual Review of Ecology, Evolution, and Systematics*, **46**, 369–392.
- Fischer, A.G. (1960) Latitudinal variations in organic diversity. *Evolution*, **14**, 64–81.
- FitzJohn, R.G. (2012) Diversitree: comparative phylogenetic analyses of diversification in R. *Methods in Ecology and Evolution*, **3**, 1084–1092.
- Floeter, S.R., Bender, M.G., Siqueira, A.C. & Cowman, P.F. (2018) Phylogenetic perspectives on reef fish functional traits. *Biological Reviews*, **93**, 131–151.
- Floeter, S.R., Rocha, L.A., Robertson, D.R., Joyeux, J.C., Smith-Vaniz, W.F., Wirtz, P., Edwards, A.J., Barreiros, J.P., Ferreira, C.E.L., Gasparini, J.L., Brito, A., Falcón, J.M., Bowen, B.W. & Bernardi, G. (2008) Atlantic reef fish biogeography and evolution. *Journal of Biogeography*, **35**, 22–47.
- Freckleton, R.P. (2012) Fast likelihood calculations for comparative analyses. *Methods in Ecology and Evolution*, **3**, 940–947.
- Frédérich, B., Sorenson, L., Santini, F., Slater, G.J. & Alfaro, M.E. (2013) Iterative ecological radiation and convergence during the evolutionary history of damselfishes (Pomacentridae). *The American Naturalist*, **181**, 94–113.
- Fricke, R., Eschmeyer, W.N. & Van der Laan, R. (2019) Eschmeyer’s catalog of fishes: genera, species,

## References

references.

- Friedman, M. (2010) Explosive morphological diversification of spiny-finned teleost fishes in the aftermath of the end-Cretaceous extinction. *Proceedings of the Royal Society B: Biological Sciences*, **277**, 1675–1683.
- Friedman, M. & Sallan, L.C. (2012) Five hundred million years of extinction and recovery: a phanerozoic survey of large-scale diversity patterns in fishes. *Palaeontology*, **55**, 707–742.
- Friedman, S.T., Price, S.A., Hoey, A.S. & Wainwright, P.C. (2016) Ecomorphological convergence in planktivorous surgeonfishes. *Journal of Evolutionary Biology*, **29**, 965–978.
- Fritz, S.A., Schnitzler, J., Eronen, J.T., Hof, C., Böhning-Gaese, K. & Graham, C.H. (2013) Diversity in time and space: wanted dead and alive. *Trends in Ecology and Evolution*, **28**, 509–516.
- Froese, R. & Pauly, D. (2019) FishBase. Available at: [www.fishbase.org](http://www.fishbase.org), version (04/2019).
- Fulton, C.J. & Bellwood, D.R. (2005) Wave-induced water motion and the functional implications for coral reef fish assemblages. *Limnology and Oceanography*, **50**, 255–264.
- Gaither, M.R. & Rocha, L.A. (2013) Origins of species richness in the Indo-Malay-Philippine biodiversity hotspot: Evidence for the centre of overlap hypothesis. *Journal of Biogeography*, **40**, 1638–1648.
- Gajdzik, L., Aguilar-Medrano, R. & Frédérick, B. (2019) Diversification and functional evolution of reef fish feeding guilds. *Ecology Letters*, **22**, 572–582.
- Gaston, K.J. (2000) Global patterns in biodiversity. *Nature*, **405**, 220–227.
- Gaudant, J., Tsaparas, N., Antonarakou, A., Drinia, H. & Dermitzakis, M.D. (2005) The Tortonian fish fauna of Gavdos Island (Greece). *Comptes Rendus - Palevol*, **4**, 687–695.
- Gavryushkina, A., Welch, D., Stadler, T. & Drummond, A.J. (2014) Bayesian inference of sampled ancestor trees for epidemiology and fossil calibration. *PLoS Computational Biology*, **10**, e1003919.
- Goatley, C.H.R. & Bellwood, D.R. (2016) Body size and mortality rates in coral reef fishes: A three-phase relationship. *Proceedings of the Royal Society B: Biological Sciences*, **283**.
- Goatley, C.H.R., Bonaldo, R.M., Fox, R.J. & Bellwood, D.R. (2016) Sediments and herbivory as sensitive indicators of coral reef degradation. *Ecology and Society*, **21**, 29.
- Goldberg, E.E., Lancaster, L.T. & Ree, R.H. (2011) Phylogenetic inference of reciprocal effects between geographic range evolution and diversification. *Systematic Biology*, **60**, 451–465.
- Gordon, A.L. (2005) Oceanography of the Indonesian seas and their throughflow. *Oceanography*, **18**, 15–27.
- Gordon, A.L. & Fine, R.A. (1996) Pathways of water between the Pacific and Indian oceans in the Indonesian seas. *Nature*, **379**, 146–149.
- Gould, S.J. (1989) *Wonderful Life*, W. W. Norton & Company, New York.
- Green, A.L. & Bellwood, D.R. (2009) *Monitoring functional groups of herbivorous reef fishes as indicators of coral reef resilience: a practical guide for coral reef managers in the Asia Pacific region*, IUCN Resilience Science Group Working Paper Series No 7, Gland, Switzerland.
- Gu, Z., Gu, L., Eils, R., Schlesner, M. & Brors, B. (2014) Circlize implements and enhances circular visualization in R. *Bioinformatics*, **30**, 2811–2812.
- Hall, R. (2002) Cenozoic geological and plate tectonic evolution of SE Asia and the SW Pacific: computer-based reconstructions, model and animations. *Journal of Asian Earth Sciences*, **20**, 353–431.

## References

- Hamner, W.M., Jones, M.S., Carleton, J.H., Hauri, I.R. & Williams, D.M. (1988) Zooplankton, planktivorous fish, and water currents on a windward reef face: Great Barrier Reef, Australia. *Bulletin of Marine Science*, **42**, 459–479.
- Harmelin-Vivien, M.L. (2002) *Energetics and fish diversity on coral reefs*. *Coral reef fishes: dynamics and diversity on a complex ecosystem* (ed. by P. Sale), pp. 265–274. Academic Press, San Diego.
- Harmon, L.J., Losos, J.B., Jonathan Davies, T., Gillespie, R.G., Gittleman, J.L., Bryan Jennings, W., Kozak, K.H., McPeck, M.A., Moreno-Roark, F., Near, T.J., Purvis, A., Ricklefs, R.E., Schluter, D., Schulte, J.A., Seehausen, O., Sidlauskas, B.L., Torres-Carvajal, O., Weir, J.T. & Mooers, A.T. (2010) Early bursts of body size and shape evolution are rare in comparative data. *Evolution*, **64**, 2385–2396.
- Hay, M.E. (1981) Spatial patterns of grazing intensity on a Caribbean barrier reef: Herbivory and algal distribution. *Aquatic Botany*, **11**, 97–109.
- Heath, T.A., Huelsenbeck, J.P. & 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**, E2957–E2966.
- Heenan, A., Williams, G.J. & Williams, I.D. (2020) Natural variation in coral reef trophic structure across environmental gradients. *Frontiers in Ecology and the Environment*, **18**, 69–75.
- Hemingson, C.R. & Bellwood, D.R. (2018) Biogeographic patterns in major marine realms: function not taxonomy unites fish assemblages in reef, seagrass and mangrove systems. *Ecography*, **41**, 174–182.
- Henao Diaz, L.F., Harmon, L.J., Sugawara, M.T.C., Miller, E.T. & Pennell, M.W. (2019) Macroevolutionary diversification rates show time dependency. *Proceedings of the National Academy of Sciences of the United States of America*, **116**, 7403–7408.
- Herrera-Alsina, L., van Els, P. & Etienne, R.S. (2019) Detecting the dependence of diversification on multiple traits from phylogenetic trees and trait data. *Systematic Biology*, **68**, 317–328.
- Hijmans, R.J. (2019) geosphere: Spherical Trigonometry. R package version 1.5-10.
- Hillebrand, H. (2004) On the generality of the latitudinal diversity gradient. *The American Naturalist*, **163**, 192–211.
- Hobson, E. & Chess, J. (1978) Trophic relationships among fishes and plankton in the lagoon at Enewetak Atoll, Marshall Islands. *Fishery bulletin*, **76**, 133–153.
- Hobson, E.S. (1991) *Trophic relationships of fishes specialized to feed on zooplankters above coral reefs*. *The Ecology of Fishes on Coral Reefs* (ed. by P.F. Sale), pp. 69–95. Academic Press, San Diego.
- Hodge, J.R. & Bellwood, D.R. (2016) The geography of speciation in coral reef fishes: the relative importance of biogeographical barriers in separating sister-species. *Journal of Biogeography*, **43**, 1324–1335.
- Hodge, J.R., van Herwerden, L. & Bellwood, D.R. (2014) Temporal evolution of coral reef fishes: global patterns and disparity in isolated locations. *Journal of Biogeography*, **41**, 2115–2127.
- Hoeksema, B.W. (2007) *Delineation of the Indo-Malayan centre of maximum marine biodiversity: the Coral Triangle*. *Biogeography, Time, and Place: Distributions, Barriers, and Islands* (ed. by W. Renema), pp. 117–178. Springer Netherlands, Dordrecht.
- Hou, Z. & Li, S. (2018) Tethyan changes shaped aquatic diversification. *Biological Reviews*, **93**, 874–896.

## References

- Huelsenbeck, J.P., Nielsen, R. & Bollback, J.P. (2003) Stochastic mapping of morphological characters. *Systematic Biology*, **52**, 131–138.
- Huertas, V. & Bellwood, D.R. (2020) Trophic separation in planktivorous reef fishes: a new role for mucus? *Oecologia*, **192**, 813–822.
- Hughes, T.P., Rodrigues, M.J., Bellwood, D.R., Ceccarelli, D., Hoegh-Guldberg, O., McCook, L., Moltschaniwskyj, N., Pratchett, M.S., Steneck, R.S. & Willis, B. (2007) Phase shifts, herbivory, and the resilience of coral reefs to climate change. *Current Biology*, **17**, 360–365.
- IUCN (2017) IUCN Red List of Threatened Species. Available at: [www.iucnredlist.org](http://www.iucnredlist.org), Version 2017.1.
- Ivany, L.C., Patterson, W.P. & Lohmann, K.C. (2000) Cooler winters as a possible cause of mass extinctions at the Eocene/Oligocene boundary. *Nature*, **407**, 887–890.
- Jablonski, D. (2008) Biotic interactions and macroevolution: extensions and mismatches across scales and levels. *Evolution*, **62**, 715–739.
- Jablonski, D. (2001) Lessons from the past: Evolutionary impacts of mass extinctions. *Proceedings of the National Academy of Sciences of the United States of America*, **98**, 5393–5398.
- Jablonski, D., Roy, K. & Valentine, J.W. (2006) Out of the tropics: evolutionary dynamics of the latitudinal diversity gradient. *Science*, **314**, 102–106.
- Jetz, W., Thomas, G.H., Joy, J.B., Hartmann, K. & Mooers, A.O. (2012) The global diversity of birds in space and time. *Nature*, **491**, 444–448.
- Johansen, J.L., Bellwood, D.R. & Fulton, C.J. (2008) Coral reef fishes exploit flow refuges in high-flow habitats. *Marine Ecology Progress Series*, **360**, 219–226.
- Johnson, K.G., Budd, A.F. & Stemann, T.A. (1995) Extinction selectivity and ecology of Neogene Caribbean reef corals. *Paleobiology*, **21**, 52–73.
- Joyeux, J., Floeter, S.R., Ferreira, C.E.L. & Gasparini, J.L. (2001) Biogeography of tropical reef fishes: the South Atlantic puzzle. *Journal of Biogeography*, **28**, 831–841.
- Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., Thierer, T., Ashton, B., Meintjes, P. & Drummond, A. (2012) Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*, **28**, 1647–1649.
- Kiessling, W. (2009) Geologic and Biologic Controls on the Evolution of Reefs. *Annual Review of Ecology and Systematics*, **40**, 173–192.
- Kiessling, W., Simpson, C. & Foote, M. (2010) Reefs as cradles of evolution and sources of biodiversity in the Phanerozoic. *Science*, **327**, 196–198.
- Knudsen, S.W. & Clements, K.D. (2016) World-wide species distributions in the family Kyphosidae (Teleostei: Perciformes). *Molecular Phylogenetics and Evolution*, **101**, 252–266.
- Konow, N. & Ferry-Graham, L.A. (2013) *Functional morphology of butterflyfishes*. *Biology of Butterflyfishes* (ed. by M.S. Pratchett, M.L. Berumen, and B.G. Kapoor), pp. 19–47. CRC Press, Boca Raton, FL.
- Kulbicki, M., Parravicini, V., Bellwood, D.R., Arias-González, E., Chabanet, P., Floeter, S.R., Friedlander, A., McPherson, J., Myers, R.E., Vigliola, L. & Mouillot, D. (2013) Global biogeography of reef fishes: a

## References

- hierarchical quantitative delineation of regions. *PLoS ONE*, **8**, e81847.
- Laliberté, A.E., Legendre, P., Shipley, B. & Laliberté, M.E. (2014) FD: measuring functional diversity from multiple traits, and other tools for functional ecology. *R package version 1.0.12*.
- Landis, M.J., Matzke, N.J., Moore, B.R. & Huelsenbeck, J.P. (2013) Bayesian analysis of biogeography when the number of areas is large. *Systematic Biology*, **62**, 789–804.
- Lanfear, R., Frandsen, P.B., Wright, A.M., Senfeld, T. & Calcott, B. (2016) PartitionFinder2 : new methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. *Molecular Biology and Evolution*, **34**, 772–773.
- Laurito, C.A., Calvo, C., Valerio, A.L., Calvo, A. & Chacón, R. (2014) Lower Miocene Ichthyofauna from the locality of Pacuare de Tres Equis, Río Banano formation, Cartago province, Costa Rica and description of a new genus and species of Scaridae. *Revista Geológica de América Central*, **50**, 153–192.
- Leprieur, F., Descombes, P., Gaboriau, T., Cowman, P.F. & Parravicini, V. (2016) Plate tectonics drive tropical reef biodiversity dynamics. *Nature Communications*, **7**, 11461.
- Lessios, H. & Robertson, D. (2006) Crossing the impassable: genetic connections in 20 reef fishes across the eastern Pacific barrier. *Proceedings of the Royal Society B: Biological Sciences*, **273**, 2201–2208.
- Lessios, H.A. (2008) The great American schism: divergence of marine organisms after the rise of the central American isthmus. *Annual Review of Ecology, Evolution, and Systematics*, **39**, 63–91.
- Lindsay, S.M. (2012) Dispersal of marine organisms and the grand challenges in biology: an introduction to the symposium. *Integrative and Comparative Biology*, **52**, 443–446.
- Lobato, F.L., Barneche, D.R., Siqueira, A.C., Liedke, A.M.R., Lindner, A., Pie, M.R., Bellwood, D.R. & Floeter, S.R. (2014) Diet and diversification in the evolution of coral reef fishes. *PLoS ONE*, **9**, e102094.
- Lohman, D.J., de Bruyn, M., Page, T., von Rintelen, K., Hall, R., Ng, P.K.L., Shih, H.-T., Carvalho, G.R. & von Rintelen, T. (2011) Biogeography of the Indo-Australian Archipelago. *Annual Review of Ecology, Evolution, and Systematics*, **42**, 205–226.
- Long, J.A. & Trinajstić, K. (2010) The late devonian gogo formation lagerstätte of western australia: Exceptional early vertebrate preservation and diversity. *Annual Review of Earth and Planetary Sciences*, **38**, 255–279.
- López-Pérez, A. (2017) *Revisiting the Cenozoic History and the Origin of the Eastern Pacific Coral Fauna*. *Coral Reefs of the Eastern Tropical Pacific. Persistence and Loss in a Dynamic Environment* (ed. by P.W. Glynn), D. Manzello, and I.C. Enochs), pp. 39–57. Springer Netherlands.
- Louca, S. & Pennell, M.W. (2020) Extant timetrees are consistent with a myriad of diversification histories. *Nature*, **580**, 502–505.
- Louca, S., Shih, P.M., Pennell, M.W., Fischer, W.W., Parfrey, L.W. & Doebeli, M. (2018) Bacterial diversification through geological time. *Nature Ecology and Evolution*, **2**, 1458–1467.
- Luiz, O.J., Allen, A.P., Robertson, D.R., Floeter, S.R., Kulbicki, M., Vigliola, L., Becheler, R. & Madin, J.S. (2013) Adult and larval traits as determinants of geographic range size among tropical reef fishes. *Proceedings of the National Academy of Sciences of the United States of America*, **110**, 16498–16502.
- MacArthur, R. (1970) Species packing and competitive equilibrium for many species. *Theoretical Population Biology*, **1**, 1–11.

## References

- Maddison, W.P. (2006) Confounding Asymmetries in Evolutionary Diversification and Character Change. *Evolution*, **60**, 1743.
- Maire, E., Grenouillet, G., Brosse, S. & Villéger, S. (2015) How many dimensions are needed to accurately assess functional diversity? A pragmatic approach for assessing the quality of functional spaces. *Global Ecology and Biogeography*, **24**, 728–740.
- Maliet, O., Hartig, F. & Morlon, H. (2019) A model with many small shifts for estimating species-specific diversification rates. *Nature Ecology and Evolution*, **3**, 1086–1092.
- Marshall, C.R. (2017) Five palaeobiological laws needed to understand the evolution of the living biota. *Nature Ecology and Evolution*, **1**, 0165.
- Martin, A.P. & Palumbi, S.R. (1993) Body size, metabolic rate, generation time, and the molecular clock. *Proceedings of the National Academy of Sciences of the United States of America*, **90**, 4087–4091.
- Matzke, N. (2013) BioGeoBEARS: BioGeography with Bayesian (and Likelihood) Evolutionary Analysis in R Scripts. *R package, version 0.2.1*.
- Matzke, N.J. (2014) Model selection in historical biogeography reveals that founder-event speciation is a crucial process in island clades. *Systematic Biology*, **63**, 951–970.
- McKenna, D.D., Shin, S., Ahrens, D., Balke, M., Beza-Beza, C., Clarke, D.J., Donath, A., Escalona, H.E., Friedrich, F., Letsch, H., Liu, S., Maddison, D., Mayer, C., Misof, B., Murin, P.J., Niehuis, O., Peters, R.S., Podsiadlowski, L., Pohl, H., Scully, E.D., Yan, E. V., Zhou, X., Ślipiński, A. & Beutel, R.G. (2019) The evolution and genomic basis of beetle diversity. *Proceedings of the National Academy of Sciences of the United States of America*, **116**, 24729–24737.
- Meade, A. & Pagel, M. (2018) BayesTraits: a computer package for analyses of trait evolution. Version 3.
- Meyer, A.L.S., Román-Palacios, C. & Wiens, J.J. (2018) BAMM gives misleading rate estimates in simulated and empirical datasets. *Evolution*, **72**, 2257–2266.
- Mihaljević, M., Renema, W., Welsh, K. & Pandolfi, J.M. (2014) Eocene-Miocene Shallow-Water Carbonate Platforms and Increased Habitat Diversity in Sarawak, Malaysia. *Palaeos*, **29**, 378–391.
- Miller, E.C., Hayashi, K.T., Song, D. & Wiens, J.J. (2018) Explaining the ocean’s richest biodiversity hotspot and global patterns of fish diversity. *Proceedings of the Royal Society B: Biological Sciences*, **285**, 20181314.
- Miller, M.A., Pfeiffer, W. & Schwartz, T. (2010) *Creating the CIPRES Science Gateway for inference of large phylogenetic trees. 2010 Gateway Computing Environments Workshop, GCE 2010*.
- Miller, T.L. & Cribb, T.H. (2007) Phylogenetic relationships of some common Indo-Pacific snappers (Perciformes: Lutjanidae) based on mitochondrial DNA sequences, with comments on the taxonomic position of the Caesioninae. *Molecular Phylogenetics and Evolution*, **44**, 450–460.
- Mitter, C., Farrell, B. & Wiegmann, B. (1988) The phylogenetic study of adaptive zones: has phytophagy promoted insect diversification? *American Naturalist*, **132**, 107–128.
- Moore, B.R., Höhna, S., May, M.R., Rannala, B. & Huelsenbeck, J.P. (2016) Critically evaluating the theory and performance of Bayesian analysis of macroevolutionary mixtures. *Proceedings of the National Academy of Sciences of the United States of America*, **113**, 9569–9574.
- Mora, C., Chittaro, P.M., Sale, P.F., Kritzer, J.P. & Ludsin, S.A. (2003) Patterns and processes in reef fish

## References

- diversity. *Nature*, **421**, 933–936.
- Morais, R.A. & Bellwood, D.R. (2018) Global drivers of reef fish growth. *Fish and Fisheries*, **19**, 874–889.
- Morais, R.A. & Bellwood, D.R. (2019) Pelagic Subsidies Underpin Fish Productivity on a Degraded Coral Reef. *Current Biology*, **29**, 1521-1527.e6.
- Mouillot, D., Graham, N.A.J., Villéger, S., Mason, N.W.H. & Bellwood, D.R. (2013) A functional approach reveals community responses to disturbances. *Trends in Ecology and Evolution*, **28**, 167–177.
- Mouillot, D., Villéger, S., Parravicini, V., Kulbicki, M., Arias-González, J.E., Bender, M., Chabanet, P., Floeter, S.R., Friedlander, A., Vigliola, L. & Bellwood, D.R. (2014) Functional over-redundancy and high functional vulnerability in global fish faunas of tropical reefs. *Proceedings of the National Academy of Sciences of the United States of America*, **111**, 13757–13762.
- Mouillot, D., Villéger, S., Scherer-Lorenzen, M. & Mason, N.W.H. (2011) Functional structure of biological communities predicts ecosystem multifunctionality. *PLoS ONE*, **6**, e17476.
- Near, T.J., Eytan, R.I., Dornburg, A., Kuhn, K.L., Moore, J.A., Davis, M.P., Wainwright, P.C., Friedman, M. & Smith, W.L. (2012) Resolution of ray-finned fish phylogeny and timing of diversification. *Proceedings of the National Academy of Sciences of the United States of America*, **109**, 13698–703.
- O’Dea, A., Jackson, J.B.C., Fortunato, H., Smith, J.T., D’Croze, L., Johnson, K.G. & Todd, J.A. (2007) Environmental change preceded Caribbean extinction by 2 million years. *Proceedings of the National Academy of Sciences*, **104**, 5501–5506.
- Paradis, E., Claude, J. & Strimmer, K. (2004) APE: Analyses of Phylogenetics and Evolution in R language. *Bioinformatics*, **20**, 289–290.
- Parravicini, V., Kulbicki, M., Bellwood, D.R., Friedlander, A.M., Arias-Gonzalez, J.E., Chabanet, P., Floeter, S.R., Myers, R., Vigliola, L., D’Agata, S. & Mouillot, D. (2013) Global patterns and predictors of tropical reef fish species richness. *Ecography*, **36**, 1254–1262.
- Pawlik, J.R., Burkepile, D.E. & Thurber, R.V. (2016) A Vicious Circle? Altered Carbon and Nutrient Cycling May Explain the Low Resilience of Caribbean Coral Reefs. *BioScience*, **66**, 470–476.
- Pellissier, L., Leprieux, F., Parravicini, V., Cowman, P.F., Kulbicki, M., Litsios, G., Olsen, S.M., Wisz, M.S., Bellwood, D.R. & Mouillot, D. (2014) Quaternary coral reef refugia preserved fish diversity. *Science*, **344**, 1016–1019.
- Plummer, M., Best, N., Cowles, K. & Vines, K. (2006) CODA: Convergence Diagnosis and Output Analysis for MCMC. *R News*, **6**, 7–11.
- Poore, A.G.B., Ahyong, S.T., Lowry, J.K. & Sotka, E.E. (2017) Plant feeding promotes diversification in the Crustacea. *Proceedings of the National Academy of Sciences of the United States of America*, **114**, 8829–8834.
- Price, S.A., Hopkins, S.S.B., Smith, K.K. & Roth, V.L. (2012) Tempo of trophic evolution and its impact on mammalian diversification. *Proceedings of the National Academy of Sciences of the United States of America*, **109**, 7008–7012.
- Price, S.A. & Schmitz, L. (2016) A promising future for integrative biodiversity research: an increased role of scale-dependency and functional biology. *Philosophical Transactions of the Royal Society B: Biological*

## References

- Sciences*, **371**, 20150228.
- Price, S.A., Wainwright, P.C., Bellwood, D.R., Kazancioglu, E., Collar, D.C. & Near, T.J. (2010) Functional Innovations and morphological diversification in parrotfish. *Evolution*, **64**, 3057–3068.
- Quental, T.B. & Marshall, C.R. (2010) Diversity dynamics: molecular phylogenies need the fossil record. *Trends in Ecology and Evolution*, **25**, 435–441.
- Quintero, I. & Jetz, W. (2018) Global elevational diversity and diversification of birds. *Nature*, **555**, 246–250.
- R Core Team (2019) R: a language and environment for statistical computing. *Version 3.5.3*.
- Rabosky, D.L. (2014) Automatic detection of key innovations, rate shifts, and diversity-dependence on phylogenetic trees. *PLoS ONE*, **9**.
- Rabosky, D.L. (2010) Extinction rates should not be estimated from molecular phylogenies. *Evolution*, **64**, 1816–1824.
- Rabosky, D.L. (2019) Phylogenies and diversification rates: Variance cannot be ignored. *Systematic Biology*, **68**, 538–550.
- Rabosky, D.L., Chang, J., Title, P.O., Cowman, P.F., Sallan, L., Friedman, M., Kaschner, K., Garilao, C., Near, T.J., Coll, M. & Alfaro, M.E. (2018) An inverse latitudinal gradient in speciation rate for marine fishes. *Nature*, **559**, 392–395.
- Rabosky, D.L., Chang, J., Title, P.O., Cowman, P.F., Sallan, L., Friedman, M., Kaschner, K., Garilao, C., Near, T.J., Coll, M. & Alfaro, M.E. (2019) Data from: An inverse latitudinal gradient in speciation rate for marine fishes. *Dryad*.
- Rabosky, D.L. & Goldberg, E.E. (2015) Model inadequacy and mistaken inferences of trait-dependent speciation. *Systematic Biology*, **64**, 340–355.
- Rabosky, D.L., Grundler, M., Anderson, C., Title, P., Shi, J.J., Brown, J.W., Huang, H. & Larson, J.G. (2014) BAMMtools: An R package for the analysis of evolutionary dynamics on phylogenetic trees. *Methods in Ecology and Evolution*, **5**, 701–707.
- Rabosky, D.L., Mitchell, J.S. & Chang, J. (2017) Is BAMM Flawed? Theoretical and Practical Concerns in the Analysis of Multi-Rate Diversification Models. *Systematic Biology*, **66**, 477–498.
- Rabosky, D.L., Santini, F., Eastman, J., Smith, S.A., Sidlauskas, B., Chang, J. & Alfaro, M.E. (2013) Rates of speciation and morphological evolution are correlated across the largest vertebrate radiation. *Nature Communications*, **4**, 1958.
- Rambaut, A., Drummond, A.J., Xie, D., Baele, G. & Suchard, M.A. (2018) Posterior Summarization in Bayesian Phylogenetics Using Tracer 1.7. *Systematic Biology*, **67**, 901–904.
- Rapacciuolo, G. & Blois, J.L. (2019) Understanding ecological change across large spatial, temporal and taxonomic scales: integrating data and methods in light of theory. *Ecography*, **42**, 1247–1266.
- Raup, D.M. & Sepkoski, J.J. (1982) Mass Extinctions in the Fossil Record. *Science*, **215**, 1501–1503.
- Ready, J., Kaschner, K., South, A.B., Eastwood, P.D., Rees, T., Rius, J., Agbayani, E., Kullander, S. & Froese, R. (2010) Predicting the distributions of marine organisms at the global scale. *Ecological Modelling*, **221**, 467–478.
- Ree, R.H. & Smith, S.A. (2008) Maximum likelihood inference of geographic range evolution by dispersal, local

## References

- extinction, and cladogenesis. *Systematic Biology*, **57**, 4–14.
- Renema, W., Bellwood, D.R. & Braga, J.C. (2008) Hopping hotspots: global shifts in marine biodiversity. *Science*, **321**, 654–657.
- Revell, L.J. (2012) Phytools: an R package for phylogenetic comparative biology (and other things). *Methods in Ecology and Evolution*, **3**, 217–223.
- Ricklefs, R.E. (2007) Estimating diversification rates from phylogenetic information. *Trends in Ecology and Evolution*, **22**, 601–610.
- Robertson, D.R., Karg, F., Moura, R.L., Victor, B.C. & Bernardi, G. (2006) Mechanisms of speciation and faunal enrichment in Atlantic parrotfishes. *Molecular Phylogenetics and Evolution*, **40**, 795–807.
- Robertson, R. & Field, A. (2005) M2 baroclinic tides in the Indonesian seas. *Oceanography*, **18**, 62–73.
- Roff, G. & Mumby, P.J. (2012) Global disparity in the resilience of coral reefs. *Trends in Ecology and Evolution*, **27**, 404–413.
- Rogl, F. (1999) Mediterranean and Paratethys. Facts and hypotheses of an Oligocene to Miocene paleogeography (short overview). *Geologica Carpathica*, **50**, 339–349.
- Ronquist, F. (1997) Dispersal-vicariance analysis: a new approach to the quantification of historical biogeography. *Systematic Botany*, **46**, 195–203.
- Russ, G.R. (2003) Grazer biomass correlates more strongly with production than with biomass of algal turfs on a coral reef. *Coral Reefs*, **22**, 63–67.
- Sala, E., Kizilkaya, Z., Yildirim, D. & Ballesteros, E. (2011) Alien marine fishes deplete algal biomass in the Eastern Mediterranean. *PLoS ONE*, **6**, e17356.
- Saladin, B., Leslie, A.B., Wüest, R.O., Litsios, G., Conti, E., Salamin, N. & Zimmermann, N.E. (2017) Fossils matter: improved estimates of divergence times in *Pinus* reveal older diversification. *BMC Evolutionary Biology*, **17**, 95.
- Sanders, H.L. (1968) Marine benthic diversity: a comparative study. *The American Naturalist*, **102**, 243–282.
- Santodomingo, N., Wallace, C.C. & Johnson, K.G. (2015) Fossils reveal a high diversity of the staghorn coral genera *Acropora* and *Isopora* (Scleractinia: Acroporidae) in the Neogene of Indonesia. *Zoological Journal of the Linnean Society*, **175**, 677–763.
- Schultz, O. (2003) The Middle Miocene fish fauna (excl. otolithes) from Mühlbach am Manhartsberg and Grund near Hollabrunn, Lower Austria. *Annalen des Naturhistorischen Museums in Wien A*, **104A**, 185–193.
- Silvestro, D., Tejedor, M.F., Serrano-Serrano, M.L., Loiseau, O., Rossier, V., Rolland, J., Zizka, A., Höhna, S., Antonelli, A. & Salamin, N. (2019) Early arrival and climatically-linked geographic expansion of New World Monkeys from tiny African ancestors. *Systematic Biology*, **68**, 78–92.
- Siqueira, A.C., Bellwood, D.R. & Cowman, P.F. (2019a) Historical biogeography of herbivorous coral reef fishes: The formation of an Atlantic fauna. *Journal of Biogeography*, **46**, 1611–1624.
- Siqueira, A.C., Bellwood, D.R. & Cowman, P.F. (2019b) The evolution of traits and functions in herbivorous coral reef fishes through space and time. *Proceedings of the Royal Society B: Biological Sciences*, **286**, 20182672.
- Siqueira, A.C., Morais, R.A., Bellwood, D.R. & Cowman, P.F. (2020) Trophic innovations fuel reef fish

## References

- diversification. *Nature Communications*, **11**, 2669.
- Siqueira, A.C., Oliveira-santos, L.G.R., Cowman, P.F. & Floeter, S.R. (2016) Evolutionary processes underlying latitudinal differences in reef fish biodiversity. *Global Ecology and Biogeography*, **25**, 1466–1476.
- Smith, S.A. & Brown, J.W. (2018) Constructing a broadly inclusive seed plant phylogeny. *American Journal of Botany*, **105**, 302–314.
- Spalding, M.D., Fox, H.E., Allen, G.R., Davidson, N., Ferdaña, Z.A., Finlayson, M., Halpern, B.S., Jorge, M.A., Lombana, A., Lourie, S.A., Martin, K.D., McManus, E., Molnar, J., Recchia, C.A. & Robertson, J. (2007) Marine ecoregions of the world: a bioregionalization of coastal and shelf areas. *BioScience*, **57**, 573–583.
- Spalding, M.D. & Grenfell, A.M. (1997) New estimates of global and regional coral reef areas. *Coral Reefs*, **16**, 225–230.
- Stadler, T. (2011) Mammalian phylogeny reveals recent diversification rate shifts. *Proceedings of the National Academy of Sciences of the United States of America*, **108**, 6187–6192.
- Steininger, F.F. & Rögl, F. (1979) The paratethys history. A contribution towards the Neogene geodynamics of the alpine orogene. *Annales Geologiques des Pays Helleniques*, **3**, 1153–1165.
- Steneck, R.S. (1983) Escalating herbivory and resulting adaptive trends in calcareous algal crusts. *Paleobiology*, **9**, 44–61.
- Steneck, R.S., Bellwood, D.R. & Hay, M.E. (2017) Herbivory in the marine realm. *Current Biology*, **27**, R484–R489.
- Streelman, J.T., Alfaro, M., Westneat, M.W., Bellwood, D.R. & Karl, S.A. (2002) Evolutionary history of the parrotfishes: biogeography, ecomorphology, and comparative diversity. *Evolution*, **56**, 961–971.
- Tavera, J., Acero, A.P. & Wainwright, P.C. (2018) Multilocus phylogeny, divergence times, and a major role for the benthic-to-pelagic axis in the diversification of grunts (Haemulidae). *Molecular Phylogenetics and Evolution*, **121**, 212–223.
- Title, P.O. & Rabosky, D.L. (2019) Tip rates, phylogenies and diversification: What are we estimating, and how good are the estimates? *Methods in Ecology and Evolution*, **10**, 821–834.
- Tittensor, D.P., Mora, C., Jetz, W., Lotze, H.K., Ricard, D., Berghe, E. Vanden & Worm, B. (2010) Global patterns and predictors of marine biodiversity across taxa. *Nature*, **466**, 1098–1101.
- Tornabene, L., Chen, Y. & Pezold, F. (2013) Gobies are deeply divided: phylogenetic evidence from nuclear DNA (Teleostei: Gobioidae: Gobiidae). *Systematics and Biodiversity*, **11**, 345–361.
- Toussaint, A., Charpin, N., Brosse, S. & Villéger, S. (2016) Global functional diversity of freshwater fish is concentrated in the Neotropics while functional vulnerability is widespread. *Scientific Reports*, **6**, 22125.
- Tripalo, K., Japundžić, S., Sremac, J. & Bošnjak, M. (2016) First record of Acanthuridae (surgeonfish) from the Miocene deposits of the Medvednica Mt. *Geologia Croatica*, **69**, 201–204.
- Tyler, J.C. (2005a) A new genus for the surgeon fish *Acanthurus gaudryi* De Zigno 1887 from the Eocene of Monte Bolca, Italy, a morphologically primitive basal taxon of Acanthuridae (Acanthuroidea, Perciformes). *Studi e Ricerche sui Giacimenti Terziari di Bolca, Museo Civico di Storia Naturale di Verona*, **11**, 149–163.
- Tyler, J.C. (2005b) Redescription and basal phylogenetic placement of the acanthurid surgeon fish

## References

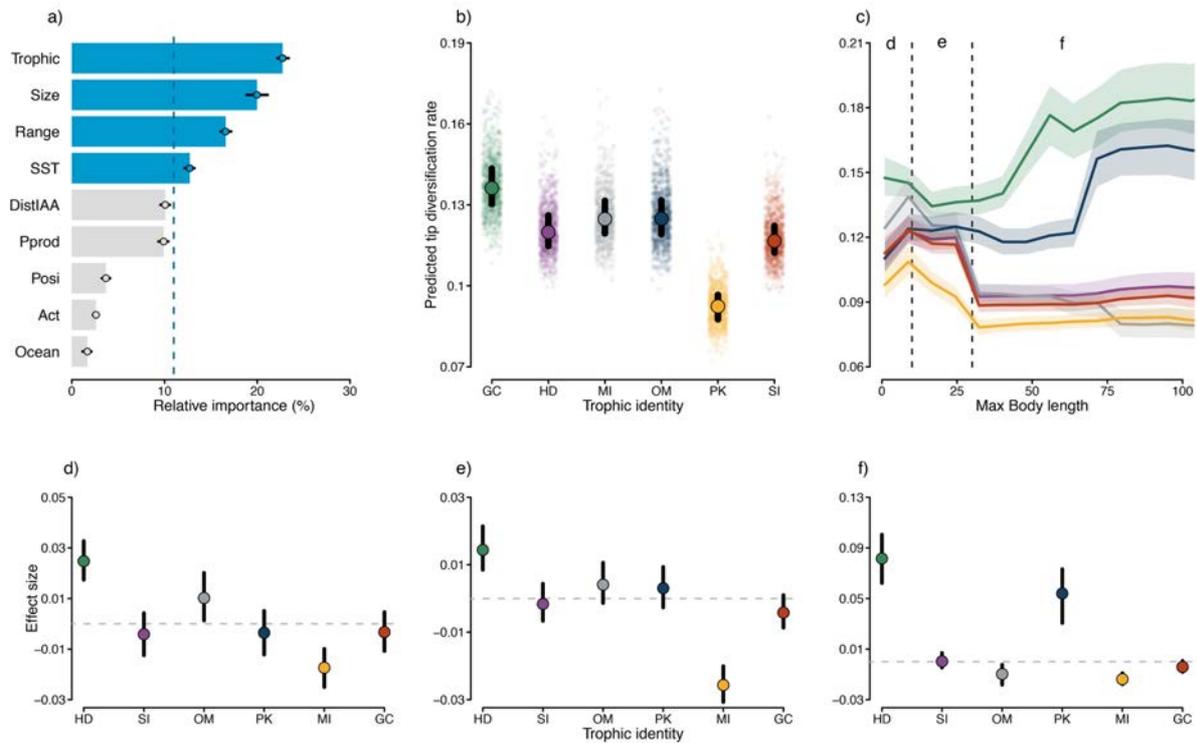
- Gazolaichthys vestenanovae from the Eocene of Monte Bolca, Italy (Perciformes; Acanthuroidea). *Studi e Ricerche sui Giacimenti Terziari di Bolca, Museo Civico di Storia Naturale di Verona*, **11**, 97–117.
- Tyler, J.C. (1997) The Miocene fish *Marosichthys*, a putative Tetraodontiform, actually a perciform surgeon fish (Acanthuridae) related to the recent *Naso*. *Beaufortia*, **47**, 1–10.
- Tyler, J.C. & Bannikov, A.F. (2000) A new species of the surgeon fish genus *Tauchthys* from the Eocene of Monte Bolca, Italy (Perciformes, Acanthuridae). *Bollettino del Museo Civico di Storia Naturale di Verona*, **24**, 29–36.
- Tyler, J.C. & Bannikov, A.F. (1997) Relationships of the fossil and recent genera of rabbitfishes (Acanthuroidei: Siganidae). *Smithsonian Contributions to Paleobiology*, **84**, 1–35.
- Tyler, J.C. & Micklich, N.R. (2011) A new genus and species of surgeon fish (Perciformes, Acanthuridae) from the Oligocene of Kanton Glarus, Switzerland. *Swiss Journal of Palaeontology*, **130**, 203–216.
- Tyler, J.C. & Sorbini, L. (1998) On the relationships of *Eonaso*, an Antillean fossil surgeon fish (Acanthuridae). *Studi e Ricerche sui Giacimenti Terziari di Bolca, Museo Civico di Storia Naturale di Verona*, **7**, 35–42.
- Upham, N.S., Esselstyn, J.A. & Jetz, W. (2019) Inferring the mammal tree: Species-level sets of phylogenies for questions in ecology, evolution, and conservation. *PLoS Biology*, **17**, e3000494.
- Varga, T., Krizsán, K., Földi, C., Dima, B., Sánchez-García, M., Sánchez-Ramírez, S., Szöllősi, G.J., Szarkándi, J.G., Papp, V., Albert, L., Andreopoulos, W., Angelini, C., Antonín, V., Barry, K.W., Bougher, N.L., Buchanan, P., Buyck, B., Bense, V., Catcheside, P., Chovatia, M., Cooper, J., Dämon, W., Desjardin, D., Finy, P., Geml, J., Haridas, S., Hughes, K., Justo, A., Karasiński, D., Kautmanova, I., Kiss, B., Kocsubé, S., Kotiranta, H., LaButti, K.M., Lechner, B.E., Liimatainen, K., Lipzen, A., Lukács, Z., Mihaltcheva, S., Morgado, L.N., Niskanen, T., Noordeloos, M.E., Ohm, R.A., Ortiz-Santana, B., Ovrebo, C., Rácz, N., Riley, R., Savchenko, A., Shiryayev, A., Soop, K., Spirin, V., Szébenyi, C., Tomšovský, M., Tulloss, R.E., Uehling, J., Grigoriev, I. V., Vágvölgyi, C., Papp, T., Martin, F.M., Miettinen, O., Hibbett, D.S. & Nagy, L.G. (2019) Megaphylogeny resolves global patterns of mushroom evolution. *Nature Ecology and Evolution*, **3**, 668–678.
- Venditti, C., Meade, A. & Pagel, M. (2011) Multiple routes to mammalian diversity. *Nature*, **479**, 393–396.
- Vermeij, G.J. (1977) The Mesozoic marine revolution: evidence from snails, predators and grazers. *Paleobiology*, **3**, 245–258.
- Villéger, S., Novack-Gottshall, P.M. & Mouillot, D. (2011) The multidimensionality of the niche reveals functional diversity changes in benthic marine biotas across geological time. *Ecology Letters*, **14**, 561–568.
- Violle, C., Navas, M.L., Vile, D., Kazakou, E., Fortunel, C., Hummel, I. & Garnier, E. (2007) Let the concept of trait be functional! *Oikos*, **116**, 882–892.
- Violle, C., Reich, P.B., Pacala, S.W., Enquist, B.J. & Kattge, J. (2014) The emergence and promise of functional biogeography. *Proceedings of the National Academy of Sciences of the United States of America*, **111**, 13690–13696.
- Wainwright, P.C. (2007) Functional Versus Morphological Diversity in Macroevolution. *Annual Review of Ecology, Evolution, and Systematics*, **38**, 381–401.
- Wainwright, P.C., Santini, F., Bellwood, D.R., Robertson, D.R., Rocha, L.A. & Alfaro, M.E. (2018) Phylogenetics

## References

- and geography of speciation in New World Halichoeres wrasses. *Molecular Phylogenetics and Evolution*, **121**, 35–45.
- Wallace, A.R. (1878) *Tropical nature and other essays*, Macmillan, London, UK, London.
- Wallace, C.C. & Rosen, B.R. (2006) Diverse staghorn corals (Acropora) in high-latitude Eocene assemblages: Implications for the evolution of modern diversity patterns of reef corals. *Proceedings of the Royal Society B: Biological Sciences*, **273**, 975–982.
- Webb, T.J. (2012) Marine and terrestrial ecology: unifying concepts, revealing differences. *Trends in Ecology and Evolution*, **27**, 535–541.
- Westneat, M.W. & Alfaro, M.E. (2005) Phylogenetic relationships and evolutionary history of the reef fish family Labridae. *Molecular phylogenetics and evolution*, **36**, 370–90.
- Whittaker, R.J., Rigal, F., Borges, P.A. V., Cardoso, P., Terzopoulou, S., Casanoves, F., Pla, L., Guilhaumon, F., Ladle, R.J. & Triantis, K.A. (2014) Functional biogeography of oceanic islands and the scaling of functional diversity in the Azores. *Proceedings of the National Academy of Sciences of the United States of America*, **111**, 13709–13714.
- Wiens, J.J., Lapoint, R.T. & Whiteman, N.K. (2015) Herbivory increases diversification across insect clades. *Nature Communications*, **6**, 8370.
- Williams, S.T. & Duda, T.F. (2008) Did tectonic activity stimulate Oligo-Miocene speciation in the Indo-West Pacific? *Evolution*, **62**, 1618–1634.
- Willig, M.R., Kaufman, D.M. & Stevens, R.D. (2003) Latitudinal Gradients Of Biodiversity: Pattern, Process, Scale, and Synthesis. *Annual Review of Ecology, Evolution, and Systematics*, **34**, 273–309.
- Wilson, S.K., Graham, N.A.J., Pratchett, M.S., Jones, G.P. & Polunin, N.V.C. (2006) Multiple disturbances and the global degradation of coral reefs: are reef fishes at risk or resilient? *Global Change Biology*, **12**, 2220–2234.
- Winterbottom, R. (1993) Myological evidence for the Phylogeny of recent genera of surgeonfishes (Percomorpha, Acanthuridae), with comments on the Acanthuroidei. *Copeia*, **1993**, 21–39.
- Winterbottom, R. & McLennan, D.A. (1993) Cladogram versatility: evolution and biogeography of Acanthuroid fishes. *Evolution*, **47**, 1557–1571.
- Womack, M.C. & Bell, R.C. (2020) Two-hundred million years of anuran body-size evolution in relation to geography, ecology and life history. *Journal of Evolutionary Biology*, 1–16.
- Wood, R. (1999) *Reef evolution*, Oxford University Press, Oxford.
- Wood, S.N. (2017) *Generalized Additive Models: an introduction with R*, Chapman and Hall/CRC.
- Woodland, D.J. (1990) *Revision of the fish family Siganidae with description of two new species and comments on distribution and biology*, Bernice Pauahi Bishop Museum, Honolulu, Hawaii.
- Xie, W., Lewis, P.O., Fan, Y., Kuo, L. & Chen, M.-H. (2011) Improving marginal likelihood estimation for Bayesian phylogenetic model selection. *Systematic Biology*, **60**, 150–160.
- Zamagni, J., Mutti, M. & Košir, A. (2012) The evolution of mid Paleocene-early Eocene coral communities: how to survive during rapid global warming. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **317–318**, 48–65.

Appendix A.  
Supplementary Material to Chapter 2

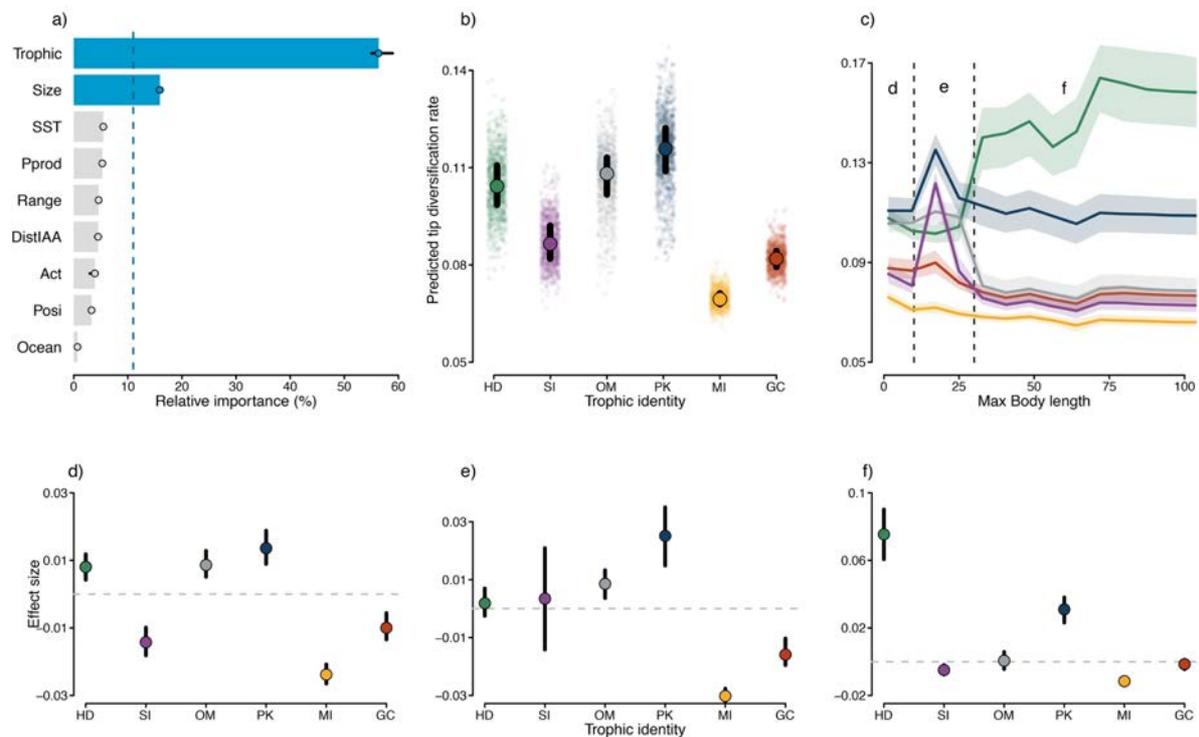
Supplementary Figures



**Supplementary Figure 2.1 |** Relative importance of ecological and geographical factors in driving reef fish tip diversification rate patterns, based on an extreme gradient boosting model using the ‘DR statistic’ estimates as response variable. (a) Mean relative importance (%) of explanatory variables. Blue bars show variables above chance expectation (dashed line). Black lines represent importance quantiles (25% and 75%) derived from 1000 model bootstraps. Trophic: trophic identity; Size: maximum body length; SST: sea surface temperature; DistIAA: distance to the Indo-Australian-Archipelago; Ocean: oceanic basin; Range: geographic range; PProd: primary productivity; Posi: position in the water column; Act: circadian activity period (see Methods section 2.3.3). (b) Predicted tip diversification rates per trophic group. In this analysis, all other continuous variables are kept at their mean values and categorical variables in the most common category. Semi-transparent dots are bootstrapped predictions ( $n = 1000$ ), with larger points representing median values with respective 25% and 75% prediction quantiles (black lines). (c) Predicted tip diversification rates for species of various maximum body lengths in different trophic groups, based on an extreme gradient boosting model ( $n = 1000$  model bootstraps). All other variables are kept at their mean values and categorical variables (except trophic identity) in the most common category. Solid lines show median predictions per trophic group with respective prediction quantile intervals (25% and 75%). Dashed line separates size classes for which we show effect sizes per trophic group: (d)

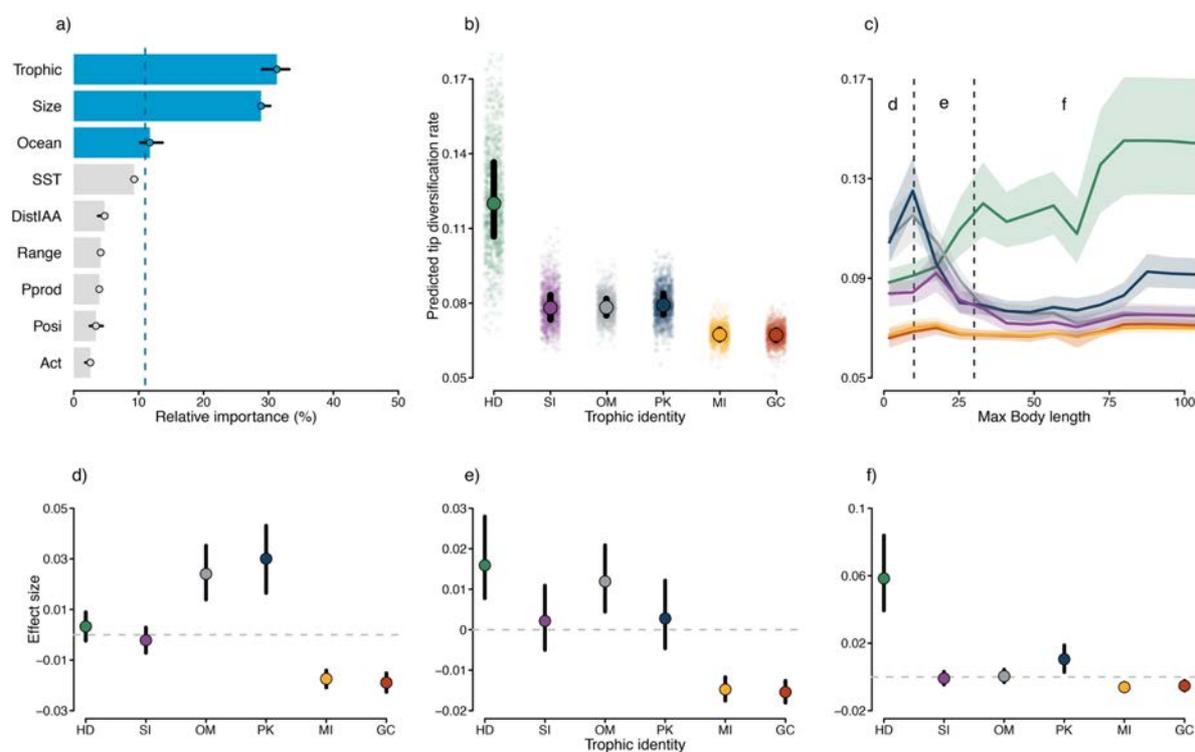
## Appendix A - Supplementary Material to Chapter 2

below 10 cm; (e) between 10 and 30 cm; (f) above 30 cm. In d-f, circles show the median effects (trophic group median minus global median in each size class) and black lines show 25% and 75% effect quantiles. HD: herbivores/detritivores (green); SI: sessile invertivores (purple); OM: omnivores (grey); PK: planktivores (blue); MI: mobile invertivores (yellow); GC: generalized carnivores (red). Source data are provided as a Source Data file.



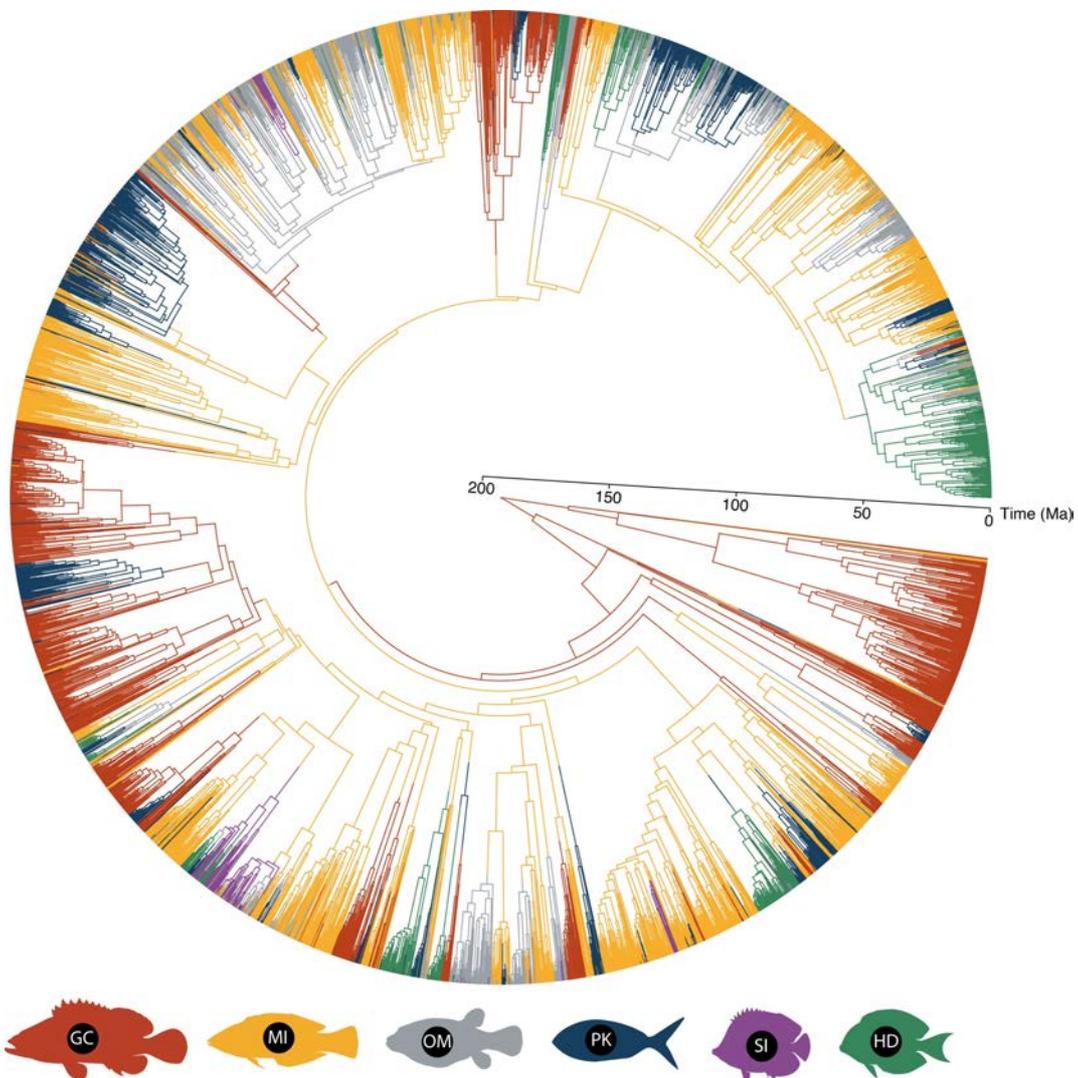
**Supplementary Figure 2.2 |** Relative importance of ecological and geographical factors in driving reef fish tip diversification rate patterns, based on an extreme gradient boosting model using only the ‘consensus’ reef fish families. (a) Mean relative importance (%) of explanatory variables. Blue bars show variables above chance expectation (dashed line). Black lines represent importance quantiles (25% and 75%) derived from 1000 model bootstraps. Trophic: trophic identity; Size: maximum body length; SST: sea surface temperature; DistIAA: distance to the Indo-Australian-Archipelago; Ocean: oceanic basin; Range: geographic range; PProd: primary productivity; Posi: position in the water column; Act: circadian activity period (see Methods section 2.3.3). (b) Predicted tip diversification rates per trophic group. In this analysis, all other continuous variables are kept at their mean values and categorical variables in the most common category. Semi-transparent dots are bootstrapped predictions ( $n = 1000$ ), with larger points representing median values with respective 25% and 75% prediction quantiles (black lines). (c) Predicted tip diversification rates for species of various maximum body lengths in different trophic groups, based on an extreme gradient boosting model ( $n = 1000$  model bootstraps). All other variables are kept at their mean values and categorical variables (except trophic identity) in the most common category. Solid lines show median predictions per trophic group with respective prediction quantile intervals (25% and 75%). Dashed line separates size classes for which we show effect sizes per trophic group: (d) below 10 cm; (e) between 10 and 30 cm; (f) above 30 cm. In d-f, circles show the median effects (trophic group median minus

global median in each size class) and black lines show 25% and 75% effect quantiles. HD: herbivores/detritivores (green); SI: sessile invertivores (purple); OM: omnivores (grey); PK: planktivores (blue); MI: mobile invertivores (yellow); GC: generalized carnivores (red). Source data are provided as a Source Data file.

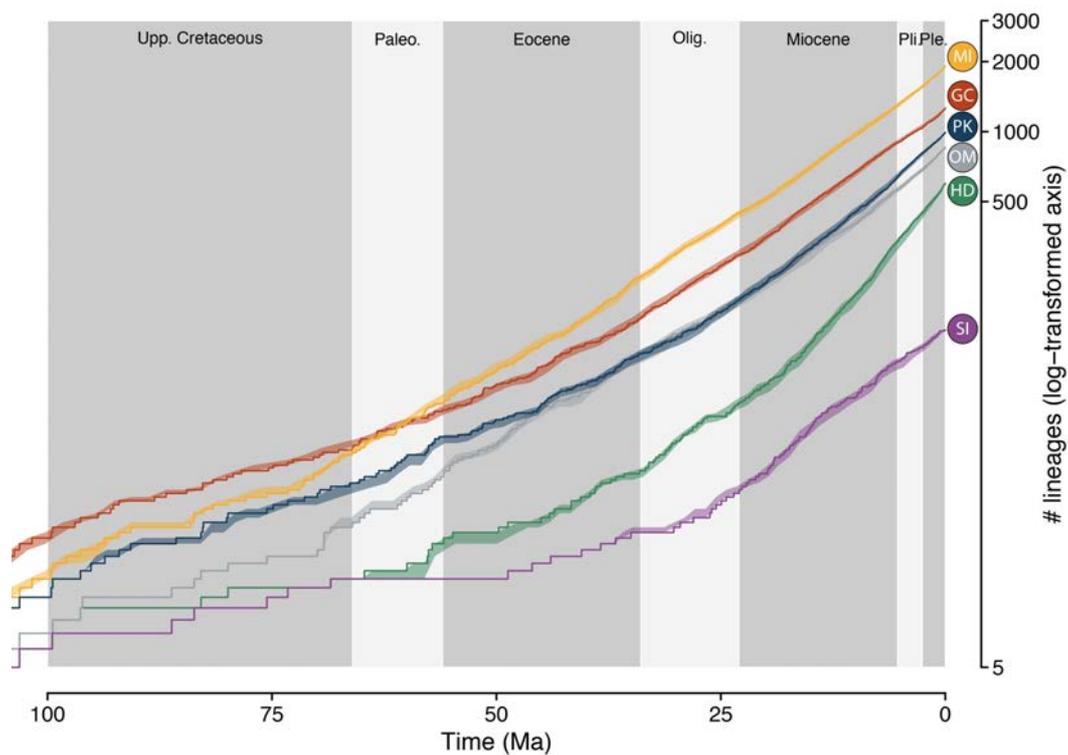


**Supplementary Figure 2.3 |** Relative importance of ecological and geographical factors in driving reef fish tip diversification rate patterns, based on an extreme gradient boosting model excluding cryptobenthic families. (a) Mean relative importance (%) of explanatory variables. Blue bars show variables above chance expectation (dashed line). Black lines represent importance quantiles (25% and 75%) derived from 1000 model bootstraps. Trophic: trophic identity; Size: maximum body length; SST: sea surface temperature; DistIAA: distance to the Indo-Australian-Archipelago; Ocean: oceanic basin; Range: geographic range; PProd: primary productivity; Posi: position in the water column; Act: circadian activity period (see Methods section 2.3.3). (b) Predicted tip diversification rates per trophic group. In this analysis, all other continuous variables are kept at their mean values and categorical variables in the most common category. Semi-transparent dots are bootstrapped predictions ( $n = 1000$ ), with larger points representing median values with respective 25% and 75% prediction quantiles (black lines). (c) Predicted tip diversification rates for species of various maximum body lengths in different trophic groups, based on an extreme gradient boosting model ( $n = 1000$  model bootstraps). All other variables are kept at their mean values and categorical variables (except trophic identity) in the most common category. Solid lines show median predictions per trophic group with respective prediction quantile intervals (25% and 75%). Dashed line separates size classes for which we show effect sizes per trophic group: (d) below 10 cm; (e) between 10 and 30 cm; (f) above 30 cm. In d-f, circles show the median effects (trophic group median minus global median in each size class) and black lines show 25% and 75% effect quantiles. HD: herbivores/detritivores

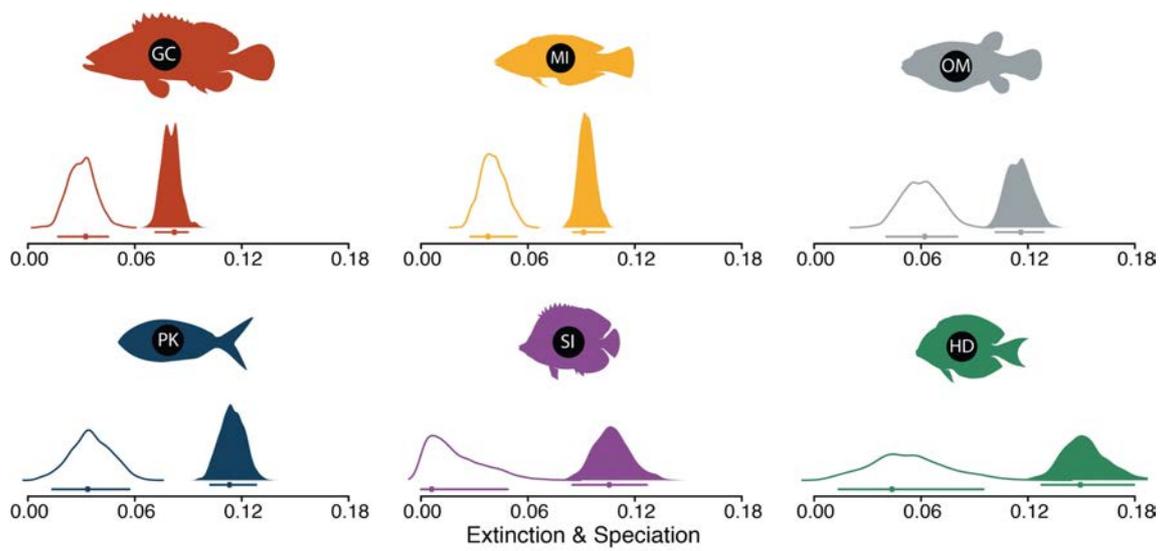
(green); SI: sessile invertivores (purple); OM: omnivores (grey); PK: planktivores (blue); MI: mobile invertivores (yellow); GC: generalized carnivores (red). Source data are provided as a Source Data file.



**Supplementary Figure 2.4** | Ancestral state reconstruction of reef fish trophic guilds. This reconstruction represents one of the stochastic character mappings. HD: herbivores/detritivores (green); SI: sessile invertivores (purple); OM: omnivores (grey); PK: planktivores (blue); MI: mobile invertivores (yellow); GC: generalized carnivores (red).



**Supplementary Figure 2.5 |** Lineage through time plot for each reef fish trophic guild. Solid line derives from a randomly selected phylogenetic tree, while semi-transparent polygons represent the 95% confidence intervals derived from 100 trees. HD: herbivores/detritivores (green); SI: sessile invertivores (purple); OM: omnivores (grey); PK: planktivores (blue); MI: mobile invertivores (yellow); GC: generalized carnivores (red).



**Supplementary Figure 2.6** | Historical extinction and speciation rate estimates for each reef fish trophic group, derived from MuSSE. Lines below the distributions show mode values (solid circles) with respective 95% credibility intervals. HD: herbivores/detritivores (green); SI: sessile invertivores (purple); OM: omnivores (grey); PK: planktivores (blue); MI: mobile invertivores (yellow); GC: generalized carnivores (red).

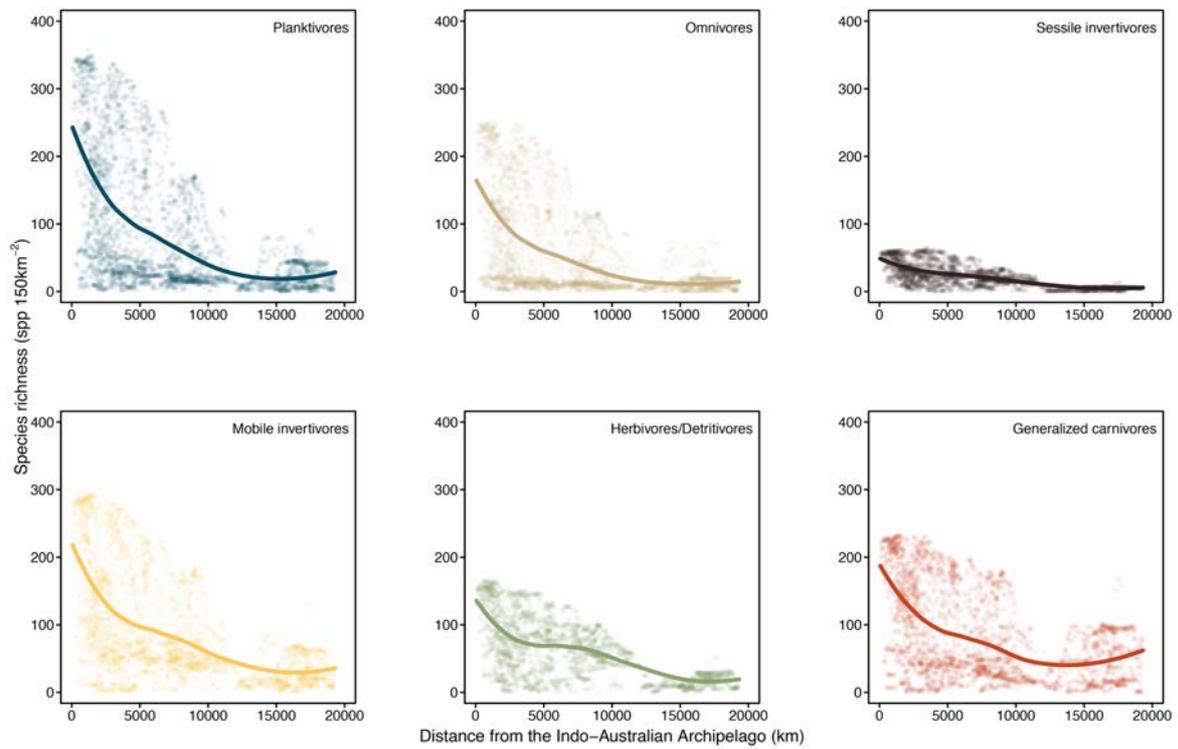
## Supplementary Tables

**Supplementary Table 2.1** | Model results derived from the HiSSE analysis ranked according to the lowest Akaike Information Criterion (AIC). The unconstrained model considered rates to be different between analysed character states (trophic group) with one hidden diversification regime per state, whereas the constrained model considers rates to be equal between analysed states but different from the hidden diversification regime. logLik: model likelihood; Net div ancs: net diversification for the ancestral trophic groups (generalized carnivores, mobile invertivores and omnivores); Net div rec: net diversification for the recently derived trophic groups (herbivores/detritivores, sessile invertivores and planktivores); Net div ancs H: hidden net diversification regime for the ancestral trophic groups; Net div rec H: hidden net diversification regime for the recently derived trophic groups.

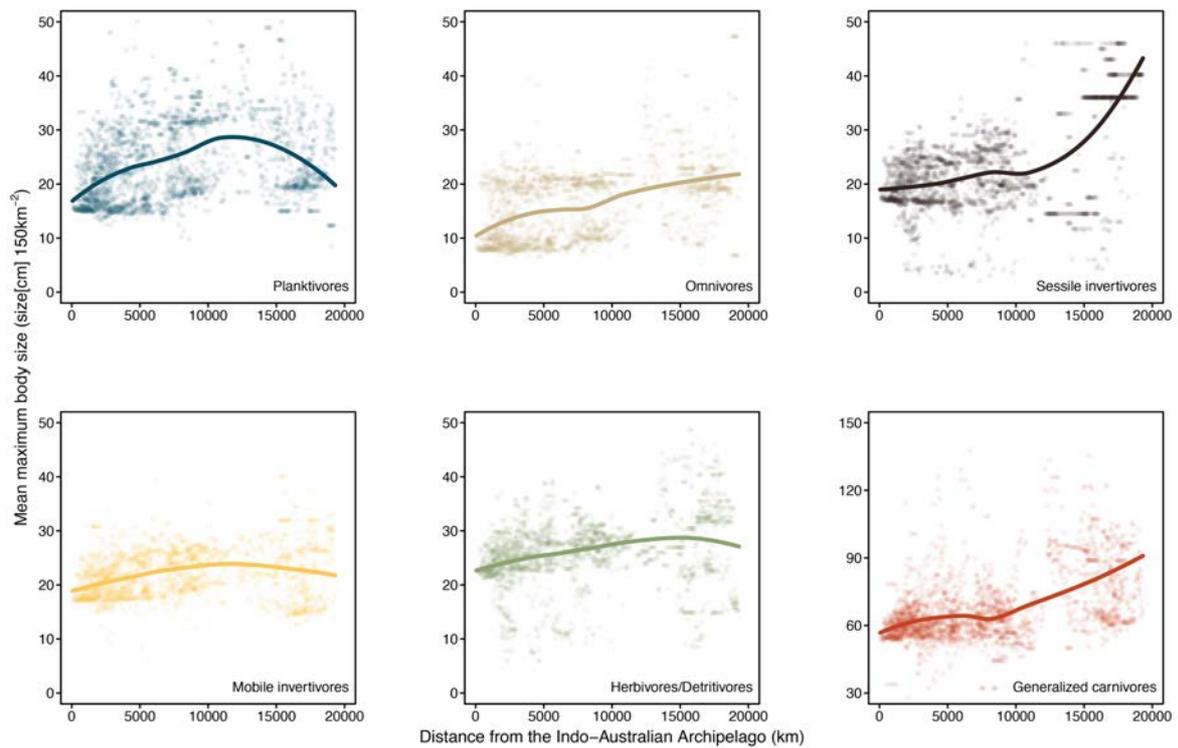
	logLik	AIC	Net div ancs	Net div rec	Net div ancs H	Net div rec H
<b>Unconstrained</b>	-21736.8	43501.7	0.0253496	0.0945460	0.1764130	0.2314718
<b>Constrained</b>	-21859.7	43739.4	0.0426502	0.0426502	0.1432365	0.1432365

Appendix B.  
Supplementary Material to Chapter 3

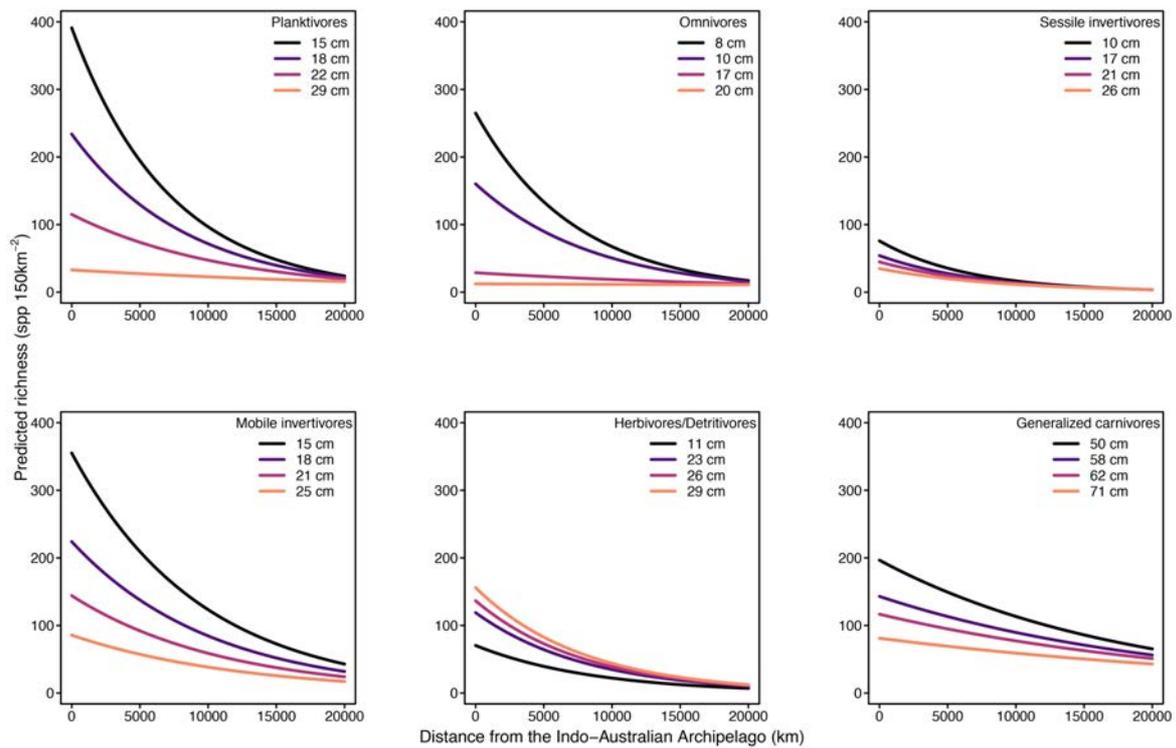
Supplementary Figures



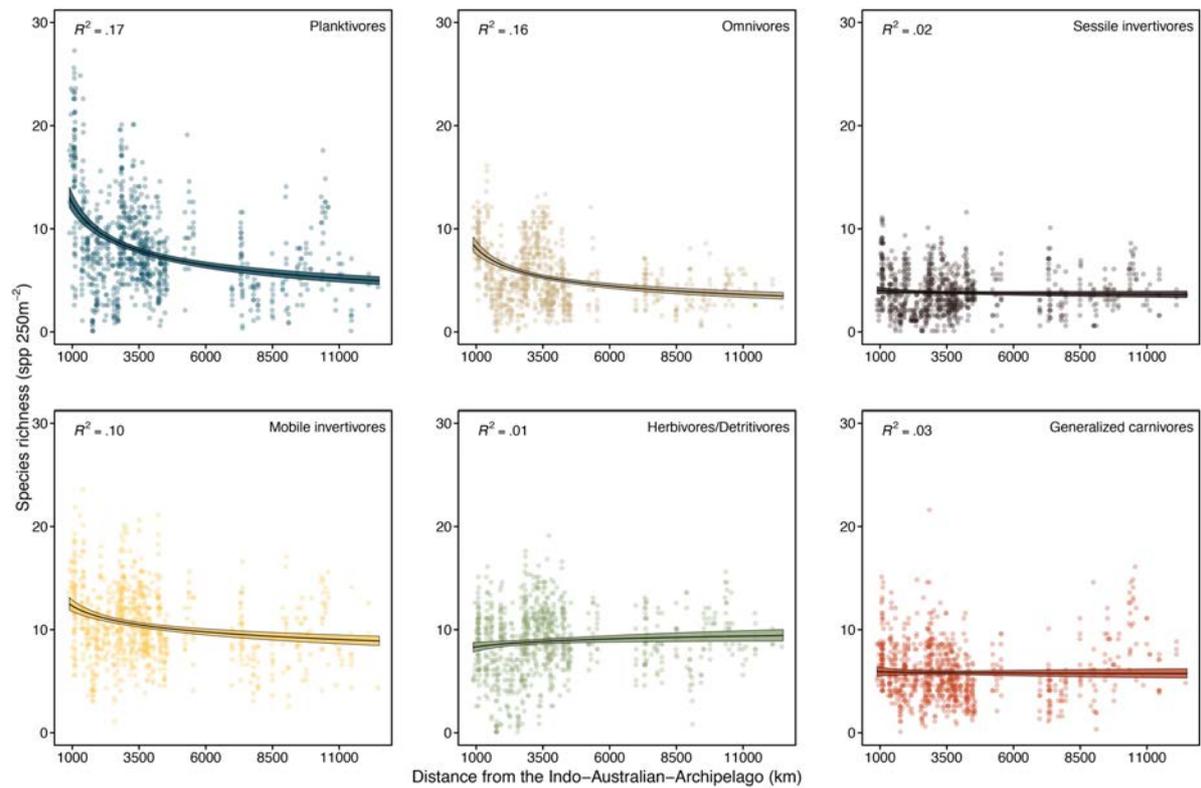
**Supplementary Figure 3.1 | Coral reef fish species richness per geographic cell with distance from the Indo-Australian Archipelago.** Semi-transparent points show the number of species per geographic cell in each trophic group. Lines represent the fitted values from a LOESS polynomial regression (with  $\alpha = 0.7$ ).



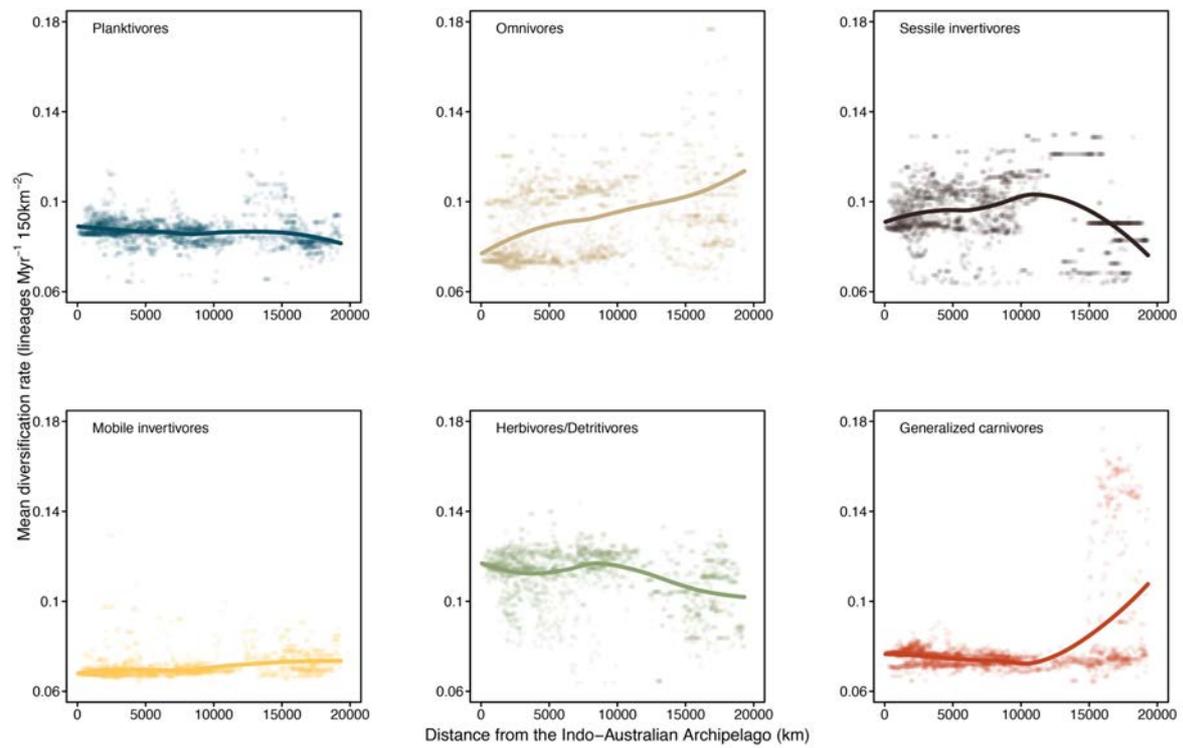
**Supplementary Figure 3.2 | Mean coral reef fish species body size per geographic cell with distance from the Indo-Australian Archipelago.** Semi-transparent points show the mean maximum body size of all species occurring in each geographic cell per trophic group. Lines represent the fitted values from a LOESS polynomial regression (with  $\alpha = 0.7$ ). Note that the y-axis scale for Generalized carnivores is different from the other trophic groups, given their generally larger body sizes.



**Supplementary Figure 3.3 | Predicted coral reef fish species richness with distance from the Indo-Australian Archipelago by varying body size.** Mean number of species per grid cell (lines) predicted from a negative binomial model per trophic group. Model predictions were made using the 2.5% quantile (black), 25% quantile (purple), median (pink), and 75% quantile (orange) of the distribution of mean species body size per geographic cell in each trophic group.



**Supplementary Figure 3.4 | Coral reef fish species richness at the site scale per trophic group.** Mean site species richness per trophic group from visual surveys (points) with distance from the Indo-Australian Archipelago (IAA). Curves show predictions from a generalized linear model (mean [black line]  $\pm$  95% confidence interval [polygons]) with respective  $R^2$  values (top-left corner). Model predictions were performed with body size fixed in the estimated value for the regions closer to the centre of the IAA.



**Supplementary Figure 3.5 | Net diversification rate per geographic cell with distance from the Indo-Australian Archipelago.** Semi-transparent points show the mean tip diversification rates of all species occurring in each geographic cell per trophic group. Lines represent the fitted values from a LOESS polynomial regression (with  $\alpha = 0.7$ ).

## Appendix C. Supplementary Material to Chapter 4

### Supplementary Methods

#### Phylogenetic inferences

For each of our herbivorous fish groups, we downloaded overlapping gene sequences for all available species from Genbank using Geneious Pro version 11.1 (Kearse *et al.*, 2012). For the Acanthuridae, we downloaded two mitochondrial (*Cox1* and *Cytb*) and seven nuclear genes (*ENC1*, *myh6*, *plagl2*, *Rag1*, *Rh*, *zic1* and *ETS2*) belonging to 72 species (~90% total diversity) from all extant genera. Two species were used as outgroups, one from the family Zanclidae (*Zanclus cornutus*) and one from the family Luvaridae (*Luvarus imperialis*). The Siganidae phylogeny was based on two mitochondrial markers (*Cytb* and *16s*) and the nuclear rRNA internal transcribed spacer 1 (*ITS1*) region. It contained 24 species (~80% total diversity) in its single genus *Siganus*, and included the species *Zanclus cornutus* (Zanclidae) and *Prionurus scalprum* (Acanthuridae) as outgroups. Finally, the Scarini phylogeny was based on five mitochondrial (*Cox1*, *Cytb*, *12s*, *16s* and control region) and six nuclear markers (*Bmp4*, *Dlx2*, *Otx1*, *Rag2*, *S7I1* and *Tmo-4C4*), for 87 species (~87% total diversity) belonging to all extant genera. Nine species from the family Labridae (two Hypsigenyines and seven Cheilines) were included to act as outgroups for the parrotfishes. The Majority of these genetic sequences have been deposited by previous studies performed with the focal taxa (Klanten *et al.*, 2004; Kuriwa *et al.*, 2007; Cowman & Bellwood, 2011; Choat *et al.*, 2012; Sorenson *et al.*, 2013). Species and accession numbers are given in Supplementary Tables 4.4-4.6. Gene datasets were aligned using the Muscle algorithm (Edgar, 2004) in Geneious Pro version 11.1 (Kearse *et al.*, 2012) and checked by eye. The resulting concatenated alignments consisted of 7,229 (Acanthuridae), 3,001 (Siganidae) and 6,296 (Scarini) base pairs. Model testing was performed using PartitionFinder2 (Lanfear *et al.*, 2016) and indicated the best gene partitioning scheme for each taxon (Supplementary Tables 4.7-4.9).

For the phylogenetic inferences, we benefited from the CIPRES Science Gateway (Miller *et al.*, 2010) computing environment. Firstly, we ran maximum likelihood (ML) analyses using the GTR + G model and 1000 bootstraps per taxa in RAxML (Stamatakis, 2014). The resulting ML tree for each taxon was converted to an ultrametric tree using the penalized likelihood method in 'ape' R package (function '*chronos*'; Paradis *et al.*, 2004), and subsequently used as a starting tree in BEAST2 (Bouckaert *et al.*, 2014) for Bayesian estimation of topology, branch lengths and node ages. For this analysis, we set gene partitions according to the results from PartitionFinder (Supplementary Tables 4.7-4.9) and we used relaxed lognormal clock priors. We also set birth-death models with node

calibration points according to fossil information for each of our groups, all of which had lognormally distributed priors with soft upper bounds. For the surgeonfishes, we placed a calibration point in the crown Acanthuridae lineage to represent the acanthurid fossils from Monte Bolca at 50 Myr (Bellwood, 1996). We also placed a prior on the stem lineage of Acanthuridae to represent the fossil *Kushlukia permira*, described as a stem Luvaridae from 55.8 Ma (Bannikov & Tyler, 1995). For the rabbitfishes, the root node representing the stem Siganidae lineage was calibrated at 55 Myr representing *Siganopygeus rarus*, the oldest fossil described for the family (Bannikov *et al.*, 2010). We also calibrated the outgroup stem Zanclidae lineage to represent the only fossil described for that family from Monte Bolca at 50 Myr (Bellwood, 1996). Finally, for the Parrotfishes, we used the only two fossils described for the group as internal calibration points: *Calotomus preisli* as a stem *Calotomus* at 14 Myr and the stem *Bolbometopon* fossil at 5 Myr (Bellwood & Schultz, 1991). We also calibrated the crown Hypsigenyines outgroup at 50 Myr to represent the labrid *Phyllopharyngodon longipinnis* from Monte Bolca (Bellwood, 1990). Five independent MCMC runs were conducted per group for 100 million generations each, storing trees every 10,000 generations (10,000 trees per run). All runs were assessed for convergence and stationarity in Tracer v1.7.0 (Rambaut *et al.*, 2018) using effective sample size (ESS) scores. After removing 20% burn in from each run, all trees were combined in LogCombiner v2.5.0 (Bouckaert *et al.*, 2014) and compiled into a maximum clade credibility (MCC) tree in TreeAnnotator v2.5.0 (Bouckaert *et al.*, 2014).

#### Herbivorous reef fish data

We categorized all parrotfish, surgeonfish and rabbitfish species present in the phylogenetic trees based on seven traits related to feeding. Firstly, we collected data on the maximum size recorded for each species in FishBase (Froese & Pauly, 2019). This was the only continuously variable trait included in our analysis. We then assigned species according to categories of tooth morphology and alimentary tract, traits that are related to food processing. To classify the types of alimentary tract found in herbivorous coral reef fishes, we considered the most important feature of internal food processing in each of our groups. In parrotfishes, the Pharyngeal Jaw Apparatus is modified in a structure specialized for grinding food, known as Pharyngeal Mill. Considering that this is a synapomorphy for the tribe Scarini (formerly the Scaridae) (Bellwood, 1994), we classified the alimentary tract of all our parrotfish species as having a Pharyngeal Mill. All parrotfishes lack a stomach and must thus rely on the pharyngeal jaws for triturating food particles. By contrast, surgeonfishes and rabbitfishes mainly rely on stomach features for initial food processing (Horn, 1989). Therefore, we divided the stomachs of these groups into two categories as being either thin-walled or gizzard-like (following Horn, 1989; Choat, 1991). Thin-walled stomachs are associated with acid lysis of food

items and is found in rabbitfishes and some surgeonfishes (*Naso* and some *Acanthurus*), while gizzard-like stomachs rely on thick, muscly walls to triturate food and is found in other surgeonfishes (*Ctenochaetus* and some *Acanthurus*). For the tooth morphology categorization, we used the most prominent feature of the tooth structure of each herbivorous group. All rabbitfishes have a bicuspid tooth, while in the surgeonfishes, tooth can be conical (*Naso*), multi-denticulate (*Acanthurus*) or brush-like (*Ctenochaetus*) (following Woodland, 1990; Purcell & Bellwood, 1993; Randall, 2001). For the parrotfishes, the tooth morphology was based on the intensity of fusion of the dental plates and on the dental margin pattern, which resulted in the following categories: not-fused (*Calotomus*, *Cryptotomus* and *Nicholsina*), weakly-fused (most *Sparisoma*), fused-crenelated (*Cetoscarus*, *Bolbometopon*, *Chlorurus* and some *Sparisoma*), and fused-even (*Hipposcarus* and *Scarus*) (following Bellwood & Choat, 1990; Bellwood, 1994).

We also classified four important behavioural traits related to food acquisition that included feeding mode, diet, feeding habitat and schooling behaviour. Each of these traits are related to different components of the feeding behaviour of herbivorous fishes on coral reefs. For instance, we classified each of our species according to the mode how they acquire food, which included browsing (i.e. species that browse on macroalgae larger than 10 mm), scraping (i.e. parrotfishes that remove the epilithic algal matrix), excavating (i.e. parrotfishes that excavate the surface of the reef matrix), planktivory (i.e. surgeonfishes that feed in the water column), cropping (i.e. surgeonfishes and rabbitfishes that crop short turf algae from the benthos), sucking (i.e. surgeonfishes that use suction to feed on particulate material on the benthos), and brushing (i.e. surgeonfishes that brush particulate material from the benthos) (classification modified after Bellwood, 1985, 1994; Bellwood & Choat, 1990; Woodland, 1990; Clements, 1991; Green & Bellwood, 2009; Tebbett *et al.*, 2017). We also classified the typical diet of each species. The food items classified included: cyanobacteria, coral, detritus, epilithic algal matrix (EAM), macroalgae, seagrass, sponges, turf-algae and zooplankton (data drawn from Hiatt & Strasburg, 1960; Randall, 1967; Choat, 1969; Bellwood, 1985, 1994; Woodland, 1990; Bellwood & Choat, 1990; Clements, 1991; Clements & Choat, 1995; Dunlap & Pawlik, 1996; Choat *et al.*, 2002, 2004; Ferreira & Gonçalves, 2006; Green & Bellwood, 2009; Francini-Filho *et al.*, 2010; Hoey *et al.*, 2013; Plass-Johnson *et al.*, 2013; Adam *et al.*, 2015; Tebbett *et al.*, 2017; Clements *et al.*, 2017). For feeding habitat, we considered the location where each species predominantly feed on, which included: concealed, open or sandy parts of the reef, the water column and off-reef species (classification modified after Bellwood, 1985; Bellwood & Choat, 1990; Brandl & Bellwood, 2013, 2014, 2016; Fox & Bellwood, 2013; Brandl *et al.*, 2014; Adam *et al.*, 2015). Finally, our classification of schooling behaviour included species that feed solitarily, in pairs or in schools (Bellwood, 1985; Woodland, 1990; Randall, 2001).

The combination of traits that we classified provides an indication of each species' ecological role in terms of ecosystem processes. However, different trait combinations can result in similar ecosystem functions (i.e. the movement or storage of energy through trophic or bioconstructional pathways; Bonaldo *et al.*, 2014) performed by herbivorous fish species on coral reefs. Therefore, we also categorized each species according to their role in ecosystem processes and consequently to the cycling of matter and nutrients in reef systems. Our ecosystem function categories included: Turf-algae removal, Macroalgae removal, Sediment removal (which includes the rework and transport of sediment particles), Zooplanktivory, Crevice cleaning (i.e. species that are capable of feeding in concealed parts of the reef), Bioerosion, Coralivory and Spongivory. We assigned these ecosystem functions to all species in our database considering that each species could potentially perform more than one function simultaneously.

#### Ancestral range estimation

To assess the ancestral ranges in each of our herbivorous fish groups, we built biogeographical models using the 'BioGeoBEARS' R package (Matzke, 2013). This package allows the comparison of candidate models for ancestral range estimation built in a maximum likelihood framework. We used this framework to build models according to the notation of the three most widely recognized models in historical biogeography: DEC (Ree & Smith, 2008); DIVA (Ronquist, 1997); and BayAREA (Landis *et al.*, 2013). We built combinations of models including the founder-speciation event parameter  $j$  (Matzke, 2014), which considers the inheritance of a new area by a daughter lineage while the sister-splitting lineage inherits the original ancestral range. All our models were built considering time-constraints from the past geological history of marine environments that are well known to influence coral reef fish biogeography (Cowman & Bellwood, 2013a,b). We constrained the root nodes of each group to be present in the ancestral (now extinct) Tethys sea, reflecting the presence of fossil species in a region that used to connect major ocean basins (Atlantic and Indo-Pacific) in the geological past (Renema *et al.*, 2008). Since our chronogram for the Siganidae shows that the extant species in the family are a product of a recent radiation (~ 25 Ma; Supplementary Figure 4.1), we included their outgroup to allow the stem node to be sampled in the biogeographical analysis. From 65 Ma to 12 Ma, we allowed the dispersal of lineages between the adjacent EA and WI regions, connected via Tethys seaway. From 12 Ma onwards, we excluded the Tethys sea from the analysis and we set very low dispersal multiplier values between EA and WI to reflect the final closure of the Tethys seaway (Steininger & Rögl, 1979), but allowing the possibility of dispersal around the South African coast (Bowen *et al.*, 2006). We also constrained the dispersal between TEP and WA from 3.1 Myr onwards,

reflecting the final closure of the Isthmus of Panama (Lessios, 2008). Finally, we assigned a low dispersal value between CP and TEP to reflect the East Pacific Barrier (Bellwood & Wainwright, 2002) in all time-slices considered, but still permitting the dispersal given the soft nature of the barrier (Lessios & Robertson, 2006; Cowman & Bellwood, 2013b). Although we set the dispersal multiplier matrices to reflect realistic dispersal probabilities relative to the presence of major biogeographical events/barriers through time, we also built models in which the matrices could be adjusted according to the data. This was achieved by setting the matrix power exponential (parameter  $w$ ) to be free and estimated with maximum likelihood, which reduces the subjectivity in user-defined values for dispersal multiplier matrices (Dupin *et al.*, 2017). In total, we fitted 12 biogeographical models to each phylogeny. These models were compared using AIC scores to assess the best estimates for ancestral range reconstructions in our fish groups.

#### Uncertainties in ancestral state reconstructions

We assessed the robustness of our trait space results against two issues that could potentially affect our ancestral state reconstructions: phylogenetic uncertainties (topology and node dates); and uncertainties of reconstructed node states. Firstly, to deal with the topological and dating uncertainties in the phylogenies, we randomly sampled 1000 trees for each group from the combined post-burn in posterior sets derived from the Bayesian inferences. In each sampled tree, we retrieved the ancestral states of all traits per time-slice (20, 15, 10 and 5 Ma). This was achieved by re-rooting (function *'reroot'* in *'phytools'* R package; Revell, 2012) the trees in all edge points cut by the time-slices and reconstructing states in each re-rooted node point. For the discrete traits, we performed reconstructions using the *'ace'* function from *'ape'* R package (Paradis *et al.*, 2004). This function implements maximum likelihood joint estimation based on a transition matrix between character states. These reconstructions were performed using single-rate models, chosen based on likelihood ratio tests over other reconstruction models. For the continuous trait, we used the function *'fastAnc'* (Revell, 2012), which estimates ancestral character states for continuous traits through maximum likelihood. The most probable state for the discrete traits and the reconstructed value for the continuous trait for each re-rooted node point (lineage) were used in subsequent analysis. After retrieving the ancestral states, we constructed 1000 trait space polygons, each based on a combined set of three sampled phylogenies (one per taxa). We then overlapped all the polygons with a high transparency value to create a 'heatmap' for the most likely area occupied in each time-slice (Supplementary Figure 4.14).

Second, we assessed the effect of uncertainty in the reconstructed states of our categorical traits. To do that, we performed two new analyses of the multidimensional trait space. In the first one, we used the results of our Bayesian ancestral state reconstructions (see Methods section 4.3.3) to reclassify the nodes in which the modal posterior probability (PP) was below 0.67. If this was the case, we classified the node with the second most likely state. This threshold means that the modal PP values of the most probable state was at least twice as likely than the second best-supported state. In the second analysis, we selected which state to consider for each node by comparing the posterior probability distributions. In case the best-supported state comprised less than 95% of the PP samples, we retrieved the second best-supported state. With the results of each analysis combined with the main results for the continuous trait reconstructions, we once again plotted the multidimensional trait space for each biogeographical realm (Supplementary Figures 4.15-4.16).

Finally, we performed ancestral reconstructions with a recently developed hidden Markov model (SecSSE - version 0.1.12, privately provided by the author; Herrera-Alsina *et al.*, 2019) for discrete trait evolution. We used this maximum likelihood framework to perform ancestral reconstructions accounting for state-dependent diversification and the existence of more than one transition matrix for each character state. For each of our classified traits, we built one model by allowing speciation and extinction rates to vary between observed and concealed states ( $\lambda_{0A} \neq \lambda_{1A} \neq \lambda_{0B} \neq \lambda_{1B}$ ;  $\mu_{0A} \neq \mu_{1A} \neq \mu_{0B} \neq \mu_{1B}$  - in a two-state example), and by setting a transition rate matrix with one concealed rate regime (B) for all character states. We then used the resulting maximum likelihood parameter estimates, to retrieve the ancestral states for each node using the 'secsse\_loglik' function in the 'SecSSE' R package (Herrera-Alsina *et al.*, 2019). These results were combined with the main results for the continuous trait reconstructions to plot the multidimensional trait space (see Methods section 4.3.4) for each biogeographical realm (Supplementary Figure 4.17). This maximum likelihood model complements the main Bayesian analysis, to ensure the robustness of our results against trait-dependent diversification and the existence of rate heterogeneity in trait evolution within the trees.

## Supplementary References

- Adam, T.C., Kelley, M., Ruttenberg, B.I. & Burkepile, D.E. (2015) Resource partitioning along multiple niche axes drives functional diversity in parrotfishes on Caribbean coral reefs. *Oecologia*, 179, 1173–1185.
- Bannikov, A.F. & Tyler, J.C. (1995) Phylogenetic revision of the fish families Luvaridae and Kushlukiidae (Acanthuroidei), with a new genus and two new species of Eocene luvarids. *Smithsonian Contributions to Paleobiology*, 81, 1–45.

- Bannikov, A.F., Tyler, J.C. & Sorbini, C. (2010) Two new taxa of Eocene rabbitfishes (Perciformes, Siganidae) from the North Caucasus (Russia), with redescription of *Acanthopygaeus agassizi* (Eastman) from Monte Bolca (Italy) and a phylogenetic analysis of the family. *Bollettino del Museo Civico di Storia Naturale di Verona, Geologia, Paleontologia, Preistoria*, 34, 3–21.
- Bellwood, D.R. (1990) A new fossil fish *Phyllopharyngodon longipinnis* gen. et sp. nov. (family Labridae) from the Eocene, Monte Bolca, Italy. *Studi e Ricerche sui Giacimenti Terziari di Bolca, Museo Civico di Storia Naturale, Verona*, 149–160.
- Bellwood, D.R. (1994) A Phylogenetic study of the parrotfishes family Scaridae (Pisces: Labroidei), with a revision of genera. *Records of the Australian Museum, Supplement*, 1–86.
- Bellwood, D.R. (1996) The Eocene fishes of Monte Bolca: the earliest coral reef fish assemblage. *Coral Reefs*, 15, 11–19.
- Bellwood, D.R. (1985) The functional morphology, systematics and behavioural ecology of parrotfishes (family Scaridae). PhD thesis, James Cook University.
- Bellwood, D.R. & Choat, J.H. (1990) A functional analysis of grazing in parrotfishes (family Scaridae): the ecological implications. *Environmental Biology of Fishes*, 28, 189–214.
- Bellwood, D.R. & Schultz, O. (1991) A Review of the fossil record of the Parrotfishes (Labroidei:Scaridae) with a description of a new *Calotomus* species from the Middle Miocene (Badenian) of Austria. *Annalen Des Naturhistorischen Museums in Wien. Serie A, Fur Mineralogie und Petrographie, Geologie Und Palaontologie, Anthropologie und Prahistorie*, 92, 55–71.
- Bellwood, D.R. & Wainwright, P.C. (2002) *The history and biogeography of fishes on coral reefs. Coral reef fishes: dynamics and diversity on a complex ecosystem* (ed. by P.F. Sale), pp. 5–32. Academic Press, San Diego.
- Bonaldo, R.M., Hoey, A.S. & Bellwood, D.R. (2014) The ecosystem roles of parrotfishes on tropical reefs. *Oceanography and Marine Biology: An Annual Review*, 52, 81–132.
- Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C.H., Xie, D., Suchard, M.A., Rambaut, A. & Drummond, A.J. (2014) BEAST 2: a software platform for Bayesian evolutionary analysis. *PLoS Computational Biology*.
- Bowen, B.W., Muss, A., Rocha, L.A. & Grant, W.S. (2006) Shallow mtDNA coalescence in Atlantic pygmy angelfishes (genus *Centropyge*) indicates a recent invasion from the Indian Ocean. *Journal of Heredity*, 97, 1–12.
- Brandl, S.J. & Bellwood, D.R. (2014) Individual-based analyses reveal limited functional overlap in a coral reef fish community. *Journal of Animal Ecology*, 83, 661–670.
- Brandl, S.J. & Bellwood, D.R. (2016) Microtopographic refuges shape consumer-producer dynamics by mediating consumer functional diversity. *Oecologia*, 182, 203–217.

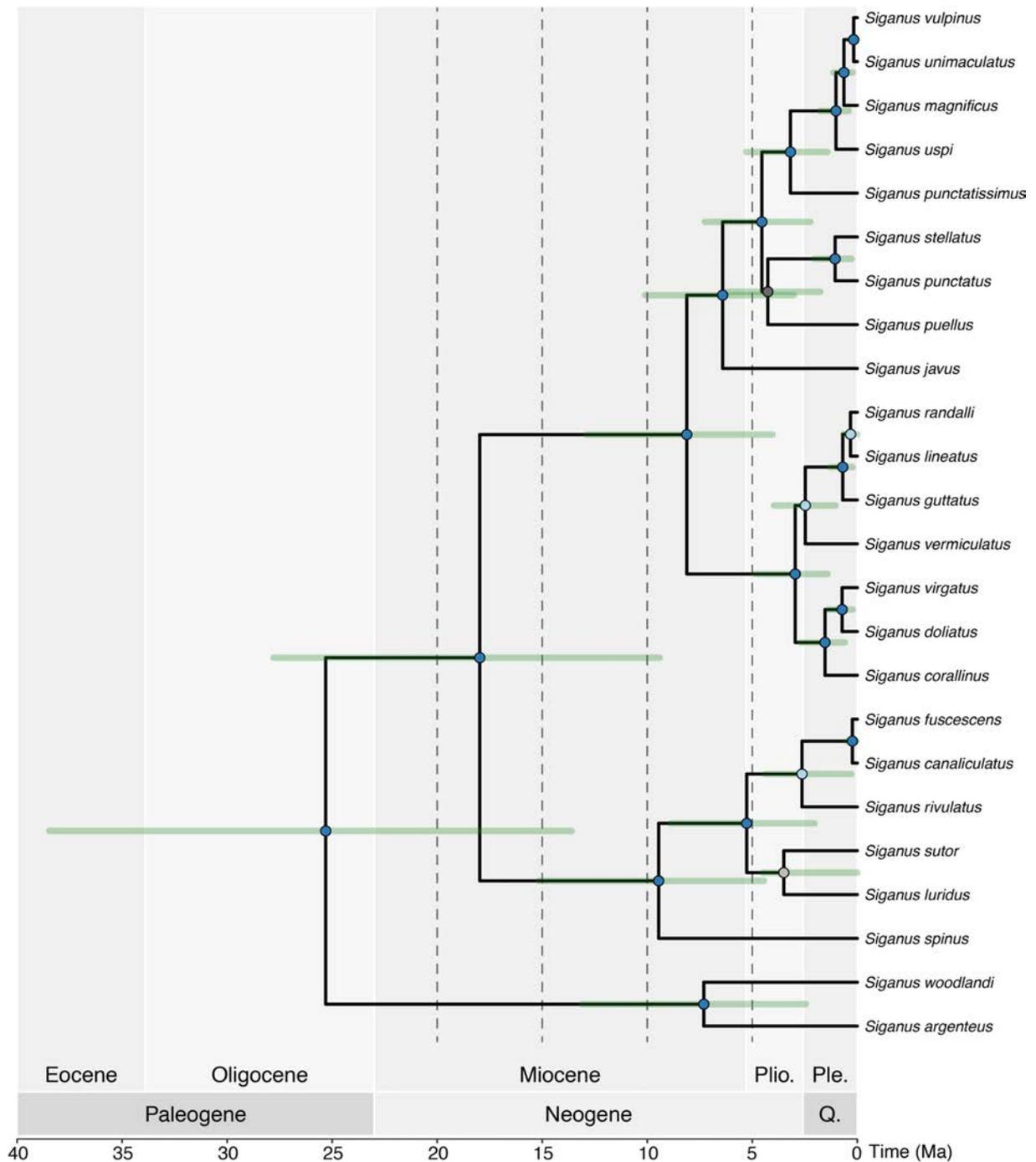
- Brandl, S.J. & Bellwood, D.R. (2013) Morphology, sociality, and ecology: can morphology predict pairing behavior in coral reef fishes? *Coral Reefs*, 32, 835–846.
- Brandl, S.J., Hoey, A.S. & Bellwood, D.R. (2014) Micro-topography mediates interactions between corals, algae, and herbivorous fishes on coral reefs. *Coral Reefs*, 33, 421–430.
- Choat, J.H. (1969) Studies on the biology of labroid fishes (Labridae and Scaridae) at Heron Island, Great Barrier Reef. PhD Thesis, School of Biological Sciences, The University of Queensland.
- Choat, J.H. (1991) *The biology of herbivorous fishes on coral reefs. The Ecology of Fishes on Coral Reefs* (ed. by P.F. Sale), pp. 120–155. Academic Press, San Diego.
- Choat, J.H., Clements, K.D. & Robbins, W.D. (2002) The trophic status of herbivorous fishes on coral reefs 1: Dietary analyses. *Marine Biology*, 140, 613–623.
- Choat, J.H., Klanten, O.S., Van Herwerden, L., Robertson, R.D. & Clements, K.D. (2012) Patterns and processes in the evolutionary history of parrotfishes (Family Labridae). *Biological Journal of the Linnean Society*, 107, 529–557.
- Choat, J.H., Robbins, W.D. & Clements, K.D. (2004) The trophic status of herbivorous fishes on coral reefs: II. Food processing modes and trophodynamics. *Marine Biology*, 145, 445–454.
- Clements, K.D. (1991) Gut microorganisms of surgeonfishes (family Acanthuridae). PhD thesis, James Cook University.
- Clements, K.D. & Choat, J.H. (1995) Fermentation in tropical marine herbivorous fishes. *Physiol. Zool.*, 68, 355–378.
- Clements, K.D., German, D.P., Piché, J., Tribollet, A. & Choat, J.H. (2017) Integrating ecological roles and trophic diversification on coral reefs: multiple lines of evidence identify parrotfishes as microphages. *Biological Journal of the Linnean Society*, 120, 729–751.
- Cowman, P.F. & Bellwood, D.R. (2011) Coral reefs as drivers of cladogenesis: expanding coral reefs, cryptic extinction events, and the development of biodiversity hotspots. *Journal of Evolutionary Biology*, 24, 2543–2562.
- Cowman, P.F. & Bellwood, D.R. (2013a) The historical biogeography of coral reef fishes: global patterns of origination and dispersal. *Journal of Biogeography*, 40, 209–224.
- Cowman, P.F. & Bellwood, D.R. (2013b) Vicariance across major marine biogeographic barriers: temporal concordance and the relative intensity of hard versus soft barriers. *Proceedings of the Royal Society B: Biological Sciences*, 280, 20131541.
- Dunlap, M. & Pawlik, J.R. (1996) Video monitored predation by caribbean reef fishes on an array of mangrove and reef sponges. *Marine Biology*, 126, 117–123.

- Dupin, J., Matzke, N.J., Särkinen, T., Knapp, S., Olmstead, R.G., Bohs, L. & Smith, S.D. (2017) Bayesian estimation of the global biogeographical history of the Solanaceae. *Journal of Biogeography*, 44, 887–899.
- Edgar, R.C. (2004) MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, 32, 1792–1797.
- Ferreira, C.E.L. & Gonçalves, J.E.A. (2006) Community structure and diet of roving herbivorous reef fishes in the Abrolhos Archipelago, south-western Atlantic. *Journal of Fish Biology*, 69, 1533–1551.
- Fox, R.J. & Bellwood, D.R. (2013) Niche partitioning of feeding microhabitats produces a unique function for herbivorous rabbitfishes (Perciformes, Siganidae) on coral reefs. *Coral Reefs*, 32, 13–23.
- Francini-Filho, R.B., Ferreira, C.M., Coni, E.O.C., De Moura, R.L. & Kaufman, L. (2010) Foraging activity of roving herbivorous reef fish (Acanthuridae and Scaridae) in eastern Brazil: influence of resource availability and interference competition. *Journal of the Marine Biological Association of the United Kingdom*, 90, 481–492.
- Froese, R. & Pauly, D. (2019) FishBase. Available at [fishbase.org](http://fishbase.org).
- Green, A.L. & Bellwood, D.R. (2009) *Monitoring functional groups of herbivorous reef fishes as indicators of coral reef resilience: a practical guide for coral reef managers in the Asia Pacific region*, IUCN Resilience Science Group Working Paper Series No 7, Gland, Switzerland.
- Herrera-Alsina, L., van Els, P. & Etienne, R.S. (2019) Detecting the dependence of diversification on multiple traits from phylogenetic trees and trait data. *Systematic Biology*, 68, 317–328.
- Hiatt, R.W. & Strasburg, D.W. (1960) Ecological relationships of the fish fauna on coral reefs of the Marshall Islands. *Ecological Monographs*, 30, 65–127.
- Hoey, A.S., Brandl, S.J. & Bellwood, D.R. (2013) Diet and cross-shelf distribution of rabbitfishes (f. Siganidae) on the northern Great Barrier Reef: implications for ecosystem function. *Coral Reefs*, 32, 973–984.
- Horn, M.H. (1989) Biology of marine herbivorous fishes. *Oceanography and Marine Biology Annual Reviews*, 27, 167–272.
- Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., Thierer, T., Ashton, B., Meintjes, P. & Drummond, A. (2012) Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*, 28, 1647–1649.
- Klanten, S.O., van Herwerden, L., Choat, J.H. & Blair, D. (2004) Patterns of lineage diversification in the genus *Naso* (Acanthuridae). *Molecular Phylogenetics and Evolution*, 32, 221–235.
- Kuriwa, K., Hanzawa, N., Yoshino, T., Kimura, S. & Nishida, M. (2007) Phylogenetic relationships and natural hybridization in rabbitfishes (Teleostei: Siganidae) inferred from mitochondrial and nuclear DNA analyses. *Molecular Phylogenetics and Evolution*, 45, 69–80.

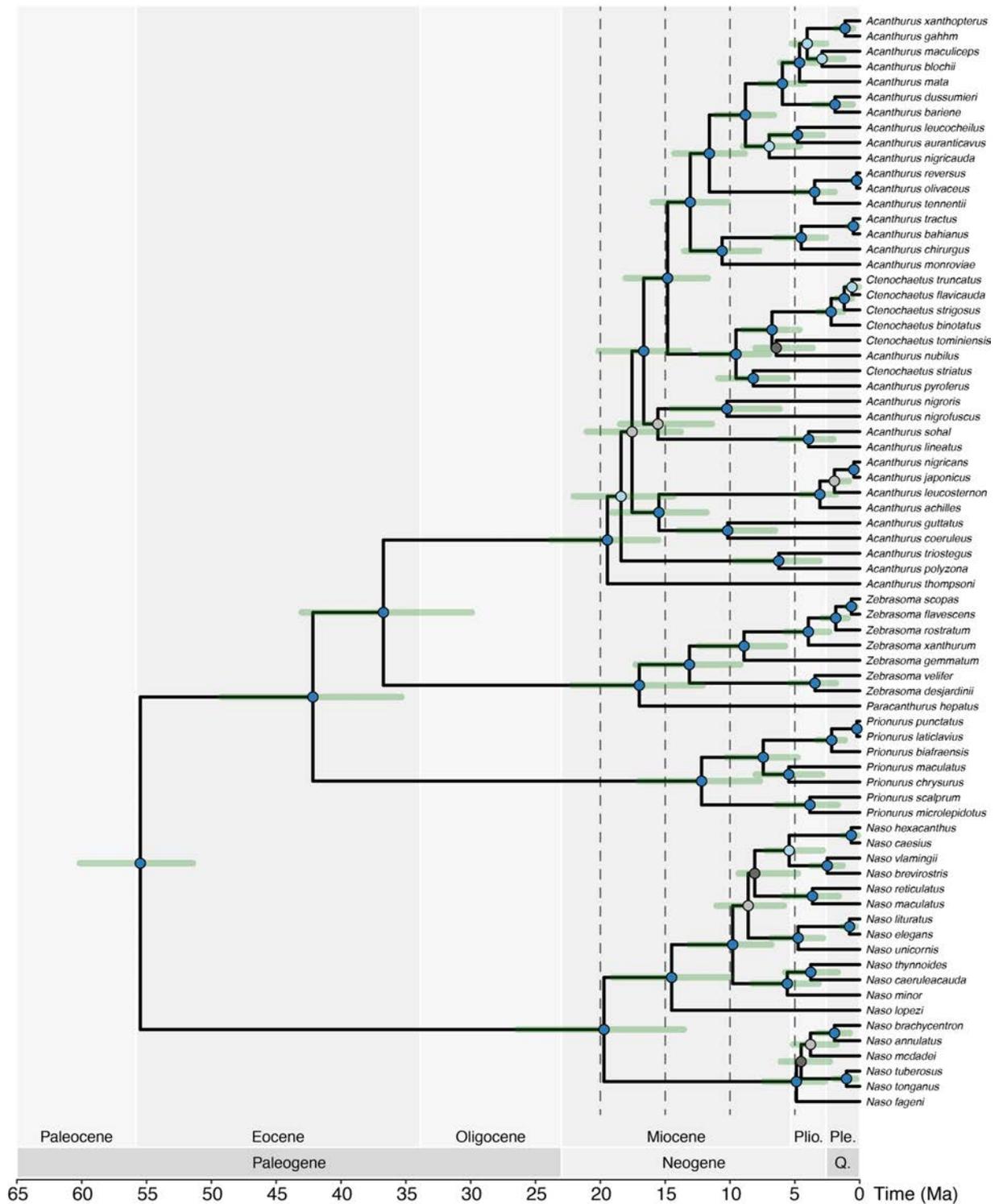
- Landis, M.J., Matzke, N.J., Moore, B.R. & Huelsenbeck, J.P. (2013) Bayesian analysis of biogeography when the number of areas is large. *Systematic Biology*, 62, 789–804.
- Lanfear, R., Frandsen, P.B., Wright, A.M., Senfeld, T. & Calcott, B. (2016) PartitionFinder2 : new methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. *Molecular Biology and Evolution*, 34, 772–773.
- Lessios, H. & Robertson, D. (2006) Crossing the impassable: genetic connections in 20 reef fishes across the eastern Pacific barrier. *Proceedings of the Royal Society B: Biological Sciences*, 273, 2201–2208.
- Lessios, H.A. (2008) The great American schism: divergence of marine organisms after the rise of the central American isthmus. *Annual Review of Ecology, Evolution, and Systematics*, 39, 63–91.
- Matzke, N. (2013) BioGeoBEARS: BioGeography with Bayesian (and Likelihood) Evolutionary Analysis in R Scripts. *R package, version 0.2.1*.
- Matzke, N.J. (2014) Model selection in historical biogeography reveals that founder-event speciation is a crucial process in island clades. *Systematic Biology*, 63, 951–970.
- Miller, M.A., Pfeiffer, W. & Schwartz, T. (2010) *Creating the CIPRES Science Gateway for inference of large phylogenetic trees. 2010 Gateway Computing Environments Workshop, GCE 2010*.
- Paradis, E., Claude, J. & Strimmer, K. (2004) APE: Analyses of Phylogenetics and Evolution in R language. *Bioinformatics*, 20, 289–290.
- Plass-Johnson, J.G., McQuaid, C.D. & Hill, J.M. (2013) Stable isotope analysis indicates a lack of inter- and intra-specific dietary redundancy among ecologically important coral reef fishes. *Coral Reefs*, 32, 429–440.
- Purcell, S.W. & Bellwood, D.R. (1993) A functional analysis of food procurement in two surgeonfish species, *Acanthurus nigrofuscus* and *Ctenochaetus striatus* (Acanthuridae). *Environmental Biology of Fishes*, 37, 139–159.
- Rambaut, A., Drummond, A.J., Xie, D., Baele, G. & Suchard, M.A. (2018) Posterior Summarization in Bayesian Phylogenetics Using Tracer 1.7. *Systematic Biology*, 67, 901–904.
- Randall, J.E. (1967) Food habits of reef fishes of the West Indies. *Stud. Trop. Oceanogr.*, 5, 665–847.
- Randall, J.E. (2001) *Surgeonfishes of the world*, Mutual Publishing, Honolulu, Hawaii.
- Ree, R.H. & Smith, S.A. (2008) Maximum likelihood inference of geographic range evolution by dispersal, local extinction, and cladogenesis. *Systematic Biology*, 57, 4–14.
- Renema, W., Bellwood, D.R. & Braga, J.C. (2008) Hopping hotspots: global shifts in marine biodiversity. *Science*, 321, 654–657.

- Revell, L.J. (2012) Phytools: an R package for phylogenetic comparative biology (and other things). *Methods in Ecology and Evolution*, 3, 217–223.
- Ronquist, F. (1997) Dispersal-vicariance analysis: a new approach to the quantification of historical biogeography. *Systematic Botany*, 46, 195–203.
- Sorenson, L., Santini, F., Carnevale, G. & Alfaro, M.E. (2013) A multi-locus timetree of surgeonfishes (Acanthuridae, Percomorpha), with revised family taxonomy. *Molecular Phylogenetics and Evolution*, 68, 150–160.
- Stamatakis, A. (2014) RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, 30, 1312–1313.
- Steininger, F.F. & Rögl, F. (1979) The paratethys history. A contribution towards the Neogene geodynamics of the alpine orogene. *Annales Geologiques des Pays Helleniques*, 3, 1153–1165.
- Tebbett, S.B., Goatley, C.H.R. & Bellwood, D.R. (2017) Clarifying functional roles: algal removal by the surgeonfishes *Ctenochaetus striatus* and *Acanthurus nigrofuscus*. *Coral Reefs*, 36, 803–813.
- Woodland, D.J. (1990) *Revision of the fish family Siganidae with description of two new species and comments on distribution and biology*, Bernice Pauahi Bishop Museum, Honolulu, Hawaii.

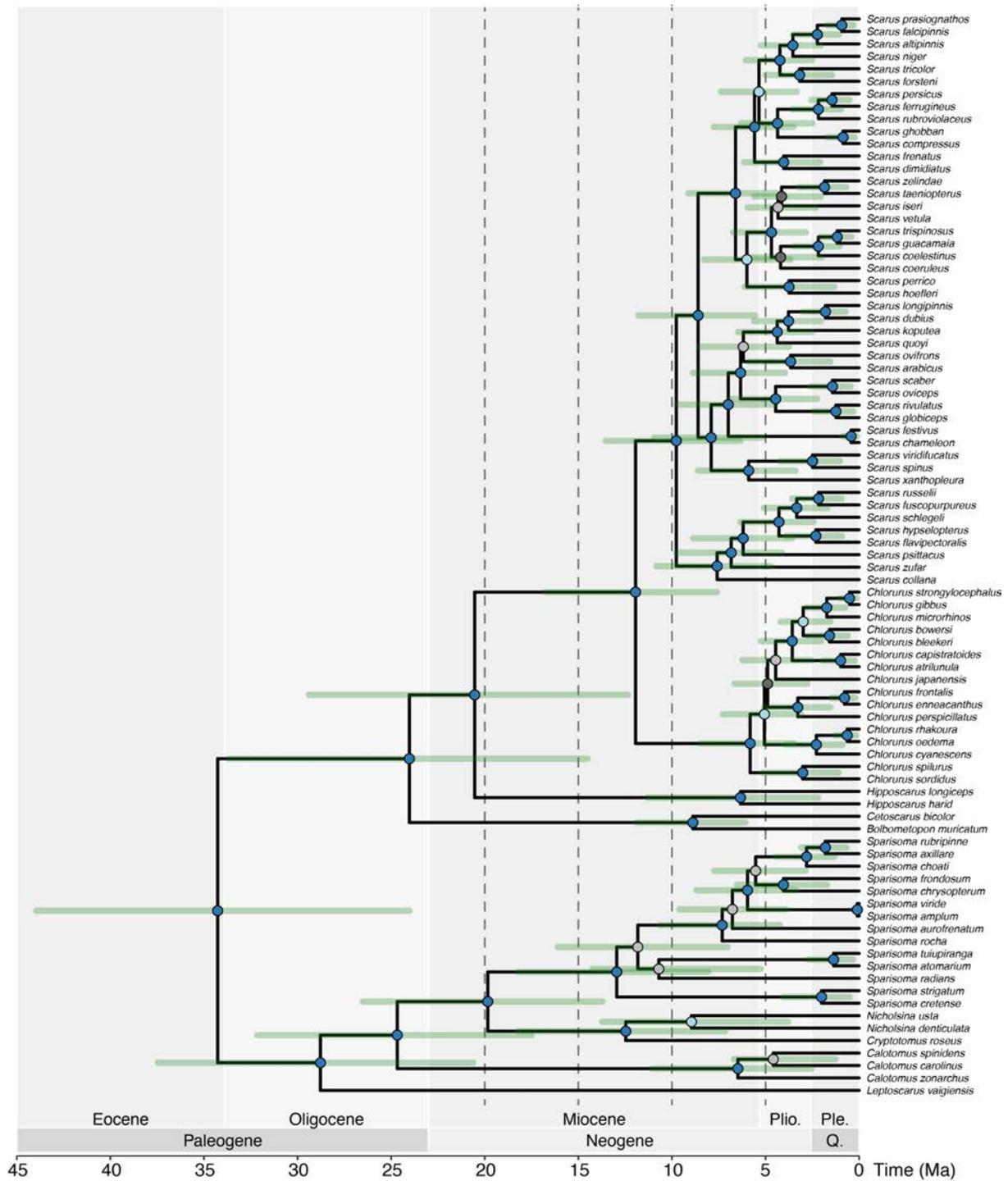
Supplementary Figures



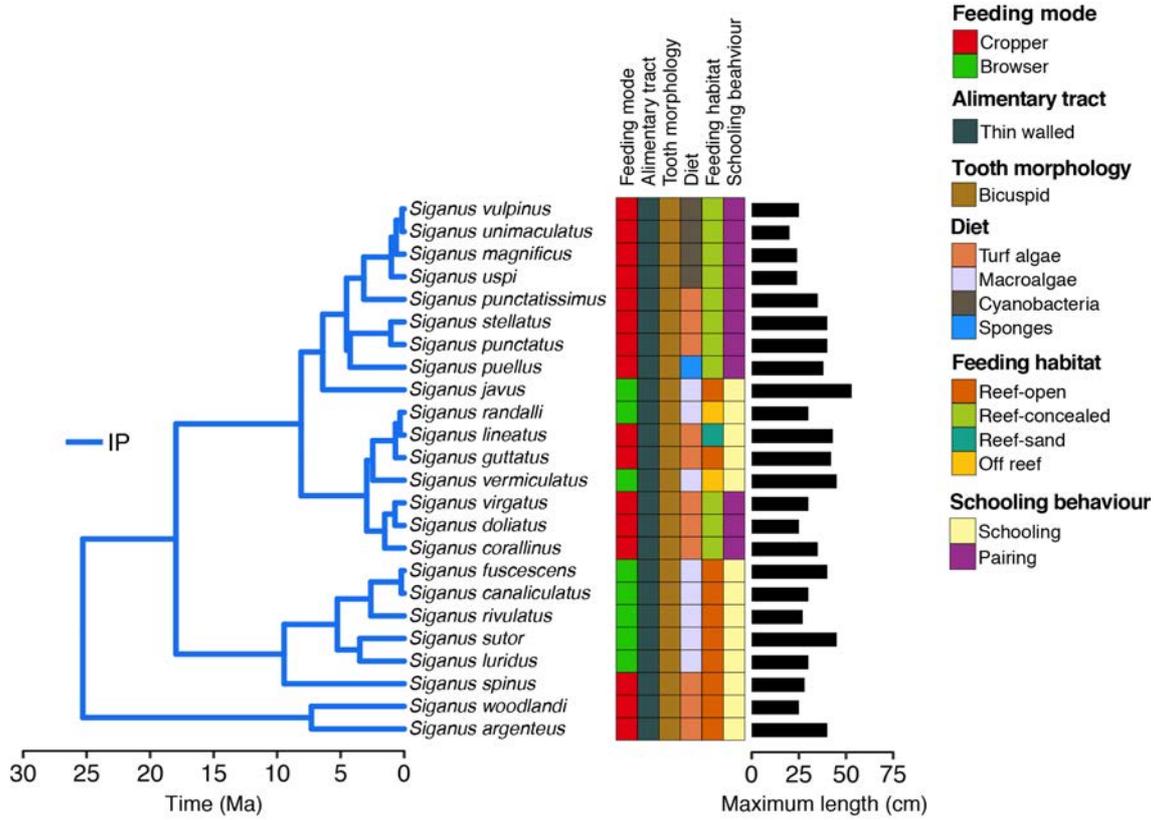
**Supplementary Figure 4.1** | Maximum Clade Credibility (MCC) tree resulting from the Bayesian Inferences for Siganiidae. Node colours represent posterior support values:  $\geq 0.9$  (dark blue);  $< 0.9 - \geq 0.7$  (light blue);  $< 0.7 - \geq 0.5$  (light grey);  $\leq 0.5$  (dark grey). Highest Posterior Density intervals for node ages are shown in green. Dashed lines represent the time-slices used to retrieve ancestral states.



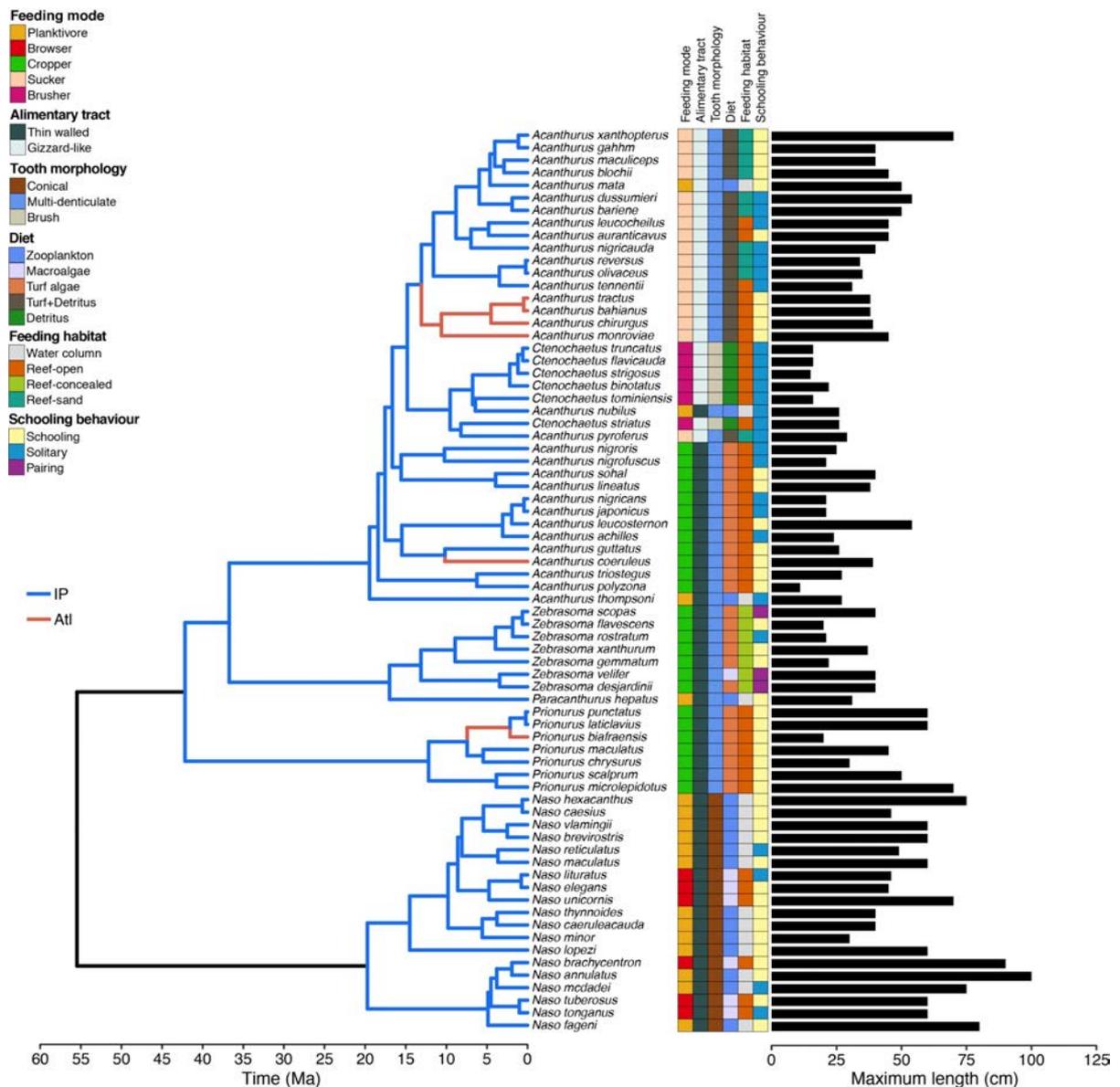
**Supplementary Figure 4.2 |** Maximum Clade Credibility (MCC) tree resulting from the Bayesian Inferences for Acanthuridae. Node colours represent posterior support values:  $\geq 0.9$  (dark blue);  $< 0.9 - \geq 0.7$  (light blue);  $< 0.7 - \geq 0.5$  (light grey);  $\leq 0.5$  (dark grey). Highest Posterior Density intervals for node ages are shown in green. Dashed lines represent the time-slices used to retrieve ancestral states.



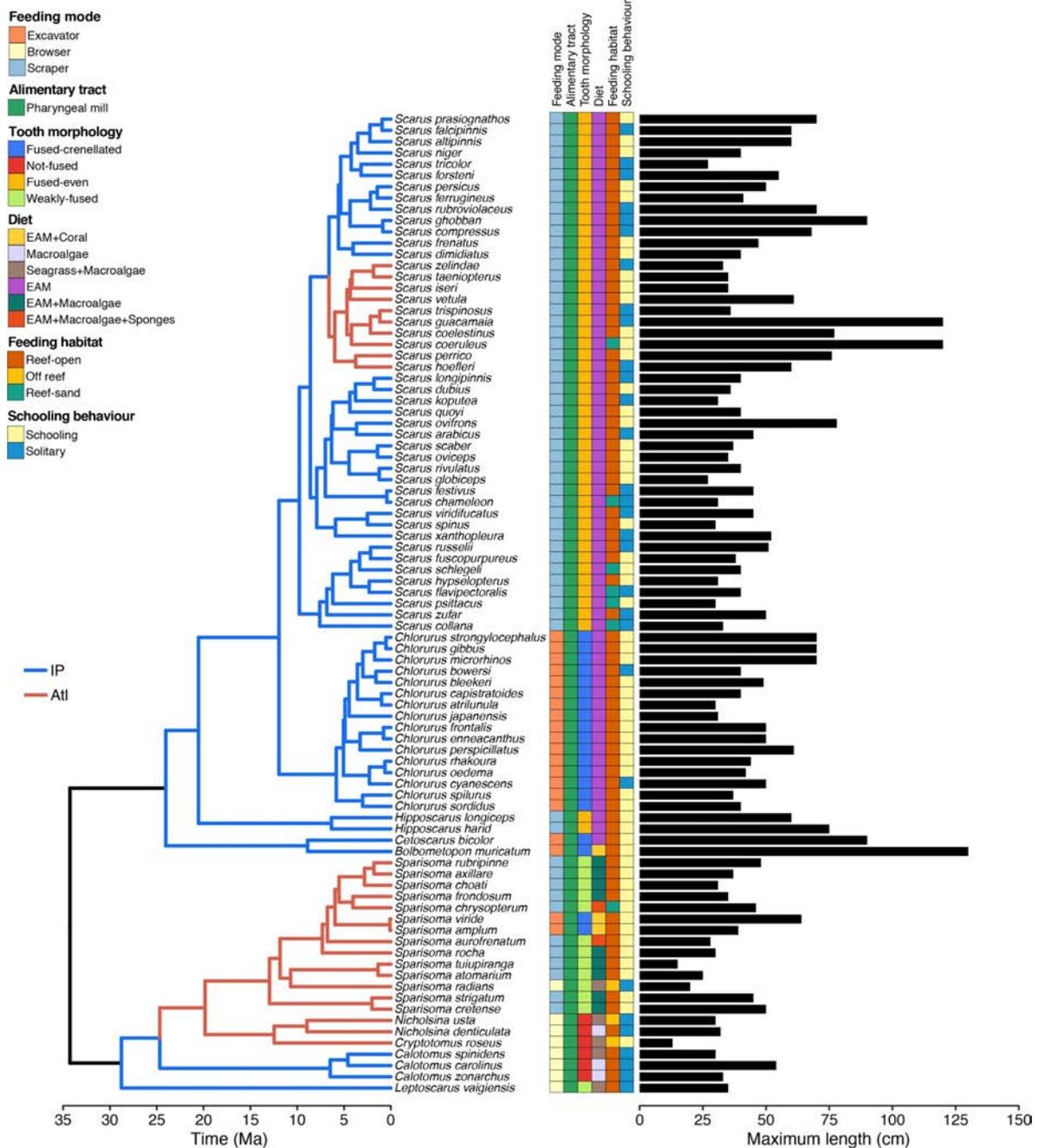
**Supplementary Figure 4.3 |** Maximum Clade Credibility (MCC) tree resulting from the Bayesian Inferences for Scarini. Node colours represent posterior support values:  $\geq 0.9$  (dark blue);  $< 0.9 - \geq 0.7$  (light blue);  $< 0.7 - \geq 0.5$  (light grey);  $\leq 0.5$  (dark grey). Highest Posterior Density intervals for node ages are shown in green. Dashed lines represent the time-slices used to retrieve ancestral states.



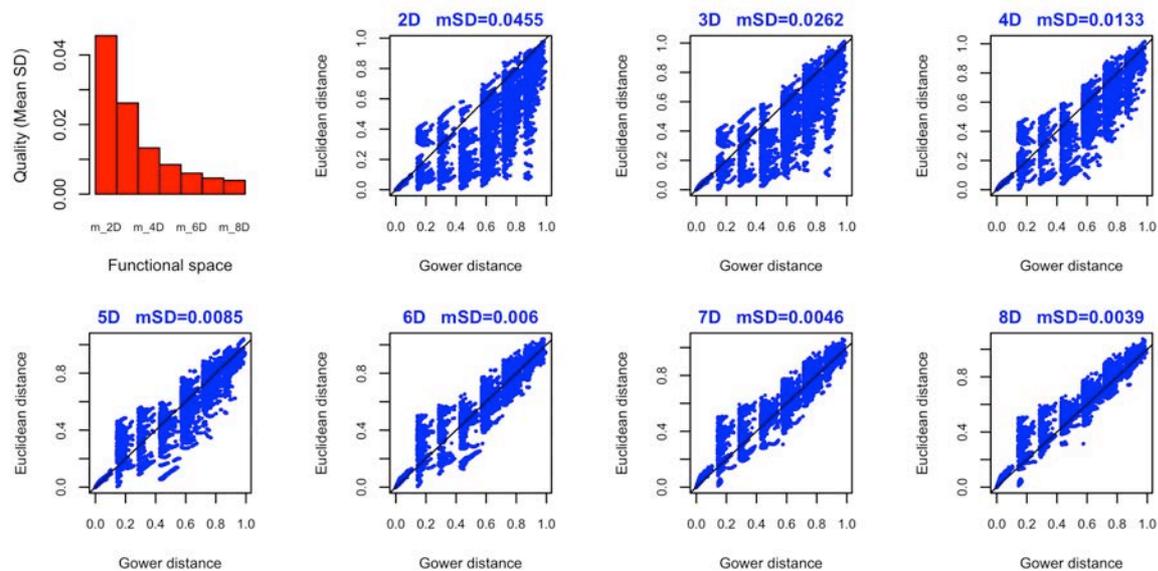
**Supplementary Figure 4.4 | Trait state distribution for Siganidae species ordered according the phylogeny.** Branches are coloured according to the results from the best ancestral range reconstruction model for the group (DEC; Supplementary Table 4.1), from which we split between areas present in the Indo-Pacific (IP) and Atlantic (Atl).



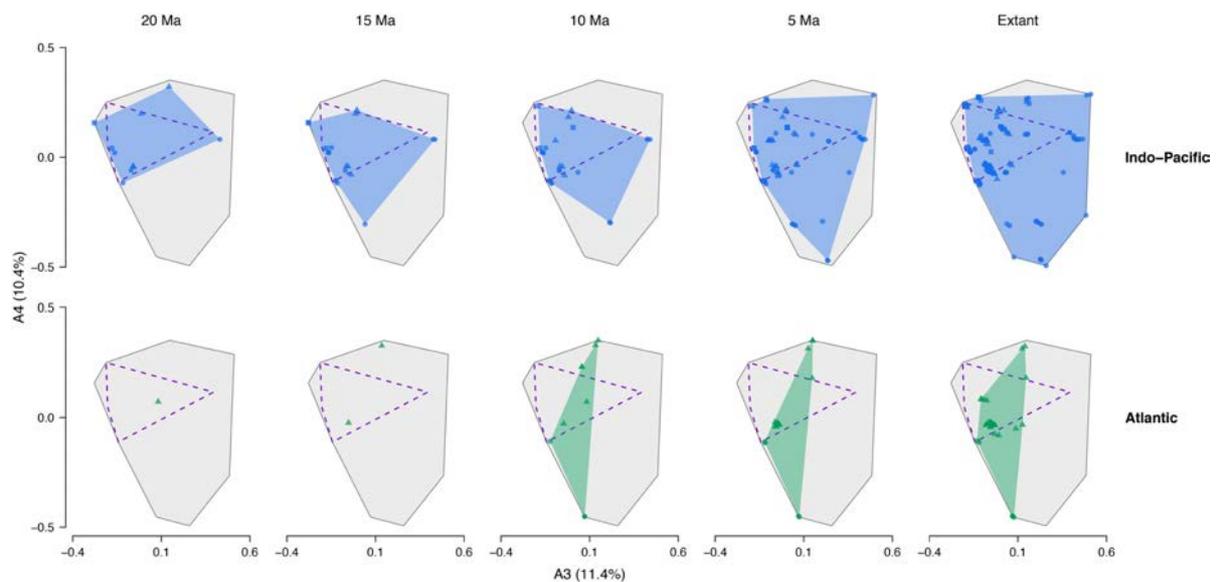
**Supplementary Figure 4.5 | Trait state distribution for Acanthuridae species ordered according the phylogeny.** Branches are coloured according to the results from the best ancestral range reconstruction model for the group (BAYAREALIKE+J+W; Supplementary Table 4.2), from which we split between areas present in the Indo-Pacific (IP) and Atlantic (Atl). Black branches represent ancestral lineages that were likely restricted to the Tethyan marine biodiversity hotspot (see Supplementary Methods).



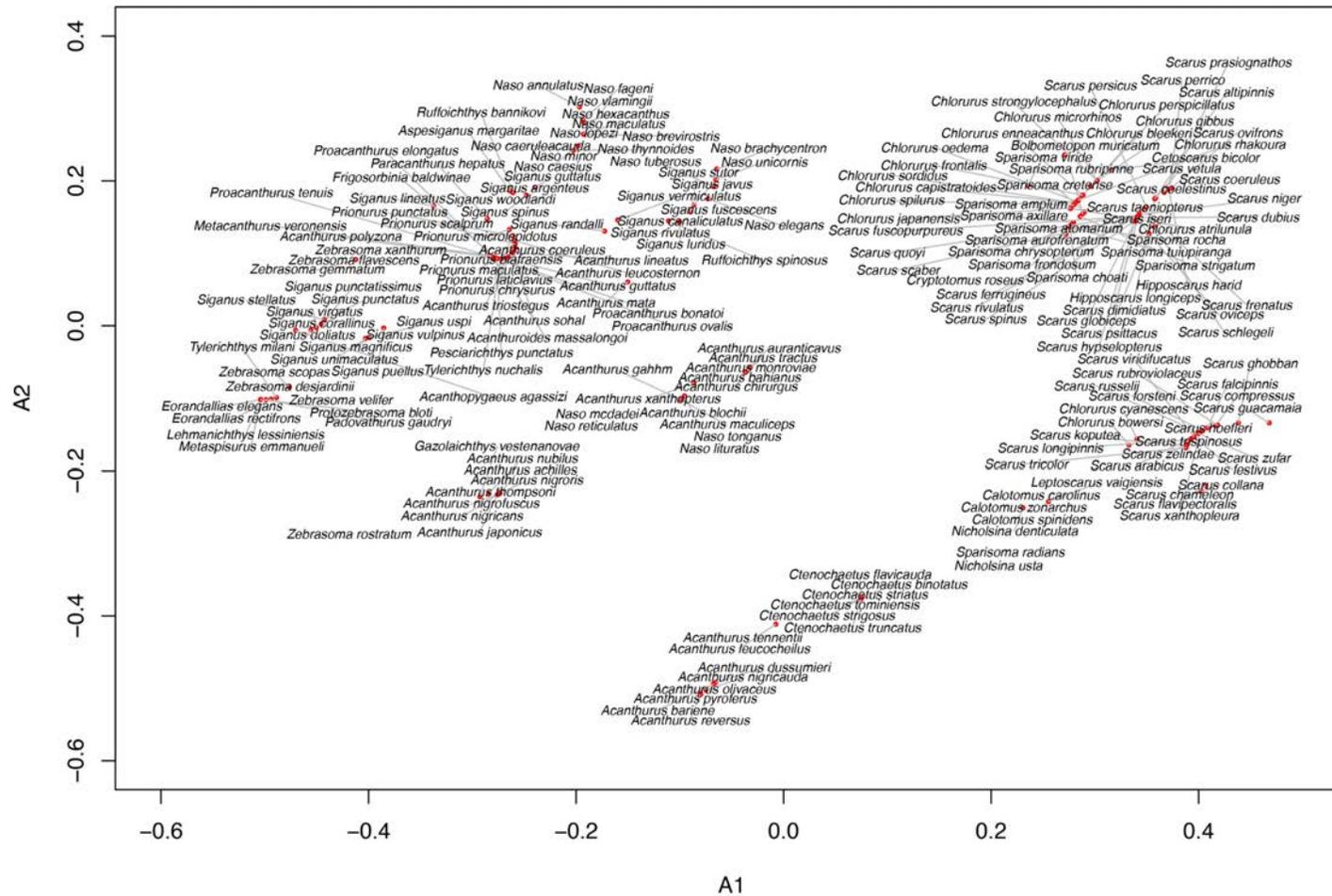
**Supplementary Figure 4.6** | Trait state distribution for Scarini species ordered according the phylogeny. Branches are coloured according the results from the best ancestral range reconstruction model for the group (DEC+J+W; Supplementary Table 4.3), from which we split between areas present in the Indo-Pacific (IP) and Atlantic (Atl). Black branches represent ancestral lineages that were likely restricted to the Tethyan marine biodiversity hotspot (see Supplementary Methods).



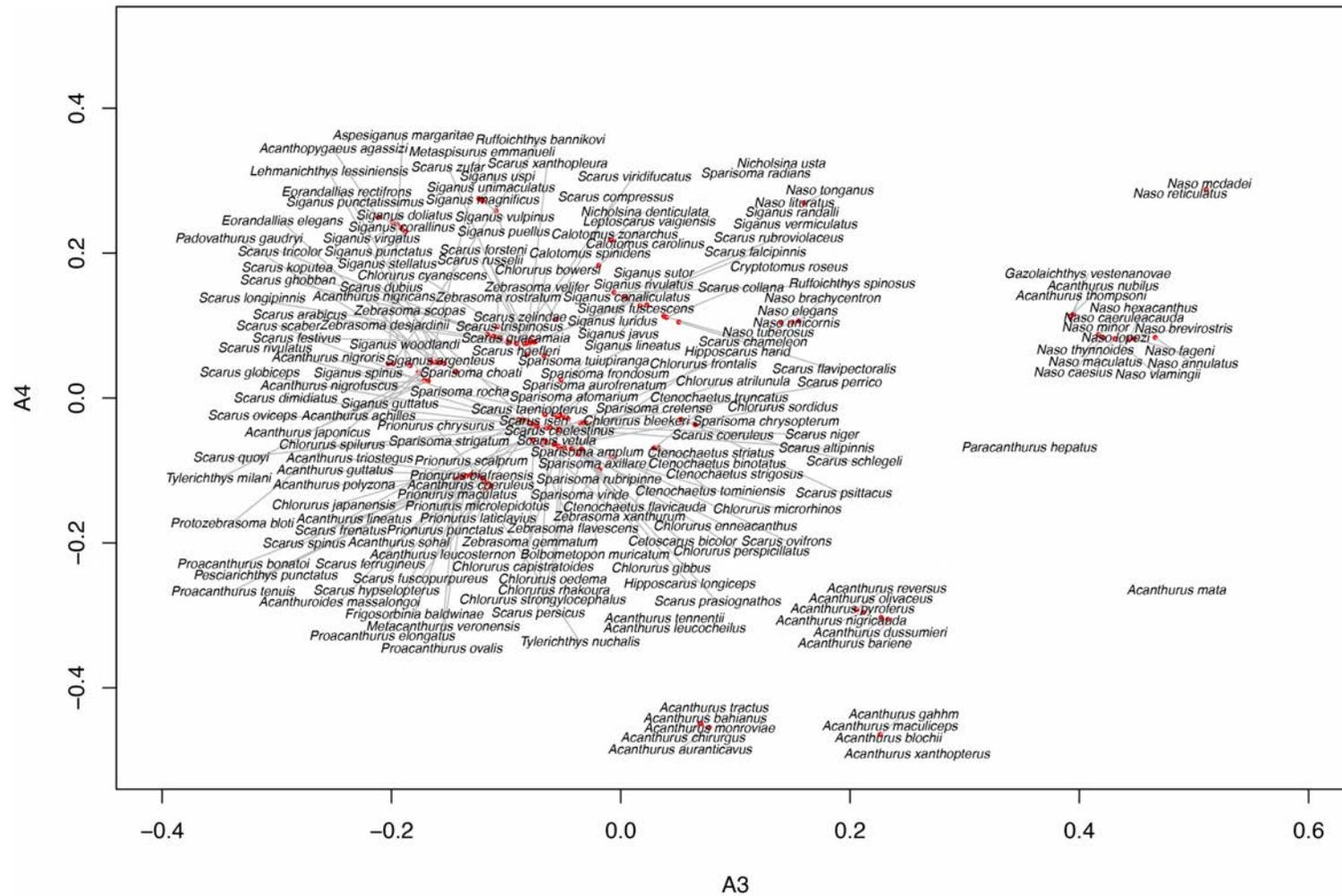
**Supplementary Figure 4.7** | Quality of representation of the multidimensional space for trait combinations in surgeonfishes, parrotfishes and rabbitfishes. Output from the 'quality\_funct\_space' R function, showing an increase in quality (lower mSD values) with concomitant increase in the number of axes.



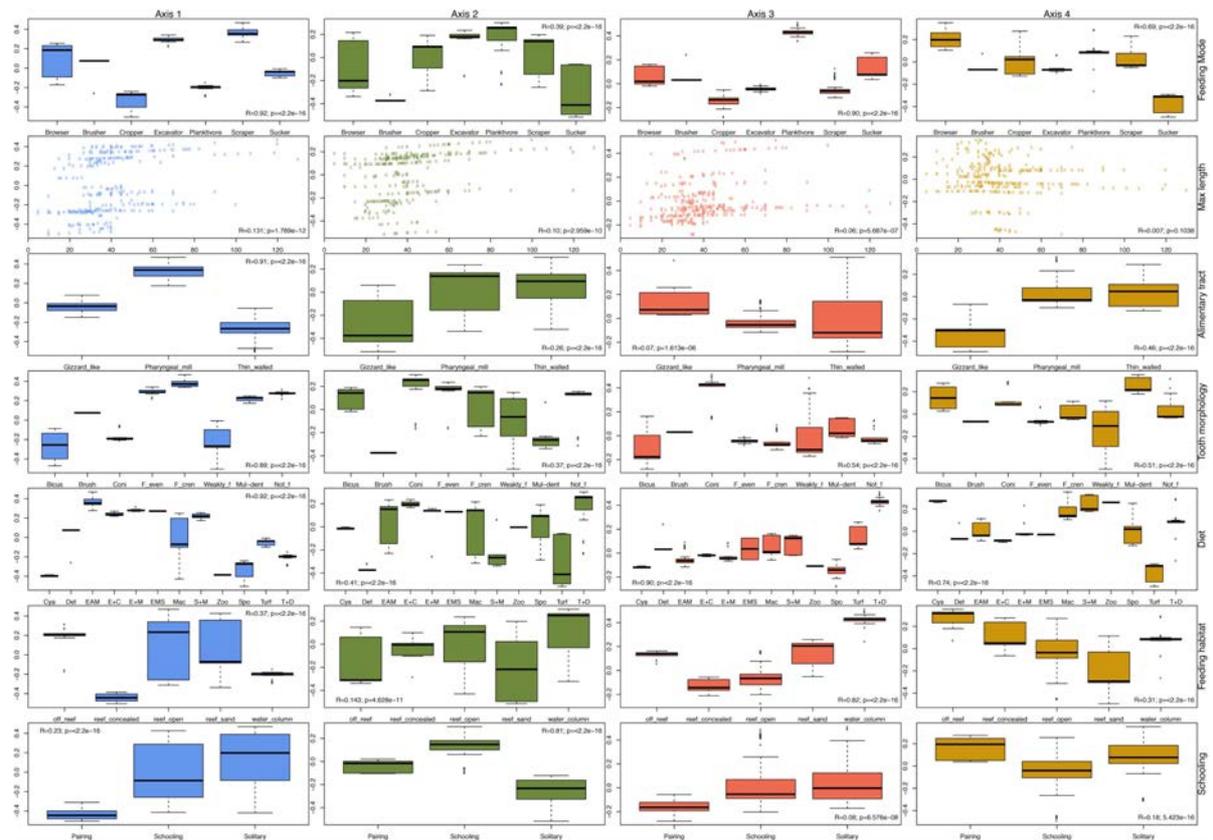
**Supplementary Figure 4.8** | Multidimensional trait space occupied by surgeonfish (circles), rabbitfish (squares) and parrotfish (triangles) lineages in two biogeographical regions through time. Plots show the third and fourth axes (A3-A4) derived from a principal coordinate analysis (PCoA) performed on seven traits related to feeding. Each column represents a time-slice (20-5 Ma) in which we assessed the traits through ancestral reconstructions. The Extant column represents trait space of extant species. Background grey area shows the total space occupied combining fossils, time-slices and extant species. Convex hulls represent space occupied by Atlantic (green), Indo-Pacific (blue) and Tethys (50 Ma; purple dashed line) lineages. Symbols represent lineages present in each point in time.



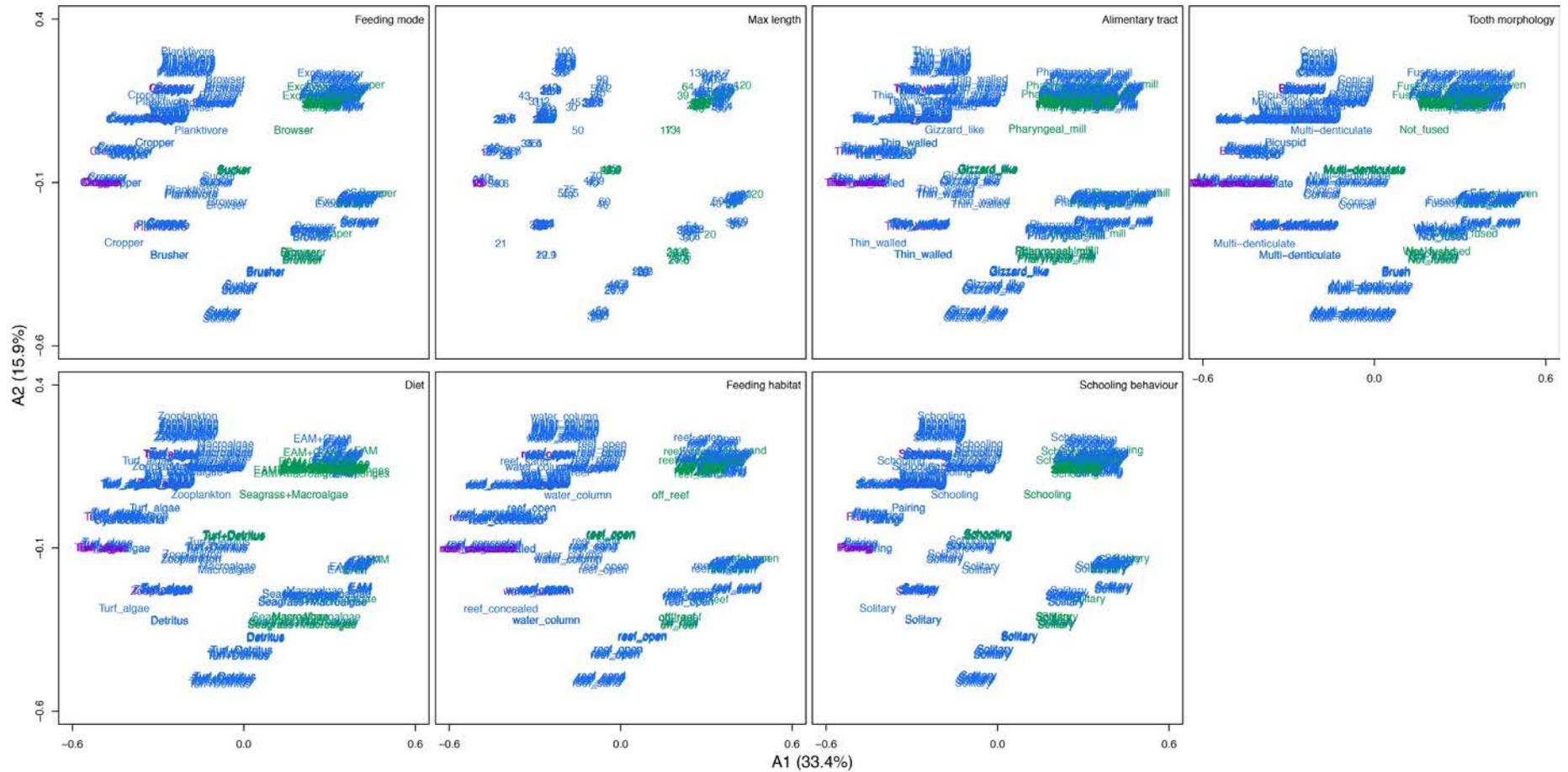
**Supplementary Figure 4.9** | Species' positions in multidimensional space for axes 1 and 2 derived from the PCoA performed on seven traits related to feeding of extant surgeonfishes, rabbitfishes and parrotfishes present in the phylogenies, and known herbivorous fish fossils from Monte Bolca.



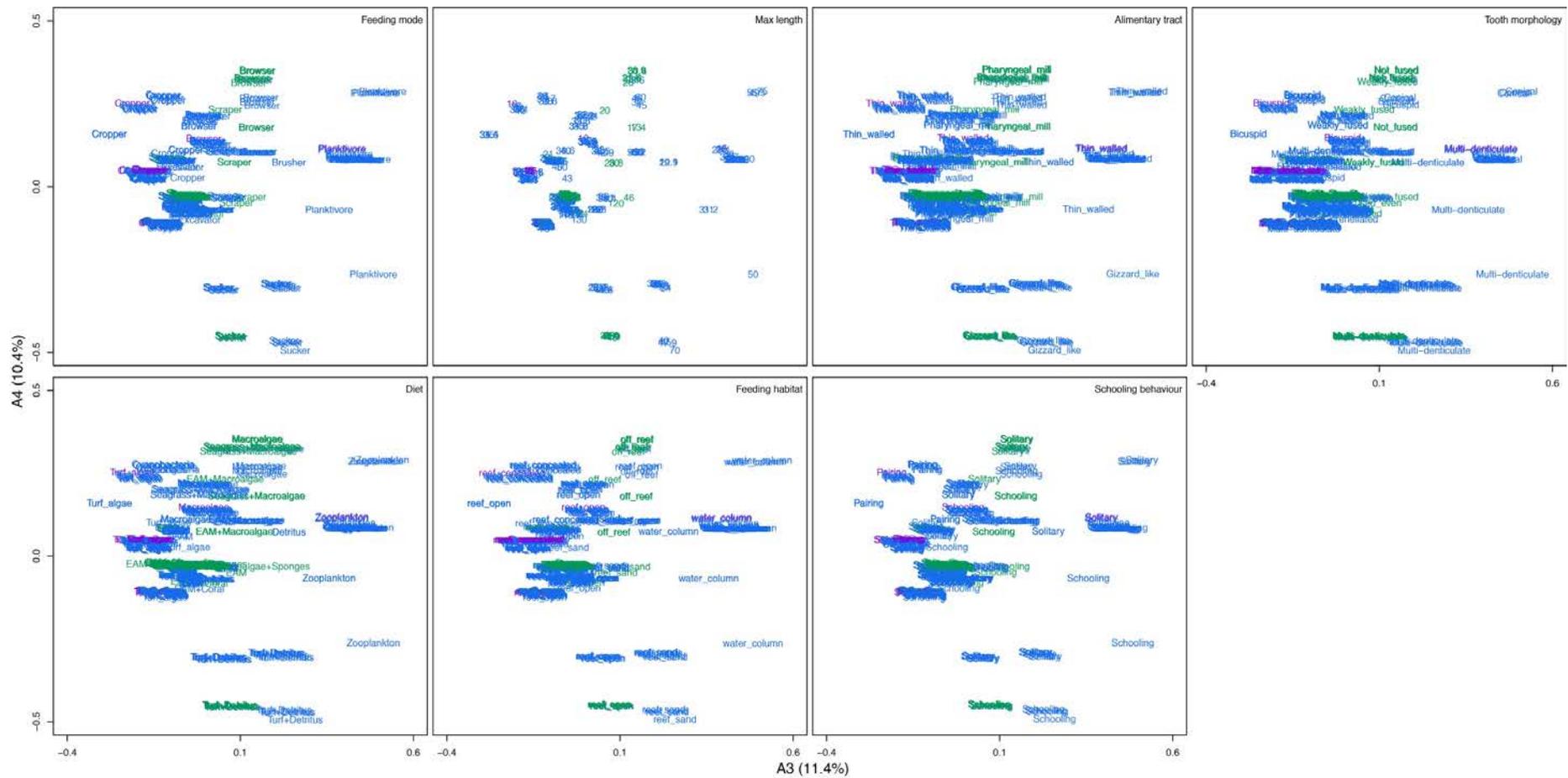
**Supplementary Figure 4.10** | Species' positions in multidimensional space for axes 3 and 4 derived from the PCoA performed on seven traits related to feeding of extant surgeonfishes, rabbitfishes and parrotfishes present in the phylogenies, and known herbivorous fish fossils from Monte Bolca.



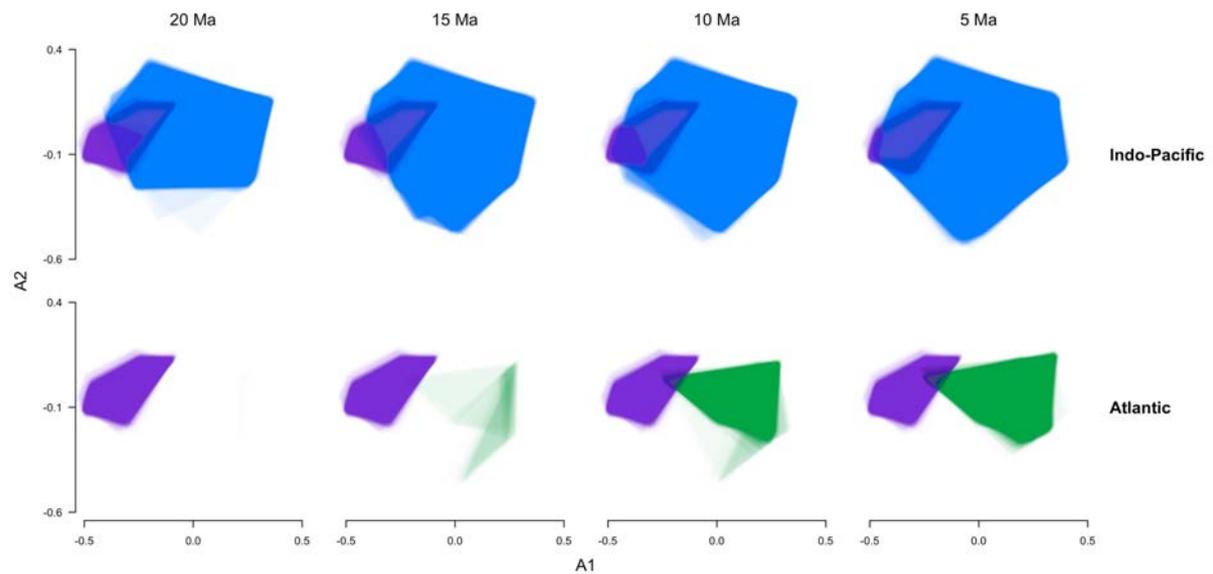
**Supplementary Figure 4.11** | Correlation between PCoA axes and traits categorized for herbivorous fishes, assessed using linear models. The different states in each trait are explained in the Supplementary methods. Abbreviations in Tooth morphology: *Bicus* - Bicuspid; *Coni* - Conical; *F\_even* - Fused-even; *F\_cren* - Fused-crenelated; *Weakly\_f* - Weakly fused; *Mul-dent* - Multi-denticulate; *Not\_f* - Not fused. Abbreviations in Diet: *Cya* - Cyanobacteria; *Det* - Detritus; *E+C* - EAM + Coral; *E+M* - EAM + Macroalgae; *EMS* - EAM + Macroalgae + Sponges; *Mac* - Macroalgae; *S+M* - Seagrass + Macroalgae; *Zoo* - Zooplankton; *Spo* - Sponges; *T+D* - Turf + Detritus.



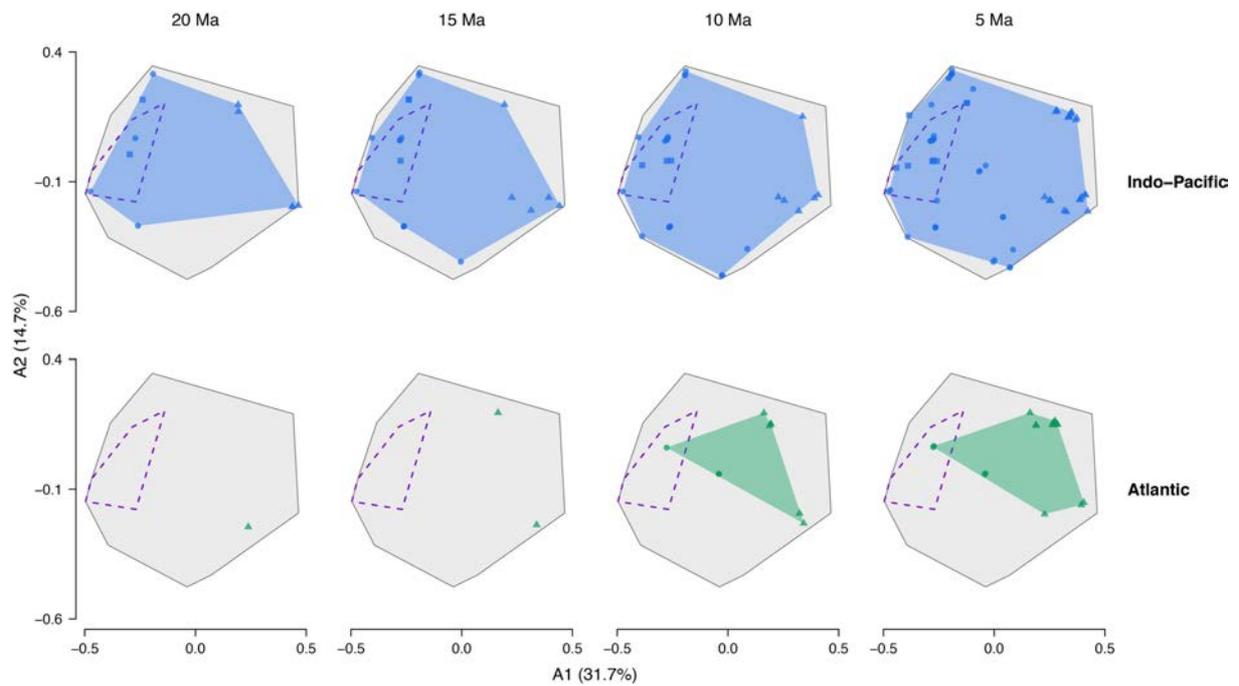
**Supplementary Figure 4.12** | Trait distribution in multidimensional space for axes 1 and 2 derived from the PCoA performed on seven traits related to feeding of extant surgeonfishes, rabbitfishes and parrotfishes present in the phylogenies, and known herbivorous fish fossils from Monte Bolca.



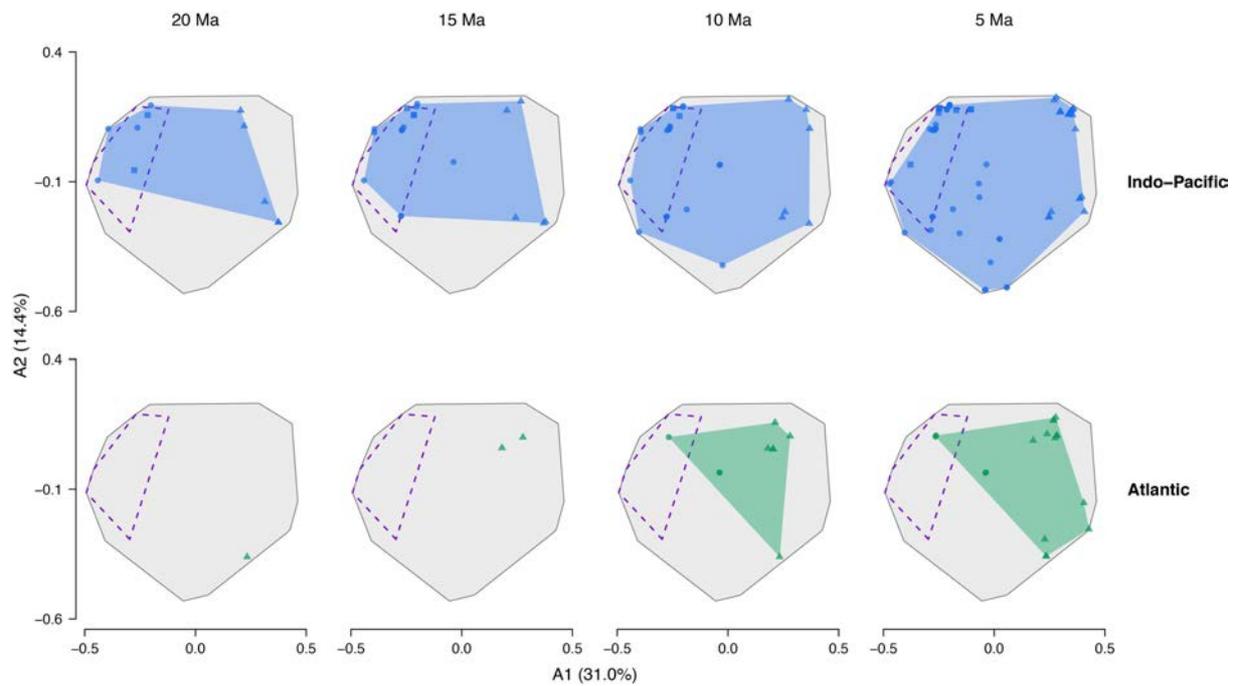
**Supplementary Figure 4.13** | Trait distribution in multidimensional space for axes 3 and 4 derived from the PCoA performed on seven traits related to feeding of extant surgeonfishes, rabbitfishes and parrotfishes present in the phylogenies, and known herbivorous fish fossils from Monte Bolca.



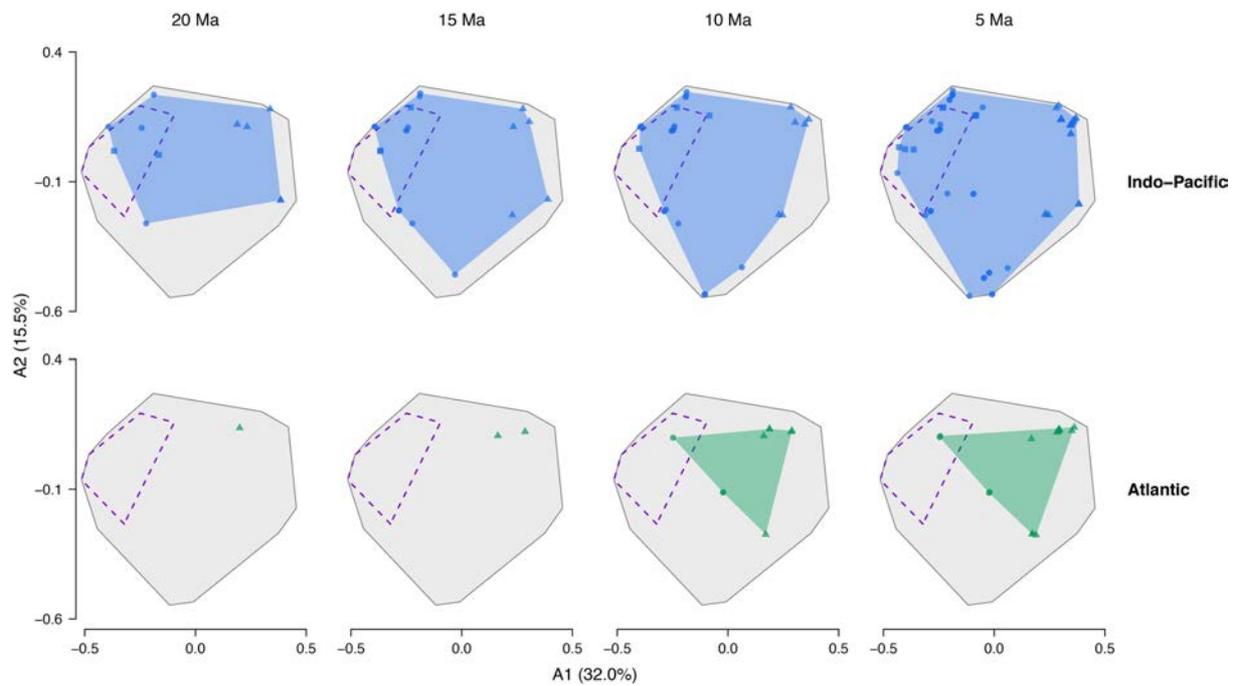
**Supplementary Figure 4.14** | *Overlapping convex hulls of 1000 PCoAs performed on seven traits related to feeding of surgeonfishes, rabbitfishes and parrotfishes. Each polygon results from a sample of three phylogenies (one per herbivorous group) from the posterior distribution derived from the Bayesian inferences (see Supplementary Methods). Each column represents a time-slice (20-5 Ma) in which we assessed the traits through ancestral reconstructions. Convex hulls represent space occupied by Atlantic (green), Indo-Pacific (blue) and Tethys (50 Ma; purple) lineages.*



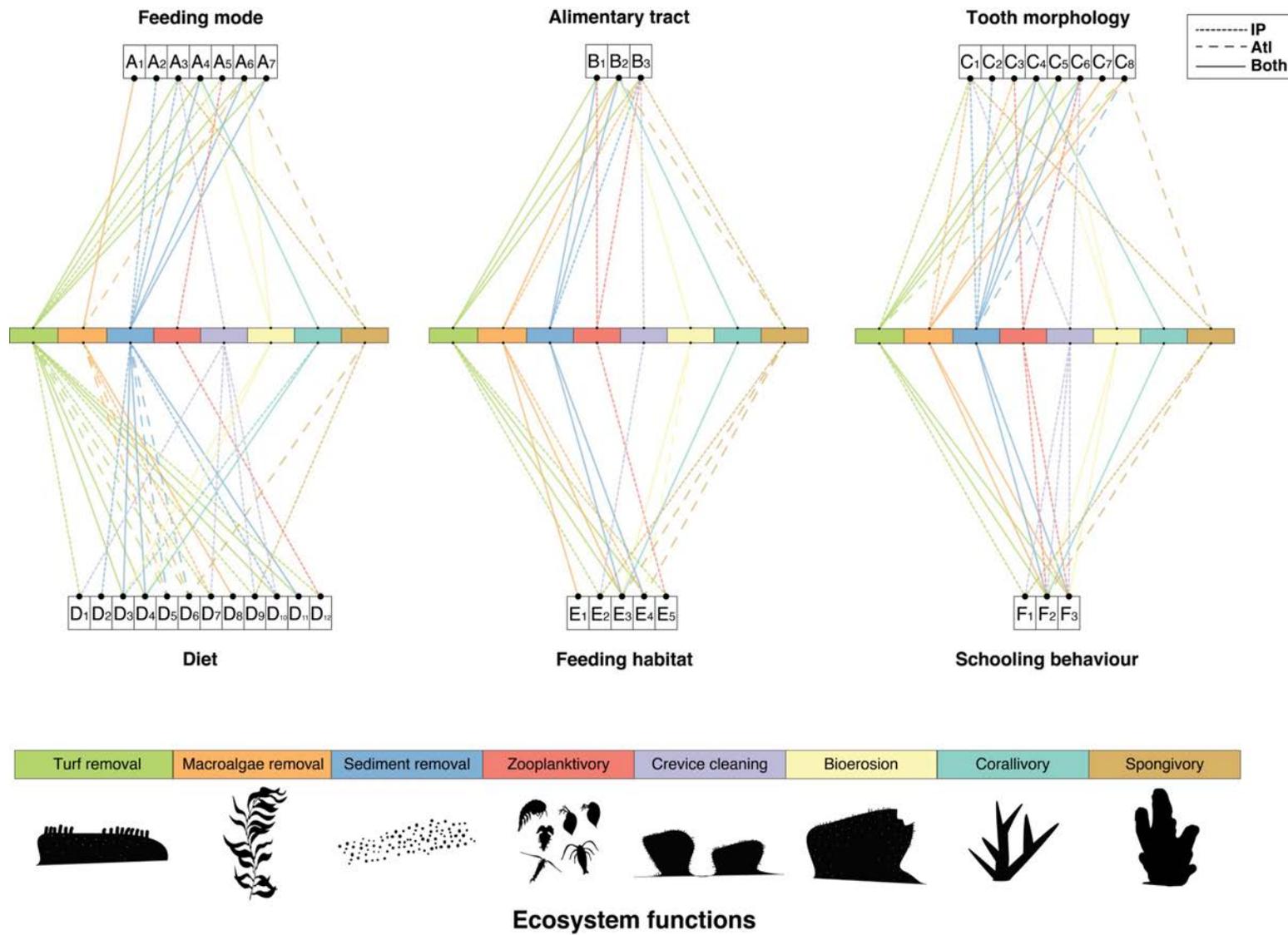
**Supplementary Figure 4.15 |** Multidimensional trait space occupied by surgeonfish (circles), rabbitfish (squares) and parrotfish (triangles) lineages in two biogeographical regions through time. Plots show the two first axes (A1-A2) derived from a PCoA performed on seven traits related to feeding. Each column represents a time-slice (20-5 Ma) in which we assessed the traits through ancestral reconstructions. Nodes that had the posterior probability for the highest state below 0.67 were classified with the second most likely state (see Supplementary Methods). The Extant column represents trait space of extant species. Background grey area shows the total space occupied combining fossils, time-slices and extant species. Convex hulls represent space occupied by Atlantic (green), Indo-Pacific (blue) and Tethys (50 Ma; purple dashed line) lineages. Symbols represent lineages present in each point in time.



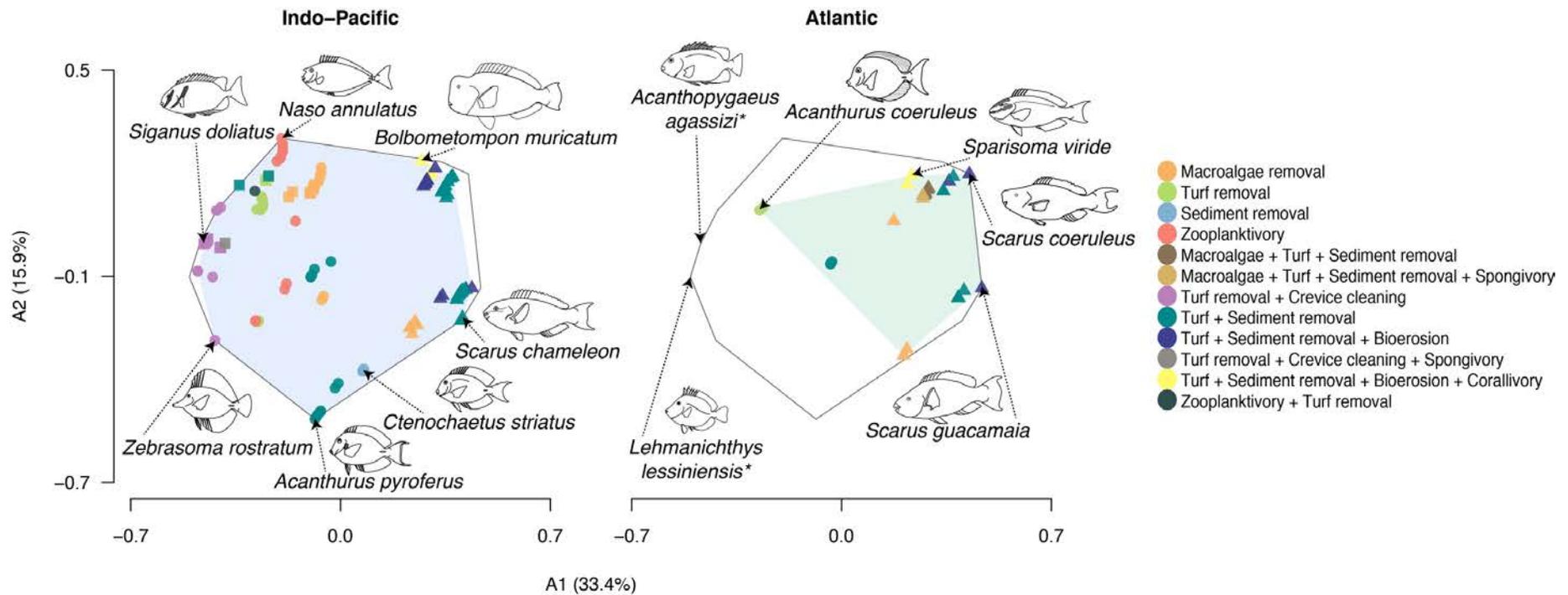
**Supplementary Figure 4.16** | Multidimensional trait space occupied by surgeonfish (circles), rabbitfish (squares) and parrotfish (triangles) lineages in two biogeographical regions through time. Plots show the two first axes (A1-A2) derived from a PCoA performed on seven traits related to feeding. Each column represents a time-slice (20-5 Ma) in which we assessed the traits through ancestral reconstructions. Nodes that had less than 95% of the higher posterior probabilities associated with the best-supported state were classified with the second most likely state (see Supplementary Methods). The Extant column represents trait space of extant species. Background grey area shows the total space occupied combining fossils, time-slices and extant species. Convex hulls represent space occupied by Atlantic (green), Indo-Pacific (blue) and Tethys (50 Ma; purple dashed line) lineages. Symbols represent lineages present in each point in time.



**Supplementary Figure 4.17 |** Multidimensional trait space occupied by surgeonfish (circles), rabbitfish (squares) and parrotfish (triangles) lineages in two biogeographical regions through time. Plots show the two first axes (A1-A2) derived from a PCoA performed on seven traits related to feeding. Each column represents a time-slice (20-5 Ma) in which we assessed the traits through ancestral reconstructions using hidden Markov models (see Supplementary Methods). The Extant column represents trait space of extant species. Background grey area shows the total space occupied combining fossils, time-slices and extant species. Convex hulls represent space occupied by Atlantic (green), Indo-Pacific (blue) and Tethys (50 Ma; purple dashed line) lineages. Symbols represent lineages present in each point in time.



**Supplementary Figure 4.18 (previous page) | Trait states of herbivorous coral reef fishes mapped with their respective ecosystem functions. Trait coding: Feeding mode (A1 - Browser; A2 - Brusher; A3 - Cropper; A4 - Excavator; A5 - Planktivore; A6 - Scraper; A7 - Sucker) ; Alimentary tract (B1 - Gizzard-like; B2 - Pharyngeal mill; B3 - Thin-walled); Tooth morphology (C1 - Bicuspid; C2 - Brush; C3 - Conical; C4 - Fused-crenelated; C5 -Fused-even; C6 - Multi-denticulate; C7 - Not fused; C8 - Weakly fused); Diet (D1 - Cyanobacteria; D2 -Detritus; D3 - EAM; D4 - EAM + Coral; D5 -EAM + Macroalgae; D6 - EAM + Macroalgae + Sponges; D7 - Macroalgae; D8 - Seagrass + Macroalgae; D9 - Sponges; D10 - Turf-algae; D11 -Turf + Detritus; D12 - Zooplankton); Feeding habitat (E1 - off-reef; E2 - reef-concealed; E3 - reef-open; E4 - reef-sand; E5 - water column); Schooling behaviour (F1 - Pairing; F2 - Schooling; F3 - Solitary).**



**Supplementary Figure 4.19** | Multidimensional trait space occupied by extant surgeonfish (circles), rabbitfish (squares) and parrotfish (triangles) species in two biogeographical regions. Plots show the first two axes (A1-A2) derived from a principal coordinate analysis (PCoA) performed on seven traits related to feeding. Points are coloured according to the ecosystem functions performed by each species. Grey contour shows the total space occupied combining fossils, time-slices and extant species. Convex hulls represent space occupied by Indo-Pacific (blue) and Atlantic (green) species. Illustrations show representatives from each biogeographical realm and two fossil species (\*).

## Supplementary Tables

**Supplementary Table 4.1** | Biogeographical model comparison for Siganidae ordered according to lowest AICc. LnL: log-likelihood; d: dispersal; e: extinction; j: founder-speciation event parameter; w: dispersal multiplier matrix power exponential parameter; AICc: Akaike Information Criteria corrected for sample size; AIC\_wt: Akaike Information Criteria weight.

Models	LnL	params	d	e	j	w	AICc	AICc wt
DEC	-67.9	2	0.16	0.042	0	1	140.3	0.62
DEC+J	-67.9	3	0.16	0.042	1.00E-05	1	142.9	0.17
DEC+W	-68.31	3	0.15	0.047	0	0.68	143.7	0.11
DEC+J+W	-67.86	4	0.16	0.042	1.00E-05	1.11	145.6	0.044
DIVALIKE	-71.57	2	0.18	0.046	0	1	147.7	0.016
BAYAREALIKE+J	-70.59	3	0.11	0.061	0.13	1	148.3	0.012
BAYAREALIKE	-72.63	2	0.13	0.083	0	1	149.8	0.0055
DIVALIKE+J	-71.49	3	0.21	0.052	1.00E-05	1	150.1	0.0048
DIVALIKE+W	-71.95	3	0.18	0.057	0	0.76	151	0.003
BAYAREALIKE+J+W	-70.58	4	0.11	0.061	0.13	1.08	151.1	0.0029
DIVALIKE+J+W	-71.55	4	0.2	0.051	1.00E-05	0.93	153	0.0011
BAYAREALIKE+W	-75.38	3	0.14	0.094	0	0.16	157.9	9.70E-05

**Supplementary Table 4.2** | Biogeographical model comparison for Acanthuridae ordered according to lowest AICc. LnL: log-likelihood; d: dispersal; e: extinction; j: founder-speciation event parameter; w: dispersal multiplier matrix power exponential parameter; AICc: Akaike Information Criteria corrected for sample size; AIC\_wt: Akaike Information Criteria weight.

Models	LnL	params	d	e	j	w	AICc	AICc wt
BAYAREALIKE+J+W	-201.9	4	0.11	0.045	0.33	0.63	412.4	0.58
BAYAREALIKE+J	-203.4	3	0.13	0.051	0.35	1	413.1	0.41
DEC+J+W	-206	4	0.22	0.036	0.095	0.57	420.6	0.0097
DEC+J	-211	3	0.26	0.04	0.14	1	428.3	0.0002
DEC	-212.5	2	0.22	0.025	0	1	429.1	0.0001
DIVALIKE+W	-213.7	3	0.3	0.051	0	0.6	433.8	1.30E-05
DIVALIKE+J+W	-213.9	4	0.29	0.05	0.11	0.62	436.3	3.70E-06
DIVALIKE+J	-217.7	3	0.33	0.052	0.036	1	441.7	2.50E-07
DEC+W	-218.2	3	0.15	0.032	0	0.0049	442.7	1.60E-07
DIVALIKE	-223.3	2	0.2	0.032	0	1	450.7	2.80E-09
BAYAREALIKE	-224	2	0.26	0.13	0	1	452.3	1.30E-09
BAYAREALIKE+W	-228	3	0.13	0.099	0	0.17	462.4	8.00E-12

**Supplementary Table 4.3** | Biogeographical model comparison for Scarini ordered according to lowest AICc. LnL: log-likelihood; d: dispersal; e: extinction; j: founder-speciation event parameter; w: dispersal multiplier matrix power exponential parameter; AICc: Akaike Information Criteria corrected for sample size; AIC\_wt: Akaike Information Criteria weight.

Models	LnL	params	d	e	j	w	AICc	AICc_wt
DEC+J+W	-197.5	4	0.091	0.0087	0.06	0.46	403.5	0.48
DEC+J	-199	3	0.098	0.0087	0.057	1	404.3	0.32
DEC	-200.6	2	0.12	0.014	0	1	405.3	0.2
BAYAREALIKE+J+W	-202	4	0.051	0.023	0.19	0.51	412.5	0.0054
BAYAREALIKE+J	-204.3	3	0.056	0.027	0.2	1	415	0.0015
DEC+W	-204.5	3	0.093	0.018	0	0.0014	415.2	0.0014
DIVALIKE+J+W	-206.4	4	0.11	0.0075	0.065	0.48	421.3	6.40E-05
DIVALIKE+J	-209.6	3	0.12	0.0078	0.079	1	425.5	7.80E-06
DIVALIKE+W	-212.8	3	0.092	0.0097	0	0.05	431.9	3.30E-07
DIVALIKE	-215.7	2	0.094	0.018	0	1	435.5	5.30E-08
BAYAREALIKE	-224.9	2	0.13	0.12	0	1	454	5.10E-12
BAYAREALIKE+W	-232.9	3	0.059	0.076	0	0.012	472.1	6.20E-16

**Supplementary Table 4.4** | Genetic accession numbers used to build the Siganiidae phylogeny.

Species	16s	Cytb	ITS1
<i>Siganus argenteus</i>	DQ898097	AB276803	AB276971
<i>Siganus canaliculatus</i>	EF210164	AB276851	AB276988
<i>Siganus corallinus</i>	DQ898105	AB276936	AB277043
<i>Siganus doliatus</i>	DQ898107	AB276957	AB277054
<i>Siganus fuscescens</i>	DQ898101	AB276831	AB276981
<i>Siganus guttatus</i>	DQ898108	AB276913	AB277025
<i>Siganus javus</i>	DQ898109	AB276852	AB276995
<i>Siganus lineatus</i>		AB276918	AB277034
<i>Siganus luridus</i>	DQ898111	AM949027	
<i>Siganus magnificus</i>		AB276882	AB277005
<i>Siganus puellus</i>	DQ898112	AB276867	AB276997
<i>Siganus punctatissimus</i>		AB276872	AB276999
<i>Siganus punctatus</i>	DQ898113	AB276879	AB277001
<i>Siganus randalli</i>		EF210188	
<i>Siganus rivulatus</i>	JX026505	EU601390	
<i>Siganus spinus</i>	DQ898117	AB276818	AB276975
<i>Siganus unimaculatus</i>	EF210170	AB276889	AB277010
<i>Siganus uspi</i>		AB276881	AB277003
<i>Siganus vermiculatus</i>	DQ898118	AB276904	AB277021
<i>Siganus virgatus</i>	EF210171	AB276949	
<i>Siganus vulpinus</i>	DQ898119	AB276902	AB277051
<i>Siganus woodlandi</i>	DQ898120	AB276793	AB276968
<i>Siganus stellatus</i>	DQ532960	KT953191	
<i>Siganus sutor</i>		MF326167	
<i>Prionurus scalprum</i>	AY264591	AB276963	AB277066
<i>Zanclus cornutus</i>	AY057282	AB276965	AB375558

**Supplementary Table 4.5 | Genetic accession numbers used to build the Acanthuridae phylogeny.**

Species	Cox1	Cytb	ENC1	ETS2	myh6	plagl2	Rag1	Rh	zic1
<i>Acanthurus achilles</i>	KC623654	KC623692	KC623730		KC623763	KC623798	KC623828	KC623863	KC623902
<i>Acanthurus auranticavus</i>	KC623655	KC623693	KC623731		KC623764	KC623799	KC623829	KC623864	KC623903
<i>Acanthurus bahianus</i>		FJ905179			JX189759	JX189441	JX189919		JX189146
<i>Acanthurus bariene</i>	KC623657	KC623695	KC623733		KC623766	KC623801	KC623831	KC623866	KC623905
<i>Acanthurus blochii</i>	HM034180	AY264632	KC623734	AY264685	KC623767	KC623802	KC623832	KC623867	KC623906
<i>Acanthurus chirurgus</i>	KC623658	KC623696	KC623735		KC623768	KC623803	KC623833	KC623868	KC623907
<i>Acanthurus coeruleus</i>	KC623659	KC623697	KC623736		KC623769	KC623804	KC623834	KC623869	KC623908
<i>Acanthurus dussumieri</i>		AY264633		AY264686					
<i>Acanthurus gahhm</i>		KT953166		KT953202					
<i>Acanthurus guttatus</i>	KC623660	KC623698	EF539241		EF536294	EF536256	EF530094	KC623870	EF533917
<i>Acanthurus japonicus</i>	KC623661	KC623699	KC623737		KC623770	KC623805	KC623835	KC623871	KC623909
<i>Acanthurus leucocheilus</i>	KC623662	KC623700				KC623806	KC623836	KC623872	KC623910
<i>Acanthurus leucosternon</i>	KC623663	EU136032	KC623738		KC623771	KC623807	KC623837	KC623873	KC623911
<i>Acanthurus lineatus</i>	KC623664	EU273284	KC623739		KC623772	KC623808	KC623838	KC623874	KC623912
<i>Acanthurus maculiceps</i>	KY683548								
<i>Acanthurus mata</i>	KC623665	KC623701	KC623740		KC623773	KC623809	KC623839	KC623875	KC623913
<i>Acanthurus monroviae</i>	KC623666	KC623702	KC623741		KC623774	KC623810	KC623840	KC623876	KC623914
<i>Acanthurus nigricans</i>	KC623667	KC623703	KC623742	AY264687	KC623775	KC623811	KC623841	KC623877	KC623915
<i>Acanthurus nigricauda</i>	HM034189	KC623704	KC623743		KC623776	KC623812	KC623842	KC623878	KC623916
<i>Acanthurus nigrofuscus</i>	KC623668	KC623705	KC623744	AY264688	KC623777	KC623813	KC623843	KC623879	KC623917
<i>Acanthurus nigroris</i>		KC623706							
<i>Acanthurus nubilus</i>	HM034193	AY264636		AY264689					
<i>Acanthurus olivaceus</i>	KC623669	KC623707	KC623745	AY264690	KC623778	KC623814	KC623844	KC623880	KC623918
<i>Acanthurus polyzona</i>	JQ349664								
<i>Acanthurus pyroferus</i>	KC623670	KC623708	KC623746	AY264691	KC623779	KC623815	KC623845	KC623881	KC623919
<i>Acanthurus reversus</i>	KY683549								
<i>Acanthurus sohal</i>	MF123727	KT953172		KT953208					
<i>Acanthurus tennentii</i>	KC623671	KC623709	KC623747		KC623780	KC623816	KC623846	KC623882	KC623920
<i>Acanthurus thompsoni</i>	KC623672	KC623710	KC623748		KC623781	KC623817	KC623847	KC623883	KC623921
<i>Acanthurus tractus</i>	KC623673	KC623694	KC623732		KC623765	KC623800	KC623830	KC623865	KC623904
<i>Acanthurus triostegus</i>	KC623674	KC623711	EF539242	AY264692	EF536295	EF536257	EF530095	KC623884	EF533918
<i>Acanthurus xanthopterus</i>	KC623674	KC623712	KC623749	AY264693	KC623782	KC623818	KC623848	KC623885	KC623922
<i>Ctenochaetus binotatus</i>		KC623713	KC623750	AY264694	KC623783		KC623849	KC623886	KC623923
<i>Ctenochaetus flavicauda</i>	HM034209								
<i>Ctenochaetus striatus</i>	KC623675	KC623714	EF539243	AY264695	EF536296	EF536258	EF530096		EF533919
<i>Ctenochaetus strigosus</i>	KC623676	FJ376811	KC623751		KC623784	KC623819	KC623850	KC623887	KC623924
<i>Ctenochaetus tominiensis</i>	KC623677	KC623715	KC623752		KC623785	KC623820	KC623851		KC623925
<i>Ctenochaetus truncatus</i>	KC623678	KC623716	KC623753		KC623786	KC623821	KC623852	KC623888	KC623926
<i>Naso annulatus</i>	HM034155	AY264643		AY264696					
<i>Naso brachycentron</i>		AY264644		AY264697					
<i>Naso brevirostris</i>	KC623679	KC623717	EF539240	AY264698	EF536293	EF536255	EF530093	KC623889	EF533916
<i>Naso caeruleacauda</i>		AY264646		AY264699					
<i>Naso caesius</i>		AY264647		AY264700					
<i>Naso elegans</i>	KC623680	KC623718	KC623754	AY264701			KC623853	KC623890	
<i>Naso fageni</i>		AY264649		AY264702					
<i>Naso hexacanthus</i>		AY264650		AY264703					
<i>Naso lituratus</i>	HM034247	AY264651	EF539239	AY264704	EF536292	EF536254	EF530092	EU637984	EF533915
<i>Naso lopezi</i>	AP009163	AY264652		AY264705					
<i>Naso maculatus</i>		AY264653		AY264706					
<i>Naso mcdadei</i>		AY264654		AY264707					
<i>Naso minor</i>		AY264655		AY264708					
<i>Naso reticulatus</i>		AY264656		AY264709					
<i>Naso thynnoides</i>		AY264657		AY264710					
<i>Naso tonganus</i>		AY264658		AY264711					
<i>Naso tuberosus</i>		AY264659		AY264712					
<i>Naso unicornis</i>	KC623681	KC623719	KC623755	AY264713	KC623787	KC623822	KC623854	KC623891	KC623927
<i>Naso vlamingii</i>	KC623682	KC623720	KC623756	AY264714	KC623788	KC623823	KC623855	KC623892	KC623928

Appendix C - Supplementary Material to Chapter 4

Species	Cox1	Cytb	ENC1	ETS2	myh6	plagl2	Rag1	Rh	zic1
<i>Paracanthurus hepatus</i>	KC623683	KC623721	EF539244	AY264680	EF536297	EF536259	EF530097	KC623893	EF533920
<i>Prionurus biafraensis</i>	KC623684	KC623722	KC623757		KC623789	KC623824	KC623856	KC623894	KP280517
<i>Prionurus chrysurus</i>	KP280488						KP280510		KP280518
<i>Prionurus laticlavus</i>	KC623685	KC623723	KC623758		KC623790	KC623825	KC623857	KC623895	KP280519
<i>Prionurus maculatus</i>	KC623687	KC623725			KC623792			KC623897	KC623932
<i>Prionurus microlepidotus</i>		KC623726		AY264681	KC623793	EF536260	KC623859		EF533921
<i>Prionurus punctatus</i>	KC623686	KC623724	KC623759		KC623791		KC623858	KC623896	KC623931
<i>Prionurus scalprum</i>		AY264629		AY264682					
<i>Zebrosoma desjardini</i>	KC623688	KC623727	KC623760		KC623794		KC623860	KC623898	KC623934
<i>Zebrosoma flavescens</i>	KC623689	AP006032	KC623761		KC623795	KC623827	KC623861	KC623899	KC623935
<i>Zebrosoma gemmatum</i>	JF494799								
<i>Zebrosoma rostratum</i>	HM034282						KF141385		KF140659
<i>Zebrosoma scopas</i>	KC623690	KC623728	EF539238	AY264683	KC623796	EF536253	AY308776	KC623900	KC623936
<i>Zebrosoma velifer</i>	KC623691	KC623729	KC623762	AY264684	KC623797		KC623862	KC623901	KC623937
<i>Zebrosoma xanthurum</i>	MF124084	KT953199		KT953235					
<i>Zanclus cornutus</i>	KC623652	AY264626	EF539247	AY264679	EF536300	EF536262	EF530100		EF533923
<i>Luvuru imperialis</i>	KC623653	AY264625	EF539246	AY264678	EF536299	EF536261	EF530099	EU637975	EF533922

Supplementary Table 4.6 | Genetic accession numbers used to build the Scarini phylogeny.

Species	12s	16s	Bmp4	Control region	Cox1	Cytb	Dlx2	Otx1	Rag2	S711	Tmo-4C4
<i>Bolbometopon muricatum</i>	EU601178	AY081091	EU601456		NC033901	NC033901	EU601506	EU601406	EU601307	JX026592	AY081108
<i>Calotomus carolinus</i>	EU601179	AY081092	EU601457		JQ349815	EU601358	EU601507	EU601407	EU601308		AY081109
<i>Calotomus spinidens</i>	EU601180	EU601228	EU601458		KJ202134	EU601359	EU601508	EU601408	EU601309		EU601265
<i>Calotomus zonarchus</i>					DQ521016						
<i>Cetoscarus bicolor</i>	EU601181	EU601229	EU601459		JQ349874	EU601360	EU601509	EU601409	EU601310	JX026593	AY081105
<i>Chlorurus atrilunula</i>		JX026457		JX026525						JX026596	
<i>Chlorurus bleekeri</i>	EU601182	EU601230	EU601460	JX026526		EU601361	EU601510	EU601410	EU601311	JX026597	EU601267
<i>Chlorurus bowersi</i>	EU601183	EU601231	EU601461	JX026527		EU601362	EU601511	EU601411	EU601312	JX026598	EU601268
<i>Chlorurus capistratoides</i>	EU601184	EU601232	EU601462	JX026528		EU601363	EU601512	EU601412	EU601313	JX026599	EU601269
<i>Chlorurus cyanescens</i>		JX026461		JX026529						JX026600	
<i>Chlorurus enneacanthus</i>		JX026462		JX026530						JX026601	
<i>Chlorurus frontalis</i>	AB974582	JX026463		JX026531	JQ431620	LC068812				JX026602	
<i>Chlorurus gibbus</i>		JX026464		JX026532							
<i>Chlorurus japonensis</i>	EU601209	EU601250	EU601487	JX026533		EU601388	EU601537	EU601437	EU601338	JX026603	EU601288
<i>Chlorurus microrhinos</i>	EU601185	EU601233	EU601463	JX026534		EU601364	EU601513	EU601413	EU601314	JX026604	EU601270
<i>Chlorurus oedema</i>	EU601186	EU601234	EU601464	JX026535		EU601365	EU601514	EU601414	EU601315	JX026605	AY081107
<i>Chlorurus perspicillatus</i>		JX026468		JX026536		KF809191				JX026606	
<i>Chlorurus rhakoura</i>		JX026469		JX026537						JX026607	
<i>Chlorurus sordidus</i>			EU601465	JX026539	EU601367	NC006355		EU601415		JX026609	AY081106
<i>Chlorurus spilurus</i>		JX026470		JX026538						JX026608	
<i>Chlorurus strongylocephalus</i>		JX026472		JX026540						JX026610	
<i>Cryptotomus roseus</i>	AY279592	AY279695	EU601466		JQ839422	EU601367	EU601516	EU601416	AY279901		AY279798
<i>Hipposcarus harid</i>		JX026455		JX026523						JX026594	
<i>Hipposcarus longiceps</i>	EU601189	AY081093	EU601467	JX026524	KF929973	EU885924	EU601517	EU885926	EU601318	JX026595	AY081110
<i>Leptoscarus vaigiensis</i>	EU601190	AY081094	EU601468		FJ237788	EU601369		EU601418	EU601319		AY081111
<i>Nicholsina denticulata</i>	U95761	U95762				DQ457021					
<i>Nicholsina usta</i>			EU601469			DQ457023		EU601419	AY279933		AY081112
<i>Scarus altipinnis</i>	EU601192	EU601237	EU601470	JX026541	JQ432095	EU601371	EU601520	EU601420	EU601321	JX026611	EU601273
<i>Scarus arabicus</i>		JX026474		JX026542						JX026612	
<i>Scarus chameleon</i>	EU601193	EU601238	EU601471	JX026613	FJ237917	EU601372	EU601521	EU601421	EU601322	JX026613	EU601274
<i>Scarus coelestinus</i>	EU601194	AY081084	EU601472	JX026544		EU601373	EU601522	EU601422	EU601323	JX026614	AY081101
<i>Scarus coeruleus</i>		JX026476		JX026545						JX026615	
<i>Scarus collana</i>			KY815801	JX026546						JX026616	KY815559
<i>Scarus compressus</i>		JX026478		JX026547						JX026617	
<i>Scarus dimidiatus</i>	AY279642	AY279745	EU601473	JX026548		EU601374	EU601523	EU601423	AY279951	JX026618	AY279848
<i>Scarus dubius</i>		JX026480		JX026549		KF809216				JX026619	
<i>Scarus falcipinnis</i>		JX026481		JX026550						JX026620	
<i>Scarus ferrugineus</i>			KY815802	JX026551						JX026621	KU862922
<i>Scarus festivus</i>	EU601196	EU601239	EU601474	JX026552		EU601375	EU601524	EU601424	EU601325	JX026622	EU601276
<i>Scarus flavipectoralis</i>	EU601197	AY081086	EU601475	JX026553	KP194579	EU601376	EU601525	EU601425	EU601326	JX026623	AY081103
<i>Scarus forsteni</i>	EU601198	EU601240	EU601476	JX026554	NC011928	NC011928	EU601526	EU601426	EU601327	JX026624	EU601278
<i>Scarus frenatus</i>	AY279643	AY279746	EU601477	JX026555	KP194827	EU601378	EU601527	EU601427	AY279952	JX026625	AY081104
<i>Scarus fuscopurpureus</i>		JX026487		JX026556						JX026626	KU862923
<i>Scarus ghabban</i>	EU601200	EU601241	EU601478	JX026558	NC011599	NC011599	EU601528	EU601428	EU601329	JX026628	KU862921
<i>Scarus globiceps</i>	EU601201	EU601242	EU601479	JX026561	JQ432103	EU601380	EU601529	EU601429	EU601330	JX026631	EU601280
<i>Scarus guacamaia</i>	EU601202	EU601243	EU601480	JX026562	JQ843039	EU601381	EU601530	EU601430	EU601331	JX026632	AY081102
<i>Scarus hoefleri</i>	AY141393	JX026493		JX026563						JX026633	

Appendix C - Supplementary Material to Chapter 4

Species	12s	16s	Bmp4	Control region	Cox1	Cytb	Dlx2	Otx1	Rag2	S711	Tmo-4C4
<i>Scarus hypselopterus</i>	EU601204	EU601245	EU601482			EU601383	EU601532	EU601432	EU601333		EU601283
<i>Scarus iseri</i>	EU601203	EU601244	EU601481	JX026564	JQ842671	EU601382	EU601531	EU601431	EU601332	JX026634	EU601282
<i>Scarus koputea</i>		JX026495		JX026565						JX026635	
<i>Scarus longipinnis</i>		JX026496		JX026566						JX026636	
<i>Scarus niger</i>	EU601205	JX026497	EU601483	JX026567	KP194654	KT953190	EU601533	EU601433	EU601334	JX026637	KU862926
<i>Scarus oviceps</i>	EU601206	EU601247	EU601484	JX026568	JQ432106	EU601385	EU601534	EU601434	EU601335	JX026638	EU601285
<i>Scarus ovifrons</i>	LC092086	JX026499		JX026569		LC068811				JX026639	
<i>Scarus perrico</i>		JX026500		JX026570						JX026640	
<i>Scarus persicus</i>		JX026501		JX026571						JX026641	
<i>Scarus prasiognathos</i>	EU601207	EU601248	EU601485	JX026572		EU601386	EU601535	EU601435	EU601336	JX026642	EU601286
<i>Scarus psittacus</i>	EU601208	EU601249	EU601486	JX026573	JQ432113	EU601387	EU601536	EU601436	EU601337	JX026643	KU862925
<i>Scarus quoyi</i>	EU601210	EU601251	EU601488	JX026574	KF930376	EU601389	EU601538	EU601438	EU601339	JX026644	EU601289
<i>Scarus rivulatus</i>	EU601211	EU601252	EU601489	JX026575		EU601390	EU601539	EU601439	EU601340	JX026645	EU601290
<i>Scarus rubroviolaceus</i>	EU601212	EU601253	EU601490	JX026577	NC011343	NC011343	EU601540	EU601440	EU601341	JX026647	EU601291
<i>Scarus russelii</i>		JX026510		JX026580	KF489744					JX026650	KU862927
<i>Scarus scaber</i>		JX026511		JX026581	JQ350333					JX026651	
<i>Scarus schlegeli</i>	EU601213	EU601254	EU601491	JX026582	NC011936	NC011936	EU601541	EU601441	EU601342	JX026652	EU601292
<i>Scarus spinus</i>	EU601214	EU601255	EU601492	JX026583	KP193990	EU601393	EU601542	EU601442	EU601343	JX026653	EU601293
<i>Scarus taeniopterus</i>	EU601215	EU601256	EU601493	JX026584	JQ842301	EU601394	EU601543	EU601443	EU601344	JX026654	EU601294
<i>Scarus tricolor</i>	EU601216	EU601257	EU601494	JX026585	JQ350335	EU601395	EU601544	EU601444	EU601345	JX026655	EU601295
<i>Scarus trispinosus</i>		JX026516		JX026586						JX026656	
<i>Scarus vetula</i>				JX026587	FJ584083					JX026657	KY815561
<i>Scarus viridifucatus</i>		JX026518		JX026588						JX026658	
<i>Scarus xanthopleura</i>		JX026519		JX026589		LC068816				JX026659	
<i>Scarus zelindae</i>		JX026520		JX026590						JX026660	
<i>Scarus zufar</i>		JX026521		JX026591						JX026661	
<i>Sparisoma amplum</i>						DQ457024					
<i>Sparisoma atomarium</i>	U95767	U95768			JQ840703	DQ457029					
<i>Sparisoma aurofrenatum</i>	EU601217	U95766	EU601495		JQ839898	DQ457031	EU601545	EU601445	EU601346		AY081099
<i>Sparisoma axillare</i>		KC526954				DQ457034					
<i>Sparisoma choati</i>	AF114483	AF115312				DQ457036					
<i>Sparisoma chrysopterum</i>	AY279645	AY279748	EU601496		GU225439	DQ457033	EU601546	EU601446	AY279954		AY081100
<i>Sparisoma cretense</i>	U95777	AF517578			KC501534	DQ457026					
<i>Sparisoma frondosum</i>	JX645341	JX645342				DQ457032					
<i>Sparisoma radians</i>	U95771	U95772				DQ457028					
<i>Sparisoma rocha</i>	GU985521	GU985520									
<i>Sparisoma rubripinne</i>	KY815337	U95774			GU225443	DQ457035					
<i>Sparisoma strigatum</i>						DQ457027					
<i>Sparisoma tuiupiranga</i>						DQ457030					
<i>Sparisoma viride</i>	EU601219	AY081081	EU601497	AF370453	JQ841013	DQ457025	EU601547	EU601447	EU601348		AY081098
<i>Achoerodus viridis</i>	AY279574	AY279677	KY815674		EF609278		KY816077	KY815925	AY279883		AY279780
<i>Cheilinus abudjubbe</i>						KY815828			KY815584		KY815482
<i>Cheilinus chlorourus</i>			KY815691		KP194859	KY815829	KY816088	KY815941	KY815585		KY815483
<i>Cheilinus fasciatus</i>	AY279580	AY279683	KY815692		KP194612	KY815830	KY816089	KY815942	AY279889		AY279786
<i>Cheilinus lunulatus</i>						KY815831			KY815586		KY815484
<i>Cheilinus oxycephalus</i>	AY279581	AY279684	KY815693		KF929730	KY815832		KY815943	AY279890		AY279787
<i>Cheilinus trilobatus</i>			KY815694		KP194018	KY815833	KY816090	KY815944	KY815587		KY815485
<i>Cheilinus undulatus</i>	AY279582	AY279685	EU601498		KM461717	KM461717	EU601548	EU601448	AY279891		AY279788
<i>Lachnolaimus maximus</i>	AY279618	AY279721	EU601503		GU224537	EU601404	EU601553	EU601453	AY279927		AY279824

Supplementary Table 4.7 | Best gene partitioning scheme resulting from PartitionFinder for Siganidae.

Subset	Model	#sites	Genes
1	SYM+I+G	512	16s
2	GTR+I+G+X	1140	Cytb
3	TRN+G+X	1349	ITS1

**Supplementary Table 4.8** | Best gene partitioning scheme resulting from PartitionFinder for Acanthuridae.

Subset	Model	#sites	Genes
1	GTR+I+G+X	651	Cox1
2	GTR+I+G+X	1047	Cytb
3	TRN+I+G+X	1498	ENC1, Myh6
4	HKY+G+X	422	ETS2
5	GTR+G+X	567	Plagl2
6	SYM+I+G	1556	Rag1
7	GTR+I+G+X	798	Rh
8	HKY+I+X	690	Zic1

**Supplementary Table 4.9** | Best gene partitioning scheme resulting from PartitionFinder for Scarini.

Subset	Model	#sites	Genes
1	GTR+I+G+X	935	12s
2	GTR+I+G+X	503	16s
3	GTR+I+G+X	1005	Cox1, Cytb
4	HKY+I+G+X	433	Control region
5	HKY+I+G+X	480	Bmp4
6	HKY+G+X	516	Dlx2
7	GTR+I+G+X	669	Otx1
8	SYM+I+G	714	Rag2
9	GTR+G+X	618	S7I1
10	GTR+I+G+X	423	Tmo-4C4

**Supplementary Table 4.10** | Summary of BayesTraits model results for discrete traits for Siganidae. Posterior Modal values (lower; upper - Highest Density Interval) of Lh: likelihood; q: transition rates between states; #Br: number of branch lengths scaled using variable rates; #Nd: number of nodes scaled using variable rates. Trait states: Diet (0 - Cyanobacteria; 1 - Macroalgae; 2 - Sponges; 3 - Turf algae); Feeding Mode (0 - Browser; 1 - Cropper); Feeding Habitat (0 - off reef; 1 - reef-concealed; 2 - reef-open; 3 - reef-sand); Schooling behaviour (0 - Pairing; 1 - Schooling).

Traits	Lh	q01	q02	q03	q10	q12	q13	q20	q21	q23	q30	q31	q32	#Br	#Nd
<b>Diet</b>	-21.68 (-24.65; -17.97)	0.09 (0.006; 0.09)	0.07 (0.006; 0.09)	0.09 (0.01; 0.09)	0.01 (3e-5; 0.08)	0.01 (1e-4; 0.08)	0.09 (0.008; 0.09)	0.08 (0.006; 0.09)	0.09 (0.007; 0.09)	0.08 (0.01; 0.09)	0.02 (1e-4; 0.07)	0.05 (0.01; 0.09)	0.01 (1e-4; 0.08)	2 (0; 4)	0 (0; 2)
<b>Feeding Mode</b>	-13.02 (-14.38; -9.11)	0.08 (0.02; 0.10)	-	-	0.04 (0.01; 0.10)	-	-	-	-	-	-	-	-	2 (0; 5)	1 (0; 3)
<b>Feeding Habitat</b>	-15.10 (-20.82; -12.54)	0.02 (9e-4; 0.02)	0.02 (0.002; 0.02)	0.02 (0.001; 0.02)	0.01 (9e-6; 0.02)	0.02 (0.002; 0.02)	0.005 (2e-6; 0.02)	0.004 (4e-4; 0.02)	0.006 (2e-4; 0.02)	0.003 (1e-5; 0.02)	0.02 (0.003; 0.02)	0.01 (8e-4; 0.02)	0.02 (0.001; 0.02)	1 (0; 4)	1 (0; 3)
<b>Schooling</b>	-8.59 (-56.40; -41.87)	0.07 (9e-4; 0.009)	-	-	0.02 (1e-6; 0.005)	-	-	-	-	-	-	-	-	1 (2; 11)	0 (2; 8)

**Supplementary Table 4.11** | Summary of BayesTraits model results for discrete traits for Acanthuridae. Posterior Modal values (lower; upper - Highest Density Interval) of Lh: likelihood; q: transition rates between states; #Br: number of branch lengths scaled using variable rates; #Nd: number of nodes scaled using variable rates. Trait states: Diet (0 - Detritus; 1 - Macroalgae; 2 - Turf algae; 3 - Turf+Detritus; 4 - Zooplankton); Feeding Mode (0 - Browser; 1 - Brusher; 2 - Cropper; 3 - Planktivore; 4 - Sucker); Feeding Habitat (0 - reef-concealed; 1 - reef-open; 2 - reef-sand; 3 - water column); Tooth morphology (0 - Brush; 1 - Conical; 2 - Multi-denticulate); Alimentary tract (0 - Gizzard-like; 1 - Thin walled); Schooling behaviour (0 - Pairing; 1 - Schooling; 2 - Solitary).

Traits	Lh	q01	q02	q03	q04	q10	q12	q13	q14	q20	q21	q23	q24
<b>Diet</b>	-48.67 (-54.50; -40.55)	1e-5 (-40.55; 9e-4)	1e-5 (1e-8; 9e-4)	8e-4 (1e-8; 9e-4)	8e-4 (1e-8; 9e-4)	1e-5 (1e-8; 9e-4)	1e-5 (1e-8; 9e-4)	1e-5 (1e-8; 9e-4)	8e-4 (1e-8; 9e-4)	1e-5 (1e-8; 9e-4)	8e-4 (1e-8; 9e-4)	8e-4 (1e-8; 9e-4)	9e-4 (2e-4; 0.001)
<b>Feeding Mode</b>	-42.93 (-50.26; -35.12)	1e-5 (1e-8; 9e-4)	1e-5 (1e-8; 9e-4)	8e-4 (1e-8; 9e-4)	1e-5 (1e-8; 9e-4)	1e-5 (1e-8; 9e-4)	1e-5 (1e-8; 9e-4)	8e-4 (1e-8; 9e-4)	8e-4 (1e-8; 9e-4)	1e-5 (1e-8; 9e-4)	8e-4 (1e-8; 9e-4)	9e-4 (3e-4; 9e-4)	8e-4 (1e-8; 9e-4)
<b>Feeding Habitat</b>	-46.13 (-52.53; -39.84)	5e-4 (1e-8; 9e-4)	3e-4 (1e-8; 9e-4)	8e-4 (7e-5; 9e-4)	-	9e-5 (1e-8; 8e-4)	8e-4 (1e-4; 0.001)	8e-4 (1e-4; 9e-4)	-	9e-5 (1e-8; 9e-4)	7e-4 (5e-5; 9e-4)	8e-4 (8e-5; 9e-4)	-
<b>Tooth Morphology</b>	-11.71 (-16.71; -8.44)	1e-4 (6e-6; 9e-4)	7e-4 (7e-5; 9e-4)	-	-	3e-4 (1e-5; 9e-4)	8e-4 (6e-5; 9e-4)	-	-	8e-4 (1e-4; 0.001)	2e-4 (6e-6; 9e-4)	-	-
<b>Alimentary tract</b>	-10.15 (-13.62; -5.64)	5e-4 (6e-5; 9e-4)	-	-	-	7e-4 (1e-4; 0.001)	-	-	-	-	-	-	-
<b>Schooling</b>	-48.26 (-56.40; -41.87)	0.008 (9e-4; 0.009)	0.005 (1e-4; 0.009)	-	-	7e-4 (1e-6; 0.005)	0.008 (0.001; 0.009)	-	-	8e-8 (1e-6; 0.007)	0.009 (0.003; 0.009)	-	-

**Supplementary Table 4.11** | *Continuation.*

<b>Traits</b>	<b>q30</b>	<b>q31</b>	<b>q32</b>	<b>q34</b>	<b>q40</b>	<b>q41</b>	<b>q42</b>	<b>q43</b>	<b>#Br</b>	<b>#Nd</b>
<b>Diet</b>	8e-4 (1e-8; 9e-4)	1e-5 (1e-8; 9e-4)	1e-5 (1e-8; 9e-4)	8e-4 (1e-8; 9e-4)	1e-5 (1e-8; 9e-4)	8e-4 (3e-4; 0.001)	1e-5 (0; 9e-4)	1e-5 (1e-8; 9e-4)	7 (1; 10)	4 (2; 7)
<b>Feeding Mode</b>	8e-4 (3e-4; 9e-4)	1e-5 (1e-8; 9e-4)	1e-5 (1e-8; 9e-4)	1e-5 (1e-8; 9e-4)	1e-5 (1e-8; 9e-4)	8e-4 (1e-8; 9e-4)	1e-5 (1e-8; 9e-4)	8e-4 (1e-8; 9e-4)	5 (2; 11)	4 (2; 8)
<b>Feeding Habitat</b>	2e-4 (2e-6; 9e-4)	9e-4 (4e-4; 0.001)	1e-4 (7e-6; 9e-4)	-	-	-	-	-	6 (1; 10)	4 (2; 7)
<b>Tooth Morphology</b>	-	-	-	-	-	-	-	-	5 (1; 10)	3 (1; 6)
<b>Alimentary tract</b>	-	-	-	-	-	-	-	-	5 (2; 10)	3 (0; 6)
<b>Schooling</b>	-	-	-	-	-	-	-	-	5 (2; 11)	5 (2; 8)

**Supplementary Table 4.12** | Summary of BayesTraits model results for discrete traits for Scarini. Posterior Modal values (lower; upper - Highest Density Interval) of Lh: likelihood; q: transition rates between states; #Br: number of branch lengths scaled using variable rates; #Nd: number of nodes scaled using variable rates. Trait states: Diet (0 - EAM; 1 - EAM+Coral; 2 - EAM+Macroalgae; 3 - EAM+Macroalgae+Sponges; 4 - Macroalgae Seagrass+Macroalgae); Feeding Mode (0 - Browser; 1 - Excavator; 2 - Scraper); Feeding Habitat (0 - off reef; 1 - reef-open; 2 - reef-sand); Tooth morphology (0 - Fused-crenellated; 1 - Fused-even; 2 - Not-fused; 3 - Weakly-fused); Schooling behaviour (0 - Schooling; 1 - Solitary).

Traits	Lh	q01	q02	q03	q04	q05	q10	q12	q13	q14	q15	q20	q21	q23	q24	q25
<b>Diet</b>	-35.19 (-40.93; -30.20)	6e-04 (1e-8; 0.009)	2e-04 (1e-8; 0.008)	1e-04 (1e-8; 0.008)	1e-04 (1e-8; 0.008)	3e-04 (1e-8; 0.009)	3e-04 (1e-8; 0.009)	4e-04 (1e-8; 0.009)	4e-04 (1e-8; 0.009)	3e-04 (1e-8; 0.009)	4e-04 (1e-8; 0.009)	2e-04 (1e-8; 0.009)	5e-04 (1e-8; 0.009)	1e-04 (1e-8; 0.009)	2e-04 (1e-8; 0.009)	5e-04 (1e-8; 0.009)
<b>Feeding Mode</b>	-22.08 (-27.66; -18.12)	0.006 (5e-5; 0.06)	0.008 (5e-5; 0.07)	-	-	-	0.01 (2e-5; 0.09)	0.07 (0.002; 0.09)	-	-	-	0.006 (4e-6; 0.06)	0.01 (9e-5; 0.07)	-	-	-
<b>Feeding Habitat</b>	-35.02 (-39.19; -31.72)	0.08 (0.01; 0.09)	0.01 (0.003; 0.09)	-	-	-	0.001 (1e-6; 0.01)	0.01 (0.002; 0.03)	-	-	-	0.01 (0.002; 0.09)	0.09 (0.02; 0.09)	-	-	-
<b>Tooth Morphology</b>	-22.44 (-26.21; -18.56)	0.005 (1e-4; 0.009)	0.002 (1e-5; 0.009)	0.001 (1e-5; 0.009)	-	-	0.003 (8e-6; 0.009)	9e-4 (5e-6; 0.008)	8e-4 (8e-6; 0.008)	-	-	0.007 (9e-6; 0.009)	0.002 (1e-4; 0.009)	0.008 (0.001; 0.009)	-	-
<b>Schooling</b>	-52.48 (-55.78; -47.87)	0.03 (0.01; 0.14)	-	-	-	-	0.04 (0.02; 0.19)	-	-	-	-	-	-	-	-	-

**Supplementary Table 4.12 | Continuation.**

<b>Traits</b>	<b>q30</b>	<b>q31</b>	<b>q32</b>	<b>q34</b>	<b>q35</b>	<b>q40</b>	<b>q41</b>	<b>q42</b>	<b>q43</b>	<b>q45</b>	<b>q50</b>	<b>q51</b>	<b>q52</b>	<b>q53</b>	<b>q54</b>	<b>#Br</b>	<b>#Nd</b>
<b>Diet</b>	3e-04 (1e-8; 0.009)	4e-04 (1e-8; 0.009)	5e-04 (1e-8; 0.009)	3e-04 (1e-8; 0.009)	4e-04 (1e-8; 0.009)	2e-04 (1e-8; 0.009)	3e-04 (1e-8; 0.009)	3e-04 (1e-8; 0.009)	3e-04 (1e-8; 0.009)	0.008 (1e-8; 0.009)	2e-04 (1e-8; 0.009)	2e-04 (1e-8; 0.009)	4e-04 (1e-8; 0.009)	2e-04 (1e-8; 0.009)	9e-04 (1e-8; 0.009)	8 (2; 12)	3 (1; 7)
<b>Feeding Mode</b>	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	6 (1; 11)	3 (1; 6)
<b>Feeding Habitat</b>	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	7 (2; 12)	3 (0; 6)
<b>Tooth Morphology</b>	0.006 (0.001; 0.009)	0.001 (2e-5; 0.009)	0.006 (5e-4; 0.009)	-	-	-	-	-	-	-	-	-	-	-	-	7 (2; 12)	2 (0; 5)
<b>Schooling</b>	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	7 (2; 12)	4 (1; 7)

**Supplementary Table 4.13** | Summary of BayesTraits model results for the continuous trait (maximum length). Posterior Modal values (lower; upper - Highest Density Interval) of Lh: likelihood; Alpha: phylogenetic mean of the trait; Sigma: Brownian motion variance for the trait; #Br: number of branch lengths scaled using variable rates; #Nd: number of nodes scaled using variable rates.

<b>Taxa</b>	<b>Lh</b>	<b>Alpha</b>	<b>Sigma</b>	<b>#Br</b>	<b>#Nd</b>
<b>Siganiidae</b>	-85.83 (-89.84; -82.91)	34.44 (28.86; 40.85)	89.37 (18.78; 99.97)	2 (0; 4)	0 (0; 2)
<b>Acanthuridae</b>	-279.42 (-290.64; -268.81)	36.54 (26.82; 63.32)	87.22 (15.24; 99.97)	7 (3; 12)	3 (0; 6)
<b>Scarini</b>	-366.91 (-373.74; -356.92)	44.95 (15.46; 80.45)	29.76 (14.20; 47.13)	7 (3; 12)	4 (1; 7)

**Supplementary Table 4.14** | Summary of BayesTraits model results for ecosystem functions. Posterior Modal values (lower; upper - Highest Density Interval) of Lh: likelihood; q: transition rates between states; #Br: number of branch lengths scaled using variable rates; #Nd: number of nodes scaled using variable rates. States: Siganidae (0 - Macroalgae removal; 1 - Turf removal; 2 - Turf removal + Crevice cleaning; 3 - Turf removal + Crevice cleaning + Spongivory; 4 - Turf + Sediment removal); Acanthuridae (0 - Macroalgae removal; 1 - Sediment removal; 2 - Turf removal; 3 - Turf removal + Crevice cleaning; 4 - Turf + Sediment removal; 5 - Zooplanktivory; 6 - Zooplanktivory + Turf removal); Scarini (0 - Macroalgae removal; 1 - Macroalgae + Turf+Sediment removal; 2 - Macroalgae + Turf+Sediment removal + Spongivory; 3 - Turf + Sediment removal; 4 - Turf + Sediment removal + Bioerosion; 5 - Turf + Sediment removal + Bioerosion + Corallivory).

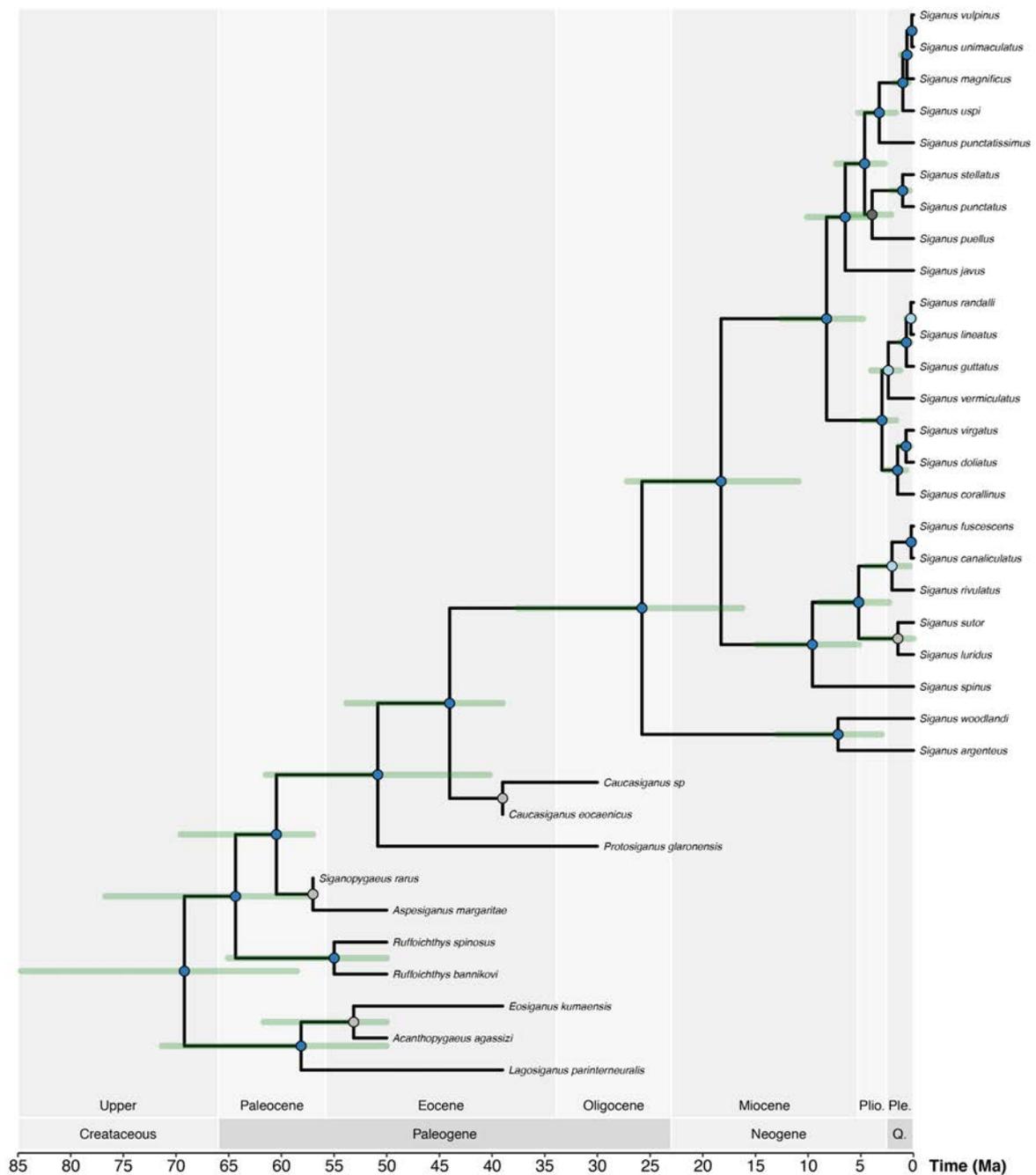
Taxa	Lh	q01	q02	q03	q04	q05	q06	q10	q12	q13	q14	q15	q16	q20	q21	q23	q24	q25	
<b>Siganidae</b>	-22.24 (-27.39; -18.90)	0.01 (1e-8; 0.01)	0.01 (1e-8; 0.01)	2e-4 (1e-8; 0.01)	0.01 (1e-8; 0.01)	-	-	0.01 (1e-8; 0.01)	0.01 (1e-8; 0.01)	2e-4 (1e-8; 0.01)	2e-4 (1e-8; 0.01)	-	-	0.01 (1e-8; 0.01)	2e-4 (1e-8; 0.01)	0.01 (1e-8; 0.01)	2e-4 (1e-8; 0.01)	-	
<b>Acanthuridae</b>	-50.73 (-56.92; -44.30)	0.002 (1e-8; 0.002)																	
<b>Scarini</b>	-58.69 (-66.58; -49.15)	0.006 (1e-8; 0.006)	2e-4 (1e-8; 0.006)	2e-4 (1e-8; 0.006)	1e-4 (1e-8; 0.006)	2e-4 (1e-8; 0.006)	-	3e-4 (1e-8; 0.006)	0.006 (1e-8; 0.006)	1e-4 (1e-8; 0.006)	1e-4 (1e-8; 0.006)	3e-4 (1e-8; 0.006)	-	1e-4 (1e-8; 0.006)	2e-4 (1e-8; 0.006)	2e-4 (1e-8; 0.006)	2e-4 (1e-8; 0.006)	2e-4 (1e-8; 0.006)	2e-4 (1e-8; 0.006)
Traits	q26	q30	q31	q32	q34	q35	q36	q40	q41	q42	q43	q45	q46	q50	q51	q52	q53	q54	
<b>Siganidae</b>	-	0.01 (1e-8; 0.01)	0.01 (1e-8; 0.01)	2e-4 (1e-8; 0.01)	0.01 (1e-8; 0.01)	-	-	0.01 (1e-8; 0.01)	3e-4 (1e-8; 0.01)	3e-4 (1e-8; 0.01)	2e-4 (1e-8; 0.01)	-	-	-	-	-	-	-	
<b>Acanthuridae</b>	0.002 (1e-8; 0.002)	0.002 (1e-8; 0.002)	0.002 (1e-8; 0.002)	0.002 (1e-8; 0.002)	0.002 (1e-8; 0.002)	0.002 (1e-8; 0.002)	0.002 (1e-8; 0.002)	0.002 (1e-8; 0.002)	0.002 (1e-8; 0.002)	0.002 (1e-8; 0.002)	0.002 (1e-8; 0.002)	0.002 (1e-8; 0.002)	0.002 (1e-8; 0.002)	0.002 (1e-8; 0.002)	0.002 (1e-8; 0.002)	0.002 (1e-8; 0.002)	0.002 (1e-8; 0.002)	0.002 (1e-8; 0.002)	
<b>Scarini</b>	-	2e-5 (1e-8; 0.006)	1e-5 (1e-8; 0.006)	2e-5 (1e-8; 0.006)	0.006 (1e-8; 0.006)	2e-5 (1e-8; 0.006)	-	1e-4 (1e-8; 0.006)	1e-4 (1e-8; 0.006)	1e-4 (1e-8; 0.006)	2e-4 (1e-8; 0.006)	0.006 (1e-8; 0.006)	-	2e-4 (1e-8; 0.006)	2e-4 (1e-8; 0.006)	2e-4 (1e-8; 0.006)	1e-4 (1e-8; 0.006)	3e-4 (1e-8; 0.006)	

**Supplementary Table 4.14** | Continuation.

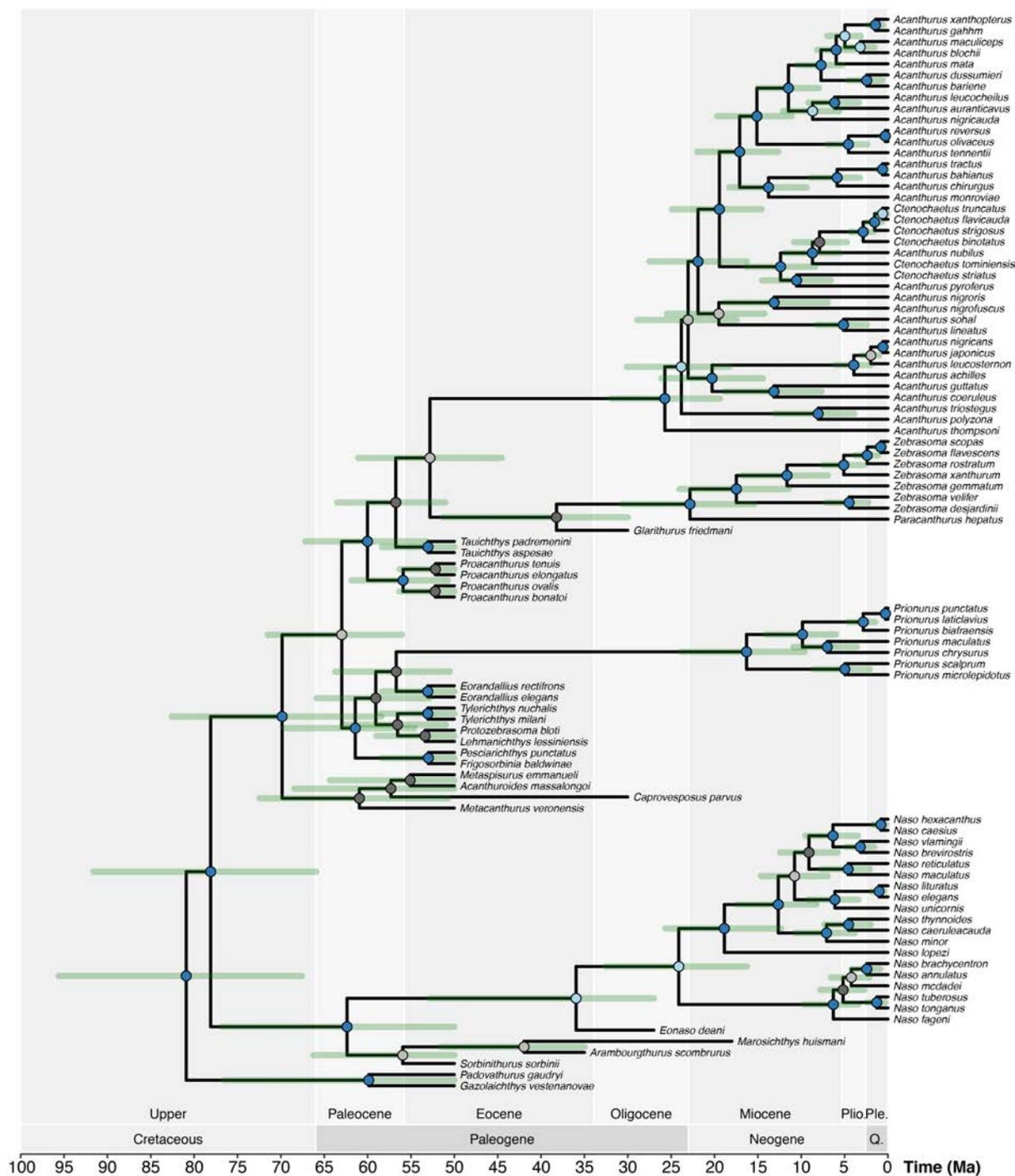
<b>Traits</b>	<b>q56</b>	<b>q60</b>	<b>q61</b>	<b>q62</b>	<b>q63</b>	<b>q64</b>	<b>q65</b>	<b>#Br</b>	<b>#Nd</b>
<b>Siganidae</b>	-	-	-	-	-	-	-	1 (0; 5)	1 (0; 3)
<b>Acanthuridae</b>	0.002 (1e-8; 0.002)	5 (2; 11)	3 (0; 5)						
<b>Scarini</b>	-	-	-	-	-	-	-	7 (3; 12)	4 (1; 7)

Appendix D.  
Supplementary Material to Chapter 5

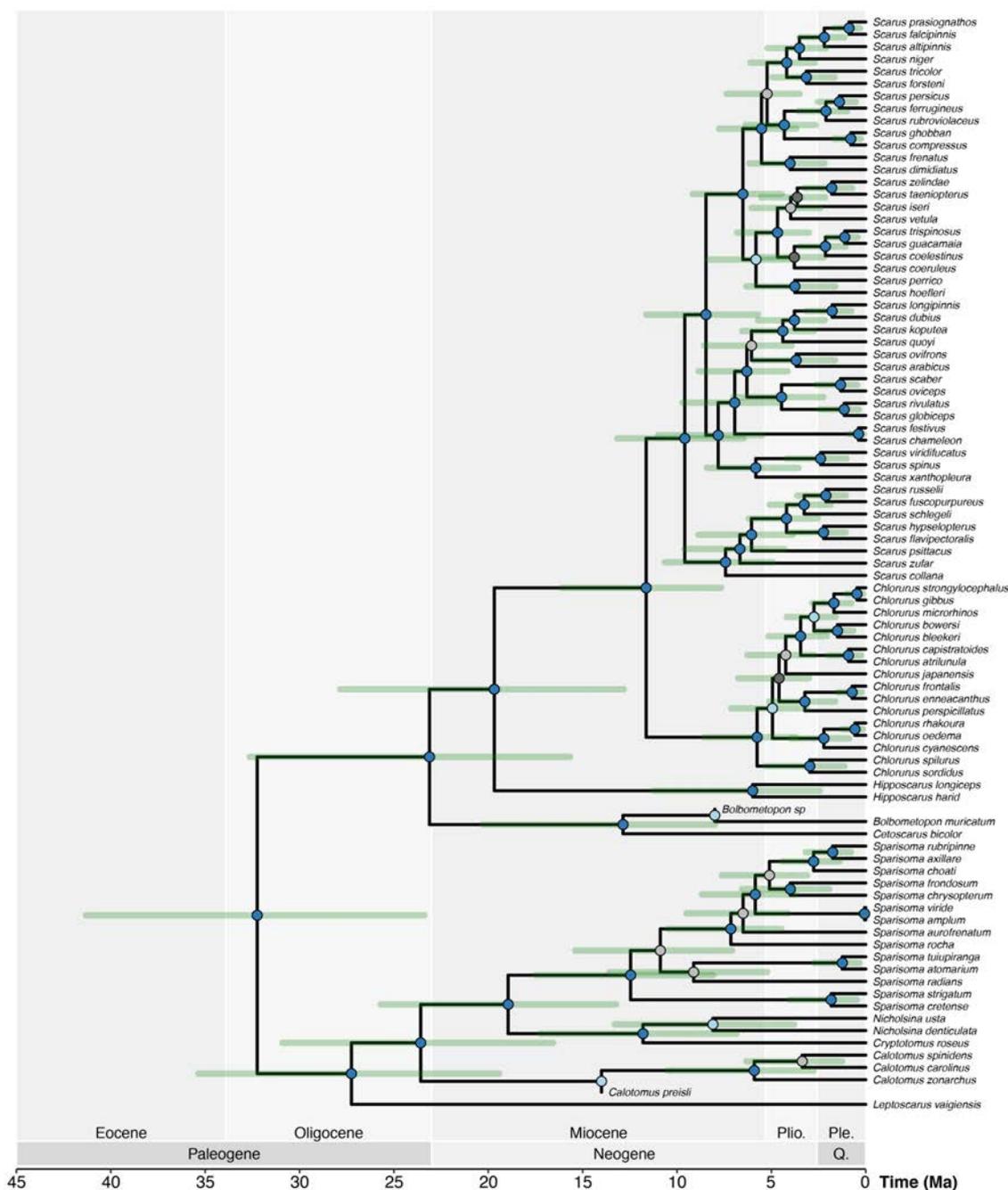
Supplementary Figures



**Supplementary Figure 5.1 | Maximum Clade Credibility (MCC) tree resulting from the Bayesian Inferences for Siganiidae.** Node colours represent posterior support values:  $\geq 0.9$  (dark blue);  $< 0.9 - \geq 0.7$  (light blue);  $< 0.7 - \geq 0.5$  (light grey);  $\leq 0.5$  (dark grey). Highest Posterior Density intervals for node ages are shown in green.



**Supplementary Figure 5.2 | Maximum Clade Credibility (MCC) tree resulting from the Bayesian Inferences for Acanthuridae.** Node colours represent posterior support values:  $\geq 0.9$  (dark blue);  $< 0.9 - \geq 0.7$  (light blue);  $< 0.7 - \geq 0.5$  (light grey);  $\leq 0.5$  (dark grey). Highest Posterior Density intervals for node ages are shown in green.



**Supplementary Figure 5.3 | Maximum Clade Credibility (MCC) tree resulting from the Bayesian Inferences for Scarini.** Node colours represent posterior support values:  $\geq 0.9$  (dark blue);  $< 0.9 - \geq 0.7$  (light blue);  $< 0.7 - \geq 0.5$  (light grey);  $\leq 0.5$  (dark grey). Highest Posterior Density intervals for node ages are shown in green.

## Supplementary Tables

**Supplementary Table 5.1** | List of fossils used in the phylogenetic analysis with respective ages, location and main reference.

<b>Siganidae</b>	<b>Epoch (Stage)</b>	<b>Location</b>	<b>Reference</b>
<i>Acanthopygaeus agassizi</i>	Eocene (Ypresian)	Italy	Bannikov <i>et al.</i> (2010)
<i>Aspesiganus margaritae</i>	Eocene (Ypresian)	Italy	Bannikov & Tyler (2002)
<i>Ruffoichthys bannikovi</i>	Eocene (Ypresian)	Italy	Tyler & Sorbini (1991)
<i>Ruffoichthys spinosus</i>	Eocene (Ypresian)	Italy	Tyler & Bannikov (1997)
<i>Eosiganus kumaensis</i>	Eocene (Bartonian)	Russia (Caucasus)	Tyler & Bannikov (1997)
<i>Lagosiganus parinterneurialis</i>	Eocene (Bartonian)	Russia (Caucasus)	Bannikov <i>et al.</i> (2010)
<i>Siganopygaeus rarus</i>	Paleocene (Thanetian)	Turkmenistan	Tyler & Bannikov (1997)
<i>Protosiganus glaronensis</i>	Oligocene (Rupelian)	Switzerland	Tyler & Bannikov (1997)
<i>Caucasiganus eocaenicus</i>	Eocene (Bartonian)	Russia (Caucasus)	Bannikov <i>et al.</i> (2010)
<i>Caucasiganus sp.</i>	Oligocene (Rupelian)	Iran	Bannikov <i>et al.</i> (2010)
<b>Acanthuridae</b>	<b>Epoch (Stage)</b>	<b>Location</b>	<b>Reference</b>
<i>Acanthuroides massalongoi</i>	Eocene (Ypresian)	Italy	Blot & Tyler (1990)
<i>Eorandallius elegans</i>	Eocene (Ypresian)	Italy	Blot & Tyler (1990)
<i>Eorandallius rectifrons</i>	Eocene (Ypresian)	Italy	Blot & Tyler (1990)
<i>Frigosorbina baldwinae</i>	Eocene (Ypresian)	Italy	Bannikov & Tyler (2012)
<i>Gazolaichthys vestenanovae</i>	Eocene (Ypresian)	Italy	Blot & Tyler (1990)
<i>Lehmanichthys lessiniensis</i>	Eocene (Ypresian)	Italy	Blot & Tyler (1990)
<i>Metacanthurus veronensis</i>	Eocene (Ypresian)	Italy	Blot & Tyler (1990)
<i>Metaspisurus emmanueli</i>	Eocene (Ypresian)	Italy	Blot & Tyler (1990)
<i>Padovathurus gaudryi</i>	Eocene (Ypresian)	Italy	Tyler (2005)
<i>Pesciarichthys punctatus</i>	Eocene (Ypresian)	Italy	Bannikov & Tyler (2012)
<i>Proacanthurus bonatoi</i>	Eocene (Ypresian)	Italy	Blot & Tyler (1990)
<i>Proacanthurus elongatus</i>	Eocene (Ypresian)	Italy	Blot & Tyler (1990)
<i>Proacanthurus ovalis</i>	Eocene (Ypresian)	Italy	Blot & Tyler (1990)
<i>Proacanthurus tenuis</i>	Eocene (Ypresian)	Italy	Blot & Tyler (1990)
<i>Protozebrasoma bloti</i>	Eocene (Ypresian)	Italy	Sorbini & Tyler (1998)
<i>Tylerichthys milani</i>	Eocene (Ypresian)	Italy	Blot & Tyler (1990)
<i>Tylerichthys nuchalis</i>	Eocene (Ypresian)	Italy	Blot & Tyler (1990)
<i>Sorbinihthys sorbinii</i>	Eocene (Ypresian)	Italy	Tyler (1999)
<i>Tauichthys padremenini</i>	Eocene (Ypresian)	Italy	Tyler (1999)
<i>Tauichthys aspesae</i>	Eocene (Ypresian)	Italy	Tyler & Bannikov (2000)
<i>Caprovesposus parvus</i>	Oligocene (Rupelian)	Russia (Caucasus)	Bannikov & Tyler (1992)
<i>Arambourgthurus scombrurus</i>	Oligocene (Rupelian)	Iran	Tyler (2000)
<i>Marosichthys huismani</i>	Miocene (Burdigalian)	Indonesia	Tyler (1997)
<i>Glarithurus friedmani</i>	Oligocene (Rupelian)	Switzerland	Tyler & Micklich (2011)
<i>Eonaso deani</i>	Oligocene (?)	Antigua and Barbuda	Tyler & Sorbini (1998)
<b>Scarini</b>	<b>Epoch (Stage)</b>	<b>Location</b>	<b>Reference</b>
<i>Bolbometopon sp.</i>	Miocene (Tortonian)	Sri Lanka	Bellwood & Schultz (1991)
<i>Calotomus preisli</i>	Miocene (Langhian)	Austria	Bellwood & Schultz (1991)
<i>Pacuarescarus kussmauli</i>	Miocene (Burdigalian)	Costa Rica	Laurito <i>et al.</i> (2014)

**Supplementary Table 5.2** | Matrices of areas allowed per time-slice (90-12 Ma; 12-3.1 Ma; 3.1-0 Ma) used for ancestral range reconstructions in BioGeoBEARS.

3.1-0 Ma	CIP	CP	WI	TEP	WA	EA	Tet
CIP	1	1	1	1	0	0	0
CP	1	1	1	1	0	0	0
WI	1	1	1	1	0	0	0
TEP	1	1	1	1	0	0	0
WA	0	0	0	0	1	1	0
EA	0	0	0	0	1	1	0
Tet	0	0	0	0	0	0	0
12-3.1 Ma	CIP	CP	WI	TEP	WA	EA	Tet
CIP	1	1	1	1	0	0	0
CP	1	1	1	1	0	0	0
WI	1	1	1	1	0	0	0
TEP	1	1	1	1	1	1	0
WA	0	0	0	1	1	1	0
EA	0	0	0	1	1	1	0
Tet	0	0	0	0	0	0	0
90-12 Ma	CIP	CP	WI	TEP	WA	EA	Tet
CIP	1	1	1	1	1	1	1
CP	1	1	1	1	1	1	1
WI	1	1	1	1	1	1	1
TEP	1	1	1	1	1	1	1
WA	1	1	1	1	1	1	1
EA	1	1	1	1	1	1	1
Tet	1	1	1	1	1	1	1

**Supplementary Table 5.3** | Matrices of dispersal multipliers per time-slice (90-12 Ma; 12-3.1 Ma; 3.1-0 Ma) used for ancestral range reconstructions in BioGeoBEARS.

3.1-0 Ma	CIP	CP	WI	TEP	WA	EA	Tet
CIP	1	1	1	0	0	0	0
CP	1	1	0	0.05	0	0	0
WI	1	0	1	0	0	0.05	0
TEP	0	0.05	0	1	0	0	0
WA	0	0	0	0	1	1	0
EA	0	0	0.05	0	1	1	0
Tet	0	0	0	0	0	0	0
12-3.1 Ma	CIP	CP	WI	TEP	WA	EA	Tet
CIP	1	1	1	0	0	0	0
CP	1	1	0	0.05	0	0	0
WI	1	0	1	0	0	0.05	0
TEP	0	0.05	0	1	1	0	0
WA	0	0	0	1	1	1	0
EA	0	0	0.05	0	1	1	0
Tet	0	0	0	0	0	0	0
90-12 Ma	CIP	CP	WI	TEP	WA	EA	Tet
CIP	1	1	1	0	0	0	1
CP	1	1	0	0.05	0	0	0
WI	1	0	1	0	0	0.05	1
TEP	0	0.05	0	1	1	0	0
WA	0	0	0	1	1	1	1
EA	0	0	0.05	0	1	1	1
Tet	1	0	1	0	1	1	1

**Supplementary Table 5.4** | Biogeographical model comparison for Acanthuridae ordered according to lowest AICc. LnL: log-likelihood; d: dispersal; e: extinction; j: founder-speciation event parameter; w: dispersal multiplier matrix power exponential parameter; AICc: Akaike Information Criteria corrected for sample size; AIC\_wt: Akaike Information Criteria weight.

Models	LnL	params	d	e	j	w	AICc	AICc_wt
BAYAREALIKE+J+W	-263.6	4	0.013	0.022	0.040	0.034	535.6	1.00
BAYAREALIKE+J	-274.6	3	0.016	0.024	0.064	1	555.4	5.1e-05
BAYAREALIKE+W	-285.5	3	0.019	0.036	0	0.039	577.2	9.4e-10
BAYAREALIKE	-295.4	2	0.024	0.039	0	1	595	1.3e-13
DEC+W	-299.8	3	0.041	0.0098	0	0.025	605.9	5.5e-16
DEC+J+W	-299.7	4	0.042	0.010	1.0e-05	0.089	607.8	2.1e-16
DEC	-313.8	2	0.049	0.011	0	1	631.8	1.3e-21
DEC+J	-313.8	3	0.049	0.011	1.0e-05	1	633.9	4.5e-22
DIVALIKE+W	-321.5	3	0.047	0.012	0	0.011	649.3	2.1e-25
DIVALIKE+J+W	-321.1	4	0.049	0.011	1.0e-05	0.099	650.6	1.1e-25
DIVALIKE	-334.1	2	0.062	0.015	0	1	672.4	2.0e-30
DIVALIKE+J	-334.1	3	0.062	0.015	1.0e-05	1	674.5	6.9e-31

**Supplementary Table 5.5** | Biogeographical model comparison for Siganidae ordered according to lowest AICc. LnL: log-likelihood; d: dispersal; e: extinction; j: founder-speciation event parameter; w: dispersal multiplier matrix power exponential parameter; AICc: Akaike Information Criteria corrected for sample size; AIC\_wt: Akaike Information Criteria weight.

Models	LnL	params	d	e	j	w	AICc	AICc_wt
DEC	-82.8	2	0.025	1.0e-12	0	1	170	0.56
DEC+J	-82.8	3	0.025	9.8e-09	1.0e-05	1	172.4	0.17
BAYAREALIKE+J	-83.52	3	0.014	0.024	0.037	1	173.8	0.081
BAYAREALIKE	-85	2	0.015	0.041	0	1	174.4	0.062
DEC+W	-83.84	3	0.023	1.0e-12	0	0.17	174.5	0.059
DEC+J+W	-82.75	4	0.025	1.0e-12	1.0e-05	1.21	174.9	0.048
BAYAREALIKE+J+W	-83.89	4	0.014	0.025	0.031	0.51	177.2	0.015
BAYAREALIKE+W	-86.16	3	0.013	0.040	0	0.061	179.1	0.0058
DIVALIKE	-91.26	2	0.029	1.0e-12	0	1	186.9	0.0001
DIVALIKE+J	-91.26	3	0.029	1.0e-12	1.0e-05	1	189.3	3.5e-05
DIVALIKE+W	-92.13	3	0.028	1.0e-12	0	0.20	191.1	1.5e-05
DIVALIKE+J+W	-91.22	4	0.029	1.0e-12	1.0e-05	1.19	191.8	1.0e-05

**Supplementary Table 5.6** | Biogeographical model comparison for Scarini ordered according to lowest AICc. LnL: log-likelihood; d: dispersal; e: extinction; j: founder-speciation event parameter; w: dispersal multiplier matrix power exponential parameter; AICc: Akaike Information Criteria corrected for sample size; AIC\_wt: Akaike Information Criteria weight.

Models	LnL	params	d	e	j	w	AICc	AICc wt
DEC+J+W	-208.3	4	0.090	0.010	0.063	0.51	425.1	0.67
DEC+J	-210.4	3	0.095	0.011	0.068	1	427.1	0.25
BAYAREALIKE+J+W	-210.8	4	0.051	0.019	0.21	0.46	430.1	0.057
BAYAREALIKE+J	-213.3	3	0.055	0.025	0.21	1	432.8	0.015
DEC+W	-213.9	3	0.091	0.018	0	0.026	434.2	0.0074
DEC	-217.8	2	0.063	0.0094	0	1	439.7	0.0005
DIVALIKE+J+W	-217.1	4	0.10	0.0065	0.041	0.35	442.8	1.0e-04
DIVALIKE+J	-220.3	3	0.11	0.0075	0.089	1	446.8	1.3e-05
DIVALIKE	-222.3	2	0.13	0.014	0	1	448.7	5.3e-06
DIVALIKE+W	-222.4	3	0.10	0.014	0	0.045	451	1.6e-06
BAYAREALIKE	-232.8	2	0.14	0.12	0	1	469.8	1.4e-10
BAYAREALIKE+W	-238.8	3	0.084	0.11	0	0.0059	483.9	1.2e-13

## Supplementary References

- Bannikov, A.F. & Tyler, J.C. (1992) *Caprovesposus* from the Oligocene of Russia: The pelagic acronurus presettlement stage of a surgeonfish (Teleostei: Acanthuridae). *Proceedings of the Biological Society of Washington*, 105, 810–820.
- Bannikov, A.F. & Tyler, J.C. (2002) A new genus and species of rabbitfish (Acanthuroidei: Siganidae) from the Eocene of Monte Bolca, Italy. *Studi e Ricerche sui Giacimenti Terziari di Bolca, Museo Civico di Storia Naturale di Verona, Miscellanea Paleontologica*, 9, 37–45.
- Bannikov, A.F. & Tyler, J.C. (2012) Redescription of the Eocene of Monte Bolca, Italy, surgeon fish *Pesciarichthys punctatus* (Perciformes, Acanthuridae), and a new genus, *Frigosorbinia*, for *P. baldwinae*. *Studi e Ricerche sui Giacimenti Terziari di Bolca, Museo Civico di Storia Naturale di Verona, Miscellanea Paleontologica*, 15, 19–30.
- Bannikov, A.F., Tyler, J.C., & Sorbini, C. (2010) Two new taxa of Eocene rabbitfishes (Perciformes, Siganidae) from the North Caucasus (Russia), with redescription of *Acanthopygaeus agassizi* (Eastman) from Monte Bolca (Italy) and a phylogenetic analysis of the family. *Bollettino del Museo Civico di Storia Naturale di Verona, Geologia, Paleontologia, Preistoria*, 34, 3–21.
- Bellwood, D.R. & Schultz, O. (1991) A Review of the fossil record of the Parrotfishes (Labroidei: Scaridae) with a description of a new *Calotomus* species from the Middle Miocene (Badenian) of Austria. *Annalen Des Naturhistorischen Museums in Wien. Serie A, Fur Mineralogie und Petrographie, Geologie Und Palaontologie, Anthropologie und Prahistorie*, 92, 55–71.

- Blot, J. & Tyler, J.C. (1990) New genera and species of fossil surgeon fishes and their relatives (Acanthuroidei, Teleostei) from the Eocene of Monte Bolca, Italy, with application of the Blot formula to both fossil and recent forms. *Studi e ricerche sui giacimenti terziari di Bolca*, 6, 13–92.
- Laurito, C.A., Calvo, C., Valerio, A.L., Calvo, A., & Chacón, R. (2014) Lower Miocene Ichthyofauna from the locality of Pacuare de Tres Equis, Río Banano formation, Cartago province, Costa Rica and description of a new genus and species of Scaridae. *Revista Geológica de América Central*, 50, 153–192.
- Sorbini, L. & Tyler, J.C. (1998) A new genus and species of Eocene surgeon fish (Acanthuridae) from Monte Bolca, Italy, with similarities to the Recent *Zebrosoma*. *Studi e Ricerche sui Giacimenti Terziari di Bolca*, 7, 7–19.
- Tyler, J.C. (1997) The Miocene fish *Marosichthys*, a putative Tetraodontiform, actually a perciform surgeon fish (Acanthuridae) related to the recent *Naso. Beaufortia*, 47, 1–10.
- Tyler, J.C. (1999) A new genus and species of surgeon fish (Acanthuridae) with four dorsal-fin spines from the Eocene of Monte Bolca, Italy. *Studi e Ricerche sui Giacimenti Terziari di Bolca*, 8, 257–268.
- Tyler, J.C. (2000) *Arambourgthurus*, a new genus of hypurostegic surgeonfish (Acanthuridae) from the Oligocene of Iran, with a phylogeny of the Nasinae. *Geodiversitas*, 22, 525–537.
- Tyler, J.C. (2005) A new genus for the surgeon fish *Acanthurus gaudryi* De Zigno 1887 from the Eocene of Monte Bolca, Italy, a morphologically primitive basal taxon of Acanthuridae (Acanthuroidea, Perciformes). *Studi e Ricerche sui Giacimenti Terziari di Bolca*, 11, 149–163.
- Tyler, J.C. & Bannikov, A.F. (1997) Relationships of the fossil and recent genera of rabbitfishes (Acanthuroidei: Siganidae). *Smithsonian Contributions to Paleobiology*, 84, 1–35.
- Tyler, J.C. & Bannikov, A.F. (2000) A new species of the surgeon fish genus *Tauichthys* from the Eocene of Monte Bolca, Italy (Perciformes, Acanthuridae). *Bollettino del Museo Civico di Storia Naturale di Verona*, 24, 29–36.
- Tyler, J.C. & Micklich, N.R. (2011) A new genus and species of surgeon fish (Perciformes, Acanthuridae) from the Oligocene of Kanton Glarus, Switzerland. *Swiss Journal of Palaeontology*, 130, 203–216.
- Tyler, J.C. & Sorbini, L. (1991) A new species of the primitive Eocene rabbitfish *Ruffoichthys*, with the genus redefined relative to the Recent forms and to the other fossil genera (Siganidae). *Studi e Ricerche sui Giacimenti Terziari di Bolca, Museo Civico di Storia Naturale di Verona, Miscellanea Paleontologica*, 6, 93–113.
- Tyler, J.C. & Sorbini, L. (1998) On the relationships of *Eonaso*, an Antillean fossil surgeon fish (Acanthuridae). *Studi e Ricerche sui Giacimenti Terziari di Bolca*, 7, 35–42.

## Appendix E. Publications during candidature

### Publications arising from thesis chapters

**Siqueira, A. C.**, Morais, R. A., Bellwood, D. R., Cowman, P. F. (2020). Trophic innovations fuel reef fish diversification. *Nature Communications*, *11*, 2669.

doi: 10.1038/s41467-020-16498-w

(Chapter 2)

**Siqueira, A. C.**, Bellwood, D. R., Cowman, P. F. (2019). The evolution of traits and functions in herbivorous coral reef fishes through space and time. *Proceedings of the Royal Society B: Biological Sciences*, *286*, 20182672.

doi: 10.1098/rspb.2018.2672

(Chapter 4)

**Siqueira, A. C.**, Bellwood, D. R., Cowman, P. F. (2019). Historical biogeography of herbivorous coral reef fishes: The formation of an Atlantic fauna. *Journal of Biogeography*, *46*, 1611–1624.

doi: 10.25903/5cd265eb0a405

(Chapter 5)

Under review:

**Siqueira, A. C.**, Morais, R. A., Bellwood, D. R., Cowman, P. F. Planktivores as trophic drivers of global coral reef fish diversity patterns. *Under review for PNAS*

(Chapter 3)

## Other peer-reviewed papers published during candidature

- Nunes, L. T., **Siqueira, A. C.**, Cord, I., Ford, B. M., Liedke, A. M. R., Ferreira, C. E. L., Floeter, S. R. (2020). The influence of species abundance, diet and phylogenetic affinity on the co-occurrence of butterflyfishes. *Marine Biology*, 167, 107.  
doi: 10.1007/s00227-020-03725-7
- Bellwood, D. R., Schultz, O., **Siqueira, A. C.**, Cowman, P. F. (2019). A review of the fossil record of the Labridae. *Annalen des Naturhistorischen Museums in Wien, Serie A*, 121, 125–193.
- Floeter, S. R., Bender, M. G., **Siqueira, A. C.**, Cowman, P. F. (2018). Phylogenetic perspectives on reef fish functional traits. *Biological Reviews*, 93, 131–151.  
doi: 10.1111/brv.12336.
- Maia, H. A., Morais, R. A., **Siqueira, A. C.**, Hanazaki, N., Floeter, S. R., Bender, M. G. (2018). Shifting baselines among traditional fishers in São Tomé and Príncipe islands, Gulf of Guinea. *Ocean and Coastal Management*, 154, 133–142.  
doi: 10.1016/j.ocecoaman.2018.01.006.
- Siqueira, A. C.**, Quimbayo, J. P. A., Cantor, M., Silveira, R., Daura-Jorge, F. G. (2017). Estimating population parameters of longsnout seahorses, *Hippocampus reidi* (Teleostei: Syngnathidae) through mark-recapture. *Neotropical Ichthyology*, 15, e170067.  
doi: 10.1590/1982-0224-20170067.